

Kapitel 1

Der Zufall in unserer Welt — Einführende Beispiele und Grundbegriffe

Jeder hat eine Vorstellung davon, was man unter **Statistik** versteht, und viele denken dabei sicherlich zunächst an umfangreiche Tabellen oder grafische Darstellungen, die bestimmte Sachverhalte in komprimierter Weise verdeutlichen. Dies ist jedoch nur ein Teil der Statistik, die sogenannte beschreibende oder **deskriptive Statistik**, die dazu dient, umfangreiche Datensätze mit Hilfe von Abbildungen und Kennzahlen anschaulich darzustellen.

In den meisten Fällen geht die Statistik jedoch weit über die reine Beschreibung von Datensätzen hinaus. In der Regel sind vorliegende Daten nur eine Stichprobe aus einer sogenannten Grundgesamtheit, und man möchte aus der Stichprobe Schlussfolgerungen für die Grundgesamtheit ziehen. Dieser Teil der Statistik wird schließende oder **induktive Statistik** genannt.

Zu Beginn dieses ersten Kapitels werden zunächst einige praktische Anwendungsbeispiele statistischer Methoden vorgestellt, um einen Eindruck von den vielfältigen Anwendungsmöglichkeiten der Statistik zu vermitteln. Im hinteren Teil des Kapitels werden dann einige wichtige Grundbegriffe der Statistik, wie zum Beispiel Stichprobe und Grundgesamtheit, eingeführt.

1.1 Deterministische und stochastische Modelle

Bevor die einführenden Anwendungsbeispiele statistischer Methoden und **Modelle** vorgestellt werden, sollen zunächst die Begriffe **deterministisches Modell** und **stochastisches Modell** erläutert werden. Dazu ist zunächst der Begriff des Modells zu definieren. Ein Modell lässt sich etwa als vereinfachte Beschreibung der Realität definieren. Ein anschauliches Beispiel ist eine Landkarte, die eine bestimmte, reale Landschaft vereinfacht auf einem Blatt Papier beschreibt.

Man hat es in der Statistik immer mit Daten zu tun, d.h. mit Größen, die gemessen, gezählt oder auf andere Art und Weise quantifiziert werden können. Statistische Modelle werden durch mathematische Formeln, durch Zahlen oder als Grafik gegeben. Auf dieser Grundlage ist eine engere Definition des Modells sinnvoll:

Ein **Modell** ist die Beschreibung eines quantitativ erfassbaren Phänomens.

Die nachfolgenden Beispiele und Bemerkungen dienen dazu, einen Eindruck von der Bedeutung stochastischer, d.h. zufallsabhängiger Sachverhalte in verschiedenen Bereichen des Lebens, sowie der Anwendung statistischer Modelle in diesem Zusammenhang zu vermitteln. Lediglich das erste Beispiel ist nicht durch zufällige Einflüsse geprägt. Es stellt eher die Ausnahmesituation als die Regel dar.

Beispiel 1.1. Schwingungsdauer eines Pendels

Zunächst wird ein Beispiel für ein deterministisches Modell betrachtet, und zwar für die Schwingungsdauer eines Pendels. Die Physik liefert dazu eine Theorie, aus der man ableiten kann, dass die Schwingungsdauer T von der Länge L des Pendels abhängt und durch die Gleichung

$$T = 2\pi\sqrt{\frac{L}{g}}$$

beschrieben wird, wobei π die Kreiszahl und g die Erdbeschleunigung bezeichnet.

Um die Schwingungsdauer T eines realen Pendels zu berechnen und somit das Modell für einen bestimmten Zweck zu verwenden, benötigt man die Länge des Pendels und die Erdbeschleunigung am Ort des Pendels. In Göttingen z.B. ist g etwa 9.81 m/s^2 (die Einheit ist hier *Meter dividiert durch Sekunde zum Quadrat*); dann ergibt sich beispielsweise für $L = 7.5\text{ m}$ die Schwingungsdauer

$$T = 2\pi\sqrt{\frac{7.5\text{ m}}{9.81\text{ m/s}^2}} = 5.5\text{ s}.$$

Dabei wird vorausgesetzt, dass der Pendelausschlag klein gegenüber der Pendellänge ist, d.h. dass das Pendel nur kleine Winkel durchläuft. Man benötigt also nur die Länge des Pendels L , um die Schwingungsdauer T zu bestimmen.

