

Haushaltsbefragungen vom KONTIV-Typ: Hochrechnung - Gewichtung - Genauigkeitsbeurteilung*)

VON HEINZ HAUTZINGER

Inhaltsübersicht

Vorwort

1. Haushaltsbefragungen vom KONTIV-Typ
 - 1.1 Grundlagen
 - 1.2 Charakterisierung der Haushaltsbefragungen vom KONTIV-Typ
 - 1.3 Alternative Stichprobendesigns
 - 1.4 Grundprinzipien der Hochrechnung und Genauigkeitsbeurteilung
 - 1.4.1 Design „Unabhängige Teilstichproben“
 - 1.4.2 Design „Partielle Rotation“
2. Erhebungsfehler und ihre Berücksichtigung bei der Hochrechnung
 - 2.1 Fehlertypologie
 - 2.1.1 Stichprobenbedingte Fehler
 - 2.1.2 Verfahrensbedingte Fehler
 - 2.2 Verfahrensbedingte Fehler bei Haushaltsbefragungen vom KONTIV-Typ
 - 2.2.1 Nichtübereinstimmung von Zielkollektiv und Auswahlgesamtheit
 - 2.2.2 Eingeschränkte Objektivität, Reliabilität und Validität des Meßverfahrens
 - 2.2.3 Nichterreichbarkeit eines Teils der Stichprobenhaushalte
 - 2.2.4 Nichtteilnahme (Antwortverweigerung) eines Teils der Stichprobenhaushalte
 - 2.2.5 Angabefehler der Befragten

Anschrift des Verfassers:

Prof. Dr. Heinz Hautzinger
Institut für angewandte Verkehrs-
und Tourismusforschung e. V.
an der Fachhochschule Heilbronn
Max-Planck-Str. 39
7100 Heilbronn

*) Die vorliegende Arbeit wurde wesentlich gefördert durch die intensiven und fruchtbaren Diskussionen, die mit Unterstützung durch das Bundesverkehrsministerium im Rahmen von zwei speziellen Fachgesprächen geführt worden sind. Teilnehmer an diesen Fachgesprächen waren die Herren Dr. J. Grevsmühl (Bundesverkehrsministerium), Dr. R. Herz (Universität Karlsruhe), Dr. M. Herry (Innovaplan, München), U. Holz (Technische Universität Berlin), H. Kuhfeld (DIW, Berlin), Prof. Dr.-Ing. E. Kutter (Technische Universität Berlin) sowie Frau B. Tassaux (IVT Heilbronn) und der Verfasser.

Der Gedankenaustausch in dieser Diskussionsrunde hat entscheidend zur Vertiefung des Problemverständnisses und damit zur Ausgewogenheit der Darstellung beigetragen. Für den Inhalt der Arbeit zeichnet allerdings allein der Autor verantwortlich.

- 2.3 Gewichtung als Mittel zur Fehlerreduzierung bzw. Fehlerkorrektur
3. Maßnahmen zur Reduzierung des Stichprobenfehlers
 - 3.1 Gebundene Hochrechnung: Regressions- und Verhältnisschätzung
 - 3.2 Nachträgliche Schichtung (Gewichtung im engeren Sinne)
4. Korrektur von ausfallbedingten Verzerrungen
 - 4.1 Verzerrungskorrektur durch nachträgliche Schichtung
 - 4.2 Verzerrungskorrektur auf der Basis von Informationen über das Verhalten von Nichtantwortern
 - 4.3 Extrapolationsmethoden zur Verzerrungskorrektur: Darstellung und Kritik
5. Korrektur von Verzerrungen durch Angabefehler
 - 5.1 Antwortvariabilität ohne Antwortverzerrung
 - 5.2 Antwortverzerrung: Korrektur auf der Basis von Nacherhebungen
 - 5.3 Antwortverzerrung: Korrektur auf der Basis exogener Daten
 - 5.4 Verzerrungskorrektur bei Angabefehlern bezüglich der Zahl der Wege
 - 5.5 Verzerrungskorrektur bei Angabefehlern bezüglich Wegelänge und Wegedauer
 - 5.6 Verzerrungskorrektur bei der Schätzung von verkehrsmittelspezifischen Verkehrsleistungswerten
6. Genauigkeitsbeurteilung bei ungewichteten Daten
 - 6.1 Varianzschätzung unter dem Design „Unabhängige Teilstichproben“
 - 6.2 Praktische Beispiele für die Berechnung von Konfidenzintervallen
 - 6.3 Quantifizierung des Design-Effekts
7. Genauigkeitsbeurteilung bei gewichteten Daten
 - 7.1 Einfluß der Gewichtung auf die Unverzerrtheit der Schätzung
 - 7.2 Einfluß der Gewichtung auf den Standardfehler der Schätzung
 - 7.3 Vereinfachte Fehlerrechnung bei komplexen Stichprobendesigns
 - 7.4 Ein Anwendungsbeispiel
8. Resümee und Ausblick

1. Haushaltsbefragungen vom KONTIV-Typ

1.1 Grundlagen

Haushaltsbefragungen zum Verkehrsverhalten haben sich im Verlauf der letzten Jahre zu dem vielleicht wichtigsten methodischen Instrument der empirischen Verkehrsforschung entwickelt. Auch für Verkehrsplanung und Verkehrspolitik sind Erhebungen dieser Art unverzichtbare Hilfsmittel zur Maßnahmenentwicklung und Erfolgskontrolle geworden. Soweit es um die Erfassung des tatsächlichen Verkehrsverhaltens von Individuen an einem bestimmten Stichtag geht, haben sich - insbesondere im Fall von Massenstichproben - schriftliche Haushaltsbefragungen in Tagebuchform eindeutig durchgesetzt.

Die meisten der in den letzten Jahren durchgeführten schriftlichen Haushaltsbefragungen zum Verkehrsverhalten orientieren sich am Design der „Kontinuierlichen Erhebung des Verkehrsverhaltens“ (KONTIV), die erstmals im Zeitraum 1975/76 stattfand und in den Jahren 1982 und 1989 wiederholt wurde (*Sozialforschung Brög*, 1975 und *Socialdata*, 1984). Die KONTIV hatte im In- und Ausland einen so nachhaltigen Einfluß auf die Methoden der empirischen Verkehrsforschung, daß die überwiegende Mehrzahl der seit Mitte der siebziger Jahre durchgeführten Haushaltsbefragungen zum Verkehrsverhalten als „Haushaltsbefragungen vom KONTIV-Typ“ apostrophiert werden können.

Anfänglich wurden Haushaltsbefragungen vom KONTIV-Typ ganz überwiegend aus dem Blickwinkel der empirischen Sozialforschung betrachtet, wobei Fragen der Genauigkeit der Merkmalerfassung, also die Validität und Reliabilität der Erfassung des Verkehrsverhaltens, im Vordergrund des Interesses standen. In dem Maße wie das Befragungsinstrument verfeinert und standardisiert wurde, trat dann die Beschäftigung mit Haushaltsbefragungen vom KONTIV-Typ unter mehr statistischen Gesichtspunkten (Stichprobenplanung und Hochrechnung) hinzu. Inzwischen haben beide Aspekte, nämlich Meßmethodik und Stichprobenverfahren, eine so breite und tiefgehende Analyse erfahren, daß eine zusammenfassende Beurteilung des erreichten Stands der Wissenschaft notwendig und hilfreich ist.

Mit der vorliegenden Arbeit wird versucht, für die Bereiche

- Hochrechnung,
- Gewichtung und
- Genauigkeitsbeurteilung

bei Haushaltsbefragungen vom KONTIV-Typ einen Überblick über die wichtigsten theoretischen Entwicklungen und praktischen Erfahrungen zu geben. Dabei wird es sich allerdings herausstellen, daß - auch in wichtigen Fragen - vorliegende methodische Ansätze, vor allem wegen mangelnder empirischer Erkenntnisse, noch nicht abschließend beurteilt werden können.

Die Arbeit konzentriert sich auf die Auswertung von Haushaltsbefragungen vom KONTIV-Typ, Probleme der Stichprobenplanung werden allenfalls gestreift. Gänzlich ausgeklammert bleiben die Fragen der Meßmethodik. Letzteres bedeutet auch, daß Alternativen zum KONTIV-Design, z. B. kombinierte telefonisch-schriftliche Befragungen, nicht explizit behandelt werden.

1.2 Charakterisierung der Haushaltsbefragungen vom KONTIV-Typ

Im folgenden wird zunächst eine kurze Charakterisierung der wesentlichen Elemente des KONTIV-Designs gegeben. Die Darstellung lehnt sich an *Socialdata* (1984) an.

Die Auswahleinheit der Befragung ist der (private) Haushalt; im Haushalt wird jede Person ab einem bestimmten Mindestalter befragt. Dabei wird die Altersgrenze in erster Linie durch den Anspruch, daß jede Person den Fragebogen selbst ausfüllen soll, bestimmt. Ziel des *Stichprobenplans* ist es eigentlich, Adressen von Haushalten so auszuwählen, daß jeder Haushalt die gleiche Auswahlchance besitzt. Praktisch ist dies allerdings meist nicht möglich, da es auf nationaler Ebene keine gültige Datei von Haushaltsadressen als Auswahlgrundlage gibt. Es wird deshalb in der Regel eine mehrstufige geschichtete Zufallsstichprobe gezogen, bei der zunächst Gemeinden ausgewählt und innerhalb dieser Gemeinden dann Adressen aus Adreßbüchern herausgegriffen oder per Random-Route ermittelt werden.

Jedem ausgewählten Haushalt wird dann nach einem Zufallsverfahren ein bestimmter *individueller Berichtszeitraum* (häufig ein Stichtag) innerhalb des gesamten Untersuchungszeitraumes (häufig ein Jahr) zugeordnet. Dieses Zuordnungsverfahren ist in der Regel so angelegt, daß die individuellen Berichtstage über den gesamten Untersuchungszeitraum verteilt sind. Die einzelnen Tage des Untersuchungszeitraums müssen dabei nicht unbedingt gleichmäßig mit individuellen Berichtstagen besetzt sein.

Ein Hauptanliegen des KONTIV-Designs ist die *Motivation der Zielpersonen*. Diese erfolgt durch ein ganzes Bündel von Maßnahmen, deren hauptsächlicher Zweck es ist, den Befragten anzusprechen und aufzuklären. Ergänzend hierzu wird vom durchführenden Institut ein Betreuungsservice eingerichtet, um dem Befragten Rückfragen zu ermöglichen.

Mit dem KONTIV-Fragebogen werden drei verschiedene Arten von Informationen erhoben:

- Merkmale des Haushalts,
- soziodemographische Merkmale der Zielpersonen,
- Merkmale des Verkehrsverhaltens der Zielpersonen am Befragungsstichtag (Berichtszeitraum).

Um eine hohe Antwortbereitschaft zu erzielen, wird der Fragebogen bewußt befragtenfreundlich gestaltet. Zur Erreichung einer hohen Antwortgenauigkeit wird versucht, den Fragebogen selbsterklärend zu gestalten. Üblicherweise werden Haushalts- und Personenfragebogen getrennt; die Personen- und die Wegemerkmale werden in Spalten (je Person bzw. je Weg) erhoben. Der individuelle Berichtszeitraum liegt beim KONTIV-Design in der Regel nicht in der Vergangenheit (recall-Technik), er wird dem zu befragenden Haushalt vielmehr im voraus mitgeteilt. Ausnahmen hiervon sind allerdings möglich, insbesondere dann, wenn es um die Erfassung seltener Ereignisse (z. B. Urlaubsreisen) geht.

