

2 Allgemeine Theorie von Random-Stichproben

2.1 Grundgesamtheit

Nehmen wir an, der Marketingleiter eines Zeitschriftenverlags möchte wissen, wie groß die Leserschaft seiner Zeitschriften im Vergleich zu seiner unmittelbaren Konkurrenz ist. Die plausibelste Vorstellung, wie diese Information zu erhalten ist, besteht darin, die gesamte Bevölkerung zu beobachten bzw. zu fragen, ob und welche Zeitschriften sie lesen. Warum man diesen Weg der vollständigen Erfassung der Gesamtheit nicht geht, sondern eine Stichprobenerhebung vorzieht, wurde schon erwähnt: Sowohl die Kosten als auch der Zeitrahmen sprechen dagegen. Man erhebt also die gewünschte Information bei einem ausgewählten Teil der Gesamtheit, man führt eine Stichprobenerhebung durch.

Die zugrundeliegende Gesamtheit, die Grundgesamtheit, muss jedoch vorher exakt beschrieben werden. Nur so kann eine Stichprobe ein adäquates Abbild der Grundgesamtheit liefern. Das ist zwar formal einfach – die Beschreibung besteht aus einem **sachlichen**, einem **räumlichen** und einem **zeitlichen Definitionsteil** – aber nicht immer trivial.

2.1.1 Räumlicher Definitionsteil

Fangen wir mit dem räumlichen Definitionsteil an: Da das Hauptverbreitungsgebiet der Zeitschriften unseres Marketingleiters die Bundesrepublik Deutschland ist, sollen auch nur **Leser, die in der Bundesrepublik wohnen**, erfasst werden. Ausgeschlossen sind also Personen, die kurzfristig oder auf Dauer ihren Wohnsitz außerhalb der Bundesrepublik haben, auch wenn sie vielleicht auf der Durchreise die Zeitschrift gekauft haben oder sie sich ins Ausland zusenden lassen.

2.1.2 *Sachlicher Definitionsteil*

Der sachliche Definitionsteil macht schon größere Schwierigkeiten: Wer gehört z.B. zu der potentiellen Leserschaft von Zeitschriften? Grundsätzlich sind das alle Personen, die in deutscher Sprache lesen können, d.h. auch Kinder und Personen in Anstaltshaushalten, aber keine Ausländer, die nicht Deutsch sprechen/lesen können. Nun macht es jedoch wenig Sinn, Personen in Anstaltshaushalten, die ihr soziales und wirtschaftliches Umfeld nicht selbst bestimmen können, sowie Kinder, in die Grundgesamtheit einzubeziehen. (Bei Kindern bestehen zudem Einschränkungen und besondere Regeln für eine Befragung.) Man wird sich also auf **deutsch-sprechende Personen ab 14 Jahren in Privathaushalten** konzentrieren. Bei Umfragen der Wahlforschung wird man die Grundgesamtheit auf alle wahlberechtigten Personen, i.d.R. also Deutsche ab 18 Jahren, eingrenzen. Und bei einem Untersuchungsziel, das sich mit den Lebensumständen der Landbevölkerung auseinandersetzt, wird man die Landbevölkerung definieren müssen.

Es würde zu weit führen, hier alle Implikationen anzuführen, die mit der Abgrenzung unterschiedlicher Grundgesamtheiten in der Bevölkerung verbunden sind. Wichtig ist zu erkennen, dass die Beschreibung der Grundgesamtheit immer am Ziel der jeweiligen Erhebung orientiert ist.

2.1.3 *Zeitlicher Definitionsteil*

Auch der zeitliche Definitionsteil ist nicht unproblematisch. Die Bevölkerung verändert sich ständig – 1.000 bis 2.000 Personen eines Jahrgangs feiern täglich Geburtstag, beginnen also ein neues Lebensjahr, und noch etwas mehr sterben an einem Tag – so dass man von „Bevölkerung“ nur im Sinne einer exakt zeitpunktbezogenen Definition sprechen kann.

