Einfache Zufallsstichprobe - Teil 1

Ist die Untersuchung aller Kinder einer definierten Grundgesamtheit mangels Ressourcen oder aus anderen Gründen nicht möglich und möchte man dennoch Informationen z.B. über den Kariesbefall dieser Kinder bekommen, so kann man eine Stichprobe, d.h., eine Teilmenge dieser Grundgesamtheit untersuchen. Dies ist normalerweise billiger und schneller durchführbar und kann unter bestimmten Umständen sogar genauer sein, da mehr Zeit für die Erhebungen zur Verfügung steht und damit systematische Fehler (Fehldiagnosen) eher vermieden werden. Die aus der Stichprobe ermittelten Zielparameter (z.B. der mittlere dmft oder der Anteil kariesfreier Kinder) sind allerdings nur Schätzungen der entsprechenden Werte der Grundgesamtheit und daher mit einer gewissen Ungenauigkeit behaftet, die sich aus zufälligen und systematischen Fehlern zusammensetzt. **Systematische Fehler** (engl. Bias) entstehen z.B. dann, wenn die Stichprobe nicht repräsentativ für die Grundgesamtheit ist (z.B. bei Untersuchungen von 6-Jährigen ausschließlich in Kindergärten) oder wenn bei schlechten Lichtverhältnissen besonders viel Karies übersehen wird. Sie lassen sich zwar minimieren (z.B. durch bessere Ausleuchtung) aber praktisch nie vollständig vermeiden. **Zufällige Fehler** resultieren hingegen aus biologischer oder messtechnischer Variabilität. Obwohl die einfache Zufallsstichprobe von Jugendzahnärzten selten oder nicht verwendet wird, sind die dort entwickelten Methoden zum Verständnis der anderen Stichprobenverfahren (z.B. Clusterstichproben) sehr nützlich.

Mittelwert und Varianz der Grundgesamtheit (Populationsvarianz)

Die Grundgesamtheit ist eine definierte, abgegrenzte Menge von Individuen (oder Objekten), über die eine Aussage gemacht werden soll hinsichtlich eines Merkmals (oder mehrerer Merkmale), das jedes Individuum trägt. Die Ausprägungen des Merkmals (qualitativ oder quantitativ) werden im Rahmen von Untersuchungen der Individuen dokumentiert. Eine Grundgesamtheit sollte so definiert werden, dass prinzipiell alle dazugehörigen Individuen auch beobachtbar sind. So ist eine Grundgesamtheit der 3- bis 6-Jährigen während eines Schuljahres nicht so einfach beobachtbar, da die 6-Jährigen sowohl in Kindergärten, als auch in Grundschulen anzutreffen sind. Dagegen ist die Grundgesamtheit der 3- bis 5-Jährigen in Deutschland durch Untersuchungen der Kinder in Kindergärten gut erreichbar.
Als **Beispiel** betrachten wir zunächst den dmft als (quasi)metrisches Merkmal in einem realen Datensatz der Reihenuntersuchung einer regionalen Grundgesamtheit von 160 Kindergärten eines Landkreises mit N = 7737 Kindern im Alter von 3 bis 5 Jahren. Systematische Fehler infolge Fehldiagnosen der Zähne sollen in diesem Beitrag unberücksichtigt bleiben. Abb. 1 zeigt die typische Häufigkeitsverteilung des dmft für diese Altersgruppe. Der mittlere dmft der Grundgesamtheit wird hier mit μ bezeichnet. Er errechnet sich aus der Formel

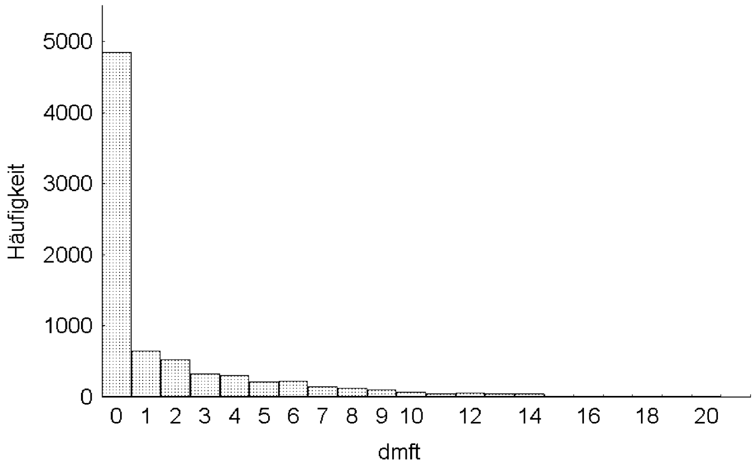
$$\mu = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i \qquad \text{(F1)}$$