Im Rahmen der *Feldorganisation* werden mehrere Versandaktionen festgelegt:

- Ankündigung der Erhebung per Postkarte
- Hauptversand der Erhebungsunterlagen
- Erste Erinnerung per Postkarte
- Zweite Erinnerung per Postkarte

- Neuversand der Erhebungsunterlagen
- Dritte Erinnerung per Postkarte.

Dieses Versandsystem hat zum Ziel, gleichzeitig hohe Ausschöpfungsraten und auf *aktuelle* Stichtage bezogene Angaben zum Verkehrsverhalten zu erhalten.

Vor Übernahme der Fragebogen auf Datenträger ist auch beim KONTIV-Design eine *Datenaufbereitung* (Signierung) erforderlich. Diese beinhaltet eine Sichtkontrolle, Verschlüsselung offener Eintragungen, Überprüfung der Wegeketten mit gegebenenfalls plausiblen Korrekturen sowie eine Komplettierung und Konfektionierung der Fragebogen für die Direkterfassung. Die *Datenerfassung* sollte eine formale Prüfung der vorkommenden Zeichen einschließen.

Die so erstellte Rohdaten-Datei wird schließlich durch Gewichtung in die Auswertungsdatei umgesetzt. Üblicherweise entsteht so eine Haushalts-, Personen- und Wegedatei. Jedem Datensatz wird hierbei der Wert einer Gewichtungsvariablen hinzugefügt. Bei der Auswertung von qualitativen Merkmalen werden dann nicht Fälle (Haushalte bzw. Personen bzw. Wege) gezählt, sondern Summen von Gewichten gebildet. Entsprechend werden bei quantitativen Merkmalen vor der Summation der Merkmalswerte diese mit dem Wert der Gewichtungsvariablen multipliziert. Diese Gewichtung ist erforderlich, da

- nicht alle Haushalte die gleiche Auswahlchance besitzen,
- die einzelnen Tage des Untersuchungszeitraums nicht unbedingt gleichmäßig stark besetzt sind,
- zufallsbedingt die Stichprobenverteilungen wichtiger soziodemographischer Merkmale von den entsprechenden Zensusverteilungen abweichen,
- durch Antwortausfälle Verzerrungen entstehen können und schließlich
- durch Angabefehler der Antwortenden Verfälschungen der Hochrechnungsergebnisse möglich sind.

Die Problematik der „richtigen“ Gewichtung ist der zentrale Gegenstand der vorliegenden Arbeit.

1.3 Alternative Stichprobendesigns

Haushaltsbefragungen vom KONTIV-Typ sind kontinuierliche Verkehrserhebungen. Die statistischen Designprinzipien für diese Klasse von Erhebungen wurden von *Hautzinger* (1987) näher untersucht. Die nachfolgenden Ausführungen basieren hierauf.

Die Untersuchungseinheiten einer Haushaltsbefragung vom KONTIV-Typ sind je nach Fragestellung meist Haushalte, Personen oder auch Wege. Als Untersuchungsmerkmale treten demgemäß Haushalts-, Personen- und Wegemerkmale auf. Zwischen den verschiedenen Typen von Untersuchungseinheiten besteht eine hierarchische Ordnung dergestalt, daß jede Person eindeutig einem bestimmten Haushalt und jeder Weg eindeutig einer bestimmten Person zugeordnet ist.

Während die soziodemographischen Merkmale eines einzelnen Haushalts oder einer bestimmten Person im Untersuchungszeitraum mehr oder weniger stabil sind, variieren die Ausprägungen der Verkehrsverhaltensmerkmale der Haushalte oder Personen von Tag zu Tag. Haushaltsbefragungen vom KONTIV-Typ sollen – wie oben ausgeführt – Informatio-

nen über das „durchschnittliche“ Verhalten einer Population von Haushalten oder Personen während eines bestimmten Zeitraums (zumeist ein Jahr) liefern. Dies bedeutet, daß eine Durchschnittsbildung sowohl über die Haushalte bzw. Personen als auch über die Tage des Untersuchungszeitraums notwendig ist. Damit Schätzungen derartiger Durchschnittswerte möglich sind, wird eine Stichprobe aus der Grundgesamtheit aller „Haushaltstage“ bzw. „Personentage“ benötigt.

Eine Haushaltsbefragung vom KONTIV-Typ wird bei gegebener Grundgesamtheit in stichprobentheoretischer Hinsicht durch die folgenden Merkmale vollständig charakterisiert:

- (1) Zahl der auszuwählenden Haushalte
- (2) Zahl der Berichtstage pro Haushalt
- (3) Verfahren der Auswahl von Haushalten
- (4) Verfahren der Zuordnung von Berichtstagen zu den Haushalten

Nach dem Zuordnungsverfahren (Zuordnung von Berichtstagen zu Haushalten) und der Länge der individuellen Berichtsperiode (Zahl der Berichtstage pro Haushalt) kann man die folgenden Grunddesigns für Haushaltsbefragungen vom KONTIV-Typ unterscheiden:

Design I Panel-Stichprobe - eine Stichprobe von Haushalten für den gesamten Untersuchungszeitraum

Bei diesem Design berichten die Mitglieder jedes ausgewählten Haushalts über ihr Verkehrsverhalten an jedem Tag des Untersuchungszeitraums. Ein Zuordnungsproblem der oben angesprochenen Art existiert hier nicht, es ist lediglich aus befragungstechnischer Sicht darüber zu entscheiden, ob der Haushalt mehrfach interviewt werden soll oder nur einmal am Ende des Untersuchungszeitraums,

Design II Unabhängige Teilstichproben - eine Stichprobe von Haushalten für jeden Abschnitt des Untersuchungszeitraums

Dieses Design besteht darin, daß zunächst die Zahl d der Berichtstage pro Haushalt festgelegt wird. Umfaßt der Untersuchungszeitraum T Tage (T sei ein ganzzahliges Vielfaches von d), so werden $m = T/d$ unabhängige Stichproben aus der Grundgesamtheit der Haushalte gezogen. Alle in die erste Stichprobe gelangten Haushalte berichten über die Tage 1 bis d des Untersuchungszeitraums. Die zur zweiten Stichprobe gehörenden Haushalte berichten über ihr Verkehrsverhalten an den Tagen $d+1$ bis $2d$ des Untersuchungszeitraums usw.

Eine für die Praxis wichtige Modifikation von Design II besteht darin, daß nicht m unabhängige Stichproben gezogen werden, sondern daß eine Gesamtstichprobe nach einem Zufallsverfahren in m Teilstichproben zerlegt wird („ineinandergreifende“ Unterstichproben). Unter diesem modifiziertem Design kann ein und derselbe Haushalt nicht gleichzeitig mehreren Teilstichproben zugeordnet sein. Bei kleinem Auswahlatz n/N kann das modifizierte Design dem Design II gleichgesetzt werden, was die analytische Behandlung ganz erheblich vereinfacht.

Design III Rotationsstichproben - teilweise Ersetzung der Haushalte im Verlauf des Untersuchungszeitraums

Dieses Design entsteht wie folgt: Eine Stichprobe von n Haushalten wird gezogen, und jeder ausgewählte Haushalt berichtet über das Verkehrsverhalten an den Tagen 1 bis d .

Dann wird eine unabhängige Stichprobe von n_2 Haushalten gezogen, und jeder Haushalt dieser Stichprobe berichtet über die Tage 2 bis $d+1$ usw. Bei diesem Design berichtet also von allen Haushalten, die über den Tag j Auskunft geben, ein Teil auch über den Tag $j-1$, ein Teil über den Tag $j-2$ usw. Da von Tag zu Tag ein Teil der Befragten durch neue ersetzt wird, spricht man von einer „Rotationsstichprobe“ bzw. von „partieller Rotation“ der Einheiten.

Die drei bekanntesten Haushaltsbefragungen vom KONTIV-Typ, die KONTIV 76, die KONTIFERN (Vorerhebung) und die KONTIV 82 lassen sich leicht in das oben vorgestellte Schema einordnen. Die KONTIV 76 (*Sozialforschung Brög*, 1975) ist eine Stichprobe mit partieller Rotation der Einheiten (Design III). Befragt wurden die Haushalte zu ihrem Verkehrsverhalten an $d = 2$ bzw. $d = 3$ aufeinanderfolgenden Tagen. Das Rotationschema geht aus Abbildung 1.1 hervor.

Abbildung 1.1 Rotationschema der Stichprobe KONTIV 76

Mo	Di	Mi	Do	Fr	Sa	So	Mo	Di
X	X							
	X	X						
		X	X					
			X	X				
				X	X	X		
					X	X	X	
							X	X

Die KONTIFERN (Vorerhebung) von 1979 wurde als kontinuierliche Verkehrserhebung in 4 Wellen im Abstand von jeweils 3 Monaten durchgeführt (*Socialdata*, 1981). Der Untersuchungszeitraum von einem Jahr war hier nicht in Tage, sondern in Monate gegliedert, und der individuelle Berichtszeitraum umfaßte 12 Monate für Urlaubsreisen und 3 Monate für sonstige Privatreisen. Wie Abbildung 1.2 zeigt, liegt damit bezüglich der Erfassung des Urlaubsreiseverkehrs das Design III (partielle Rotation) und bezüglich der Erfassung des sonstigen Privatreiseverkehrs das Design II (unabhängige Teilstichproben) vor.

Abbildung 1.2 Rotationsschema der Stichprobe KONTIFERN (Vorerhebung)

- Erfassung Urlaubsreisen -

Welle	Berichtsmonat							
	1	3	6	9	12	15	18	21
1	X	X	X	X	X	X	X	X
2		X	X	X	X	X	X	X
3			X	X	X	X	X	X
4				X	X	X	X	X

- Erfassung sonstige Privatreisen -

Welle	Berichtsmonat											
	1	2	3	4	5	6	7	8	9	10	11	12
1	X	X	X									
2				X	X	X						
3							X	X	X			
4										X	X	X

Der Stichprobenplan der KONTIV 82 schließlich ist als eine modifizierte Version von Design II anzusehen (nachträgliche Zuordnung der insgesamt ausgewählten Haushalte zu Berichtstagen). Jeder einzelne ausgewählte Haushalt berichtete über einen Tag des Jahres ($d=1$). Die Umfänge der „Tagesstichproben“ waren dabei nicht gleich, sondern je nach Wochentagstyp unterschiedlich.

Über die relativen Vor- und Nachteile der verschiedenen Designs läßt sich aus stichprobentheoretischer Sicht ganz allgemein folgendes sagen:

1. Zur Schätzung der Veränderung eines Total- oder Mittelwerts von einer Periode zur nächsten ist Design I optimal.
2. Zur Schätzung des Total- oder Mittelwerts über alle Perioden des Untersuchungszeitraums hinweg ist Design II am besten geeignet.
3. Zur Schätzung des jeweils aktuellsten Total- oder Mittelwerts sollte man Design III verwenden.

