

Einführung in Grundbegriffe der Statistik

Von Stichproben Rückschlüsse auf
Grundgesamtheiten ziehen

Wahrscheinlichkeiten kennenlernen

Hypothesen testen

Zwei Arten von Fehlern

Kapitel 1

Daten, Statistiken und Entscheidungen

Statistik? Dabei geht es doch vor allem um das Berechnen von Werten mithilfe von geheimnisvoll anmutenden Formeln, oder? Nicht wirklich. Im Rahmen der Statistik geht es vor allem darum, *Entscheidungen zu fällen*. Natürlich ist auch das Rechnen wichtig, jedoch besteht das primäre Ziel darin, diese Zahlen zu verwenden, um Entscheidungen zu treffen. Statistiker schauen sich Daten an und fragen sich, was diese Zahlen aussagen. Welche Trends zeichnen sich ab? Welche Vorhersagen können wir treffen? Welche Schlüsse können wir ziehen?

Um die Bedeutung der Daten zu verstehen und um diese Fragen zu beantworten, haben Statistiker eine beeindruckende Vielfalt an Analysetools entwickelt.

Was die Berechnungen angeht: Falls Sie diese mit Stift und Papier (oder mithilfe eines Taschenrechners) durchführen müssten, wären Sie vermutlich schnell entmutigt. Die durchzuführenden Berechnungen sind sehr umfangreich, und es würden sich schnell Fehler einschleichen. Softwaretools wie R unterstützen Sie bei der Rechenarbeit. Als Dreingabe kann R helfen, statistische Konzepte zu verstehen.

R wurde speziell für statistische Analysen entwickelt. Diese Programmiersprache implementiert viele der analytischen Werkzeuge, die Statistiker für das Füllen von Entscheidungen entwickelt haben. Ich habe dieses Buch geschrieben, um zu zeigen, wie Sie diese Werkzeuge bei Ihrer Arbeit einsetzen können.

Die statistischen (und damit verwandten) Begriffe, die Sie einfach kennen müssen

Die analytischen Werkzeuge, die R bereitstellt, basieren auf statistischen Konzepten, die ich im weiteren Verlauf dieses Kapitels vorstelle. Wie Sie sehen werden, basieren diese Konzepte auf dem gesunden Menschenverstand.

Stichproben und Grundgesamtheiten

Wenn Sie am Wahlabend fernsehen, wissen Sie, dass eines der wichtigen Ereignisse die Veröffentlichung der Prognose ist. Diese wird nach der Schließung der Wahllokale und lange, bevor die Stimmen ausgezählt sind, bekannt gegeben. Meistens liegen die Prognose und das Wahlergebnis sehr nahe beieinander. Woran liegt das?

Ganz einfach: Eine Stichprobe von Wählern wird nach der Stimmabgabe gefragt, wen oder welche Partei sie gewählt haben. Vorausgesetzt, dass die Wähler ehrlich sagen, wen sie gewählt haben, und vorausgesetzt, die Stichprobe ist für die Grundgesamtheit (oder Population) repräsentativ, können die Analysten aufgrund der Stichprobendaten Rückschlüsse auf die Grundgesamtheit der Wähler ziehen.

Das ist kurz das, worum es bei der Statistik geht: aufgrund der Daten von Stichproben Rückschlüsse auf die Grundgesamtheit zu ziehen, aus der die Stichprobe entnommen wurde.

Noch ein Beispiel. Nehmen wir einmal an, Ihre Aufgabe besteht darin, die durchschnittliche Größe aller zehn Jahre alten Kinder in den USA herauszufinden. Sie verfügen jedoch weder über die Zeit noch über die finanziellen Mittel, die Größe jedes einzelnen Kinds zu messen. Stattdessen messen Sie die Größe einer repräsentativen Stichprobe und berechnen dann den Mittelwert. Sie können dann den Mittelwert der Stichprobe zur Schätzung des Mittelwerts der Grundgesamtheit heranziehen.

Das Schließen auf den Mittelwert einer Grundgesamtheit ist eine Art *Inferenz*, die Statistiker aus Stichprobendaten ziehen. Die Inferenz wird im Abschnitt »Inferenzstatistik: Testen von Hypothesen« ausführlicher beschrieben.