Aber auch, wenn ein zeitlich exakter Definitionsteil sinnvoll wäre, ist niemand in der Lage, die Bevölkerung oder z.B. die Leser einer Zeitschrift zu einem bestimmten Zeitpunkt (z.B. am 16.02.2013 um 09.35.27 Uhr) exakt abzubilden. Man geht deshalb davon aus, dass die Veränderungen einer Grundgesamtheit während eines relativ kurzen Zeitraums nur marginal sind. Ein daran orientierter zeitraumbezogener Definitionsteil kann dann z.B. ein Monat sein, oder – an der Praxis ausgerichtet – lauten: **Während der Feldzeit der Untersuchung**. Es ist klar, dass die mit einer solchen Zeitraumorientierung erhobenen Informationen Durchschnittscharakter haben. So werden Untersuchungen auch ganz bewusst über einen längeren Zeitraum ausgedehnt, um z.B. saisonale Schwankungen auszugleichen.

2.1.4 *Adäquationsproblem*

Oben wurde geschrieben, dass eine Stichprobe ein adäquates Abbild der Grundgesamtheit liefern soll. Jetzt ist erkennbar geworden, dass bereits bei der Definition der Grundgesamtheit Adäquationsprobleme auftreten können. Tatsächlich wird man sich mit solchen Problemen immer auseinandersetzen müssen, weil die theoretische Definition einer Grundgesamtheit niemals so formuliert werden kann, dass sie sich in der Praxis exakt erreichen lässt.

Adäquationsprobleme (= Auseinanderklaffen der theoretischen Ziele einer „Statistik“ und deren Realisierung in der Praxis) treten bei der Bildung einer Stichprobe im Prinzip in allen Auswahlstufen und in allen Bereichen auf. Das gilt nicht nur bei der Realisierung einer Grundgesamtheit im Vergleich zur theoretischen Grundgesamtheit, wie oben dargestellt, sondern auch z.B. zwischen der theoretischen Grundgesamtheit und den real existierenden Auswahlgrundlagen, wie im nächsten Abschnitt beschrieben wird; also im Grunde genommen für die realisierte Stichprobe im Vergleich zur theoretischen Stichprobe. Fortgesetzt wird das Adäquationsproblem dann im weiteren Forschungsverlauf mit der Bildung eines Hypothesenraums und dem realisierten Abfrageinstrument sowie theoretischen Messparametern und realisierten Messgrößen. Auch Reliabilität und Validität von Daten versus einer i.d.R. einmaligen Messung stellen ein Adäquationsproblem dar, das dann mit dem theoretischen Repräsentationsschluss im Vergleich zu der in der Praxis durchführbaren Hochrechnung den Abschluss findet.

2.2 *Auswahlgrundlage*

2.2.1 *Adäquation von Auswahlgrundlagen*

Zur „Realisierung“ der Grundgesamtheit und damit als Auswahlgrundlage bieten sich bei der Definition „Bevölkerung ab x Jahren in der Bundesrepublik“ die Einwohner-Melderegister an. Leider sind diese Bevölkerungsregister, die alle Personen mit Wohnsitz in der BRD enthalten sollen, nicht vollständig und nicht aktuell. Das liegt nicht nur an der zuweilen laxen Einstellung zur Meldepflicht (vor allem bei Ausländern und jungen Personen in Ausbildung), sondern auch am Aufwand, der betrieben werden muss, bis eine Änderung ins Melderegister eingetragen ist. Zwar sind im Zusammenhang mit der Volkszählung 2011 die Melderegister „ertüchtigt“, also mit einer Reihe von Maßnahmen verbessert worden, man muss aber trotzdem auch jetzt noch damit rechnen, dass die Melderegister hinter der Realität „hinterherhinken“, dass also Geburten, Sterbefälle, Zu- und Fortzüge erst nach einer gewissen Zeit im Melderegister realisiert sind. Und zu-

sätzlich enthalten die Register alle Personen, die in der Gemeinde mit Haupt- oder Nebenwohnsitz gemeldet sind; Personen mit mehreren Wohnsitzen sind also in mehreren Dateien aufgeführt.