Bei Verkehrserhebungen können alle drei eben genannten Aufgabenstellungen auftreten. Insofern kann man nicht generell einem bestimmten Design den Vorzug geben. Wichtig ist aber, daß bei der Datenauswertung - und hier insbesondere bei der Fehlerrechnung - das jeweilige Stichprobendesign explizit berücksichtigt wird.

1.4 Grundprinzipien der Hochrechnung und Genauigkeitsbeurteilung

1.4.1 Design „Unabhängige Teilstichproben“

Die Hochrechnung von der Stichprobe auf die Gesamtheit kann als Umkehrung aller „Verkleinerungsprozeduren“ durch die Auswahl betrachtet werden. Sie hat deshalb alle Besonderheiten der Auswahl wie Mehrstufigkeit, Anwendung verschiedener Auswahlsätze in den einzelnen Schichten, variierende Auswahlwahrscheinlichkeiten der Haushalte usw. genau - nur in umgekehrter Richtung - zu wiederholen (*Statistisches Bundesamt, 1960*).

Das Grundkonzept der Hochrechnung und Genauigkeitsbeurteilung bei Haushaltsbefragungen vom KONTIV-Typ soll am Beispiel des Designs II (unabhängige Teilstichproben) dargestellt werden. Es wird dabei von einem Untersuchungszeitraum der Länge $T = 365$ Tage ausgegangen, und es wird angenommen, daß jeder Haushalt über einen Stichtag ($d=1$) berichtet. Die Zahl der insgesamt zu befragenden Haushalte wird mit n bezeichnet (n sei ein ganzzahliges Vielfaches von 365).

In diesem Fall sind $T = 365$ voneinander unabhängige Stichproben von Haushalten zu ziehen. Der Stichprobenumfang n_j sei dabei für jeden Tag j des Jahres derselbe:

$$n_j = n/365 \quad (j=1, \dots, 365)$$

Die Auswahl der Haushalte erfolgt jeweils uneingeschränkt zufällig nach der Ziehungsvorschrift „ohne Zurücklegen“. In diesem Fall hat man es also mit 365 voneinander unabhängigen einfachen Klumpenstichproben zu tun, wobei die einem Haushalt zugeordneten Personen bzw. deren Wege jeweils einen Klumpen bilden.

Bezeichnet man mit X_{jk} die Ausprägung des Untersuchungsmerkmals bei dem an k -ter Stelle ausgewählten Haushalt in der Stichprobe für den j -ten Tag des Jahres, so ist

$$\hat{X}_j = \frac{N}{n_j} \sum_{k=1}^{n_j} X_{jk} = 365 \frac{N}{n} \sum_{k=1}^{n_j} X_{jk}$$

eine erwartungstreue Schätzung des Totals des Untersuchungsmerkmals für den j -ten Tag. Beispielsweise könnte man so die Gesamtlänge aller Pkw-Fahrten an einem bestimmten Tag j

des Jahres schätzen; X_{jk} wäre in diesem Fall die Gesamtlänge aller Pkw-Fahrten, die am j -ten Tag des Jahres von den Mitgliedern des k -ten Stichprobenhaushalts durchgeführt wurden.

Das Jahrestotal des Untersuchungsmerkmals, also die aggregierte Verkehrsleistung der Pkw-Benutzer, wäre durch

$$(1.4.1) \quad \hat{X} = \sum_{j=1}^{365} \hat{X}_j = 365 \frac{N}{n} \sum_{j=1}^{365} \sum_{k=1}^{n_j} X_{jk}$$

erwartungstreu zu schätzen.

Ganz analog könnte man eine Schätzfunktion \hat{Y} für die Gesamtlänge aller Wege einschließlich der Pkw-Fahrten bilden. Der Quotient

$$(1.4.2) \quad \hat{R} = \hat{X} / \hat{Y}$$

wäre dann eine Schätzfunktion für die Maßzahl „Anteil der Pkw-Verkehrsleistung an der gesamten Verkehrsleistung“. Bei der ganz überwiegenden Zahl von Schätzproblemen in der Verkehrsforschung geht es um die Schätzung eines Totals gemäß (1.4.1) bzw. einer Verhältniszahl gemäß (1.4.2).

Die Varianzen

$$\text{var}(\hat{X}) \text{ und } \text{var}(\hat{R})$$

der Schätzfunktionen für das Total bzw. das Verhältnis können ebenfalls aus der Stichprobe geschätzt werden, so daß die Berechnung von Konfidenzintervallen möglich wird. Näheres hierzu findet man z. B. bei *Cochran* (1977).

Unter dem Design II ist die Hochrechnung (einschließlich Abschätzung des Stichprobenfehlers) von Haushaltsbefragungen vom KONTIV-Typ auch dann mit Standardverfahren der Stichprobentheorie möglich, wenn

- die Auswahl der Haushalte mit ungleichen Wahrscheinlichkeiten erfolgt (z. B. Auswahl von Haushalten aus einer Personenkartei) oder wenn
- eine geschichtete Auswahl von Haushalten - mit gleichen oder variierenden Auswahlchancen - vorgenommen wird (z. B. Schichtung nach dem Haushaltstyp bei Auswahl aus einer Haushaltskartei).

Für die Varianzschätzung in den eben genannten Fällen liegt inzwischen auch schon spezielle Statistik-Software vor (*Hautzinger* und *Schorer*, 1985). Die Schätzungen \hat{X} und \hat{R} können mit Hilfe herkömmlicher Statistik-Programmpakete berechnet werden. Jedem Haushalt in der Stichprobe ist dabei ein Hochrechnungsfaktor zuzuordnen, in der Sprache der Statistik-Programmpakete also eine „Gewichtungsvariable“ (caseweight). Unter dem Design II und bei gleichem Stichprobenumfang pro Tag des Jahres hat - gleiche Auswahlchancen vorausgesetzt - jeder Haushalt denselben Hochrechnungsfaktor, nämlich

$$(1.4.3) \quad 365 N/n.$$

Wird der Stichprobenumfang n nicht gleichmäßig auf die 365 Tage des Jahres verteilt, so hängt der Hochrechnungsfaktor vom Tag innerhalb des Jahres ab. Alle Haushalte, die am j -ten Tag befragt werden, erhalten hierbei denselben Faktor, nämlich

$$(1.4.4) \quad N/n_j \quad (j=1, \dots, 365).$$

In jedem Fall ist aber die Summe aller n Hochrechnungsfaktoren gleich $365N$, d. h. gleich der Zahl der „Haushaltstage“ in der Grundgesamtheit (N Haushalte werden über 365 Tage hinweg betrachtet).

Wenn unter dem Design II die Auswahl der Haushalte nach der Ziehungsvorschrift „mit Zurücklegen“ mit variierenden Auswahlchancen erfolgt, etwa mit Auswahlchancen proportional zur Haushaltsgröße, so hängt der Hochrechnungsfaktor von der Haushaltsgröße ab (bei ungleichen Stichprobenumfängen n_j zusätzlich noch vom Befragungstag innerhalb des Jahres).

Der k -te Haushalt in der Stichprobe für den j -ten Tag des Jahres hat dann den Hochrechnungsfaktor

$$(1.4.5) \quad 365 \frac{N}{n} \frac{\bar{z}}{z_k} \quad \text{bzw.} \quad \frac{N}{n_j} \frac{\bar{z}}{z_k},$$

wobei z_k die Größe (Anzahl Personen) des k -ten Stichprobenhaushalts und \bar{z} die mittlere Haushaltsgröße in der Grundgesamtheit ist. Man beachte, daß

$$(1.4.6) \quad N\bar{z}/z_k$$

der Kehrwert der Wahrscheinlichkeit ist, mit welcher der k -te Haushalt bei einem beliebigen Zug gewählt wird. Die Einführung von Hochrechnungsfaktoren dieser Art wird in der Praxis der empirischen Sozialforschung auch „Gewichtung aus methodischen Gründen“ genannt. Eine Hochrechnung mit Varianzschätzung ist hierbei - wie bereits ausgeführt - problemlos möglich.

Wenn in einem Anwendungsfall unter dem Design II eine einstufige Auswahl von Haushalten nicht durchführbar ist und man mehrstufig vorgehen muß, wird das Hochrechnungsproblem komplizierter. Bei zweistufiger Auswahl beispielsweise, wenn in Stufe 1 Gemeinden und in Stufe 2 Haushalte (aus den ausgewählten Gemeinden) ausgewählt werden, ist eine Punkt- und Intervallschätzung von Totalwerten und Verhältniszahlen jedoch immer noch möglich, wenn

- die Gemeinden mit größenproportionalen Auswahlwahrscheinlichkeiten nach der Ziehungsvorschrift mit Zurücklegen und
- die Haushalte innerhalb der Stichprobengemeinden uneingeschränkt zufällig nach der Ziehungsvorschrift ohne Zurücklegen ausgewählt werden.

Das Schätzproblem bleibt schließlich auch dann handhabbar, wenn sowohl auf der ersten als auch auf der zweiten Stufe geschichtet wird (*Hansen u. a.*, 1953).

Insgesamt, so kann man zusammenfassend feststellen, bietet das Design II gute Voraussetzungen für eine Hochrechnung einschließlich Genauigkeitsbeurteilung, wenn man sich nur an die durch die Stichprobentheorie nahegelegten Vorgehensweisen bei der Auswahl der Haushalte hält.

1.4.2 Design „Partielle Rotation“

Wird eine Haushaltsbefragung vom KONTIV-Typ nach dem Design III („partielle Rotation“) durchgeführt, so gestaltet sich zwar nicht die Punkt- wohl aber die Intervallschätzung etwas komplizierter. Dies hängt damit zusammen, daß die Korrelation zwischen dem Verkehrsverhalten eines Haushalts bzw. einer Person in aufeinanderfolgenden Zeitabschnitten (Tagen, Monaten) explizit berücksichtigt werden muß. Bei partieller Rotation der Stichprobenhaushalte sind die Schätzungen für die Tages- oder Monatstotalwerte nicht voneinander unabhängig. Der Grund hierfür ist, daß sich die Schätzungen für zwei verschiedene Tage bzw. Monate auf teilweise dieselben Befragten stützen. Als Folge hiervon ist die Varianz des Schätzers für ein Jahrestotal oder einen Jahresmittelwert nicht einfach gleich der Summe der Varianzen der entsprechenden Tages- oder Monatsschätzer, man muß vielmehr die Kovarianzen zwischen den Tages- oder Monatsschätzern mit einbeziehen. Dieses Grundprinzip der Schätzung gilt unabhängig vom Auswahlverfahren für die Haushalte.

Das Design „partielle Rotation“ bei Haushaltsbefragungen vom KONTIV-Typ wurde von *Hautzinger* und *Stenger* (1986) näher untersucht. Dort werden Verfahren für die Punkt- und Intervallschätzung

- bei einstufiger geschichteter Zufallsauswahl von Einheiten sowie
- bei zweistufiger Zufallsauswahl von Einheiten mit Schichtung auf der ersten Stufe (Gemeinden)

behandelt.

