

## 2 Grundlagen sozialwissenschaftlicher Forschung

### 2.1 Wissenschaftstheoretische Grundlagen

#### 2.1.1 Wissenschaftstheoretische Grundpositionen

Im Mittelpunkt der empirischen Marketingforschung steht die Erlangung von Erkenntnisfortschritten in Bezug auf die Wahrnehmung und das Verhalten von Marktteilnehmern. Bezüglich der Generierung von Erkenntnisfortschritten sind unterschiedliche Grundpositionen relevant, die sich im Laufe der Zeit in der Wissenschaftstheorie herauskristallisiert haben. Die prominentesten sind der durch den *Wiener Kreis* begründete logische Empirismus sowie der kritische Rationalismus von *Popper*.<sup>5</sup>

Der logische Empirismus geht davon aus, dass nur existent sein kann, was sich auch in der Sinneswahrnehmung des Menschen manifestiert. Dementsprechend ist dieser Grundposition folgend das Anliegen der Wissenschaft, auf Basis von Beobachtungen formulierte Hypothesen zu verifizieren und auf diese Weise Theorien zur allgemeingültigen Beschreibung der Realität zu konstruieren. Die Verifizierbarkeit einer Hypothese ist es also, was dieser ihren Sinn verleiht.<sup>6</sup> Das zentrale Prinzip des logischen Empirismus ist damit die Induktion: Anhand von Daten über die Realität werden Hypothesen verifiziert, was letztlich zur Formulierung einer Theorie führt. Mit anderen Worten werden an einzelnen Sachverhalten oder Objekten gemachte Erfahrungen (Daten) zur Überprüfung von Annahmen über die Welt (Hypothesen) genutzt, um auf dieser Basis zu allgemeingültigen Aussagen (Theorie) zu kommen. Die Theorienbildung bezieht sich damit ausschließlich auf gemachte Beobachtungen, die die Realität jedoch nicht unbedingt vollständig abdecken: So kommt ein Beobachter, der bisher lediglich weiße Schwäne gesehen hat, zu dem Schluss, dass alle Schwäne weiß sein müssen und bezeichnet die Aussage, es gäbe lediglich weiße Schwäne, als allgemeingültig. Dieser induktive Schluss ist aber letztlich nicht logisch begründbar, was die größte Schwäche des logischen Empirismus darstellt. Die Anwendung der Grundposition des logischen Empirismus lässt sich trotz alledem relativ einfach rechtfertigen: Erstens verfügt die empirische Wissenschaft über keine besseren Methoden als die Überprüfung von Hypothesen an Einzelfällen und zweites sind die abgeleiteten und als allgemeingültig erklärten Gesetzmäßigkeiten immerhin sehr wahrscheinlich.<sup>7</sup>

Der von *Popper* begründete kritische Rationalismus kann als Gegenentwurf zum logischen Empirismus angesehen werden. Für ihn beginnt der Erkenntnisgewinn nicht mit Beobach-

---

<sup>5</sup> Vgl. Heller (2012), S. 12ff.

<sup>6</sup> Vgl. Carrier (2009), S. 23.

<sup>7</sup> Vgl. Westermann (1987).

tungen, sondern mit einer Theorie und der daraus abgeleiteten Hypothese. Durch Deduktion, also dem Schluss vom Allgemeinen auf das Spezielle, kommt der Forscher zu einem die Realität abbildenden Modell in einem bestimmten Anwendungsfall. Dieses Modell ist letztlich einer empirischen Überprüfung zu unterziehen, um die Gültigkeit der Theorie in diesem konkreten Fall zu testen.<sup>8</sup> Die Theorie wird dabei als Vermutung angesehen, die es in Bezug auf konkrete reale Phänomene zu testen gilt und deren Gültigkeit in bestimmten Anwendungsfeldern zu prüfen ist. Folglich werden Voraussagen hier nicht verifiziert, sondern ausschließlich falsifiziert bzw. nicht falsifiziert. Die Falsifikation von Theorien erfolgt anhand empirischer Daten mit dem Ziel, falsche Theorien in Bezug auf einen Anwendungsfall auszusortieren. *Popper* folgend, kann der Erkenntnisgewinn damit ausschließlich durch das Falsifikationsprinzip erzielt werden. Damit einher geht auch die Notwendigkeit der ständig wiederholten Überprüfung von Theorien, da ein empirischer Test niemals abschließend theoretisch fundierte Hypothesen verifizieren kann, sondern sich der Forscher durch wiederholte Nicht-Falsifikation schrittweise der Wahrheit annähert. Auch ist damit ein induktives oder exploratives Vorgehen im Forschungsprozess nicht zulässig, lediglich das deduktive Vorgehen genügt den Ansprüchen des Falsifikationsprinzips. Allerdings führt die konsequente Ausrichtung der empirischen Forschung an deduktiven und damit logischen Schlüssen zu einer Einschränkung hinsichtlich der Mehrung von Erkenntnissen, da lediglich Hypothesen entlang bestehender Theorien akzeptiert sind.