und beträgt μ = 1,66. Dabei stehen die x_i für die einzelnen dmft-Werte der 7737 Kinder. Sie liegen in mehr oder weniger großem Abstand zu μ, d.h., die Einzelwerte „streuen“ um den Mittelwert.
Als Maß für diese Streuung der 7737 Einzelwerte um ihren Mittelwert ist die so genannte Varianz σ^2 üblich, die man für die Grundgesamtheit aus der Formel F2 berechnet,

$$\sigma^2 = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(x_i - \mu\right)^2 \qquad \text{(F2)}$$

mit N = 7737 und μ = 1,66 erhält man 9,386. Mittelwert und Varianz der Grundgesamtheit sind feste Größen. Normalerweise werden solche Berechnungen mit Statistikprogrammen (z.B. SPSS) oder in Tabellenkalkulationen (z.B. Excel) durchgeführt, in denen die Daten auch gespeichert sind. Solche Programme arbeiten notwendigerweise mit vielen Stellen nach dem Komma und zeigen diese Dezimalstellen auch an. Bei der Dokumentation der gefundenen Parameter in Berichten jedoch ist die Darstellung der Genauigkeit nicht nur von der Rechenleistung solcher Programme, sondern auch von inhaltlichen Überlegungen abhängig.

Abb. 1: Verteilung der dmft-Werte aus dem Beispiel im Text



Mittelwert und Varianz von Zufallsstichproben

Zieht man (mit dem Computer) aus der Grundgesamtheit von 7737 Kindern eine einfache Zufallsstichprobe von 10%, was der Anzahl von 774 Kindern entspricht, dann hat jedes Kind die gleiche Wahrscheinlichkeit, in die Stichprobe zu gelangen. Solche Stichproben sind normalerweise repräsentativ für die Grundgesamtheit. Der Stichprobenmittelwert soll wie üblich mit $\overline{x}$ bezeichnet werden. Die Varianz der 774 Einzelwerte des dmft, die Stichprobenvarianz, wird hier als s² bezeichnet.

$$\overline{x} = \frac{1}{(n-1)} \cdot \sum_{i=1}^{n} x_i \qquad \text{(F3)}$$

$$s^2 = \frac{1}{(n-1)} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 \qquad \text{(F4)}$$

Mittelwert und Varianz sind Schätzung der entsprechenden Parameter der Grundgesamtheit und Zufallsvariable. Die Stichprobenvarianz kennzeichnet die Streuung der n = 774 Stichprobenwerte um den Stichprobenmittelwert $\overline{x}$. Bei der Berechnung von s² ist die Division durch (n-1) zu beachten.

Die Wurzel aus s^2 ($s = \sqrt{s^2}$) wird als Standardabweichung (SD = Standard Deviation) der Einzelwerte der Stichprobe bezeichnet. Die Formeln 3 und 4 ähneln den Formeln 1 und 2, wobei aber für das unbekannte μ jetzt das bekannte $\bar{x}$ gesetzt und durch (n-1) statt durch N geteilt wird.

Schätzung des dmft-Mittelwertes

Eine dieser einfachen Zufallsstichprobe liefert als Schätzung des Wertes der Grundgesamtheit einen mittleren dmft von 1,82 (Tab. 1, erste Stichprobe) mit einer Varianz von 10,556 (SD = 3,249) und man stellt sich die Frage, wie genau diese Schätzung ist. Als Antwort lässt sich nicht sagen, wie groß der Abstand des Schätzwertes $\bar{x}$ zum tatsächlichen Wert μ ist (in diesem Beispiel 0,16), da man μ normalerweise nicht kennt. Es besteht jedoch die Möglichkeit, mittels statistischer Methoden eine Wahrscheinlichkeit dafür anzugeben, den wahren Wert in einem vorher definierten Intervall um den Stichprobenmittelwert, dem sogenannten Konfidenzintervall (oder Vertrauensbereich), zu finden. Mit einer Wahrscheinlichkeit von üblicherweise 95% überdeckt (enthält) das Konfidenzintervall F5 den wahren Wert der Grundgesamtheit. Die Formel lautet:

$$\left(\bar{x} - 1{,}96 \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{1-\frac{n}{N}}\ ;\ \bar{x} + 1{,}96 \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{1-\frac{n}{N}}\right) \qquad \text{(F5)}$$