Diese Formel ist ein Beispiel für ein Modell, das den quantitativen Zusammenhang zwischen zwei Größen, der Länge eines Pendels und der Schwingungsdauer, beschreibt (bei gegebener Erdbeschleunigung). Grafisch ist dieser Zusammenhang in Abb. 1.1 dargestellt. Sowohl die Grafik als auch die Formel sind als Modell für das Pendel zu verstehen. Lediglich die Darstellungsform ist eine andere. Die Anwendungsmöglichkeit des Modells ist offensichtlich: wir können mit Hilfe dieses Modells die Schwingungsdauer T für verschiedene Werte von L bestimmen.

Falls jetzt die Frage aufkommt, aus welchem Grund man sich für die Schwingungsdauer eines Pendels interessieren sollte: bis zur Erfindung der Quarz-Uhren in den 1930er Jahren waren Pendeluhrn über Jahrhunderte das genaueste Mittel der Zeitmessung, und auch heute noch hat sicher der eine oder andere eine Pendeluhr zu Hause.

Das mathematische Pendel ist ein Beispiel für eine **deterministische Beziehung** zwischen zwei Größen. Mit *deterministisch* ist gemeint, dass es für jeden Wert der

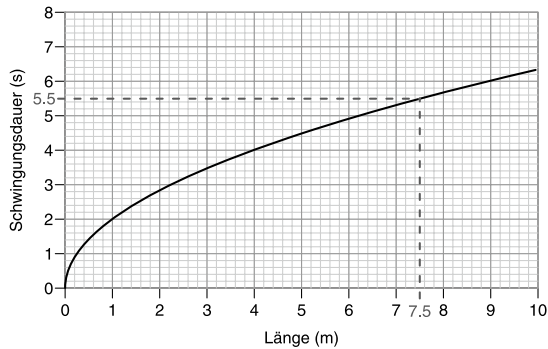


Abb. 1.1 Modell für die Schwingungsdauer eines Pendels in Abhängigkeit von der Länge

Länge L *genau einen* Wert für die Schwingungsdauer T gibt. Die Schwingungsdauer ist durch die Länge des Pendels determiniert. Es gibt hier keine Unsicherheit oder Unbestimmtheit. Wenn die Länge des Pendels bekannt ist, kennt man auch die Schwingungsdauer. Ganz anders ist die Beziehung zwischen zwei Größen in dem folgenden Beispiel.

Beispiel 1.2. Blockzeit eines Linienfluges

Wer privat oder geschäftlich mit dem Flugzeug reist, ist sicherlich nicht nur daran interessiert, sicher am Ziel anzukommen, sondern auch möglichst schnell und pünktlich. Dabei hängt die Dauer eines Linienfluges in erster Linie von der Länge der Flugstrecke ab.

Tabelle 1.1 enthält die Strecke d in nautischen Meilen (entspricht 1 852 Metern) sowie die dazugehörige Dauer t in Minuten für 100 zufällig ausgewählte inneramerikanische Flüge (mit einer Flugstrecke von maximal 1 500 Meilen) der Fluggesellschaft American Airlines im Februar 2006.¹ Die Dauer umfasst dabei die Zeit vom Losrollen eines Flugzeugs von der Start-Position bis zum Stillstand auf der Ziel-Position (*on blocks*) und wird daher auch *Blockzeit* genannt. Sie beinhaltet neben der reinen Flugzeit auch die sogenannte *Taxi-Out-Zeit* (Zeit vom Losrollen bis zum Abheben) sowie die *Taxi-In-Zeit* (Zeit vom Aufsetzen bis zum Stillstand). Die im Flugplan einer Fluggesellschaft angegebene Dauer eines Fluges stellt immer die geplante Blockzeit dar, und nicht die reine Flugzeit.

Wenn man Tabelle 1.1 betrachtet, stellt man fest, dass es den erwarteten Zusammenhang zwischen der Flugstrecke und der Blockzeit gibt; je länger die Flugstrecke, desto länger ist tendenziell die Blockzeit. Die Beziehung zwischen Flugstrecke und Blockzeit ist jedoch von anderer Art als die oben betrachtete Beziehung zwischen der Länge und der Schwingungsdauer eines Pendels. Im Falle des Pendels gehört

¹ Die Original-Daten der American Airlines Flüge im Februar 2006, die im Rahmen dieses Beispiels betrachtet werden, stammen aus der *Airline On-Time Performance Data* Datenbank, die das US-amerikanische Bureau of Transportation Statistics auf seiner Internetseite <http://www.transtats.bts.gov> zur Verfügung stellt (Stand 24. April 2008).