Folglich wird die Sinnhaftigkeit einer konsequenten Anwendung des kritischen Rationalismus in der sozialwissenschaftlichen Forschung angezweifelt.<sup>9</sup> Die strenge Orientierung an den kritisch-rationalistischen Prinzipien gilt hier als realitätsfern, da eine schlüssige Falsifikation in den sozialwissenschaftlichen Disziplinen nicht zu realisieren ist.<sup>10</sup> Zurückzuführen ist dies auf die höhere Komplexität in Bezug auf die zu berücksichtigenden Einflussfaktoren im Vergleich zu naturwissenschaftlichen Disziplinen. Darüber hinaus sind empirische Überprüfungen in den Sozialwissenschaften stets mit Messfehlern in den berücksichtigten Größen verbunden, was eine Falsifikation unmöglich macht.<sup>11</sup> Viele Forschungsbestrebungen bringen es zudem mit sich, dass sie theoretisch zu wenig durchdrungen sind, als dass ein deduktives Vorgehen erfolgen könnte.<sup>12</sup>

Der von *Leplin* und *Hunt* geprägte wissenschaftliche Realismus trägt diesen Kritikpunkten Rechnung.<sup>13</sup> Induktive Schlussfolgerungen sind dementsprechend zu akzeptieren und die empirische Überprüfung theoretisch postulierter Annahmen kann auch in deren Bestätigung gipfeln. Dies bedeutet allerdings nicht, dass der wissenschaftliche Realismus die Existenz universeller Wahrheiten vertritt. Vielmehr geht es um die schrittweise Annäherung an die Wahrheit, was durch den mehrfachen Test von Hypothesen erfolgt.<sup>14</sup> Dieser

---

<sup>8</sup> Vgl. *Popper* (1973), S. 213 ff.

<sup>9</sup> Vgl. u. a. *Kubiczek* (1975), S. 48 ff.; *Deshpande* (1983), S. 105 ff.

<sup>10</sup> Vgl. *Kubiczek* (1975), S. 49; *Witte* (1981), S. 18.

<sup>11</sup> Vgl. *Anderson* (1983), S. 21; *Martin* (1989), S. 23.

<sup>12</sup> Vgl. *Homburg* (2000), S. 56.

<sup>13</sup> Vgl. *Leplin* (1981); *Hunt* (1984).

<sup>14</sup> Vgl. *Peter* (1997), S. 71.

Orientierungsrahmen lässt eine induktive Vorgehensweise zu,<sup>15</sup> was auch mit einem größeren Spektrum möglicher Erkenntnisfortschritte im Vergleich zum kritischen Rationalismus verbunden ist. Der wissenschaftliche Realismus ist folglich ein für die empirische Marketingforschung geeigneter Orientierungsrahmen. Es ist allerdings festzuhalten, dass auch bei diesem Prinzip nicht von einer universellen Gültigkeit der erlangten Erkenntnisse auszugehen ist.<sup>16</sup>

### 2.1.2 Arbeiten mit Theorien zur Hypothesenbildung

Der Erkenntnisfortschritt in der empirischen Marketingforschung beruht auf der Formulierung von Hypothesen. Die interessierende wissenschaftliche Fragestellung wird dabei zunächst in wissenschaftliche Hypothesen überführt. Diese sind letztlich Vermutungen über die Beziehung zwischen zwei oder mehr Variablen, die das betrachtete Phänomen charakterisieren. Die Gültigkeit der angenommenen Beziehung wird dabei für eine definierte Population vergleichbarer Objekte oder Ereignisse angenommen. Relevant für die Formulierung von Hypothesen ist, dass diese empirisch überprüfbar und als allgemeingültige Vermutung formuliert sind. Die Formulierung erfolgt dabei als Konditionalsatz in Form einer Wenn-Dann- bzw. Je-Desto-Aussage. Eine Hypothese wird folglich durch die Verknüpfung einer Bedingung (Wenn- bzw. Je-Satz) und einer Konsequenz (Dann- bzw. Desto-Satz) aufgestellt.