Einige Begriffe, die Sie kennen sollten: Die Eigenschaften einer Grundgesamtheit (wie der Mittelwert einer Grundgesamtheit) werden als *Parameter* bezeichnet und die Eigenschaften einer Stichprobe (wie der Mittelwert einer Stichprobe) als *Statistiken*. Wenn Sie sich bei Ihren Betrachtungen auf die Eigenschaften der Stichproben beschränken (wie die Größe der Kinder in Ihrer Stichprobe), sind Ihre Statistiken *deskriptiv* oder *beschreibend*. Wenn Sie Ihren Horizont erweitern und sich mit Grundgesamtheiten beschäftigen, sind Ihre Statistiken *inferenziell*.



Einige Schreibweisen, die Sie kennen sollten: Statistiker verwenden griechische Buchstaben (μ , σ , ρ) für Parameter und lateinische Buchstaben (\bar{x} , s , r) für Statistiken. In Abbildung 1.2 sind die Beziehungen zwischen Grundgesamtheiten und Stichproben sowie zwischen Parametern und Statistiken dargestellt.

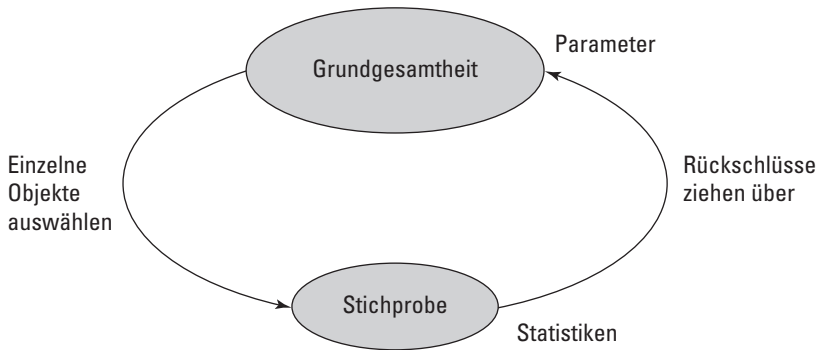


Abbildung 1.1: Die Beziehung zwischen Grundgesamtheit, Stichprobe, Parametern und Statistiken

Abhängige und unabhängige Variablen

Eine *Variable* ist etwas, das mehr als einen Wert annehmen kann – wie beispielsweise Ihr Alter, der Wert des Dollars in Euro oder die Anzahl der Spiele, die Ihre Lieblingsfußballmannschaft gewonnen hat. Etwas, das nur einen Wert annehmen kann, wird *Konstante* genannt. Wissenschaftler behaupten, dass die Lichtgeschwindigkeit eine Konstante ist. Wir verwenden die Konstante π , um beispielsweise den Flächeninhalt eines Kreises zu berechnen.

Für Statistiker sind zwei Arten von Variablen wichtig: *unabhängige Variablen* und *abhängige Variablen*. Beide Variablen tauchen in jeder Studie und Untersuchung auf, und Statistiker bewerten die Beziehung zwischen beiden.

Stellen Sie sich beispielsweise vor, es gebe eine neue computerbasierte Trainingsmethode, mit der der IQ einer Person verbessert werden kann. Wie würde ein Forscher untersuchen, ob diese Methode hält, was sie verspricht? Dazu würde er zuerst eine Stichprobe von Personen nach dem Zufallsprinzip in zwei Gruppen einteilen. Eine Gruppe verwendet die neue Trainingsmethode und die andere würde eine andere, ebenfalls computerbasierte Aktivität verwenden, beispielsweise das Lesen eines Texts auf einer Website. Der Forscher misst vor und nach dem Durchführen der jeweiligen Aktivität den IQ jeder Person. Was passiert dann? Dies beschreibe ich in einem der nächsten Abschnitte (»Inferenzstatistik: Testen von Hypothesen«).

Hier geht es zunächst darum, dass Sie wissen, dass die unabhängige Variable in diesem Beispiel die *jeweilige Aktivität* ist. Die beiden möglichen Werte dieser Variablen sind *IQ-Training* und *Text lesen*. Die abhängige Variable ist die Änderung des IQ, wie er vor und nach der jeweiligen Aktivität gemessen wurde.