Das Melderegister ist nur ein Beispiel für die Probleme, die grundsätzlich immer mit der Realisierung einer Grundgesamtheit verknüpft sind: Register haben leider grundsätzlich einen Time-lag. Hinzu kommen wesentlich gravierendere inhaltliche Probleme von Registern: Einzelne Gruppen der Grundgesamtheit können beispielsweise unter- oder überrepräsentiert sein, andere sind vielleicht gar nicht enthalten, schließlich können Gruppen in den Registern enthalten sein, die gar nicht zur Grundgesamtheit gehören.

Auch hierfür können die Melderegister als Beispiel dienen, wenn auch deren Qualität durch die „Ertüchtigung“ zugenommen hat:

- Ausländer werden im Melderegister selten völlig korrekt abgebildet. Einerseits halten sich nicht alle konsequent an die Meldevorschriften, so dass beispielsweise häufig die Abmeldung unterlassen wird, wenn man in sein Heimatland zurückkehrt. Die Folge ist ein überhöhter Adressenbestand von Ausländern. Andererseits wird bei einzelnen Ausländergruppen (Einwanderer aus osteuropäischen Ländern oder aus Entwicklungsländern) in vielen Fällen bewusst keine Anmeldung vorgenommen, um die Einwanderung zu verschleiern.
- Junge Personen in Ausbildung oder junge Erwerbstätige geben häufig als Hauptwohnsitz nicht ihren Arbeits- oder Ausbildungsort an, sondern den Wohnsitz der Eltern, weil sie an ihrem derzeitigen Wohnort nur kurze Zeit bleiben werden. Die Folge ist eine Untererfassung der Hauptwohnsitze junger Erwachsener.
- Personen, deren Hauptwohnsitz im Ausland liegt, die aber in der Bundesrepublik arbeiten (erwerbstätige Grenzpendler), sind in den Registern gar nicht enthalten, obwohl sie beispielsweise zu der Grundgesamtheit der Personen gehören würden, die das wirtschaftliche Leben der Bundesrepublik mitgestalten.

Es ließe sich noch eine Fülle ähnlicher Beispiele für die eingeschränkte Abbildungstreue von Einwohner-Melderegistern anführen. Es sind aber nicht diese Details, auf die es hier ankommt, sondern das daraus ableitbare **Adäquationsproblem** zwischen Grundgesamtheit und Auswahlgrundlage. Dieses Problem gibt es nicht nur bei Grundgesamtheiten, die sich auf die Bevölkerung beziehen, sondern bei praktisch allen Grundgesamtheiten mit einer Ausnahme: Die Grundgesamtheit ist als real existierende Adressendatei definiert. Diese Ausnahme ist

aber äußerst selten, denn normalerweise will man eine Gruppe von Personen (z.B. die eigenen Kunden) abbilden, die man in einer Kartei erfasst hat. Auch eine solche Kartei leidet jedoch unter den oben angesprochenen Problemen (z.B. Karteileichen, noch nicht registrierten Neukunden, Umzügen, Todesfällen usw.).

2.2.2 *Einwohner-Melderegister*

Aber selbst wenn man geneigt ist, die Unzulänglichkeiten der Melderegister zu akzeptieren, darf und kann man sie nur in den seltensten Fällen als Auswahlgrundlage verwenden. Der Grund liegt in dem Melderecht der Bundesrepublik. Das schreibt vor, dass Adressen aus dem Melderegister nur dann zur Verfügung gestellt werden dürfen, wenn die Befragung „in öffentlichem Interesse“ ist. Da jedoch die Aufträge der privaten Forschungsinstitute zumeist für die freie Wirtschaft durchgeführt werden, liegen sie außerhalb des (eng verstandenen) öffentlichen Interesses.

