

2.3.3.3. Stichproben-Umfang

Zu den wichtigen Stichproben-Entscheidungen gehören auch die Überlegungen zum Stichproben-Umfang. Um die Fragen danach, wie viele Teilnehmer eine Studie haben sollte, beantworten zu können, sind eine Reihe von Informationen nötig. In der Realität sind diese meist nicht bekannt, sondern einige davon - wie z. B. die Kenntnisse von den Verteilungen bestimmter Merkmale - werden erst durch die geplanten Studien ermittelt. Der Stichprobenfehler, der z. B. aus den Verweigerungsraten entsteht, kann zwar aufgrund vorheriger Erfahrungen abgeschätzt werden, die exakte Größe ergibt jedoch erst die Erhebung. So erfordern exakte Stichproben-Berechnungen im Prinzip aufwendige Vorstudien.

Ein wesentlicher Sinn der Überlegungen zur Stichprobengröße liegt darin, ein ökonomisches Optimum herauszufinden. Je größer die Stichprobe, desto repräsentativer sind die Ergebnisse. Doch die Erhöhung der Aussagefähigkeit einer Studie bedeutet gesteigerte Kosten. Eine zu knappe, sparsame Auswahl vermindert den Wert der Arbeit, eventuell ist die Mühe ohne jede Aussagekraft, dann wurde nichts gespart, sondern alles verschwendet.

Es gibt eine Reihe von allgemeinen Grundsätzen, die bei den Überlegungen zur Abschätzung des Stichprobenumfanges berücksichtigt werden sollten:

- Der Stichproben-Umfang wächst mit der Zahl der zu untersuchen den Variablen bzw. Merkmale. Hierbei kann folgende "Faustregel" angewandt werden. Im Hinblick auf das Untersuchungsziel und den dazu notwendigen Merkmalsgegenüberstellungen sollten in Kreuztabellen in jedem Feld mindestens 10 Fälle erscheinen. Für ein relativ einfaches Beispiel aus der Ernährungsepidemiologie mit nur fünf Merkmalen, die jeweils auch nur in zwei Ausprägungen auftreten - d. h. bei dichotomer Verteilung von z. B. Nahrungsmenge (Viel- und Wenig-Esser), Geschlecht (weiblich/männlich), Alter (jung/alt), Gewicht (Normal- und Übergewicht) und Konstitution (lepto- und pyknomorph) - ergeben sich 32 Zellen; so sollte der entsprechende Stichprobenumfang nicht 320 Personen unterschreiten. Dabei sind optimal-utopische Annahmen bereits die Voraussetzung, dass sich die Merkmale gleich verteilen und dass es keine Ausfälle gibt.

Die allgemeine Formel für eine Mindest-Stichprobengröße N lautet:

$$N = 10 \times K^v$$

wobei

K = Zahl der Merkmalsausprägungen ist; dichotom = 2

v = Zahl der betrachteten Merkmale bei der Hypothesen-Prüfung, im obigen Beispiel = 5

- Je breiter die Merkmalsausprägungen der Variablen sind, je größer die Variationsbreiten der Merkmale in der Grundgesamtheit sind und bei größerem Messfehler, umso größer muss die Stichprobe werden.

- Je kleiner der zu erwartende Unterschied, je genauer die Hypothese geprüft werden soll, d. h. mit höherer statistischer Absicherung, desto größer muss die Stichprobe werden.

Für die Berechnung der Stichprobengröße gibt es eine Reihe von Formeln, deren Ursprünge in einfachen Wahrscheinlichkeitsberechnungen liegen, wie z. B. der Normalverteilung von einem Merkmal. Diese Situation ist weit von der entfernt, die für ernährungsepidemiologische Studien charakteristisch ist. Selbstverständlich hat sich dieser mathematische Spezialzweig weiterentwickelt, doch dies kann hier nicht dargestellt werden. Es wird auf die entsprechende Fachliteratur hingewiesen (z. B. Hall 1983; Leiberich 1981; Tull, Hawkins 1980).

Wie selbst aus den einfachen Abschätzungen leicht nachzuvollziehen ist, werden recht schnell große Stichprobenumfänge erreicht. Daraus lässt sich leicht ableiten, welche enorme Erhebungskosten dafür notwendig wären; sie übersteigen in der Regel deutlich die

verfügbaren Mittel. So begrenzen die ökonomischen Gegebenheiten die Stichprobengröße und damit aber auch die Aussage-Qualität der geplanten Studie. Als absolute Untergrenze sollte jedoch die bereits erwähnte Faustregel (s. o.) gelten, dass die notwendigen Matrix-Zellen mit nicht weniger als 10 Fällen besetzt werden sollten; darunter sind keine Aussagen mehr aus den Datenmaterial abzuleiten.

Die Erfahrungen für die Bundesrepublik und ähnlich großer anderer Länder zeigen, dass ein guter Kompromiss zwischen Aufwand und Stichprobenfehler bei einer Stichprobengröße $N = 2000$ liegt, um damit noch repräsentative Aussagen ableiten zu können. Eine im folgenden wiedergegebene Überlegung kommt zur selben Größenordnung. Für explorative statistische, insbesondere multivariate Auswertungen ist eine Mindestgröße von 50 - besser noch 100 - anzusetzen. Wenn also aus einer Grundgesamtheit eine entsprechende Anzahl von interessanten Konstellationen identifiziert werden sollen und die Grenze in Analogie zur statistischen Signifikanzgrenze oder auch zur Prozentklausel bei politischen Wahlen bei 5% festgelegt wird, so wird auch eine Stichprobengröße von 1000-2000 erreicht (Bodenstedt, Oltersdorf et al 1983). In größeren Ländern - wie z. B. den USA, und wenn die Merkmale noch detaillierter dargestellt werden sollen, z. B. weitere regionale und auch zeitliche bzw. saisonale Verteilungen, dann muss von noch größeren Stichprobenumfängen ausgegangen werden. Dazu einige Beispiele:

- bei HANES - der National Health and Nutrition Examination Survey der USA - beträgt der Stichprobenumfang über 20.000;
- beim Mikrozensus und auch bei der Einkommens- und Verbrauchsstichprobe (EVS) des statistischen Bundesamtes Wiesbaden sind 50.000 Haushalte beteiligt; bei der EVS schreiben im monatlichen Wechsel je 4000 Haushalte detailliert - sog. "Feinanschreibung" - ihre Haushaltsausgaben auf;
- die Nationale Verzehrsstudie in der Bundesrepublik umfasste 10.800 Privathaushalte mit ca. 24.000 Personen;
- die Begleit- und Evaluierungsstudien zur Deutschen Herz-Kreislauf-Präventionsstudie (DHP) haben im regionalen Maßstab 7200 Personen und bei der nationalen Erhebung 16.000 Studienteilnehmer;
- die berühmte Framingham-Studie (bei Boston/USA) hat 6000 Teilnehmer;
- und die Lipid Research Clinics Program Population Studies (LRCPP) in den USA hat über 18.000 Studienteilnehmer.