Die abhängige Variable ist das, was der Forscher misst. In einem Experiment ist eine unabhängige Variable das, was der Forscher manipuliert, was er ändert. In anderen Zusammenhängen kann der Forscher eine unabhängige Variable nicht manipulieren. Stattdessen notiert er die sich natürlich ergebenden Werte der unabhängigen Variablen und wie sich diese auf eine abhängige Variable auswirken.



Grundsätzlich geht es darum, herauszufinden, ob Änderungen der unabhängigen Variablen mit Änderungen der abhängigen Variablen zusammenhängen.



In den Beispielen in diesem Buch erfahren Sie, wie Sie mit R verschiedene Eigenschaften von Wertegruppen berechnen oder wie Sie Wertegruppen vergleichen können. Denken Sie immer daran, dass ich mit einer Wertegruppe stets die Werte einer abhängigen Variablen meine.

Arten von Daten

Es gibt vier verschiedene Arten von Daten. Wenn Sie mit einer Variablen arbeiten, hängt es von der Datenart ab, wie Sie mit der Variablen arbeiten. Die erste Art wird als *nominalskalierte* oder *nominale* Daten bezeichnet. Wenn eine Zahl eine nominale Variable ist, handelt es sich lediglich um einen Namen. Der Zahlenwert bedeutet nichts. Ein gutes Beispiel hierfür ist die Zahl auf dem Trikot eines Sportlers. Sie dient lediglich der Identifizierung des Sportlers, um ihn von den anderen Mitgliedern seines Teams unterscheiden zu können. Die Zahl ist kein Hinweis auf das Können des Sportlers.

Als Nächstes kommen die *ordinalskalierten* oder *ordinalen* Daten. Bei ordinalen Daten geht es um Ordnung. Die Zahlen enthalten eine Bedeutung, die über die bloße Identifizierung hinausgeht. Wenn ich Sie bitte, zehn Lebensmittel zu bewerten, und zwar angefangen bei dem, das Sie am meisten mögen (eins), bis zu dem, das Sie am wenigsten mögen (zehn), dann haben wir einen Satz ordinalskalierter Daten.

Jedoch kann es sein, dass bei den Lebensmitteln die Differenz zwischen Ihrem dritten und vierten Platz nicht mit der Differenz zwischen Ihrem neunten und letzten Platz identisch ist. Was bei diesem Datentyp fehlt, ist das Konzept von gleichen Intervallen oder gleichen Differenzen.

Intervallskalierte Daten verwenden identische Differenzen. Temperaturangaben in Celsius und Fahrenheit sind Beispiele für intervallskalierte Daten. Die Differenz zwischen 10 °C und 20 °C ist dieselbe wie die Differenz zwischen 30 °C und 40 °C. Jedes Grad ist somit ein Intervall.

Menschen sind oft überrascht, wenn sie hören, dass 20 °C nicht doppelt so warm wie 10 °C ist. Um eine Aussage hinsichtlich der Relation (doppelt so viel wie, halb so viel wie) machen zu können, muss null bedeuten, dass vom gemessenen Attribut absolut nichts vorhanden ist. Ein Temperaturwert von 0 °C bedeutet jedoch nicht, dass keine Wärme vorhanden ist. 0 °C ist lediglich ein willkürlicher Punkt auf der Celsius-Skala. (Dies gilt übrigens auch für die Fahrenheit-Skala.)

Zum letzten Datentyp zählen *verhältnisskalierte* Daten. Hier ist ein sinnvoll interpretierbarer Nullpunkt vorhanden. Bei Temperaturangaben liefert die Kelvin-Skala verhältnisskalierte Daten. 100 °K ist doppelt so warm wie 50 °K. Dies liegt daran, dass der Nullpunkt der Kelvin-Skala ein *absoluter Nullpunkt* ist, bei dem es keine molekulare Bewegung (die Voraussetzung für Wärme) mehr gibt. Ein weiteres Beispiel ist das Lineal. 8 cm ist doppelt so lang wie 4 cm. Der Wert 0 cm bedeutet, dass keine Länge vorhanden ist.



All diese Datenarten können die Basis einer unabhängigen oder einer abhängigen Variablen bilden. Welche Analysetools Sie verwenden, hängt von der Art der Daten ab, mit denen Sie zu tun haben.