Die Auswahl der zu berücksichtigenden Variablen und damit die Beschreibung des zu untersuchenden Sachverhalts in Form eines Modells erfolgt in einem ersten Schritt typischerweise durch den Rückgriff auf Theorien, die im Falle der Marketingforschung in den meisten Fällen der Psychologie oder Sozialpsychologie entstammen. Beim Theoriemonismus wird lediglich eine einzige Theorie zur Erklärung eines Phänomens herangezogen.<sup>17</sup> Der auf *Feyerabend* zurückgehende Theorienpluralismus sieht hingegen die Anwendung und Kombination mehrerer Theorien vor.<sup>18</sup> Dieser Auffassung liegt die Idee zugrunde, dass sich einzelne theoretische Konzepte sowie die daraus erzielten Erkenntnisse gegenseitig ergänzen können und daraus ein umfassenderes Verständnis des interessierenden Phänomens resultiert.<sup>19</sup> Folglich ist der Rückgriff auf mehrere Theorien zur Identifikation relevanter Variablen nicht nur legitim, sondern kann vielmehr für den Erkenntnisgewinn äußerst förderlich sein. Die Ermittlung der relevanten Variablen erfolgt in der Form, dass der Forscher nach einer auf das interessierende Phänomen übertragbaren Theorie sucht. Entweder abstrahiert er dabei vom konkreten Marketingphänomen, um eine passende Theorie zu identifizieren, z. B. sieht er das Verhältnis von Marke zu Konsument als Beziehung und zieht dementsprechend Beziehungstheorien heran, oder aber der Forscher identifiziert einzelne Elemente des zu untersuchenden Phänomens (durch Beobachtung oder

---

<sup>15</sup> Vgl. Zaltman/LeMasters/Heffring (1982), S. 97 ff. zur komplementären Anwendung deduktiver und induktiver Methoden.

<sup>16</sup> Vgl. Hunt (1994), S. 134 ff.

<sup>17</sup> Vgl. Huber (2004), S. 74.

<sup>18</sup> Vgl. Feyerabend (1965). Zur Kritik am theoretischen Pluralismus vgl. Staehle (1990).

<sup>19</sup> Vgl. Peter (1997), S. 72 sowie von der Oelsnitz (1997), S. 24.

vorliegende empirische Erkenntnisse) und identifiziert dann solche Theorien, welche unter anderem dieses Element berücksichtigen. Im zweitgenannten Fall beobachtet der Forscher bspw. dass sich der Konsum nach dem sozialen Umfeld des einzelnen Konsumenten richtet. Diese als soziale Norm bezeichnete Variable ist Teil der Theorie des vernünftigen Handelns, des geplanten Verhaltens oder aber auch des zielorientierten Verhaltens, die der Forscher somit für die Beschreibung des Phänomens heranziehen könnte.

*Merton* nennt in Bezug auf die Beschreibung sozialer Phänomene Theorien mit mittlerer Reichweite als erstrebenswert.<sup>20</sup> Mit der Kategorisierung von Theorien in Haupttheorien, Theorien mittlerer Reichweite und Mikrotheorien unterscheidet *Merton* das Abstraktionsniveau einer Theorie. Haupttheorien sind endlos weitreichende Theorien, wie sie in den Naturwissenschaften angestrebt werden und die für sämtliche beobachtbaren Phänomene Gültigkeit besitzen. Mikrotheorien stellen hingegen das Gegenteil dar. Die Reichweite dieser Theorien ist auf fallspezifische Fakten begrenzt. Die Allgemeingültigkeit des Phänomens spielt damit eine untergeordnete Rolle (radikaler Empirismus). Gleichwohl können auch der radikale Empirismus bzw. die mikrotheoretische Begründung einen Erkenntnisfortschritt ermöglichen (z.B. Erfolgsfaktorenforschung). Theorien mittlerer Reichweite besitzen einen engen Geltungsbereich und weisen eine hohe Operationalität auf.<sup>21</sup> Die Diffusionstheorie ist bspw. eine solche Theorie, da sie konkrete und allgemeingültige Muster bei der Durchsetzung von Marktneuheiten postuliert. Häufig beinhalten Theorien mittlerer Reichweite auch die zu berücksichtigenden Konstrukte bzw. Variablen, die zur Erklärung eines Phänomens von Relevanz sind. Eine konsequente Anwendung einer solchen Theorie im Rahmen verschiedener Studien ist für den Erkenntnisfortschritt deshalb förderlich, da dies ihre stetig wiederholte, empirische Überprüfung forciert und damit mit zunehmender Anzahl von Überprüfungen eine fundierte Aussage zu ihrer Gültigkeit überhaupt erst möglich macht. Die Marketingwissenschaft sieht sich bei der Anwendung von Theorien immer wieder größerer Kritik ausgesetzt, da sie vor der Herausforderung steht Theorien aus anderen Disziplinen, z. B. Psychologie oder Sozialpsychologie, entleihen und transferieren zu müssen. Ein Auslassen von zunächst einmal vielversprechenden Theorien, die also in irgendeiner Weise Erkenntnisse beitragen können, ist allerdings nicht begründbar.<sup>22</sup> Wegen des neu postulierten Gültigkeitsbereichs ist eine empirische Testung solcher Theorien unbedingt erforderlich. Vielmehr sollten die anzuwendenden Theorien möglichst verstanden werden, um diese adäquat auf das Marketingphänomen übertragen zu können.