Aber auch bei Studien, deren öffentliches Interesse keinem Zweifel unterliegt, ist die Nutzung von Adressen aus Einwohner-Melderegistern nicht einfach. Da sind zum einen die von den Gemeinden erhobenen Gebühren für die Adressen und zum anderen der hohe Zeitbedarf, der für die Ziehung benötigt wird. Hinzu kommt ein erheblicher Aufbereitungsaufwand, da jede Gemeinde eigene Satzformate, Satzinhalte und Dateiformate bei der Lieferung hat. (siehe auch Kapitel 9)

Diese rechtlichen und praktischen Einschränkungen der Nutzung von Einwohnerdateien waren mitentscheidend für die Gestaltung und Einführung des ADM-Stichproben-Systems für Face-to-Face-Befragungen. Dieses ist nämlich – darauf wird später ausführlich eingegangen – unabhängig von den Einwohner-Melderegistern und basiert in wesentlichen Teilen auf Quellen, die qua Gesetz öffentlich zugänglich sind.

2.3 Gütekriterien für Stichproben

2.3.1 *Repräsentativität*

Eine Stichprobe soll – das wurde schon mehrfach erwähnt – ein adäquates Abbild der Grundgesamtheit liefern. Dazu gehören zunächst und unabdingbar eine Definition der Grundgesamtheit und eine dazu adäquate Auswahlgrundlage – wie oben beschrieben. Erst wenn beides vorhanden ist, wird das Abbildungsverfahren relevant.

Was sind nun Verfahren, die adäquate Abbildungen liefern? Was heißt adäquat in diesem Sinn? In der Theorie ist diese Frage einfach zu beantworten: Adäquat ist eine Abbildung, wenn die Grundgesamtheit in allen wesentlichen Details richtig abgebildet wird. Die Abbildung darf zwar vergrößert sein (was offenbar eine Frage der Stichprobengröße ist), sie muss aber die Details ähnlich wie eine Fotografie in den richtigen Proportionen wiedergeben.

In der Stichprobentheorie wurde für diese Abbildungstreue der Begriff Repräsentativität gewählt und häufig wie folgt definiert, „Eine Teilmasse heißt repräsentativ, wenn sie in der Verteilung aller interessierenden statistischen Merkmale der Gesamtmasse entspricht, d.h. ein verkleinertes aber sonst wirklichkeitsgetreues Abbild der Gesamtheit darstellt“ (Statistisches Bundesamt 1960, S. 13).

Der Hintergrund dieser und ähnlicher Definitionen ist leicht zu erkennen: Das Ziel jeder Stichprobenerhebung ist es, aus den bei einer ausgewählten Menge von Befragten gewonnenen Informationen auf die analogen Gegebenheiten (Meinungen, Einstellungen, Sachverhalte) in der Grundgesamtheit zu schließen bzw. hochzurechnen. Wenn die Stichprobe im obigen Sinn repräsentativ ist, kann man diesen Schluss von der Stichprobe auf die Grundgesamtheit (Inferenz-, Repräsentations- oder Rückschluss) natürlich machen.

Man ist also darauf angewiesen, den Stichprobenplan und seine Realisierung so zu gestalten, dass man Repräsentativität erwarten kann. Die angegebene Definition von Repräsentativität ist deshalb im eigentlichen Sinn eine Definition der **Erwartungstreue**. Die Auftraggeber von Forschungsprojekten stellen dem Institut implizit immer die Forderung nach Erwartungstreue, denn sie erwarten, dass die vom Institut gelieferten Ergebnisse auf die Grundgesamtheit unverzerrt übertragen werden können.

Nicht erwartungstreue Stichproben generieren einen sog. **Bias** oder systematischen Fehler. Ihre Ergebnisse streuen also nicht nur zufällig um die gesuchten Parameter der Grundgesamtheit, sondern sie weichen systematisch davon ab. Sämtliche im Zusammenhang mit dem Adäquationsproblem genannten Anstrengungen zielen darauf ab, systematische Fehler zu vermeiden, also erwartungstreue Stichproben zu realisieren.