Ein bisschen Wahrscheinlichkeit

Wenn Statistiker Rückschlüsse ziehen, drücken sie ihre Einschätzung der Glaubwürdigkeit dieser Rückschlüsse in Form von Wahrscheinlichkeiten aus. Sie können sich ihrer Rückschlüsse nie sicher sein. Sie können nur sagen, für wie *wahrscheinlich* sie ihre Rückschlüsse halten.

Was also ist Wahrscheinlichkeit? Von Mathematikern und Philosophen erhalten Sie hierfür komplexere Definitionen. Meiner Meinung nach ist es jedoch am einfachsten, das Konzept der Wahrscheinlichkeit anhand von ein paar Beispielen zu erläutern.

Hier ein einfaches Beispiel: Wie groß ist die Wahrscheinlichkeit, dass beim Werfen einer Münze Kopf geworfen wird? Wenn sich die Münze symmetrisch verhält, wissen Sie intuitiv, dass die Chancen für Kopf ebenso wie die für Zahl 50 : 50 stehen. Im Hinblick auf die zur Wahrscheinlichkeit gehörende Art der Zahlen ist das $1/2$.

Und wie ist das beim Würfeln? Wie groß ist die Wahrscheinlichkeit, dass Sie eine 4 würfeln? Hmm ... Ein Würfel hat sechs Flächen, und eine davon zeigt die 4, also sollte die Wahrscheinlichkeit bei 1 zu 6 liegen, richtig? Richtig. Noch ein Beispiel. Sie ziehen aus einem Stapel Spielkarten wahllos eine Karte heraus. Wie groß ist die Wahrscheinlichkeit, dass Sie Kreuz ziehen? Nun, ein Kartenspiel hat vier Farben, also lautet die Antwort 1 zu 4.

Ich glaube, Sie verstehen, worum es geht. Wenn Sie ermitteln möchten, mit welcher Wahrscheinlichkeit ein Ereignis eintritt, müssen Sie herausfinden, wie häufig dieses Ereignis eintreten kann, und Sie müssen diese Anzahl durch die Gesamtzahl aller möglichen Ereignisse teilen. Bei den ersten beiden Beispielen (Kopf, 4) tritt das Ereignis, an dem Sie interessiert sind, lediglich einmal ein (eine Münze, ein Würfel). Daher teilen wir bei der Münze 1 durch 2 (zwei Seiten) und beim Würfel teilen wir 1 durch 6 (sechs Flächen). Beim dritten Beispiel (Kreuz) kann das Ereignis durch eine von 13 Karten ausgelöst werden (Ass bis König). Daher teilen wir 13 durch 52 (die Anzahl der Karten) und erhalten $1/4$.

Das Ganze kann jedoch noch etwas komplexer werden. Hinsichtlich der Wahrscheinlichkeit stellen sich noch weitere Fragen. Nehmen wir einmal an, Sie würfeln und werfen gleichzeitig eine Münze. Wie groß ist die Wahrscheinlichkeit, dass Sie eine 4 würfeln und Kopf werfen? Berücksichtigen Sie alle möglichen Ereignisse, die eintreten können, wenn Sie würfeln und gleichzeitig eine Münze werfen. Sie können Kopf und die Zahlen 1 bis 6 oder Zahl und die Zahlen 1 bis 6 werfen. Das ergibt insgesamt 12 Möglichkeiten. Für Kopf und 4 gibt es nur eine Möglichkeit. Also lautet die Antwort $1/12$.

Die Formel für die Wahrscheinlichkeit, mit der ein bestimmtes Ereignis eintritt, lautet wie folgt:

$$P(\text{Ereignis}) = \frac{\text{Anzahl der Möglichkeiten, mit denen ein Ereignis eintreten kann}}{\text{Gesamtzahl der möglichen Ereignisse}}$$

Ich habe diesen Abschnitt mit der Feststellung begonnen, dass Statistiker ihre Einschätzung der Glaubwürdigkeit von Rückschlüssen in Form von Wahrscheinlichkeiten ausdrücken, weshalb ich eigentlich auf dieses Thema gekommen bin. Wenn wir in diese Richtung weiterdenken, stoßen wir auf den Begriff der *bedingten* Wahrscheinlichkeit, also der Wahrscheinlichkeit, mit der ein Ereignis eintritt, wenn ein anderes Ereignis eintritt. Angenommen, ich würfle, schaue mir das Ergebnis an (sodass Sie es nicht sehen können) und sage Ihnen, dass ich eine ungerade Zahl gewürfelt habe. Wie groß ist die Wahrscheinlichkeit, dass ich eine 5 gewürfelt habe? Eigentlich beträgt die Wahrscheinlichkeit einer 5 $1/6$, aber ich habe die Auswahl begrenzt. Ich habe die drei geraden Zahlen (2, 4 und 6) als Möglichkeiten ausgeschlossen. Somit sind nur noch die drei ungeraden Zahlen (1, 3 und 5) möglich, sodass die Wahrscheinlichkeit, dass eine 5 gewürfelt wird, nun $1/3$ beträgt.