Es gibt zwei Gründe, weshalb die Hochrechnung und Genauigkeitsbeurteilung in der Praxis meist nicht in reiner Form nach den eben vorgestellten Grundprinzipien durchgeführt werden kann bzw. durchgeführt werden sollte:

1. Die eben skizzierten Standardverfahren berücksichtigen nicht eventuell vorliegende Zusatzinformationen über die Grundgesamtheit.
2. Die Erhebung verläuft in der Regel nicht genau nach Plan, es treten vielmehr Fehler der verschiedensten Art auf, so daß die Voraussetzungen der Standardverfahren nicht erfüllt sind.

Im nachfolgenden Kapitel 2 wird auf die bei Haushaltsbefragungen vom KONTIV-Typ vorkommenden Fehler eingegangen.

2. Erhebungsfehler und ihre Berücksichtigung bei der Hochrechnung

2.1 Fehlertypologie

2.1.1 Stichprobenbedingte Fehler

Wie bei allen Erhebungen treten auch bei Haushaltsbefragungen vom KONTIV-Typ Fehler auf, die dazu führen, daß die gewonnenen Ergebnisse nicht genau mit dem übereinstimmen, was die Befragung eigentlich hätte erbringen sollen. Eine übersichtliche Darstellung zum Thema „Fehler bei Erhebungen“ findet man z. B. bei *Krug* und *Nourney* (1982), auf die im folgenden Bezug genommen wird. Im Hinblick auf Fehler bei Haushaltsbefragungen vom KONTIV-Typ kann auf die Arbeit von *Wermuth* (1983) verwiesen werden.

Im Idealfall sollte die Befragung die genauen Werte der interessierenden Kennziffern der Verkehrsnachfrage für die Grundgesamtheit aller Haushalte und den gesamten Untersuchungszeitraum (meist 1 Jahr) liefern. Daß dies nicht gelingt, liegt zum einen natürlich daran, daß nur ein sehr kleiner Teil der Haushalte befragt wird und dies auch nur zu einem (oder höchstens einigen wenigen) Tagen des Jahres. Bei 25 Millionen Haushalten und einem einjährigen Untersuchungszeitraum gibt es 9,1 Milliarden „Haushaltstage“ in der Grundgesamtheit. Wenn man 15.000 Haushalte auswählt und zum Verkehrsverhalten an zwei aufeinanderfolgenden Tagen befragt, hat man eine Stichprobe von 30.000 Haushaltstagen. Dies entspricht einem Auswahlsatz von 0,0000032: Von je einer Million Haushaltstagen in der Grundgesamtheit werden in der Stichprobe also nur rund drei Haushaltstage erfaßt.

Die stichprobenbedingte Nichtübereinstimmung zwischen Erhebungsergebnis und eigentlich interessierender Größe wird „Zufallsfehler“, „Auswahlfehler“ oder auch „Stichprobenfehler“ genannt. Präziser wäre es, von „Zufallsschwankung des Schätzwerts“ zu sprechen. Über die Bestimmung des Stichprobenfehlers bei Haushaltsbefragungen vom KONTIV-Typ ist im vorausgegangenen Abschnitt bereits einiges gesagt worden.

Neben dem Zufalls- oder Auswahlfehler (random sampling error) gibt es weitere stichprobenbedingte Fehler, nämlich die Auswahlverzerrungen (sampling biases). Diese sind zurückzuführen auf

- Mängel hinsichtlich der Erhebungsgesamtheit
- Verzerrung durch die Auswahlmethode
- Verzerrung durch die Auswahltechnik
- Verzerrung durch das Schätzverfahren.

Die erstgenannten Mängel gehen vor allem darauf zurück, daß die Auswahlgrundlage (z. B. Personen- oder Haushaltsdatei) nicht ausreichend aktuell ist und so die inzwischen neu hinzugekommenen Erhebungseinheiten zumindest teilweise keine Erfassungschance haben.

Während Verzerrungen durch die Auswahlmethode als solche weniger zu befürchten sind (es wird ja stets nach einem Zufallsauswahlverfahren vorgegangen und nicht etwa nach einer bewußten Auswahl), sind Verzerrungen durch die Auswahltechnik zu erwarten. Dies liegt vor allem daran, daß vielfach – beispielsweise wegen Unzulänglichkeiten der als Auswahlgrundlage dienenden Adreßbücher – die tatsächliche Auswahlwahrscheinlichkeit der einzelnen Haushalte nicht angegeben werden kann.

Verzerrungen können auch vom verwendeten Schätzverfahren herrühren. So liefert bekanntermaßen z. B. das Verfahren der Verhältnisschätzung keine unverzerrten Schätzwerte; das Ausmaß der Verzerrung nimmt allerdings mit wachsendem Stichprobenumfang rasch ab. Wichtig erscheint hier der Hinweis, daß Verzerrungen der Ergebnisse auch durch Einführung von Gewichtungsfaktoren in das Hochrechnungsverfahren entstehen können und zwar auch dann, wenn die Zielsetzung der Gewichtung gerade der Ausgleich von Verzerrungen (z. B. durch Antwortausfälle) ist. Vgl. hierzu Abschnitt 7.1.

2.1.2 Verfahrensbedingte Fehler

Neben den stichprobenbedingten Fehlern gibt es eine nahezu unbegrenzte Zahl weiterer Fehlerquellen. Als besonders hilfreich erweist sich im vorliegenden Fall eine Typisierung der Fehler nach der Untersuchungsstufe, in welcher sie auftreten (Böltken, 1976):

1. Fehler bei der Planung und Vorbereitung
 - a) Fehler bei der Bestimmung des Forschungsproblems, des Forschungsobjekts und des Forschungsziels
 - b) Fehler bei der Definition einer dem Forschungsproblem, dem Forschungsobjekt und dem Forschungsziel sowohl räumlich und zeitlich als auch sachlich adäquaten Grundgesamtheit
 - c) Fehler bei der Abfassung des Fragebogens
 - d) Fehler bei der Zusammenstellung der Aufbereitungs- und Auswertungspläne
 - e) Fehler durch schlechte Auswahl sowie unzureichende Ausbildung und Information der Mitarbeiter
 - f) Fehler durch mangelhafte oder sorglose Organisation der Feldarbeit, der Aufbereitungs- und der Auswertungsarbeiten
2. Fehler bei der Auswahl der Erhebungseinheiten
 - a) Verzerrungen der Erhebungsauswahl durch bewußte oder unbeabsichtigte Abweichungen der Mitarbeiter vom festgelegten Auswahlplan
 - b) Verzerrungen der Erhebungsauswahl durch Ausfälle ausgewählter Erhebungseinheiten infolge Nichterreichbarkeit oder Verweigerung
3. Fehler bei der Befragung
 - a) Fehler durch bewußte oder unbeabsichtigte falsche Antworten und Angaben des Befragten
 - b) Fehler durch Mißverständnisse bei der Fragestellung und bei den Antworten sowie durch falsche Einordnung der Antworten und Angaben auf dem Fragebogen
 - c) Fehler durch Verweigerung der Antwort auf einzelne Fragen oder Teile von Fragen
4. Fehler bei der Aufbereitung und Auswertung
 - a) Fehler bei der Verschlüsselung und Übertragung der Angaben sowie der Erstellung von Dateien
 - b) Fehler bei der Datenauswertung, insbesondere durch falsche Anwendung von Statistik-Programmpaketen
 - c) Fehler durch Anwendung falscher oder fehlerhafter Formeln für die Berechnung der Schätzwerte und der Standardfehler
 - d) Fehler bei der Interpretation der Erhebungsergebnisse

Die wichtigsten dieser nicht stichproben-, sondern allgemein verfahrensbedingten Fehler (procedural biases) werden im folgenden Abschnitt näher behandelt.

2.2 *Verfahrensbedingte Fehler bei Haushaltsbefragungen vom KONTIV-Typ*

2.2.1 Nichtübereinstimmung von Zielkollektiv und Auswahlgesamtheit

Fehler und Ungenauigkeiten resultieren zwangsläufig aus der praktischen Schwierigkeit der Identifizierung von Haushalten als Auswahlseinheiten in einer nicht klar abzugrenzenden Grundgesamtheit. Die Ergebnisse der Volkszählung von 1987 werden zeigen, in welcher Größenordnung die Abweichungen zwischen den registrierten und vorhandenen Einwohnern bzw. Haushalten liegen. Untererfassungen, teilweise aber auch Doppelerfassungen,

insbesondere durch Haushalte mit häufigem Wohnortwechsel, ergeben sich vor allem durch Personen, die sich in Ausbildung oder beim Militärdienst befinden, sowie durch Zweitwohnsitze, die nicht korrekt gemeldet werden.

Daß gerade bei Schülern, Auszubildenden und Studenten ab 18 Jahren Erfassungsprobleme zu befürchten sind, wird bei Verkehrserhebungen durch den in dieser Gruppe besonders hohen Anteil der Angabe „am Stichtag nicht am Wohnort“ unterstrichen. Die Auswahl wird weiterhin verkompliziert durch die steigende Anzahl von alten Menschen, die in Heimen untergebracht sind oder sich größere Teile des Jahres insbesondere im Winter im Ausland aufhalten.

Ungenauigkeiten sind stets auch mit der unscharfen Definition des Begriffs „Haushalt“, der z. B. bei der Haushaltsgründung im elterlichen Haus, beim Zusammenleben unverheirateter Paare oder bei Wohngemeinschaften nicht richtig zu erfassen bzw. auf dem jeweiligen Stand zu halten ist, verbunden. Personen wären so gesehen die besser geeigneten Auswahlmöglichkeiten.

2.2.2 Eingeschränkte Objektivität, Reliabilität und Validität des Meßverfahrens

Als wichtigste Gütekriterien für die Qualität eines wissenschaftlich fundierten Meßverfahrens werden üblicherweise

- Objektivität
- Reliabilität (Zuverlässigkeit) und
- Validität (Gültigkeit)

genannt (Roth, 1987). Objektivität wird dabei bestimmt als Grad, in welchem die Meßergebnisse hinsichtlich ihres Zustandekommens, ihrer Auswertung und ihrer Interpretation vom Wissenschaftler, der die Messung durchführt, sowie von der Meßsituation unabhängig sind. Unter der Reliabilität wird der Grad der Genauigkeit verstanden, mit der ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal gemessen wird, gleichgültig, ob das Meßverfahren die Messung eben dieses Merkmals für sich beansprucht; es geht also um die formale Exaktheit der Merkmalerfassung. Die Validität schließlich gibt den Grad der Genauigkeit an, mit welchem ein Meßverfahren gerade dasjenige mißt, was es zu messen vorgibt; hierbei wird meist zwischen Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität unterschieden.

Es kann an dieser Stelle nicht in allen Einzelheiten auf die eben angesprochenen Fehlerquellen eingegangen werden. Als wichtiges Problemfeld muß jedoch die „wegeorientierte“ Erfassung des individuellen Verkehrsverhaltens mit Hilfe des KONTIV-Fragebogens erwähnt werden. Zwar liefert der Bezug zum Tagesablauf (1. Weg, 2. Weg usw.) für den Befragten eine gute Orientierungshilfe beim Ausfüllen des Fragebogens, doch läßt der Begriff „Weg“ ohne präzisierende Hinweise auf dem Fragebogen beim Befragten zweifellos einen gewissen Interpretationsspielraum offen.