$\sqrt{n}$ ist die Wurzel aus der Anzahl der Stichprobenelemente (hier der Anzahl der Kinder in der Stichprobe) und $\left(1-\frac{n}{N}\right)$ ist die sog. Endlichkeitskorrektur (EK), die besonders in regionalen Stichproben von Bedeutung ist. Der Ausdruck $\frac{s}{\sqrt{n}} \cdot \sqrt{1-\frac{n}{N}}$ wird Standardfehler (SE = Standard Error) des Mittelwertes genannt (Erklärung unten).

Tab. 1: Mittlerer dmft, Standardabweichung (SD) und Konfidenzintervalle (Erklärung im Text).

Grundgesamtheit

Mittlerer dmft μ	Varianz σ^2	SD s
1,662	9,386	3,064

10 einfache Zufallsstichproben

Mittlerer dmf-t $\bar{x}$	SD s	Konfidenzintervall
1,82	3,249	1,60 ; 2,04
1,77	3,368	1,54 ; 1,99
1,94	**3,378**	**1,71 ; 2,17**
1,79	3,259	1,57 ; 2,01
1,87	3,237	1,65 ; 2,09
1,53	2,905	1,34 ; 1,72
1,46	**2,821**	**1,27 ; 1,65**
1,66	3,037	1,46 ; 1,86
1,75	3,262	1,53 ; 1,97
1,68	3,142	1,47 ; 1,89

Setzt man die Werte der ersten Stichprobe aus Tab. 1 in Formel F5 ein, so erhält man mit der (computerberechneten) Standardabweichung s = 3,249 , mit $\sqrt{n} = \sqrt{774}$ = 27,82 und $\sqrt{EK}$ = 0,949 das folgende, einfach nachzuvollziehende Konfidenzintervall der ersten Stichprobe (gerundet):

(1,82 - 1,96 * (3,249*0,949) / 27,82 ; 1,82 + 1,96 * (3,249*0,949 / 27,82) = **(1,60 ; 2,04)**.

Der mittlere dmft der Grundgesamtheit mit 7737 Kindern (μ = 1,66) liegt mit 95%iger Wahrscheinlichkeit zwischen 1,60 und 2,04 (was der Wahrheit entspricht). Die Wahrscheinlichkeit, dass der tatsächliche Wert außerhalb dieses Intervalls liegt, die sogenannte Irrtumswahrscheinlichkeit, beträgt 5%. Für kleine Fallzahlen n ist das Konfidenzintervall breit (da $\sqrt{n}$ im Nenner steht) und daher die Schätzung unsicher, für große Fallzahlen ist es schmal und damit die Schätzung genauer. Je mehr Kinder in die Stichprobe kommen, desto genauer kann der mittlere dmft der Grundgesamtheit geschätzt werden. Soll die Breite des Konfidenzintervalls halbiert und damit die Genauigkeit der Schätzung erhöht werden, so benötigt man dafür unter sonst gleichen Bedingungen die vierfache Fallzahl (Kinderzahl).

Zieht man eine andere Zufallsstichprobe, so wird sich ein anderer Mittelwert mit einem anderen Konfidenzintervall zufällig ergeben (Tab. 1). Insgesamt gibt es (N über n) = $\frac{N!}{n!(N-n)!}$ Möglichkeiten, einfache Zufallsstichproben vom Umfang n aus einer Grundgesamtheit von N Individuen (oder Objekten) zu ziehen. Beispielsweise gibt es 120 Möglichkeiten, aus einer Gesamtheit von 10 Objekten 3 auszuwählen: (10 über 3) = $\frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3}$ = 120. Für die Auswahl von 30 aus 100 Objekten gibt es bereits $2{,}9 \cdot 10^{25}$ Möglichkeiten. Diese Größenordnung ist uns noch aus dem Physikunterricht bekannt (Anzahl der Moleküle in 1m³ eines Gases unter Normzustand beträgt $2{,}7 \cdot 10^{25}$ [Loschmidt-Konstante]). In unserem Fall einer Auswahl von 774 aus 7737 Kindern ergeben sich etwa 10^{1091} mögliche Stichproben, von denen wir normalerweise lediglich eine einzige auswählen (!).