Die Überprüfung der Repräsentativität einer Stichprobe durch einen Vergleich von Merkmalen aus der Stichprobe mit der Grundgesamtheit ist meistens nur sehr eingeschränkt möglich, da man von der Grundgesamtheit, auf die man schließen will, nur wenige Merkmale kennt. Vor allem kennt man sie nicht in den Merkmalen, Meinungen, Verhaltensweisen, die das eigentliche Untersuchungsziel sind (sonst brauchte man die Untersuchung ja nicht durchzuführen). Zur Prüfung der Repräsentativität benötigt man also andere, eher formale Kriterien, was

bedeutet, dass man sich in erster Linie auf die Methode der Stichprobenbildung stützen muss und erst in zweiter Linie auf deren kontrollierbare Ergebnisse.

Für die Überprüfung der Methode der Stichprobenbildung stellt die Statistik aus der Wahrscheinlichkeitstheorie ein Kriterium zur Verfügung:

Eine Stichprobe ist repräsentativ, wenn jedem Element der Grundgesamtheit die gleiche, berechenbare und von Null verschiedene Inklusionswahrscheinlichkeit zugeordnet werden kann, und die einzelnen Stichprobenelemente zufällig ausgewählt werden. Dabei ist es unerheblich, ob die Nivellierung der Inklusionswahrscheinlichkeiten a priori in einem proportionalen Stichprobenansatz oder a posteriori durch Gewichtung eines disproportionalen Ansatzes erreicht wird.

Wie solche Stichproben realisiert werden, d.h. mit welchen Verfahren und Techniken man vorgehen kann, werden die nächsten Kapitel zeigen.

(Vorab noch ein Hinweis: Der für die Beurteilung von Stichproben und ihren Ergebnissen notwendige formalmathematische Rahmen ist in praktisch jedem Lehrbuch der Stichprobentheorie und -praxis zu finden; einige sind im Literaturverzeichnis genannt. Besondere Erwähnung für den Praktiker im deutschsprachigen Raum verdienen aber nach Meinung der Autoren die Bücher von Esser/Grohmann/Müller/Schäffer 1989 und Krug/Nourney/Schmidt 2001. Sie zeigen anhand konkreter statistischer Erhebungen, welche Probleme auf den Stichprobenpraktiker zukommen können, und liefern in leicht lesbarer Form den theoretischen Hintergrund dazu. Im vorliegenden Buch werden hingegen nur die in der Praxis am häufigsten verwendeten Formeln aufgeführt.)

2.3.2 Genauigkeit

Eine Stichprobe kann – wie schon beschrieben – nur ein Abbild der Grundgesamtheit liefern. Sie liefert also Schätzwerte für die Gegebenheiten in der Grundgesamtheit. Diese Schätzwerte treffen die gesuchten wahren Werte der Grundgesamtheit nur selten exakt, d.h. sie streuen zufällig um diese Werte. Diese Streuung folgt, wie man unterstellt, angenähert bei – fast – stetigen Variablen (z.B. Einkommen, Größe, Miethöhe) der Normalverteilung und bei kategorialen Variablen (z.B. Leser einer Zeitschrift, Geschlecht, Pkw-Besitz) der Hypergeometrischen- oder der Binomial-Verteilung. Die letzteren gehen bei genügend großen Stichproben in die Normalverteilung über, so dass diese die in der Stichprobenpraxis am häufigsten gebrauchte Verteilung ist.

Auf der Basis dieser Verteilung lässt sich berechnen, wie hoch die Wahrscheinlichkeit dafür ist, dass die Stichprobe exakt die gesuchten Werte der Grundgesamtheit trifft; sie ist i.d.R. leider sehr, sehr klein. Viel wichtiger ist deshalb die Möglichkeit zu beurteilen, ob die Stichprobe die Grundgesamtheit adäquat abbilden wird (**Repräsentativität/Erwartungstreue**), und zu berechnen, wie groß der Bereich ist, in dem mit bestimmbarer Wahrscheinlichkeit der gesuchte Wert der Grundgesamtheit liegt (**Präzision**). Diese beiden Kriterien bilden die Messlatte für die Genauigkeit von Stichprobenergebnissen.