Grundsätzlich kann wohl davon ausgegangen werden, daß der Begriff „Weg“ von den Befragten durchaus unterschiedlich interpretiert wird, was eine Einschränkung der Validität des Erhebungsverfahrens darstellt. Die Einkaufs-„Rundtour“ mit 7 Wegen, von denen 5 Wege lediglich über kurze Distanzen von Geschäft zu Geschäft führten, wird wohl selten vollständig berichtet werden. Neben den Wegen beim Verbinden mehrerer Einkaufsaktivi-

täten - insbesondere zu Fuß und zu Läden, in denen dann doch kein Einkauf getätigt wurde, - dürften Wege im Zuge längerer dienstlich-geschäftlicher Wegeketten solche sein, die im Fragebogen häufig nicht angegeben werden.

Obwohl also mit der wegeorientierten Erfassung des Verkehrsverhaltens offenkundig bestimmte Probleme verbunden sind, ist derzeit für Befragungen vom KONTIV-Typ kein tragfähiges alternatives Erhebungskonzept in Sicht. Als Konsequenz hieraus bietet es sich an, bei der Datenauswertung anstelle von einzelnen Wegen nur noch ganze „Ausgänge“, also Folgen von Wegen mit der Eigenschaft, daß die Quelle des ersten Wegs der Folge mit dem Ziel des letzten Wegs der Folge identisch ist, zu betrachten. Die wichtigsten Merkmale wie Uhrzeit Beginn/Ende, Länge und Dauer dürften aus den Angaben im KONTIV-Fragebogen zuverlässiger für Ausgänge als für Wege zu ermitteln sein.

Wählt man Ausgänge anstelle von Wegen als Analyseeinheiten, so wird allerdings das Merkmal „Zweck“ ein Merkmal mit Mehrfachnennungen, was die Auswertungen komplizierter werden läßt. Beim Merkmal „Benutztes Verkehrsmittel“ verschärft sich die auch schon beim Wegekonzept bestehende Mehrfachnennungsproblematik. Patentlösungen gibt es hier also nicht. Welchem Konzept der Vorzug zu geben ist, muß vor dem Hintergrund der jeweiligen Fragestellung entschieden werden.

Außer durch die eingeschränkte Validität der Befragungsergebnisse - auf Reliabilitätsfragen wird im Zusammenhang mit den sogenannten Angabefehlern noch eingegangen - entstehen Komplikationen bei der Hochrechnung dadurch, daß bei einem Teil der in die Stichprobe gelangten Haushalte die Ausprägungen der Untersuchungsmerkmale überhaupt nicht festgestellt werden können, weil die betreffenden Haushalte entweder nicht erreichbar sind oder die Teilnahme an der Befragung verweigern. Diese beiden Fehlerquellen werden nachfolgend untersucht.

2.2.3 Nichterreichbarkeit eines Teils der Stichprobenhaushalte

Bei einem Teil der in die Stichprobe gelangten Haushalte ist eine Feststellung der Ausprägungen der Untersuchungsmerkmale nicht möglich, weil die betreffenden Haushalte nicht erreichbar sind. In der KONTIV 82 waren dies immerhin 14 Prozent aller Haushalte. Man spricht in diesem Zusammenhang auch von „unechten“ oder „qualitätsneutralen“ Ausfällen, wohinter sich die Hoffnung verbirgt, daß die Wahrscheinlichkeit der Nichterreichbarkeit für alle Haushalte in der Grundgesamtheit dieselbe ist (unter der Bedingung, daß der Haushalt in die Stichprobe gelangt). Wie empirische Ergebnisse von *Wermuth* und *Maerschalk* (1983) zeigen, ist diese Wahrscheinlichkeit aber zumindest von der Haushaltsgröße abhängig: Als Folge einer mit abnehmender Haushaltsgröße steigenden Wohnstandortmobilität weisen insbesondere Einpersonenhaushalte eine überdurchschnittliche Nichterreichbarkeit auf.

Da es sich hier um ein allgemeines Problem jeder Befragung handelt, kann auf die entsprechende Literatur verwiesen werden. Soweit die Verteilungen bestimmter soziodemographischer Merkmale in der Grundgesamtheit (in guter Näherung) bekannt sind, können diese zur Bestimmung der wirklichen Auswahlätze für die Hochrechnung benutzt werden (vgl. Abschnitt 3.2).

Daß dem Problem der Nichterreichbarkeit Aufmerksamkeit zu schenken ist, zeigt folgendes Beispiel. In einer Berliner Befragung von 1986 wurde ein Test der Zuverlässigkeit der Post durchgeführt. Eine Reihe von bewußt falsch adressierten Schreiben (einschließlich Erinnerungsschreiben) wurden dabei von der Post *nicht* an den Absender zurückgesandt. Aus diesem und ähnlichen anderen „Versuchen“ folgt die Notwendigkeit von Kontrollmechanismen und Kontrollerhebungen. Derzeit scheint es wenig sinnvoll, in diesem Bereich mit speziellen Korrekturfaktoren zu arbeiten, da die empirische Basis hierfür völlig unzureichend ist (Kutter und Holz, 1987).

2.2.4 Nichtteilnahme (Antwortverweigerung) eines Teils der Stichprobenhaushalte

Von einem weiteren Teil der in die Stichprobe gelangten Haushalte sind die Ausprägungen der Untersuchungsmerkmale wegen Verweigerung der Teilnahme an der Befragung nicht bekannt. Bei der KONTIV 82 hat letztlich rund ein Drittel aller erreichbaren Haushalte die Teilnahme an der Erhebung verweigert.

Geht man - unter Vernachlässigung des stochastischen Charakters der Antwortausfälle - vereinfachend von der Modellvorstellung aus, daß die Grundgesamtheit G der Haushalte in 3 Schichten gegliedert ist, nämlich

- nicht erreichbare Haushalte (G_1)
- erreichbare, aber nicht teilnahmebereite Haushalte (G_2)
- erreichbare und teilnahmebereite Haushalte (G_3)

so liefert eine Zufallsstichprobe mit fest vorgegebenem Umfang n insgesamt

- n_1 Haushalte aus Schicht G_1
- n_2 Haushalte aus Schicht G_2
- n_3 Haushalte aus Schicht G_3

wobei n_1 , n_2 und n_3 zufallsabhängige Größen sind, die sich aber zu n aufsummieren (Cochran, 1977, S. 359 f).

Informationen zum Verkehrsverhalten liegen nach Durchführung der Erhebung nur für die n_3 Haushalte aus Schicht G_3 vor. Statistische Schlußfolgerungen sind also nur in bezug auf G_3 , nicht aber in bezug auf die Grundgesamtheit G möglich. Diese Einschränkung wäre unproblematisch, wenn die Untersuchungsmerkmale - insbesondere natürlich die Verkehrsverhaltensmerkmale - in den Teilgesamtheiten G_1 , G_2 und G_3 jeweils dieselbe Verteilung besäßen. Dies kann aber nicht ohne weiteres angenommen werden.

In der bereits zitierten Arbeit von Wermuth und Maerschalk (1983) wird u. a. dargestellt, daß

- antwortende Haushalte eine überdurchschnittliche Haushaltsgröße aufweisen,
- Personen aus antwortenden Haushalten tendenziell älter sind als Personen aus nicht antwortenden Haushalten und
- Personen aus antwortenden Haushalten häufiger erwerbstätig sind als Personen aus nicht antwortenden Haushalten.

Solche Abhängigkeiten der Antwortwahrscheinlichkeit von soziodemographischen Strukturmerkmalen der Haushalte können durch Vergleich der Stichprobenverteilungen dieser Merkmale mit den entsprechenden Zensusverteilungen entdeckt werden. Wenn die Ant-

wortbereitschaft in dieser Weise zwischen verschiedenen Teilgruppen von Haushalten variiert, wird in der Regel die Stichprobe verzerrt sein: Haushaltsgruppen mit hoher Antwortbereitschaft sind überrepräsentiert, solche mit niedriger Antwortbereitschaft unterrepräsentiert. In dem Maße, wie die Merkmale, welche die Antwortbereitschaft beeinflussen, mit den Verkehrsverhaltensmerkmalen im Zusammenhang stehen, sind natürlich auch die Stichprobenergebnisse zum Verkehrsverhalten verzerrt. Wenn es zutrifft, daß die Antwortbereitschaft von Alter und Erwerbstätigkeit der Haushaltsmitglieder abhängt, wäre dies zweifellos eine Quelle von Verzerrungen der oben genannten Art.

Weitere Verzerrungen können entstehen, wenn die Antwortbereitschaft nicht nur von soziodemographischen Haushaltsmerkmalen, sondern zusätzlich noch von Verkehrsverhaltensmerkmalen der Haushaltsmitglieder abhängt. Ob es bei gegebenen soziodemographischen Merkmalen einen Zusammenhang zwischen Verkehrsverhalten und Antwortbereitschaft gibt, kann nur durch entsprechende Befragungen von Nichtantwortern geklärt werden. Solche Kontrollerhebungen liegen derzeit aber zumindest für die Bundesrepublik Deutschland noch nicht vor. Insofern haben gelegentlich geäußerte Vermutungen, wonach

- Haushalte mit extremen „Vielfahrern“ und andererseits
- Haushalte mit vorwiegend immobilen Personen

überdurchschnittlich häufig die Teilnahme verweigern (*Wermuth*, 1985) derzeit noch vorläufigen Charakter.

Das weitgehende Fehlen gesicherter empirischer Erkenntnisse über Verhaltensunterschiede zwischen Personen aus antwortenden und nicht antwortenden Haushalten macht die Berücksichtigung des Non-Response-Problems bei der Hochrechnung schwierig. Soweit es sich um Verzerrungen handelt, die durch Abhängigkeiten zwischen Antwortbereitschaft und soziodemographischen Haushaltsmerkmalen entstehen, kann - zumindest teilweise - durch nachträgliche Schichtung (vgl. Abschnitt 3.2) Abhilfe geschaffen werden. Gibt es zusätzlich noch direkte Zusammenhänge zwischen Verkehrsverhalten und Antwortbereitschaft, könnte man an die Einführung geeigneter, empirisch ermittelter Korrekturfaktoren denken (vgl. Abschnitt 4). In dieser durch erhebliche Unsicherheiten gekennzeichneten Situation wurde von *Wermuth* und *Maerschalk* (1983) eine Extrapolationsmethode vorgeschlagen, welche von einem Zusammenhang zwischen Verkehrsverhalten und Antwortgeschwindigkeit ausgeht. Von *Socialdata* wurde dieses Verfahren bei der Hochrechnung der KONTIV 82 (und nachträglich auch bei der KONTIV 76) in extensiver Form angewandt. Näheres hierzu findet sich in Abschnitt 4.3.