### 2.1.3 Unterschiedshypothesen und Kausalhypothesen

In der Marketingforschung sind die Unterschiedshypothese und die Kausalhypothesen von großer Bedeutung. Unterschiedshypothesen machen eine Aussage über die Unterschiede zwischen zwei (oder mehreren) Gruppen bezüglich einer Variablen. In Kausalhy-

---

<sup>20</sup> Vgl. Merton (1998).

<sup>21</sup> Vgl. Franke (2002), S. 196.

<sup>22</sup> Vgl. Franke (2002), S. 198.

pothesen sind hingegen Aussagen über Zusammenhänge zwischen zwei oder mehreren Variablen in einer Gruppe formuliert.

Häufig ist von Interesse, welche Wirkung eine Veränderung der Marketingaktivitäten und/oder -umwelt entfaltet. So hat sich die Markentransferforschung bspw. die Frage gestellt, wie Konsumenten ein neues Produkt eines etablierten Markenherstellers bewerten (Einstellung gegenüber dem neuen Produkt als Wirkung). Und zwar in Abhängigkeit davon, wie stark sich dieses neue Produkt von den bisher angebotenen Produkten des Herstellers unterscheidet (Produktfit als Ursache). Folglich geht es hier um die Prüfung folgender Hypothesen: „Je höher der Fit zwischen dem neuen Produkt und den bisherigen Produkten der Marke, desto positiver ist die Einstellung gegenüber diesem neuen Produkt“. Zur Untersuchung solcher Fragestellungen bietet sich vor allem das Experiment an, da hier das zu untersuchende Phänomen (Produktfit) absichtlich herbeigeführt (manipuliert) und die Wirkung durch Beobachtung oder Befragung gemessen werden kann. Den Einfluss möglicher Störgrößen gilt es zu kontrollieren. Überprüft wird die Hypothese dann anhand der Mittelwerte, welche die abhängige Variable bei den unterschiedlichen Bedingungen annimmt. Damit sind der Unterschied der Werte sowie die Signifikanz dieses Unterschieds von Relevanz.

In wissenschaftlichen Publikationen der Marketingforschung wird in der Regel die sogenannte Alternativhypothese verfasst, in der die Annahme der Existenz eines Unterschieds bzw. Zusammenhangs zum Ausdruck kommt. Das Gegenteil zur Alternativhypothese ist die Nullhypothese, die von der Nicht-Existenz eines Unterschieds oder Zusammenhangs ausgeht. Bei der Verbalisierung von Hypothesen ist darauf zu achten, dass diese möglichst präzise und damit auch als eindeutig überprüfbar formuliert werden. Im Hinblick auf die Eindeutigkeit sind einseitige Hypothesen zu bevorzugen. Einseitige Hypothesen machen eine klare Aussage darüber, in welcher Richtung ein Zusammenhang verläuft (positiv oder negativ) bzw. in welcher Gruppe eine Variable stärker ausgeprägt ist und nicht nur, dass ein Zusammenhang vorliegt oder es Unterschiede in verschiedenen Gruppen gibt. Folglich muss einer Bedingung bzw. Ursache eine eindeutige Konsequenz bzw. Wirkung zugeordnet werden, was deren Messung (Operationalisierung) erforderlich macht. Dafür muss der zu untersuchende Sachverhalt zunächst identifiziert und dann genau beschrieben, definiert, wertfrei benannt sowie für die statistische Prüfung letztlich auch quantifiziert werden.