Was hat nun die bedingte Wahrscheinlichkeit mit statistischer Analyse zu tun? Lesen Sie weiter.

Inferenzstatistik: Testen von Hypothesen

Vor dem Durchführen einer Studie formuliert ein Statistiker eine *Hypothese*, das heißt, er stellt eine vorsichtige Prognose auf, welches bestimmte Ergebnis zu erwarten ist. Wenn nach Abschluss der Untersuchung die Stichprobendaten in einer Tabelle erfasst sind, trifft er die zentrale Entscheidung, die ein Statistiker treffen muss: Er entscheidet, ob die Hypothese verworfen oder nicht verworfen wird.

Diese Entscheidung hängt von der Frage nach der bedingten Wahrscheinlichkeit ab: Wie groß ist die Wahrscheinlichkeit, dass sich diese Daten unter der Voraussetzung ergeben, dass die Hypothese zutrifft? Die statistische Analyse stellt Tools zum Berechnen der Wahrscheinlichkeit bereit. Wenn sich die Wahrscheinlichkeit als gering erweist, verwirft der Statistiker die Hypothese.

Zurück zum Beispiel mit dem Werfen von Münzen. Nehmen wir einmal an, Sie möchten wissen, ob eine bestimmte Münze symmetrisch ist, das heißt, ob Kopf ebenso häufig geworfen wird wie Zahl. Lassen Sie uns zuerst von der Hypothese ausgehen, dass die Münze symmetrisch ist.

Um diese Hypothese zu testen, werfen Sie die Münze beispielsweise 100-mal. Diese 100 Würfe stellen Ihre Stichprobendaten dar. Wenn Sie von der Hypothese ausgehen, dass sich die Münze symmetrisch verhält, erwarten Sie, dass die Daten in Ihrer Stichprobe mit 100 Würfeln 50-mal Kopf und 50-mal Zahl ergeben.

Wenn sich herausstellt, dass Sie 99-mal Kopf und 1-mal Zahl werfen, werden Sie die Hypothese von der symmetrischen Münze zweifellos verwerfen. Warum? Die bedingte Wahrscheinlichkeit, dass mit einer symmetrischen Münze 99-mal Kopf und 1-mal Zahl geworfen wird, ist sehr gering. Aber einen Moment. Die Münze kann symmetrisch sein, und Sie können dennoch 99-mal Kopf und 1-mal Zahl werfen, richtig? Absolut. Das weiß man nie so genau. Sie müssen Stichprobendaten sammeln (das Ergebnis aus 100 Würfeln) und Rückschlüsse ziehen. Die Rückschlüsse können richtig sein oder auch nicht.

Geschworene stehen ständig vor dieser Frage. Sie müssen zwischen widersprüchlichen Hypothesen entscheiden und die Indizien vor Gericht begründen. (Stellen Sie sich die Indizien als Daten vor.) Eine Hypothese lautet, dass der Angeklagte schuldig ist. Die andere Hypothese lautet, dass der Angeklagte unschuldig ist. Die Geschworenen müssen unter Berücksichtigung der Indizien im Prinzip die Frage nach der bedingten Wahrscheinlichkeit beantworten. Wie groß ist die Wahrscheinlichkeit des Indizes, vorausgesetzt, der Angeklagte ist nicht schuldig? Diese Frage wird durch den Urteilsspruch beantwortet.