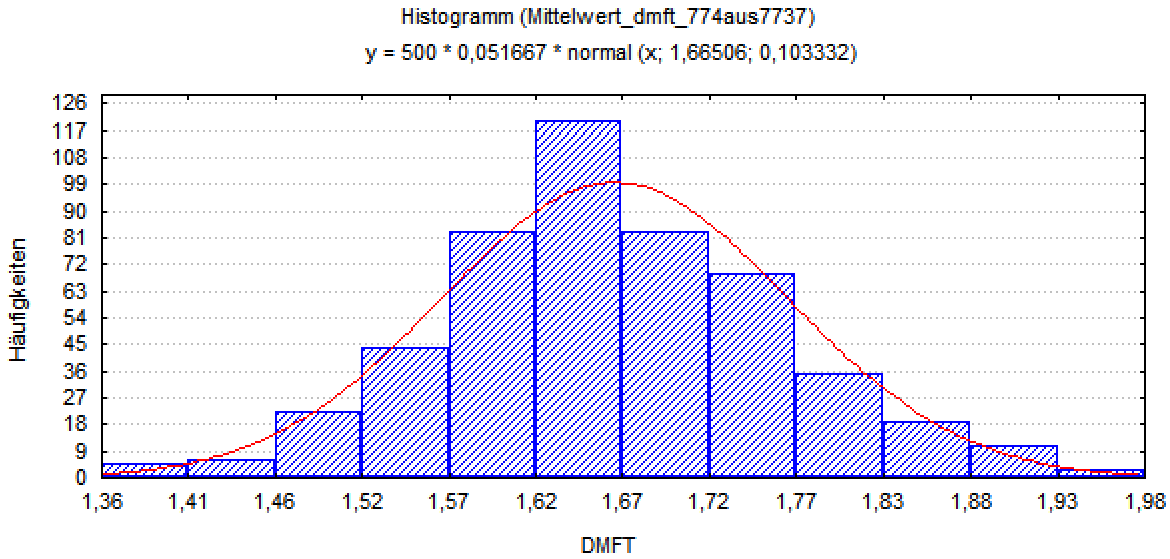
Von jeweils 100 Zufallsstichproben werden etwa 5 Konfidenzintervalle (5%) den wahren, unbekannten Mittelwert der Grundgesamtheit nicht enthalten. Bei den Stichproben mit dem mittleren dmft-Wert 1,94 und 1,46 in Tab. 1 können wir das erkennen, da in unserem simulierten Beispiel der wahre Wert der Grundgesamtheit (1,66) bekannt ist. Hier haben wir irrtümlich eine Stichprobe gezogen, deren Mittelwert nicht die Grundgesamtheit unseres Beispiels repräsentiert. In der Praxis müssen wir mit dieser Unsicherheit, in 5% der Fälle eine „falsche" Stichprobe zu ziehen, leben und auch solche mittlere dmft-Werte als Ergebnis akzeptieren.

Die vielen möglichen Mittelwerte, die man erhält, wenn man aus einer Grundgesamtheit immer neue Stichproben gleichen Umfangs zieht, besitzen eine Verteilung (normalerweise eine Normalverteilung) und eine Varianz. Abb. 2 zeigt die Verteilung von 500 solcher Stichprobenmittelwerte $\bar{x}_j$ (j = 1 bis 500) des dmft aus ebenso vielen Zufallsstichproben des Beispiels (774 aus 7737). Die Werte sind annähernd normalverteilt mit einem Maximum bei 1,665, was dem Wert μ = 1,662 der Grundgesamtheit etwa entspricht, und einer Standardabweichung von 0,1033.
Diese Standardabweichung der Verteilung der 500 Stichprobenmittelwerte in Abb.2 wird im Falle einer einzelnen Stichprobe durch den Standardfehler SE des Mittelwertes dieser Stichprobe geschätzt.

$$SE = \frac{s}{\sqrt{n}} \cdot \sqrt{1 - \frac{n}{N}} \qquad (F6)$$

Für den Mittelwert der ersten Stichprobe in Tabelle 1 beispielsweise erhält man den Standardfehler von $SE = \frac{3{,}249}{\sqrt{774}} \cdot \sqrt{\left(1 - \frac{774}{7737}\right)}$ = 0,1108 als geschätzten Wert für die Standardabweichung der Verteilung in Abbildung 2.