Die Vielschichtigkeit des Non-Response-Problems sei noch an folgendem Beispiel demonstriert. In der Schweiz fand im Zusammenhang mit dem nach KONTIV-Design durchgeführten Mikrozensus „Verkehr und Umwelt“ 1984 eine telefonische Befragung eines Teils der Haushalte, die die Teilnahme an der schriftlichen Befragung verweigert hatten, statt. Hierbei zeigte sich nach Angaben des Schweizerischen Bundesamts für Statistik, daß viele dieser „Antwortverweigerer“ die Teilnahme an der Befragung gar nicht prinzipiell ablehnten, sondern ganz einfach mit der schriftlichen Erhebungsform Schwierigkeiten hatten: Im Telefoninterview waren rund zwei Drittel der „Antwortverweigerer“ auskunftsbereit. Diese Beobachtung unterstreicht eindringlich die Notwendigkeit empirischer Studien zur Antwortbereitschaft bei Verkehrserhebungen vom KONTIV-Typ.

2.2.5 Angabefehler der Befragten

Idealerweise sollten die teilnahmebereiten (d. h. antwortenden) Haushalte zu allen Fragen im Fragebogen vollständige und korrekte Angaben machen. Dies ist, wie man weiß, nicht der Fall. Solche Nichtübereinstimmungen zwischen der angegebenen und der „wahren“ Merkmalsausprägung werden Angabefehler (response error) genannt. Auch das Fehlen oder die Unvollständigkeit von Angaben zu Untersuchungsmerkmalen wird man als Angabefehler betrachten.

Um zu einer Klassifikation der Angabefehler bei Haushaltsbefragungen vom KONTIV-Typ zu kommen, ist es zweckmäßig, von den hierarchisch geordneten Untersuchungseinheiten

- Haushalt
- Person
- Ausgang
- Weg (Aktivität)

auszugehen: Zum Haushalt gehört mindestens eine Person, jede Person hat am Stichtag eine bestimmte Anzahl (einschließlich Null) Ausgänge durchgeführt und jeder Ausgang umfaßt eine bestimmte Anzahl von Wegen bzw. Aktivitäten. Da die Untersuchungseinheiten Träger bestimmter Merkmale sind, hat man es also mit Haushalts-, Personen-, Ausgangs- oder Wegemerkmalen zu tun. Bei den Haushalts- und Personenmerkmalen muß noch zwischen soziodemographischen Variablen und Verkehrsverhaltensmerkmalen unterschieden werden.

Angabefehler können nun auf zwei Arten entstehen: Von „einheitenbezogenen“ Angabefehlern wird man dann sprechen, wenn innerhalb eines Haushalts bestimmte Einheiten (Personen, Ausgänge, Wege) nicht angegeben werden, obwohl sie eigentlich angegeben werden müßten, bzw. wenn – was aber wohl nur selten vorkommt – bestimmte Einheiten angegeben werden, obwohl sie nicht angegeben werden dürften. Wenn innerhalb eines Haushalts eine Person fehlt, so fehlen automatisch auch deren Ausgänge und Wege; fehlt bei einer bestimmten Person ein Ausgang, so fehlen auch die diesem Ausgang zugeordneten Wege. Nach derzeitigem Kenntnisstand dürften

- das Fehlen von Ausgängen (und dazu gehörenden Wegen) sowie
- das Fehlen von Wegen

innerhalb eines Haushalts bzw. seiner Personen die häufigsten einheitenbezogenen Angabefehler sein.

Ein „merkmalsbezogener“ Angabefehler liegt demgegenüber vor, wenn bei einem Haushalts-, Personen-, Ausgangs- oder Wegemerkmal die Merkmalsausprägung gemäß Antwort im Fragebogen nicht mit der tatsächlich richtigen oder „wahren“ Merkmalsausprägung übereinstimmt. Das Fehlen einer eigentlich notwendigen Angabe zu einem Merkmal sowie die Unterlassung eigentlich erforderlicher Nennungen bei Merkmalen mit Mehrfachnennung wird, wie bereits ausgeführt, auch als Angabefehler betrachtet.

Merkmalsbezogene Angabefehler sind naturgemäß bei allen Merkmalen möglich. Von besonderer praktischer Bedeutung sind bei Haushaltsbefragungen vom KONTIV-Typ aber — falsche bzw. fehlende Angaben zur Wegelänge sowie zur Wegedauer (*Wermuth und Maerschalk*, 1981) sowie das

– Unterlassen der Nennung „zu Fuß“ beim Wegemerkmals „Benutztes Verkehrsmittel“, bei welchem Mehrfachnennungen möglich sind (Hautzinger und Tassaux, 1989).

Im folgenden wird gezeigt, welche prinzipiellen Möglichkeiten der Fehlerreduzierung bzw. Fehlerkorrektur zur Verfügung stehen.

2.3 Gewichtung als Mittel zur Fehlerreduzierung bzw. Fehlerkorrektur

Nach diesen Ausführungen über Erhebungsfehler ist es naheliegend, das Hochrechnungsverfahren so zu gestalten, daß einerseits der stichprobenbedingte Zufallsfehler der Schätzung reduziert und andererseits Auswahlverzerrungen sowie Verzerrungen durch verfahrensbedingte Fehler nach Möglichkeit korrigiert werden. Beide Forderungen führen dazu, daß Hochrechnung und Genauigkeitsbeurteilung nicht nach den in Abschnitt 1.4 dargestellten Standardverfahren durchgeführt werden können.

Die Reduzierung des Zufallsfehlers bzw. die Korrektur von Verzerrungen setzt die Verfügbarkeit von Zusatzinformationen voraus. Hier kann zwischen

- Zusatzinformationen über die Grundgesamtheit und
- Zusatzinformationen über die Natur der Erhebungsfehler

unterschieden werden. Erstere stammen aus Totalerhebungen der amtlichen Statistik („Zensus-Zahlen“), letztere aus entsprechenden Zusatz- oder Kontrollerhebungen zur Haushaltsbefragung.

Geht man einmal vom Erhebungsdesign II (unabhängige Tagesstichproben) aus, so führt rechenstechnisch die Einbeziehung von Zusatzinformationen mit dem Ziel der Fehlerreduzierung bzw. Fehlerkorrektur dazu, daß in Formel (1.3.1) anstelle der Stichprobenwerte

$$X_{jk} \quad (k=1, \dots, n_j; j=1, \dots, 365)$$

(X_{jk} bezeichnet die Ausprägung des Untersuchungsmerkmals - z. B. Zahl der Pkw-Fahrten - bei dem an k -ter Stelle ausgewählten Haushalt in der Stichprobe für den j -ten Tag des Jahres) modifizierte Werte

$$(2.3.1) \quad X_{jk}^* = w_{jk} X_{jk}$$

verwendet werden. Die Faktoren w_{jk} werden „Gewichte“ oder „Gewichtungsfaktoren“ genannt. Die Einbeziehung solcher Faktoren in das Hochrechnungsverfahren kann allgemein als „Gewichtung“ bezeichnet werden.

Sei x ein zu schätzender Totalwert, beispielsweise das Jahrestotal der Fußgänger-Verkehrsleistung. Indem man von der ungewichteten Schätzung \hat{X} zur gewichteten Schätzung \hat{X}^* übergeht, erhofft man einerseits eine Reduzierung des Zufallsfehlers: Man versucht die Gewichte w_{jk} so zu wählen, daß

$$(2.3.2) \quad \text{var}(\hat{X}^*) < \text{var}(\hat{X})$$

gilt. Andererseits soll eine eventuelle Verzerrung der Schätzung zumindest teilweise korrigiert werden: Die Gewichtung soll also dazu führen, daß

$$(2.3.3) \quad |E(\hat{X}^*) - x| < |E(\hat{X}) - x|.$$

Im Idealfall sollte die gewichtete Schätzung unverzerrt sein, d. h. es sollte $E(\hat{X}^*) - x = 0$ gelten.

Es ist wichtig, darauf hinzuweisen, daß die Verzerrung

$$(2.3.4) \quad B = E(\hat{X}) - x$$

immer im Zusammenhang mit dem Standardfehler der Schätzung

$$(2.3.5) \quad \sigma_{\hat{X}} = \sqrt{\text{var}(\hat{X})}$$

gesehen werden muß. Generell gilt, daß bei normalverteiltem Schätzer der Einfluß der Verzerrung auf die Schätzung vernachlässigt werden kann, solange

$$(2.3.6) \quad |B| < 0.1 \sigma_{\hat{X}}$$

ist. Selbst wenn $|B| < 0.2 \sigma_{\hat{X}}$ gilt, sind die Auswirkungen der Verzerrung noch vergleichsweise begrenzt (Cochran, 1977, S. 13-15).

Wird die Verzerrung durch das Schätzverfahren (Verhältnisschätzung) verursacht, so kann eine obere Schranke für $B/\sigma_{\hat{X}}$ mathematisch bestimmt werden und bei hinreichend großem Stichprobenumfang ist sichergestellt, daß Bedingung (2.3.6) erfüllt ist. Bei ausfallbedingten Verzerrungen und Verzerrungen durch Angabefehler ist es dagegen meist nicht möglich zu entscheiden, ob Bedingung (2.3.6) eingehalten wird oder nicht.

Im folgenden wird nun gezeigt, durch welche Arten von Gewichtung Genauigkeitsgewinne bei der Hochrechnung von Haushaltsbefragungen vom KONTIV-Typ möglich sind.

3. Maßnahmen zur Reduzierung des Stichprobenfehlers

3.1 Gebundene Hochrechnung: Regressions- und Verhältnisschätzung

Eine gebundene Hochrechnung ist möglich, wenn bei den Untersuchungseinheiten zusätzlich zum Untersuchungsmerkmal noch ein weiteres Merkmal (Hilfs- oder Basismerkmal) erfaßt wurde, dessen Totalwert aus einer anderen Erhebung bekannt ist. Sofern Untersuchungs- und Basismerkmal miteinander korrelieren, liefert eine gebundene Hochrechnung Schätzwerte für das arithmetische Mittel bzw. das Total des Untersuchungsmerkmals, welche gegenüber den Schätzwerten bei freier Hochrechnung einen kleineren Zufallsfehler besitzen (Stenger, 1986, S. 61 ff).

Die gebundene Hochrechnung kann immer in Form einer *Regressionschätzung* durchgeführt werden. Neben dem Totalwert des Basismerkmals in der Grundgesamtheit wird hierfür noch die Steigung der Regressionsgeraden (Untersuchungsmerkmal in Abhängigkeit vom Basismerkmal) benötigt. Vielfach ist diese Steigung nicht genau bekannt, sondern muß vielmehr aus der Stichprobe selbst geschätzt werden.

Wenn der korrelative Zusammenhang zwischen Untersuchungs- und Basismerkmal positiv ist und außerdem die Regressionsgerade durch den Ursprung des Koordinatensystems geht, bringt eine *Verhältnisschätzung* bei geringerem technischen Aufwand dieselben Genauigkeitsvorteile wie die Regressionschätzung. Damit die Verhältnisschätzung gegenüber einer freien Hochrechnung vorteilhaft ist, darf allerdings die Variabilität des Basismerkmals

(gemessen durch den Variationskoeffizienten) nicht wesentlich größer sein als die Variabilität des Untersuchungsmerkmals (Barnett, 1984, S. 61 f). Wenn man im Rahmen des Erhebungsdesigns II die Ausprägungen des Untersuchungsmerkmals mit $X_{j,k}$ und die Ausprägungen des Basismerkmals mit $Y_{j,k}$ bezeichnet, so ist der Verhältnisschätzwert für das Total x des Untersuchungsmerkmals durch

$$(3.1.1) \quad \hat{X}_{\text{c r h}} = (\hat{X}/\hat{Y})y = \hat{R}y$$

gegeben, wobei y den als bekannt vorausgesetzten Grundgesamtheitstotalwert des Basismerkmals bezeichnet.