Nullhypothese und Alternativhypothese

Betrachten wir noch einmal das eben beschriebene Experiment mit dem Münzenwerfen. Die Ergebnisse aus 100 Würfeln stellen die Stichprobendaten dar. Wie erwähnt können wir vor dem Werfen der Münze die Hypothese formulieren, dass die Münze symmetrisch ist, das heißt, Sie erwarten, dass Kopf und Zahl gleich häufig geworfen werden. Dieser Ausgangspunkt wird als *Nullhypothese* bezeichnet. In der Statistik wird für die Nullhypothese die Notation H_0 verwendet. Nach dieser Hypothese ist jede Kopf-Zahl-Verteilung in den Daten mit einer symmetrischen Münze vereinbar. Stellen Sie sich das Ganze so vor, dass nichts in den Ergebnissen der Untersuchung außer der Reihe ist.

Eine alternative Hypothese ist möglich, nämlich dass die Münze nicht symmetrisch ist und daher Kopf und Zahl nicht gleich häufig geworfen werden. Diese Hypothese besagt, dass mit einer nicht symmetrischen Münze jede Kopf-Zahl-Verteilung vereinbar ist. Ob Sie es glauben oder nicht: Die alternative Hypothese wird *Alternativhypothese* oder *Gegenhypothese* genannt. In der Statistik wird für die Alternativhypothese die Notation H_1 oder H_A verwendet.

Werfen Sie, diese Hypothesen vorausgesetzt, die Münze 100-mal und notieren Sie die Anzahl der Kopf- und Zahl-Würfe. Wenn sich dabei ergibt, dass etwa 90-mal Kopf und 10-mal Zahl geworfen wird, sollten Sie H_0 verwerfen. Wenn sich ergibt, dass Kopf und Zahl jeweils etwa 50-mal geworfen werden, sollten Sie H_0 nicht verwerfen.

Ähnliches gilt für das Beispiel mit dem IQ weiter vorne in diesem Kapitel. Eine Stichprobe verwendet die computerbasierte IQ-Trainingsmethode, während die andere Stichprobe an einer anderen, ebenfalls computerbasierten Aktivität teilnimmt, nämlich dem Lesen von Text auf einer Website. Der IQ der Teilnehmer wird von den Forschern vor und nach der jeweiligen Aktivität gemessen. Die Nullhypothese H_0 besagt, dass sich bei der Verbesserung des IQ der beiden Gruppen keine Unterschiede ergeben. Wenn der Fortschritt mit der IQ-Trainingsmethode größer ist als der mit der anderen Aktivität, und zwar so viel größer, dass es unwahrscheinlich ist, dass sich die Methoden nicht voneinander unterscheiden, verwerfen Sie H_0 . Wenn nicht, dann verwerfen Sie H_0 nicht.



Ist Ihnen aufgefallen, dass ich *nicht* gesagt habe: »Nehmen Sie H_0 an«? So, wie die Logik nun mal funktioniert, können Sie eine Hypothese *niemals* annehmen. Sie können H_0 verwerfen oder Sie können H_0 nicht verwerfen.

Hier noch ein Beispiel aus dem echten Leben, das helfen soll, dieses Konzept zu verstehen. Wenn ein Angeklagter vor Gericht erscheint, gilt die Unschuldsvermutung, bis das

Gegenteil bewiesen ist. Sie können sich »nicht schuldig« als H_0 vorstellen. Die Aufgabe des Staatsanwalts besteht darin, das Gericht dazu zu bringen, H_0 zu verwerfen. Wenn das Gericht die Unschuldsvermutung ablehnt, wird der Angeklagte für schuldig befunden. Wenn es die Unschuldsvermutung nicht verwirft, dann ist der Angeklagte nicht schuldig. Es geht also immer um »schuldig« und »nicht schuldig« und nicht um »schuldig« und »unschuldig«. »Unschuldig« wäre gleichbedeutend damit, H_0 zu akzeptieren.

Lassen Sie uns zum Beispiel mit dem Münzenwerfen zurückkehren. Ist Ihnen aufgefallen, dass ich beim Beispiel mit dem Münzenwerfen »etwa 50-mal« geschrieben hatte? Was bedeutet dieses »etwa«? Außerdem habe ich geschrieben, dass Sie H_0 verwerfen sollen, wenn Kopf und Zahl im Verhältnis 90 : 10 geworfen wird. Aber was ist, wenn 85 : 15 geworfen wird? 80 : 20? 70 : 30? Wie groß muss die Differenz zur Verteilung 50 : 50 sein, damit H_0 verworfen wird? Um wie viel größer muss beim Beispiel mit der IQ-Trainingsmethode die Verbesserung sein, damit H_0 verworfen wird?