2.3.3 Präzision

Im Präzisionsmaß für eine Stichprobe, dem sog. **Konfidenzintervall**, kommt zum Ausdruck, wie genau im Sinne von Zufallsabweichung die Ergebnisse einer Stichprobenerhebung sind. Im nachstehenden Kapitel wird das Konfidenzintervall für einstufige Stichproben beschrieben. Auf das Konfidenzintervall für mehrstufige, häufig mehrfach geschichtete Stichproben wird ausführlich im Kapitel 3.3.1.2 eingegangen.

Das Konfidenzintervall gibt also die Größe des Unsicherheitsbereichs an, mit dem jede Stichprobe behaftet ist, weil in ihr nicht alle Einheiten der Grundgesamtheit vertreten sind. Seine Berechnung basiert auf der Variabilität der einzelnen Elemente in der Grundgesamtheit. Diese wird bestimmt als Varianz der Einzelwerte. Das ist die mittlere quadratische Abweichung der Einzelwerte um ihr gemeinsames arithmetisches Mittel. Die Varianz wird nach folgender Formel berechnet:

$$\text{Varianz der Grundgesamtheit: } \sigma_x^2 = \frac{1}{N} \sum (x_i - \bar{x})^2$$

mit: N = Zahl der Einheiten der Auswahlgrundlage
 x_i = Erhobene Information der Einheit i der Grundgesamtheit
 $\bar{x}$ = Durchschnitt dieser Messwerte in der Grundgesamtheit

Hier und in den weiteren Formeln wird folgende Nomenklatur verwendet:

Großbuchstaben ($N, X, P, \dots$) bezeichnen Werte der **Grundgesamtheit**,
Kleinbuchstaben ($n, x, p, \dots$) bezeichnen Werte der **Stichprobe**.

Tatsächlich kennt man die Varianz der Elemente der Grundgesamtheit nicht, so dass man sie aus der Stichprobe schätzen muss. Das ist sehr einfach, weil re-

präsentative Stichproben, die die Grundgesamtheit in allen wesentlichen Details korrekt abbilden, auch die Varianz richtig schätzen. Die Schätzung für die Varianz der Grundgesamtheit ist also die Varianz der Stichprobe allerdings mit einer um 1 reduzierten Fallzahl. Sie hat folgende Form:

$$\text{Aus der Stichprobe geschätzte Varianz: } s_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

mit: n = Zahl der Einheiten in der Stichprobe

x_i = Erhobene Information der Einheit i der Stichprobe

$\bar{x}$ = Durchschnitt dieser Messwerte in der Stichprobe

Normalerweise sind es nicht die Einzeldaten der Stichprobe, die für den Forscher von Interesse sind, sondern daraus abgeleitete Durchschnitte, Summen oder Anteilswerte und deren Konfidenzintervalle. Auch diese ergeben sich prinzipiell aus der unbekannten Varianz in der Grundgesamtheit, müssen also ebenfalls aus der Varianz der Stichprobe geschätzt werden. Auch hier wird auf eine Ableitung der Formeln verzichtet:

$$\text{Varianz arithmetisches Mittel: } s_{\bar{x}}^2 = \frac{1}{n} \cdot \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n} s_x^2$$

$$\text{Varianz Anteilswert: } s_p^2 = \frac{1}{n-1} p (1-p)$$

mit: n = Zahl der Einheiten in der Stichprobe

p = In der Stichprobe erhobener Anteilswert

Man sieht, dass die Varianz direkt von der Variabilität der Merkmalsträger bzw. von der Größe des Anteilswertes abhängt und umso kleiner wird, je größer die Stichprobe ist. Damit wird auch deutlich, dass die Größe der Stichprobe nichts mit ihrer Repräsentativität zu tun hat, wohl aber mit der Messpräzision zusammenhängt. Die Frage danach, ab welcher Fallzahl eine Stichprobe repräsentativ sei, ist demnach falsch gestellt, dagegen kann man durchaus danach fragen, welchen Gewinn an Präzision man durch Vergrößerung der Stichprobe erhält.