In der empirischen Marketingforschung stellen die beiden Bestandteile einer Hypothese (Bedingung und Konsequenz) konkrete Ausprägungen von Variablen dar, mit deren Hilfe das untersuchte Phänomen messbar gemacht wird (Operationalisierung). Die Bedingung wird demnach repräsentiert durch die unabhängige Variable, die Konsequenz wird durch die abhängige Variable widergespiegelt. Die empirische Überprüfung von wissenschaftlichen Hypothesen erfolgt durch deren Überführung in statistische Hypothesen, die wiederum Wahrscheinlichkeitsaussagen darstellen. Hintergrund dessen ist die Erkenntnis, dass eine Aussage nicht einfach durch das Anführen von Gegenbeispielen falsifiziert oder durch das Anführen von Positivbeispielen verifiziert werden kann. Aus diesem Grunde macht man sich das Prüfkriterium der statistischen Signifikanz zunutze. Dieses liefert eine

Aussage darüber, mit welcher Wahrscheinlichkeit davon auszugehen ist, dass die empirischen Daten die Hypothese stützen und folglich nicht der Zufall für das Vorliegen eines postulierten Phänomens verantwortlich gemacht werden kann. Dabei ist allerdings zu konstatieren, dass die Bestätigung von Hypothesen auf einem festgelegten Signifikanzniveau nicht zur Erklärung der endgültigen Gültigkeit verleiten darf.

## 2.2 Statistische Grundlagen

### 2.2.1 Ein intuitives Beispiel zur statistischen Signifikanz

Um den Unterschied der Mittelwerte verschiedener Gruppen zu untersuchen, ist das Konzept der statistischen Signifikanz von Bedeutung. Ein intuitives Beispiel schafft hierzu zunächst das notwendige Grundverständnis. Die Konsumentengruppen von zwei neuen Getränken sollen anhand von jeweils 50 Probanden im Hinblick auf Unterschiede beim Alter überprüft werden. Zur Verdeutlichung seien die beiden Produkte mit „Traditionsgetränke“ und „Junge Brause“ bezeichnet. Insgesamt konsumieren deutschlandweit bereits jeweils 500 Personen diese neuen Getränke (Grundgesamtheit). Es wird nun das Alter der 50 anwesenden Probanden erhoben, um ein durchschnittliches Alter aller 500 Konsumenten des jeweiligen Produkts zu schätzen. Im Vergleich sind nun folgende drei Varianten der Altersverteilung denkbar:

1. Für das Traditionsgetränk beträgt das Durchschnittsalter der Konsumenten 37 Jahre. Für die Junge Brause beträgt es dagegen nur 22 Jahre. Intuitiv ist sofort ersichtlich, dass diese Differenz kein Zufall sein kann. Die 500 Konsumenten des Traditionsgetränks sind durchschnittlich mit hoher Sicherheit älter als die 500 Konsumenten der Jungen Brause.
2. Für das Traditionsgetränk beträgt das Durchschnittsalter der Konsumenten 22,5 Jahre. Für die Junge Brause beträgt es 22,4 Jahre. Intuitiv erscheint diese Differenz zu klein zu sein, um mit Sicherheit von einer unterschiedlichen Altersverteilung zu sprechen. Die jeweils 500 Konsumenten beider Getränke können somit als etwa gleich alt angenommen werden.
3. Für das Traditionsgetränk beträgt das Durchschnittsalter der Konsumenten 25 Jahre. Für die Junge Brause beträgt es 22 Jahre. Kann man hier schon mit Sicherheit sagen, dass das Durchschnittsalter aller Konsumenten sich unterscheidet? Oder kommt die Differenz zufällig zustande, da gerade heute nur jüngere Konsumenten der Jungen Brause bei der Befragung anwesend sind? Sind die Konsumenten des Traditionsgetränks nun älter oder nicht? Um diese Frage zu klären, reicht die Intuition nicht mehr aus. Es ist eine mathematische Kennzahl notwendig, und zwar die statistische Signifikanz.