Ich werde diese Fragen hier nicht beantworten. Statistiker haben Entscheidungsregeln für Situationen wie diese entwickelt, und Sie werden diese Regeln im Verlauf dieses Buchs kennenlernen.

Zwei Arten von Fehlern

Beim Auswerten der Daten aus einer Untersuchung und beim Entscheiden, ob H_0 verworfen werden soll oder nicht, können Sie nie absolut sicher sein. Sie wissen nie, wie die Realität wirklich aussieht. Im Zusammenhang mit dem Münzwurfbeispiel bedeutet das, dass Sie nie sicher wissen, ob die Münze symmetrisch ist. Ihnen bleibt nur, eine Entscheidung anhand der gesammelten Stichprobendaten zu treffen. Wenn Sie, was die Münze betrifft, sichergehen möchten, müssen Sie alle Daten für die gesamte Grundgesamtheit der Würfe sammeln. Das bedeutet, Sie müssten die Münze bis ans Ende aller Tage werfen.

Da Ihre Entscheidung nie sicher ist, ist es möglich, dass Sie einen Fehler machen, gleichgültig, wie Sie entscheiden. Wie bereits erwähnt, kann die Münze symmetrisch sein, und Sie können bei 100 Würfeln dennoch ein Ergebnis von 99 : 1 erhalten. Das ist nicht wahrscheinlich, weshalb Sie H_0 verwerfen. Es ist außerdem möglich, dass die Münze nicht symmetrisch ist, und bei 100 Würfeln dennoch 50-mal Kopf geworfen wird. Auch das ist nicht wahrscheinlich, weshalb Sie H_0 in diesem Fall nicht verwerfen.

Obwohl diese Fehler nicht wahrscheinlich sind, sind sie dennoch möglich. Sie kommen in jeder Untersuchung vor, bei der Inferenzstatistik im Spiel ist. Statistiker nennen diese Fehler *Fehler 1. Art* (oder auch Alphafehler) und *Fehler 2. Art* (oder auch Betafehler).

Wenn Sie H_0 verwerfen, obwohl Sie das nicht sollten, dann ist das ein Fehler 1. Art. Das wäre bei dem Beispiel mit der Münze das Verwerfen der Hypothese, die besagt, dass die Münze symmetrisch ist, obwohl die Münze tatsächlich symmetrisch ist.

Wenn Sie H_0 nicht verwerfen, obwohl Sie das sollten, dann ist das ein Fehler 2. Art. Das ist dann der Fall, wenn Sie die Hypothese, die besagt, dass die Münze symmetrisch ist, nicht verwerfen, obwohl die Münze tatsächlich nicht symmetrisch ist.

Woher wissen Sie, ob Sie einen dieser Fehler gemacht haben? Das können Sie nicht wissen, zumindest nicht gleich, nachdem Sie entschieden haben, ob Sie H_0 verwerfen oder nicht. (Wenn es möglich wäre, das zu wissen, würden Sie den Fehler erst gar nicht machen!) Ihnen bleibt nur, weitere Daten zu sammeln und zu prüfen, ob die zusätzlichen Daten mit Ihrer Entscheidung vereinbar sind.

Wenn Sie meinen, H_0 neige dazu, den Status quo zu erhalten, und nichts als außergewöhnlich interpretieren (gleichgültig, wie es aussieht), bedeutet ein Fehler 2. Art, dass Sie etwas Wichtiges übersehen haben. So betrachtet basieren viele ironische Ereignisse in der Geschichte auf Fehlern 2. Art.

Hierfür ein Beispiel: Am Neujahrstag 1962 hört sich ein Musikproduzent die Demokassette einer Band an, die aus drei Gitarristen und einem Schlagzeuger besteht. Dem Musikproduzenten gefiel das, was er hörte, überhaupt nicht, er mochte nicht, was er sah, und er war der festen Überzeugung, dass die Zeit von Gitarrenbands vorbei ist. Obwohl die Musiker ihr Bestes gaben, wurde die Gruppe (zumindest von dieser Plattenfirma) nicht unter Vertrag genommen.

Wissen Sie, um welche Band es sich handelt? The Beatles!

Das war in der Tat ein Fehler 2. Art.