Alle bisher verwendeten Formeln beziehen sich auf sog. **Stichproben mit Zurücklegen** und reine Zufallsauswahl. Das sind Stichproben, bei denen jede Einheit der Auswahlgrundlage **bei jedem Ziehungsvorgang** die gleiche Chance hat, in die Stichprobe zu gelangen. Eine gezogene Einheit wird registriert aber nicht aus der Auswahlgrundlage entfernt, so dass sie beim nächsten Ziehungsvorgang erneut gezogen werden kann.

Häufiger in der Markt- und Meinungsforschung sind **Stichproben ohne Zurücklegen**. Bei diesem Verfahren werden einmal gezogene Einheiten nicht mehr zurückgelegt, sie haben also auch keine Mehrfachchance gezogen zu werden (Denn sonst bestünde ja die Möglichkeit, mehrfach die gleiche Person in der Stichprobe zu befragen). Solche Stichproben haben prinzipiell eine geringere Varianz als Stichproben mit Zurücklegen, wie man sich unschwer mit einem Gedankenexperiment erklären kann: Würde man die Stichprobe aus einer endlichen Auswahlgesamtheit immer mehr vergrößern, bis alle Einheiten der Auswahlgrundlage gezogen sind, so würde diese Stichprobe die Auswahlgesamtheit vollständig, d.h. ohne jede zufallsbedingte Varianz abbilden. Würde man zwar nicht die ganze, wohl aber einen sehr großen Teil der Auswahlgrundlage ziehen, wäre die zufallsbedingte Varianz immer noch recht klein. Die Größe des Zufallsfehlers hängt also offenbar von dem Auswahlatz ab. Tatsächlich gibt es einen sog. Korrekturfaktor für Ziehung ohne Zurücklegen, der folgende Form hat:

$$\text{Korrekturfaktor der Varianz bei Ziehung ohne Zurücklegen: } k = 1 - \frac{n}{N}$$

Die Quadratwurzel der Varianz ist die sog. Standardabweichung. Diese hat bei der Normalverteilung eine besondere Eigenschaft: Sie lässt Aussagen über die Wahrscheinlichkeit zu, mit der ein Stichprobenergebnis in einem aus der Standardabweichung und einem Multiplikator gebildeten Bereich um den analogen Wert der Grundgesamtheit liegen wird. Dieser sog. direkte Schluss (von der Grundgesamtheit auf die Stichprobe) lässt sich bei genügend großen Stichproben auch umkehren in den indirekten Schluss (von der Stichprobe auf die Grundgesamtheit). Es lässt sich also eine nur von einem Multiplikator (t) abhängige Wahrscheinlichkeit dafür angeben, dass der gesuchte Wert der Grundgesamtheit in einem bestimmten, von der Standardabweichung (s) und dem Multiplikator (t) gebildeten Intervall um das analoge Stichprobenergebnis liegt. Das ist das **Konfidenzintervall** (hier formuliert für Stichproben ohne Zurücklegen), mit dem folgende Aussage verbunden ist: „Die Wahrscheinlichkeit dafür, dass das gesuchte arithmetische Mittel/der gesuchte Anteilswert in der Grundgesamtheit im Bereich t mal Standardabweichung um das analoge Stichprobenergebnis liegt, ist eine Funktion von t .“

Formal wird diese Aussage folgendermaßen ausgedrückt:

$$\text{Arithmetisches Mittel: } W \{ \bar{x} - ts_{\bar{x}} \leq \bar{X} \leq \bar{x} + ts_{\bar{x}} \} = f(t)$$

$$\text{mit } s_{\bar{x}} = \sqrt{\frac{1}{n} \cdot \frac{1}{n-1} \sum (x_i - \bar{x})^2} \sqrt{1 - \frac{n}{N}}$$

Stichproben-Verfahren in der Umfrageforschung

Eine Darstellung für die Praxis

2014, XVIII, 292 S. 20 Abb., Softcover

ISBN: 978-3-531-16445-8