Zur Ermittlung der statistischen Signifikanz sind drei Komponenten erforderlich. Erstens finden die Mittelwerte Berücksichtigung, die bereits intuitiv einen Einblick in das Phänomen geben können. Zweitens spielt die Varianz eine Rolle. Diese gibt an, wie groß die Abweichung vom Mittelwert der interessierenden Variablen ist. Intuitiv macht auch diese Größe Sinn, da eine größere Streuung zusätzliche Unsicherheit in der Beurteilung des durchschnittlichen Alters schafft. Drittens ist die Größe der Stichproben von Bedeutung. Soll von 50 anwesenden Probanden auf alle 500 Konsumenten geschlossen werden, so ist die Beurteilung des durchschnittlichen Alters mit einer Unsicherheit behaftet. Soll jedoch von beispielsweise 480 anwesenden Probanden auf alle 500 Konsumenten geschlossen werden, so ist die Unsicherheit deutlich geringer. Ein mathematischer Vergleich der Mittelwerte verschiedener Stichproben beinhaltet somit die Mittelwerte, die Varianzen und die Stichprobengrößen der Untersuchung. Die Sicherheit, mit der man eine Entscheidung treffen will, beträgt in der Statistik und im Bereich wissenschaftlicher Forschung meist 95% (Signifikanzniveau). Die übrigen 5% repräsentieren den Fehler erster Art und beschreiben einen Fall, in welchem die beiden Altersdurchschnitte in der Grundgesamtheit tatsächlich gleich sind, aber fälschlicherweise auf unterschiedliche Altersdurchschnitte geschlossen wird.

2.2.2 Fehler erster Art

Anders ausgedrückt besagt die Nullhypothese hinter einem Test auf Gleichheit der Mittelwerte, dass kein Unterschied zwischen beiden Mittelwerten besteht. Ein Fehler erster Art tritt auf, wenn eine korrekte Nullhypothese fälschlicherweise abgelehnt wird. Dies bedeutet im obigen Beispiel zu dem Schluss zu kommen, dass der Altersdurchschnitt der Konsumenten beider Produkte unterschiedlich ist, obwohl dies gar nicht der Fall ist. Ein Fehler zweiter Art bedeutet dagegen, eine falsche Nullhypothese fälschlicherweise beizubehalten. Einen Fehler zweiter Art zu begehen hieße also im obigen Beispiel, zu dem Schluss zu kommen, dass der Altersdurchschnitt der Konsumenten beider Produkte gleich ist, obwohl dies gar nicht der Fall ist.

Tabelle 2.1 Fehler 1. Art und Fehler 2. Art im Überblick

	Wahrer Sachverhalt: $H_0$	Wahrer Sachverhalt: $H_1$
durch einen statistischen Test fällt eine Entscheidung für: $H_0$	richtige Entscheidung Wahrscheinlichkeit: $1-\alpha$ (Spezifität)	Fehler 2. Art Wahrscheinlichkeit: $\beta$
durch einen statistischen Test fällt eine Entscheidung für: $H_1$	Fehler 1. Art Wahrscheinlichkeit: $\alpha$	richtige Entscheidung Wahrscheinlichkeit: $1-\beta$ (Power, Teststärke, Sensitivität)

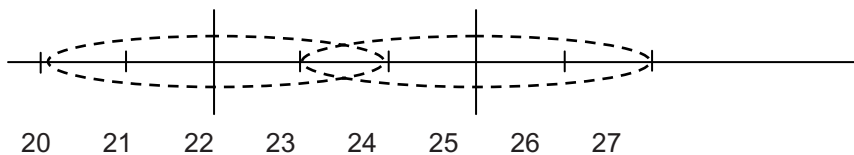


Da der Fehler erster Art in der Wissenschaft als „schlimmer“ angesehen wird als ein Fehler zweiter Art, richtet sich die angestrebte Sicherheit bei einer statistischen Schätzung meist nach dem Fehler erster Art. Diese wird als Spezifität bezeichnet. **Tabelle 2.1** gibt einen Überblick auf die beiden Fehler im Rahmen statistischer Signifikanzschätzungen.