Abb. 2: Verteilung der Stichprobenmittelwerte von 500 Stichproben (774 aus 7737) des Beispiels im Text.



Bei großen Grundgesamtheiten und im Vergleich dazu kleinen Stichproben kann die Endlichkeitskorrektur (EK) vernachlässigt werden (Anhaltspunkt: $n/N < 0{,}05$). Ist es geplant, die gesamte Population zu untersuchen, d.h., eine Vollerhebung durchzuführen, dann ist n = N und der Standardfehler des Mittelwertes ist gleich Null. In diesem Fall gibt es kein Konfidenzintervall für einen geschätzten Mittelwert, denn jetzt kennen wir den wahren Mittelwert der Grundgesamtheit. Sofern bei einer geplanten Vollerhebung einige Kinder fehlen (aus verschiedenen Gründen), so ist es nicht angebracht, die untersuchten Kinder als große Stichprobe aufzufassen und Konfidenzintervalle zu berechnen. In diesen Fällen sollte die Grundgesamtheit neu definiert und wenn möglich die Auswirkung der fehlenden Kinder auf den Zielparameter (z.B. den mittleren dmft) diskutiert werden. Das Ergebnis einer Vollerhebung zum Beispiel in Kindergärten eines Landkreises bezieht sich immer nur auf die untersuchten Kinder, nicht auf die in den Einrichtungen gemeldeten Kinder.

Schätzung des Anteils kariesfreier Kinder

Mit einer einfachen Zufallsstichprobe lassen sich nicht nur Mittelwerte sondern auch Anteile schätzen. Wir betrachten dazu wieder den realen Datensatz der Reihenuntersuchung in 160 Kindergärten eines Landkreises mit 7737 Kindern. In dieser Grundgesamtheit haben 4842 Kinder das Merkmal dmft = 0 (kariesfrei) und 2895 das Merkmal dmft > 0. Der Anteil kariesfreier Kinder in der Grundgesamtheit, der normalerweise unbekannt ist, beträgt demnach P = 4842 / 7737 = 0,626 oder 62,6%. Die Varianz der Grundgesamtheit beträgt 0,234. Man erhält sie nach der Formel F7. P (Großbuchstabe) bezeichnet den Anteil in der Grundgesamtheit und p (Kleinbuchstabe) in der Stichprobe.

$\sigma^2 = P \cdot (1 - P)$ (F7)

Wird aus dieser Grundgesamtheit (mit dem Computer) eine einfache Zufallsstichprobe vom Umfang 10% gezogen (also 774 aus 7737 Kindern zufällig ausgewählt), so erhält man mit p eine Schätzung für den Anteil der kariesfreien Kinder und mit s² die Varianz der Stichprobenwerte $s^2 = p \cdot (1-p)$. Der Standardfehler ist gegeben zu:

$$s_p = \sqrt{\frac{p \cdot (1-p)}{n}} \cdot \sqrt{\left(1 - \frac{n}{N}\right)} \quad \text{(F8)}$$

und für die Grenzen des Konfidenzintervalls gelten folgende Formeln:

$$\{\left(p - \frac{1}{2n}\right) - 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \cdot \sqrt{1 - \frac{n}{N}} \;;\; \left(p - \frac{1}{2n}\right) + 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \cdot \sqrt{1 - \frac{n}{N}}\} \quad \text{(F9)}$$

Bei großen Grundgesamtheiten und im Vergleich dazu kleinen Stichproben kann die Endlichkeitskorrektur (EK) in den Formeln vernachlässigt werden, was sie geringfügig vereinfacht. Der Faktor $\frac{1}{2n}$ in Formel F9 wird als Stetigkeitskorrektur bezeichnet. Bei der üblichen Anzahl von den zahnärztlichen Diensten untersuchter Kinder kann diese Korrektur unberücksichtigt bleiben kann. Schon bei der Untersuchung von nur 1000 Kindern nämlich beträgt die Stetigkeitskorrektur 1/2000 = 0,0005 und im Vergleich dazu der Anteil kariesfreier Kinder in unserem Beispiel 0,626 (62,6%).