Durch gebundene Hochrechnung kann die Schätzgenauigkeit natürlich nur dann wirklich verbessert werden, wenn der Totalwert des Basismerkmals genau bekannt ist. Hier ergeben sich in der Praxis der Verkehrsstatistik aber nicht unerhebliche Probleme:

- a) Die Zahl der möglichen Basismerkmale ist nicht sehr groß, da nur für relativ wenige in Betracht kommende Merkmale der Totalwert mit hinreichender Genauigkeit aus der amtlichen oder halbamtlichen Statistik bekannt ist. In dieser Hinsicht kann man vor allem an den Pkw-Besitz als Basismerkmal denken, aber selbst hier gibt es nicht zu unterschätzende Abgrenzungsprobleme (Hautzinger, 1986). Totalwerte aus VÖV- oder DB-Statistiken sind wegen spezieller Modalitäten der Erhebungen wohl nur in Ausnahmefällen verwertbar.
- b) Die interessierenden Total- oder Mittelwerte sind in der Regel nach bestimmten Aufgliederungsmerkmalen differenziert zu schätzen. Dies bedeutet aber, daß auch das Total des Basismerkmals in der entsprechenden Aufgliederung vorliegen muß. Eine Verhältnisschätzung, z. B. der jährlichen Verkehrsleistung von Pkw-Benutzern gegliedert nach der Gemeindegröße des Wohnorts, setzt also die Kenntnisse des Pkw-Bestands gegliedert nach Gemeindegrößenklassen voraus. Totalwerte von Basismerkmalen sind jedoch, wenn überhaupt, meist nur in wenigen Aufgliederungen sachlicher, räumlicher oder zeitlicher Art verfügbar.

Insgesamt, so kann man feststellen, sind der Verhältnisschätzung bei Haushaltsbefragungen vom KONTIV-Typ vergleichsweise enge Grenzen gesetzt, da nur selten die für die Verhältnisschätzung benötigten Zusatzinformationen (Total des Basismerkmals, gegebenenfalls in geeigneter Aufgliederung) verfügbar sind. Dies kann nicht weiter verwundern, da ja die genannten Haushaltsbefragungen gerade wegen des eklatanten Informationsdefizits der amtlichen Personenverkehrsstatistik durchgeführt werden.

3.2 Nachträgliche Schichtung (Gewichtung im engeren Sinne)

Eine Verbesserung der Hochrechnungsgenauigkeit ist weiterhin möglich, wenn Zusatzinformationen über den Umfang bestimmter Teilgruppen von Haushalten bzw. Personen (sog. „Hochrechnungsgruppen“) aus der amtlichen Statistik vorliegen. Beim Verfahren der *nachträglichen Schichtung* wird wie folgt vorgegangen:

1. Die Stichprobe wird in Hochrechnungsgruppen gegliedert und es wird der Umfang dieser Gruppen nach dem in Abschnitt 1.4 beschriebenen Verfahren geschätzt. Die geschätzte Zahl der Haushalte in der i -ten Gruppe wird mit $\hat{H}(i)$ bezeichnet (vorläufige Schätzung).

2. Die (vorläufige) Schätzung $\hat{H}(i)$ der Besetzungshäufigkeit der i -ten Gruppe stimmt im allgemeinen nicht mit der aus der amtlichen Statistik bekannten wahren Häufigkeit $N(i)$ überein. Wenn man aber jeden Stichprobenhaushalt, der zur i -ten Gruppe gehört, mit dem Gewichtungsfaktor

$$N(i) / \hat{H}(i)$$

versieht, so ist nach der Gewichtung für alle Hochrechnungsgruppen die Übereinstimmung der (endgültigen) Schätzung mit der entsprechenden Zensuszahl sichergestellt.

Statt von nachträglicher Schichtung spricht man auch von „Anpassung der Randverteilungen soziodemographischer Merkmale“. Wenn in der Praxis der empirischen Sozialforschung von Gewichtung gesprochen wird, so ist damit in der Regel die nachträgliche Schichtung gemeint. Wir bezeichnen sie deshalb auch als „Gewichtung im engeren Sinne“.

Das Verfahren der nachträglichen Schichtung hat für die übrigen Merkmale – also die Merkmale, die nicht zur Gruppierung benutzt wurden, asymptotisch die gleiche Wirkung wie eine geschichtete Auswahl von Haushalten mit gleichem Auswahlatz: Sofern ein Untersuchungsmerkmal von den verwendeten Gruppierungsmerkmalen abhängig ist, kann mit einer Verkleinerung des Zufallsfehlers der Schätzung („Schichtungseffekt“) gerechnet werden. Durch nachträgliche Schichtung lassen sich aber nicht nur, wie man gelegentlich sagt, „zufallsbedingte Unvollkommenheiten der Stichprobe“ ausgleichen: Die Anpassung von Randverteilungen soziodemographischer Merkmale korrigiert zugleich unerwünschte Non-Response-Effekte. Auf diesen Aspekt wird in Kapitel 4 näher eingegangen.

Da die zu schätzenden Total- oder Mittelwerte der Untersuchungsmerkmale sich auf die Grundgesamtheit aller „Haushaltstage“ des Untersuchungszeitraums (Jahr) beziehen, sollte die nachträgliche Schichtung der Haushaltstage in der Stichprobe simultan nach soziodemographischen und zeitlichen Merkmalen erfolgen. Wenn man beispielsweise nach den Merkmalen

- Größe des Wohnorts (5 Größenklassen)
- Quartal des Befragungsstichtags (4 Quartale)
- Wochentag des Befragungsstichtags (7 Wochentage)

gleichzeitig schichtet, so entstehen $5 \cdot 4 \cdot 7 = 140$ Hochrechnungsgruppen. Geht man der Einfachheit halber von einem 364-tägigen Untersuchungszeitraum aus, so kommt in der Grundgesamtheit der Haushaltstage jede der 28 Kombinationen von Quartal und Wochentag genau 13mal vor. Bezeichnet man in der Grundgesamtheit mit $N(g)$ die Zahl der Haushalte in Gemeindegrößenklasse g , so ist

$$(3.2.1) \quad N(g, h, i) = 13 N(g)$$

die Zensuszahl der Haushaltstage von Haushalten aus Gemeindegrößenklasse g , welche im Quartal h liegen und bei denen es sich um den Wochentag i handelt. Die entsprechende hochgerechnete Zahl $\hat{H}(g, h, i)$ (vorläufige Schätzung) von Haushaltstagen des Typs (g, h, i) erhält man, indem man in Formel (1.4.1) die Variable X_{jk} gleich 1 setzt, falls der k -te Haushalt in der j -ten Tagesstichprobe zur Gemeindegrößenklasse g gehört und es sich beim j -ten Tag des Untersuchungszeitraums um einen Tag im Quartal h und einen Wochentag i handelt; trifft dies nicht zu, so wird X_{jk} gleich 0 gesetzt. Die nachträgliche Schichtung nach

den drei genannten Merkmalen besteht nun darin, daß jeder Haushaltstag in der Stichprobe, der vom Typ (g, h, i) ist, den Gewichtungsfaktor

$$(3.2.2) \quad N(g, h, i) / \hat{H}(g, h, i)$$

erhält. Auf diese Weise wird jedem Stichprobenhaushalt ein Gewichtungsfaktor zugeordnet. Für die endgültige Hochrechnung wird der Gewichtungsfaktor (3.2.2) multiplikativ mit dem Hochrechnungsfaktor (1.4.3) - bzw. (1.4.4) oder (1.4.5) - verknüpft. Dieser kombinierte Hochrechnungs- und Gewichtungsfaktor ist selbstverständlich auch dann anzuwenden, wenn Auswertungen nicht auf der Haushaltsebene, sondern auf der Ebene von Personen oder Wegen erfolgen: Allen Personen eines Haushalts und allen Wegen dieser Personen ist der jeweilige haushaltsspezifische Hochrechnungs- und Gewichtungsfaktor zugeordnet.

Die zur nachträglichen Schichtung (Bildung der Hochrechnungsgruppen) verwendeten Merkmale müssen folgende Eigenschaften besitzen:

- a) Die Schichtungsmerkmale sollen in einem möglichst engen Zusammenhang zu den Merkmalen des Verkehrsverhaltens stehen.
- b) Die Besetzungshäufigkeit der Hochrechnungsgruppen in der Grundgesamtheit muß bekannt sein.
- c) Die Besetzungshäufigkeit der Hochrechnungsgruppen in Stichproben muß größer als Null und im übrigen nicht „zu klein“ sein.

Nach den Erkenntnissen der empirischen Verkehrsforschung sind vor allem die Merkmale

- Haushaltsgröße
- Pkw-Besitz
- Gemeindegröße (wegen der Korrelation mit Aktivitäten- und ÖV-Angebot)
- Wochentag

zur nachträglichen Schichtung der Haushalte bzw. Haushaltstage geeignet. Die Verteilung der Haushalte nach Haushalts- und Gemeindegröße steht für die Bundesrepublik Deutschland als Ganzes aus dem Mikrozensus (demnächst auch aus der Volkszählung 1987) zur Verfügung. Der Pkw-Besitz wurde dagegen bisher im Mikrozensus nicht erfaßt, auch nicht in der Volkszählung 1987. Auf Daten der Einkommens- und Verbrauchsstichprobe (EVS) zurückzugreifen kann wegen der Besonderheiten des EVS-Stichprobenverfahrens nur mit einigen Einschränkungen empfohlen werden.

Bei statistischen Datenanalysen zeigt sich regelmäßig, daß zwar die genannten Gruppierungsmerkmale mit den meisten Verkehrsverhaltensmerkmalen in einem eindeutigen Zusammenhang stehen, daß ihr Beitrag zur Erklärung der Variabilität der Verhaltensmerkmale vielfach aber doch nur begrenzt ist. Insofern sollten die Erwartungen hinsichtlich des Schichtungseffekts generell nicht zu hoch angesetzt werden. Eine differenzierte Beurteilung muß getrennt nach Verhaltensmerkmalen erfolgen. Hingewiesen sei auch noch auf praktische Probleme, die dadurch entstehen, daß bei einem - wenn auch meist kleinen - Teil der Haushalte die Angaben zu einzelnen Gruppierungsmerkmalen fehlen (z. B. keine Angaben zum Pkw-Besitz).