### 2.2.3 Konfidenzintervall

Das Konfidenzintervall ist die erste von drei Möglichkeiten, die Mittelwerte zweier Stichproben auf Unterschiedlichkeit zu untersuchen und stellt den ersten Schritt zum Verständnis der statistischen Signifikanz dar. Das Wort Konfidenzintervall beschreibt eine Art Vertrauensbereich, der um den Mittelwert gebildet werden soll. Es wird also berücksichtigt, dass von 50 anwesenden Probanden nicht mit Sicherheit auf alle Konsumenten des Produkts geschlossen werden kann. Sehr wohl kann jedoch geschätzt werden, dass das durchschnittliche Alter aller Konsumenten in etwa so hoch ist wie bei den 50 anwesenden Probanden bzw. nur plus/minus einen gewissen Betrag um den Wert in der Stichprobe schwankt. Diesen Bereich um den Wert aus der Stichprobe bildet das Konfidenzintervall ab. Es berechnet sich aus den oben genannten Werten und nutzt die Standardnormalverteilung, um eine Wahrscheinlichkeit für die Lage des Mittelwerts der Grundgesamtheit zu ermitteln. Auf diese Weise wird ein Intervall rund um den Mittelwert aus der Stichprobe gebildet, in welchem der tatsächliche Wert der Grundgesamtheit voraussichtlich liegt. In unserem Beispiel 3 hieße dies, dass rund um das durchschnittliche Alter von 25 ein Vertrauensbereich gezogen wird, in welchem der tatsächliche Wert der Grundgesamtheit mit 95-prozentiger Sicherheit liegt. Die folgende **Abbildung 2.1** verdeutlicht diesen Zusammenhang und zeigt zwei fiktive Konfidenzintervalle rund um die Mittelwerte von 22 und 25.

**Abbildung 2.1** Beispielhaftes Konfidenzintervall 1



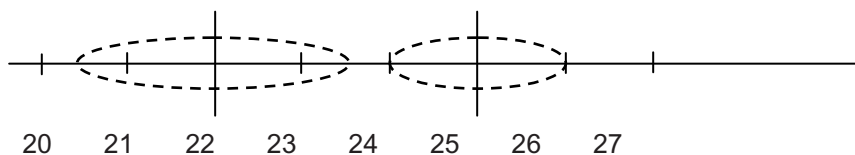
Da die Erläuterung des Konfidenzintervalls lediglich dem Verständnis der statistischen Signifikanz dienen soll, wird auf die Erklärung der Berechnung hier verzichtet. Vielmehr sollen die beiden Konfidenzintervalle zeigen, welche Aussagen in Bezug auf Mittelwertunterschiede getroffen werden können. Das Konfidenzintervall rund um den Mittelwert 22 zeigt, dass das wahre Durchschnittsalter aller Konsumenten der Jungen Brause mit 95-prozentiger Wahrscheinlichkeit zwischen 20 und 24 Jahren liegt. Das Konfidenzintervall rund um den Mittelwert 25 zeigt, dass das wahre Durchschnittsalter aller Konsumenten des Traditionsgetränks mit 95-prozentiger Wahrscheinlichkeit zwischen 23 und 27 Jahren



liegt. Die Konsumenten beider Produkte könnten somit beispielsweise ein Durchschnittsalter von 23,5 haben. Da dies nicht mit 95-prozentiger Wahrscheinlichkeit ausgeschlossen werden kann – oder anders: da sich beide Konfidenzintervalle überschneiden – kann man *nicht* mit 95-prozentiger Sicherheit sagen, dass die beiden Mittelwerte sich signifikant voneinander unterscheiden.

**Abbildung 2.2** zeigt hingegen eine Situation, in welcher beide Konfidenzintervalle sich nicht überschneiden. Hier kann mit 95-prozentiger Sicherheit festgestellt werden, dass das durchschnittliche Alter der Konsumenten beider Produkte auch in der Grundgesamtheit voneinander verschieden ist. Der Altersunterschied in den Stichproben ist somit nicht auf zufällige Effekte zurückzuführen. Oder anders gesagt: Konsumenten des Traditionsgetränks sind durchschnittlich signifikant älter als Konsumenten der Jungen Brause. Es sei an dieser Stelle darauf hingewiesen, dass die Konfidenzintervalle zweier Stichproben nicht gleich groß sein müssen. Die Größe der Konfidenzintervalle hängt vom Stichprobenumfang und der Varianz innerhalb der Stichprobe ab. Die Varianz bei den Konsumenten des Traditionsgetränks scheint gemäß **Abbildung 2.2** weniger groß zu sein als bei den Konsumenten der Jungen Brause. Das Konfidenzintervall ist daher kleiner. Mit 95-prozentiger Sicherheit kann hier davon ausgegangen werden, dass das Alter der Konsumenten des Traditionsgetränks zwischen 24 und 26 Jahren liegt, wohingegen das Alter der Konsumenten der Jungen Brause zwischen 20,5 und 23,5 liegt.