Die Werte für 10 einfache Zufallsstichproben (ausgewählt aus einer größeren Anzahl Simulationen) sind in Tabelle 2 dargestellt. Hier überdecken (enthalten) beispielhaft zwei Konfidenzintervalle (fett gedruckt) nicht den wahren mittleren Anteil der Grundgesamtheit von 0,626. Die grafische Darstellung der Werte von Tab. 2 ist in Abb. 3 zu sehen.

Tab. 2: Anteile kariesfreier Kinder, Varianz und Konfidenzintervalle des Beispiels im Text.

Grundgesamtheit

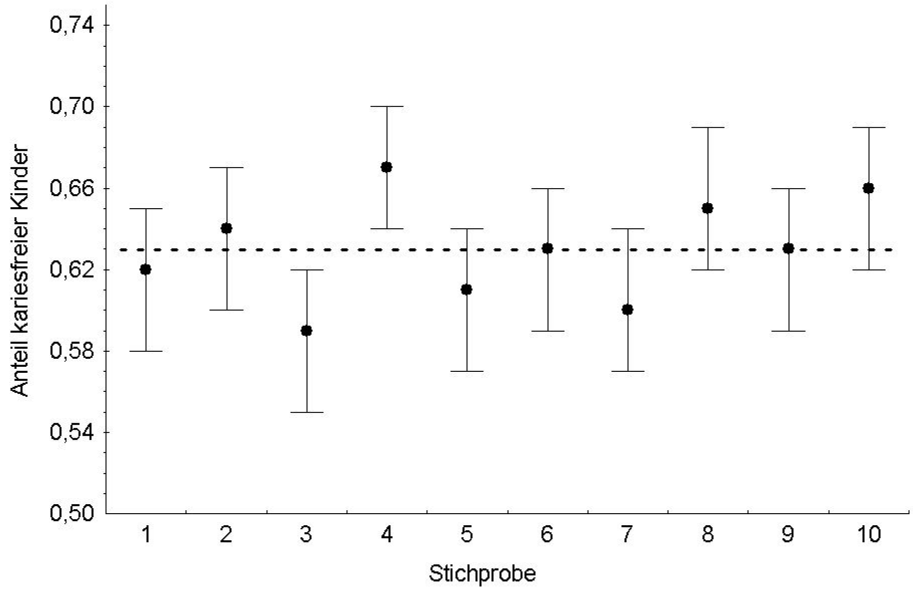
Anteil P	Varianz σ^2	SD s
0,626	0,234	0,484

10 einfache Zufallsstichproben

Anteil p	SD s	Konfidenzintervall
0,62	0,486	0,58 ; 0,65
0,64	0,481	0,60 ; 0,67
0,59	**0,492**	**0,55 ; 0,62**
0,67	**0,470**	**0,64 ; 0,70**
0,61	0,489	0,57 ; 0,64
0,63	0,484	0,59 ; 0,66
0,60	0,490	0,57 ; 0,64
0,65	0,476	0,62 ; 0,69
0,63	0,484	0,59 ; 0,66
0,66	0,474	0,62 ; 0,69

Beim Nachrechnen beachte man, dass die Zahlenangaben gerundet sind.

Abb. 3: Anteile kariesfreier Kinder mit Konfidenzintervallen. Darstellung der 10 Zufallsstichproben aus Tabelle 2 (Anteil in der Grundgesamtheit gestrichelt).



Auf den ersten Blick sieht es so aus, als ob alle 10 Konfidenzintervalle in Abbildung 3 die gleiche Breite hätten. Annähernd stimmt das auch, obgleich es geringe Unterschiede gibt, die aus den etwas variierenden Standardabweichungen (SD) zwischen den Stichproben resultieren (Tab. 1 und 2). Die Breite ist Ausdruck für die Genauigkeit der jeweiligen Schätzung des Anteils kariesfreier Kinder in der Grundgesamtheit. Dieser Wert sollte vor der Untersuchung (Studie) zahlenmäßig festgelegt werden. Im Verfahren der Fallzahlplanung wird dann mit dieser vorgegebenen Genauigkeit die hierfür erforderliche Anzahl der Individuen (Kinder) in der Stichprobe bestimmt nach der Formel:

$$n = \frac{N}{1+B} \text{ , mit } B = \frac{e^2 \cdot N}{u_{\alpha/2}^2 \cdot s^2}.$$

Dabei sind e die halbe Breite des gewünschten Konfidenzintervalls, s² die in früheren Untersuchungen ermittelte Varianz der Stichprobenwerte (für Anteilsschätzung $s^2 = p \cdot (1-p)$) und $u_{\alpha/2}^2$ = 1,96 das $(1-\alpha/2)$- Quantil der Standardnormalverteilung (α =5%).