Bisher wurde stets davon ausgegangen, daß durch nachträgliche Schichtung die gemeinsame Verteilung der Schichtungsmerkmale angepaßt wird. Wenn nun aber diese gemeinsame Verteilung nicht bekannt ist, und nur die jeweiligen Randverteilungen der einzelnen Schich-

tungsmerkmale vorliegen, so kann man z. B. im Fall von drei Schichtungsmerkmalen daran denken, anstelle von

$$N(g, h, i) / \hat{H}(g, h, i)$$

das Produkt

$$(3.2.3) \quad \frac{N(g, \cdot, \cdot)}{\hat{H}(g, \cdot, \cdot)} \cdot \frac{N(\cdot, h, \cdot)}{\hat{H}(\cdot, h, \cdot)} \cdot \frac{N(\cdot, \cdot, i)}{\hat{H}(\cdot, \cdot, i)}$$

als Gewichtungsfaktor zu verwenden. Man spricht hier auch von „Gewichten mit Hilfe marginaler Häufigkeiten“. In (3.2.3) bedeutet das Symbol „ \cdot “ Summation über den entsprechenden Index.

Das Gewichten mit Hilfe marginaler Häufigkeit ist in der Praxis der Markt- und Meinungsforschung durchaus üblich. Von dieser Gewichtungstechnik muß aber unbedingt abgeraten werden, da immer dann, wenn die Antwortbereitschaft der Haushalte von Hochrechnungsgruppe zu Hochrechnungsgruppe variiert, die Verwendung von (3.2.3) anstelle von (3.2.2) zu Verzerrungen der Schätzung führt. Vgl. hierzu auch Abschnitt 7.1.

Auswertungen von Haushaltsbefragungen vom KONTIV-Typ erfolgen häufiger auf der Personen- als auf der Haushaltsebene. Es kann deshalb durchaus sinnvoll sein, den in die Stichprobe gelangten Personen Gewichtungsfaktoren zuzuordnen, welche sicherstellen, daß die hochgerechnete Stichprobenverteilung bestimmter soziodemographischer Personenmerkmale mit der entsprechenden Zensusverteilung übereinstimmt.

Möchte man beispielsweise die gemeinsame Verteilung der Personenmerkmale Alter und Geschlecht anpassen, so sind zunächst entsprechende Hochrechnungsgruppen in Form von Alters- und Geschlechtskohorten zu bilden. Die Zensuszahl der Personen in der i -ten Kohorte bezeichnen wir mit $M(i)$, die entsprechende hochgerechnete Zahl mit $\hat{P}(i)$. Bei der Hochrechnung der Besetzungshäufigkeiten der einzelnen Alters- und Geschlechtskohorten sind natürlich die haushaltsspezifischen Hochrechnungs- und Gewichtungsfaktoren zu verwenden. Alle Personen in der Stichprobe, welche zur i -ten Kohorte gehören, erhalten den Gewichtungsfaktor

$$(3.2.4) \quad M(i) / \hat{P}(i)$$

Bei Auswertungen auf der Personenebene ist der Personengewichtungsfaktor (3.2.4) multiplikativ mit dem haushaltsspezifischen Hochrechnungs- und Gewichtungsfaktor zu einem Gesamtfaktor zu verknüpfen. Der einer Person zugeordnete Gesamtfaktor ist also davon abhängig, aus welchem Haushalt die betreffende Person stammt und zu welcher Alters- und Geschlechtskohorte sie gehört. Der Vollständigkeit halber sei auch an dieser Stelle erwähnt, daß der personenspezifische Hochrechnungs- und Gewichtungsfaktor (Gesamtfaktor) auch allen Wegen der Person zugeordnet wird.

Die Anpassung auf der Personenebene kann auch zeitliche und räumliche Merkmale mit einbeziehen. Selbst wenn auf der Ebene der Haushaltstage beispielsweise die Verteilung des Merkmals „Wochentag“ angepaßt ist, ist dadurch nicht sichergestellt, daß dies auf der Ebene der Personentage auch innerhalb jeder einzelnen Alters- und Geschlechtskohorte gilt. Ähnliches ist hinsichtlich der Gemeindegrößenklasse zu erwarten. Nachträgliche Schichtung ist

auf der Wegebene nicht möglich, da keine Zensusverteilungen von Wegmerkmalen vorliegen.

Durch nachträgliche Schichtung zur Anpassung der Verteilung von Personenmerkmalen kann es bei Auswertungen zu gewissen Inkonsistenzen kommen, auf die hingewiesen werden muß. Wenn beispielsweise das Total des Haushaltsmerkmals „Zahl der Personen ab 10 Jahren im Haushalt“ unter Verwendung der haushaltsspezifischen Hochrechnungs- und Gewichtungsfaktoren geschätzt wird, so stimmt das Hochrechnungsergebnis im allgemeinen natürlich nicht genau mit der aus der amtlichen Statistik bekannten Gesamtzahl der Einwohner ab 10 Jahren überein. Wertet man die Stichprobe dagegen auf der Personenebene aus, so ist nach entsprechender Personengewichtung gemäß (3.2.4) diese Übereinstimmung gegeben. Um zu einer Schätzung der durchschnittlichen Haushaltsgröße (Personen ab 10 Jahren) zu kommen, sollte man hier also nicht den Mittelwert des Haushaltsmerkmals „Zahl der Personen ab 10 Jahren im Haushalt“ bilden, sondern vielmehr auf Personenebene die Gesamtzahl der Personen ab 10 Jahren hochrechnen und diesen hochgerechneten Wert anschließend durch die Zahl der Haushalte dividieren.

[Fortsetzung in Heft 1/1990]

Literatur

- Barnett, V.: Elements of Sampling Theory, Hodder and Stoughton, London, 1984
- Böltkén, F.: Auswahlverfahren, Teubner-Verlag, Stuttgart, 1976
- Brög, W., Meyer, N.: From Price Elasticity to Price Susceptibility, Socialdata, München, 1978
- Brög, W., Hewinkel, D., Neumann, K.H.: Psychological Determinants of User Behaviour, 34th Round Table on Transport Economics, CEMT, Paris, 1976
- Cochran, W.G.: Sampling Techniques, Wiley, New York, 1977
- Deming, W.E.: Sampling Design in Business Research, Wiley, New York, 1960
- Hansen, M.H., Hurwitz, W.N.: The Problem of Non-Response in Sample Surveys, Journal of the American Statistical Association, Vol. 41, 1946, S. 517-529
- Hansen, M.H., Hurwitz, W.N., Madow, W.G.: Sample Survey Methods and Theory, Bd. I und II, Wiley, New York, 1953
- Hautzinger, H.: Pkw-Besitz der deutschen Wohnbevölkerung: Ein kritischer Vergleich von EVS- und KONTIV-Ergebnissen, Internationales Verkehrswesen, 38. Jg., Heft 3, 1986, S. 184-188
- Hautzinger, H.: Statistische Designprinzipien für kontinuierliche Verkehrserhebungen, Schriftenreihe der DVWG, Band B97, Bergisch Gladbach, 1987, S. 5-25
- Hautzinger, H., Schorer, B.: Statistik-Programmsystem zur Schätzung von Merkmalssummen und Verhältnis-zahlen sowie deren Varianzen und Variationskoeffizienten (1985), in: Hautzinger und Tassaux (1989), Methoden-anhang
- Hautzinger, H., Stenger, H.: Hochrechnungsverfahren zur Infratest-Erhebung „Reiseausgaben im Ausland“, Gutachten, Heilbronn, 1986
- Hautzinger, H., Tassaux, B.: Verkehrsmobilität und Unfallrisiko in der Bundesrepublik Deutschland, Schlussbericht zum BASt-Forschungsprojekt FP 8303, Forschungsberichte der BASt, Band 195, Bergisch Gladbach, 1989
- Hendricks, W.A.: Adjustment for Bias by Non-Response in Mailed Surveys, Agr. Econ. Res., Vol. 1, 1949, S. 52-56
- Hengst, M.: Einführung in die mathematische Statistik und ihre Anwendung, Mannheim, 1967
- Herz, R.: Diskussionspapier zur Projektgruppensitzung „Verbesserung der Gewichtung und Hochrechnung von Haushaltsbefragungen vom KONTIV-Typ“, unveröffentlichtes Manuskript, Karlsruhe, 1987
- Krug, W., Nourney, M.: Wirtschafts- und Sozialstatistik: Gewinnung von Daten, Oldenbourg-Verlag, München, 1982
- Kutter, E., Holz, R.: Diskussionspapier zur Projektgruppensitzung „Verbesserung der Gewichtung und Hochrechnung von Haushaltsbefragungen vom KONTIV-Typ“, unveröffentlichtes Manuskript, Berlin, 1987
- McCarthy, P.J.: Replication: An Approach to the Analysis of Data from Complex Surveys, Nation Center for Health Statistics, Washington, D.C., Series 2, 14, 1966
- McCarthy, P.J.: Pseudo Replication: Half-Samples, Rev. Int. Stat. Inst., 37, 1969, S. 239-264
- Morris, N.: The Accuracy of Commuter's Reported Walking Times: A Comparison of First World and Third World Commuters. Beitrag zur Konferenz "New Survey Methods in Transport", Hungerford Hill (Australien), 1983
- Roth, E.: Sozialwissenschaftliche Methoden, Oldenbourg-Verlag, München, 1987
- Socialdata: Marktanalyse des Verkehrsverhaltens im Personenfernverkehr, Expertenband, München, 1981
- Socialdata: KONTIV 82 - Bericht, München, 1984
- Socialdata: Erhebungen zur Ermittlung von Fußwegen und "non-reported trips", Untersuchung im Auftrag des BMV, München, 1986
- Sozialforschung Brög: KONTIV - Methoden, Band I, München, 1975
- Statistisches Bundesamt (Hrsg): Stichproben in der Amtlichen Statistik, Kohlhammer-Verlag, Stuttgart, 1960
- Stenger, H.: Varianzschätzung bei komplexen Stichprobenerhebungen, Schriftenreihe der DVWG, Band B81, Bergisch Gladbach, 1985, S. 1243-150
- Stenger, H.: Stichproben, Physica-Verlag, Heidelberg/Wien, 1986
- Stenger, H.: Zur heutigen Gewichtungspraxis in Markt- und Meinungsforschung, Manuskript (erscheint demnächst), 1986
- Streckler, H. u.a.: Messung der Antwortvariabilität auf Grund von Erhebungsmodellen mit Wiederholungszählungen, Vandenhoeck & Ruprecht, Göttingen, 1983
- Swov (Hrsg.): The Extent to Which People Can Estimate Their Trip Distances, Voorburg (Niederlande), 1978
- Wermuth, M.: Datenerhebung und Datenanalyse im Verkehrswesen, Schriftenreihe der DVWG, Band B66, 1983, S. 13-43

- Wermuth, M.*: Schriftliche Haushaltsbefragungen zum Verkehrsverhalten – Stichprobenauswahl und Nonresponse-problematik, Schriftenreihe der DVWG, Band B81, Bergisch Gladbach, 1985, S. 151-175
- Wermuth, M.*: Theoretische Anforderungen in der Stichprobenplanung, Schriftenreihe der DVWG, Band B85, Bergisch Gladbach, 1986, S. 137-163
- Wermuth, M., Maerschalk, G.*: Zur Repräsentanz schriftlicher Haushaltsbefragungen zum Verkehrsverhalten, Untersuchung im Auftrag des BMV, München, 1981
- Wermuth, M., Maerschalk, G.*: Entwicklung von Verfahren für Verkehrserhebungen aus der Sicht der Empirischen Sozialforschung und der Ingenieurwissenschaften, Untersuchung im Auftrag des BMV, München, 1983