**Abbildung 2.2** Beispielhaftes Konfidenzintervall 2



## 2.2.4 t-Test und p-Wert

Der t-Test umgeht die Komplexität der Berechnung eines Konfidenzintervalls. Stattdessen nutzt er die Verteilungsfunktion, um einen kritischen Grenzwert zu ermitteln, der nach Berechnung des sogenannten t-Wertes nicht überschritten werden darf. Die Frage, ob sich die Konfidenzintervalle überschneiden, wird somit durch die Berechnung eines einzigen Wertes beantwortet. Dieser t-Wert berechnet sich genau wie das Konfidenzintervall aus den drei Komponenten Mittelwert, Varianz und Stichprobenumfang mit Hilfe der folgenden Formel:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

mit  $\bar{X}_1$  = Mittelwert in Stichprobe 1       $\bar{X}_2$  = Mittelwert in Stichprobe 2

$S_1^2$  = Varianz in Stichprobe 1       $S_2^2$  = Varianz in Stichprobe 2

$N_1$  = Größe der Stichprobe 1       $N_2$  = Größe der Stichprobe 2

Als Vergleichswert dient der sogenannte kritische t-Wert. Dieser ist je nach Stichprobengröße unterschiedlich hoch und kann in entsprechenden Tabellen zur t-Statistik abgelesen werden. Sind die Stichproben ausreichend groß, so beträgt für das Signifikanzniveau von 95% der kritische t-Wert 1,96. Es ergeben sich somit zwei mögliche Fälle: Ist der berechnete t-Wert nach obiger Formel größer als dieser kritische Wert von 1,96, dann kann davon ausgegangen werden, dass die Mittelwerte auch in den Grundgesamtheiten voneinander verschieden sind. Ist der berechnete t-Wert kleiner als der kritische Wert von 1,96, so kann *nicht* von unterschiedlichen Mittelwerten ausgegangen werden. Um festzustellen, ob die Mittelwerte von 25 und 22 mit 95-prozentiger Sicherheit auch in der Grundgesamtheit voneinander verschieden sind, müssen also lediglich Mittelwerte (25 und 22), Stichprobengrößen (jeweils 50) und Varianzen (bisher nicht gegeben) aus beiden Kursen bekannt sein und ein t-Wert gemäß der Formel berechnet werden. Nimmt man an dieser Stelle beispielhaft Varianzen von 5,5 und 6,1 an, so ergibt sich nach obiger Formel der folgende t-Wert:

$$t = \frac{25 - 22}{\sqrt{\frac{5,5^2}{50} + \frac{6,1^2}{50}}} = \frac{3}{\sqrt{1,35}} = 2,58$$

Der berechnete t-Wert von 2,58 liegt über dem kritischen t-Wert von 1,96. Somit kann davon ausgegangen werden, dass das durchschnittliche Alter aller Konsumenten der beiden Produkte unterschiedlich hoch ist. Die Konsumenten des Traditionsgetränks sind also im Durchschnitt älter als die der Jungen Brause. In der Stichprobe war dies ohnehin bereits der Fall, aber es war bisher unklar, ob dieser Unterschied zufällig zustande kam. Anhand des t-Wertes kann nun aber auch darauf geschlossen werden, dass der Unterschied für alle Konsumenten tatsächlich Gültigkeit hat. Und zwar zu einer Sicherheit bzw. einem Signifikanzniveau von 95 %.

Die meisten Statistikprogramme geben neben den t-Statistiken auch direkt das Signifikanzniveau an, bei dem sich zwei Mittelwerte signifikant voneinander unterscheiden. Dies erspart den Vergleich mit kritischen t-Werten. Stattdessen wird der Vergleich von berechnetem und kritischem t-Wert vom Programm (bspw. SPSS) vorgenommen. Als Ergebnis bleibt dann der sogenannte p-Wert (oder p-value), der angibt zu welchem Fehler erster Art der Mittelwert als unterschiedlich angenommen werden kann. Ein p-Wert von 0,05 oder geringer bedeutet demnach, dass die beiden Mittelwerte in den beiden Grundgesamt-

Grundlagen der Varianzanalyse

Konzeption - Durchführung - Auswertung

Huber, F.; Meyer, F.; Lenzen, M.

2014, VIII, 109 S. 51 Abb., Softcover

ISBN: 978-3-658-05665-0