Möchte man beispielsweise in unserer Grundgesamtheit von N = 7737 Kindern den Anteil kariesfreier Kinder mit einer Genauigkeit von ± 0,03 schätzen und betrug dieser Anteil in einer zeitnahen früheren Untersuchung z.B. 60 %, so ergibt sich für

$B = \frac{(0{,}03)^2 \cdot 7737}{(1{,}96)^2 \cdot 0{,}6 \cdot 0{,}4}$ = 7,5525 und damit n = 7737 / 8,5525 = 904 . Mit dieser Anzahl Kinder sollte man bei zahnärztlichen Untersuchungen dieser Grundgesamtheit für jede der möglichen Zufallsstichproben die gewünschte Genauigkeit von ± 0,03 erreichen. In Tabelle 2 erkennt man übrigens, dass die etwas niedrigere Fallzahl von 774 aus dem obigen Beispiel lediglich eine Genauigkeit von 0,035 bewirkt.

Diese hier kurz beschriebene Fallzahlplanung für einfache Zufallsstichproben kann man gern auch einem Statistikprogramm, wie z.B. WinPepi (www.brixtonhealth.com) oder Open Epi (www.open-epi.com) überlassen.

In der Gesundheitsberichterstattung (GBE) wäre es wichtig, nicht nur den Stichprobenmittelwert, etwa der ersten Stichprobe von Tab. 2 , p = 0,62 sondern auch das dazugehörige Konfidenzintervall (0,58 ; 0,65) anzugeben. Hatte man etwa im Vorjahr einen Anteil Kariesfreier von 59% und dieses Jahr 62%, so sollte dargelegt werden, dass dieses Ergebnis zwar erfreulich ist und möglicherweise auch im Trend liegt (falls man das aus früheren Daten ablesen kann), dass die Verbesserung gegenüber dem Vorjahr statistisch aber nicht abgesichert werden kann. Zur besseren Absicherung wäre zwar eine Erhöhung der Präzision von ± 4% auf ± 2% möglich, dies könnte aber u.U. mit den vierfachen Kosten verbunden sein. Auch sollte den Entscheidungsträgern vermittelt werden, welche Präzision in der wissenschaftlichen Literatur für Anteilsschätzungen üblich ist, um hieraus eine Ressourcenplanung zu begründen. Beim Vergleich mit den Ergebnissen anderer Kommunen ist Vorsicht geboten. Liegt deren Mittelwert im eigenen Konfidenzintervall, hat z.B. der Nachbarkreis in der fraglichen Altersgruppe bei gleicher Zusammensetzung der Bevölkerung (!) 64% Kariesfreie, so ist dieser Unterschied als zufällig zu interpretieren.

Die in den beiden Beispielen vorgestellten Berechnungen funktionieren unabhängig davon, ob die Stichprobe repräsentativ für die Grundgesamtheit ist, doch nur wenn das der Fall ist, erhält man brauchbare Ergebnisse. Eine wichtige Voraussetzung für die Repräsentativität ist die zufällige Auswahl der Stichprobenelemente aus der Grundgesamtheit, ganz gleich, um welches Auswahlverfahren es sich handelt. Einfache Zufallsstichproben sind nicht immer möglich. Wollte man eine der Simulationen des Beispiels in der Realität durchführen, so müssten aus einer Liste der 7737 Kinder mittels Zufallsgenerator 774 Kinder und deren Kindergärten ermittelt werden. Man müsste eine große Anzahl von Kindergärten aufsuchen, um dort nur einige wenige Kinder zu untersuchen – ein Procedere, das so nicht durchführbar ist. Stattdessen ist es üblich, ganze Kindergärten auszuwählen und dort alle Kinder zu untersuchen. Dieses Auswahlverfahren erfordert andere Rechenvorschriften für die Planung und Auswertung der Daten, denn hierbei handelt es sich um Clusterstichproben.

Anschriften der Autoren:
Dr. Michael Herzog
Höhenweg 32
78315 Radolfzell

Prof. Dr. Rafael Weißbach
Universität Rostock, Lehrstuhl Statistik und Ökonometrie,
Wirtschafts- und Sozialwissenschaftliche Fakultät
D - 18051 Rostock