Tabelle 1.1 Flugstrecke d in nautischen Meilen und Blockzeit t in Minuten für 100 zufällig ausgewählte inneramerikanische American Airlines Flüge (mit einer maximalen Flugstrecke von 1 500 Meilen) im Februar 2006

d	258	1 189	1 145	258	403	612	175	733	337	761	783	468	762	1 017	888
t	64	195	178	72	78	146	46	138	70	144	100	79	175	137	150
d	748	733	416	1 437	950	888	1 121	1 235	988	1 055	583	1 217	868	1 235	1 171
t	126	105	98	220	154	143	193	193	168	174	106	207	160	170	182
d	1 145	1 062	1 389	733	1 045	1 440	190	175	1 313	175	950	868	190	1 205	551
t	204	203	197	148	158	210	67	50	182	53	147	155	63	199	115
d	1 171	1 045	236	583	1 035	1 471	867	1 162	1 017	1 055	1 171	551	1 235	641	1 068
t	173	142	65	124	179	195	126	185	172	183	196	102	181	118	168
d	569	1 431	190	733	1 464	1 235	177	190	247	786	551	1 055	592	1 182	1 213
t	89	243	49	131	199	165	62	59	82	124	96	162	115	189	166
d	551	1 302	1 372	448	190	867	762	987	678	334	964	612	1 144	177	551
t	82	182	197	86	58	167	128	164	110	86	140	142	167	59	96
d	762	612	603	1 456	1 189	861	522	1 005	733	1 438					
t	141	128	95	222	177	149	114	159	149	212					

zu jedem Wert der Länge genau ein Wert für die Schwingungsdauer. Die Schwingungsdauer ist durch die Länge eindeutig bestimmt. Im Beispiel mit der Blockzeit ist das anders. Es gibt z.B. 5 Flüge mit einer Flugstrecke von 733 Meilen und dazu-gehörigen Blockzeiten von 138, 105, 148, 131 und 149 Minuten. Die Blockzeit ist also nicht eindeutig durch die Flugstrecke bestimmt. Vielmehr scheint es zufällige Schwankungen zu geben. Eine solche Beziehung nennt man stochastisch. Eine grafische Darstellung der Beziehung zwischen Flugstrecke und Blockzeit gibt Abb. 1.2.

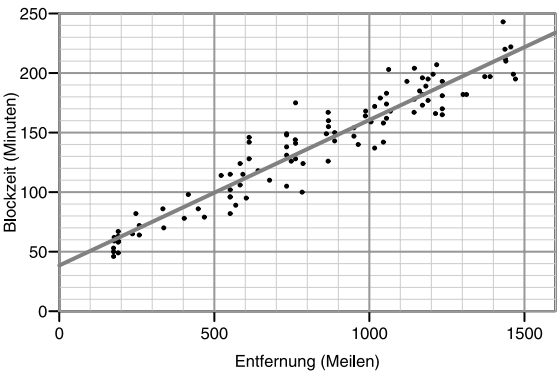


Abb. 1.2 Flugstrecke d in nautischen Meilen und Blockzeit t in Minuten für 100 zufällig ausgewählte inneramerikanische American Airlines Flüge (mit einer maximalen Flugstrecke von 1 500 Meilen) im Februar 2006 sowie angepasste Gerade

Neben den Beobachtungen ist in Abb. 1.2 auch eine an die Daten angepasste Gerade eingezeichnet (wie man eine solche Gerade bestimmt, wird in Kapitel 11 gezeigt). Die eingetragenen Punkte liegen nur annähernd auf der Geraden. Das ist der entscheidende Unterschied zum vorigen Beispiel. Hier tritt eine zufällige Variation oder Streuung auf. Die Blockzeit ist nicht nur durch die Flugstrecke bestimmt, sondern es treten noch weitere Einflussfaktoren auf, die für die zufällige Variation verantwortlich sind; beispielsweise spielt die Verkehrsdichte am Start- und Zielflughafen, sowie im durchfliegenden Luftraum eine Rolle. Die Beziehung ist damit nicht **deterministisch**, sondern **stochastisch**.

Auch wenn die tatsächlichen Blockzeiten schwanken, ist es dennoch von Nutzen, die annähernde Gerade zu kennen, um die Blockzeit für eine bestimmte Strecke, beispielsweise für die Flugplanung, zumindest ungefähr abschätzen zu können. Die wichtigste Einflussgröße auf die Blockzeit ist neben der Flugstrecke die Flugrichtung; so ist z.B. ein Flug von Los Angeles nach New York auf Grund des Rückenwindes in der Regel deutlich kürzer als der entsprechende Flug von New York nach Los Angeles. Doch auch wenn man die Flugdauer für eine ganz bestimmte Flugstrecke betrachtet, kann man deutliche zufällige Schwankungen beobachten. Dies verdeutlicht Tabelle 1.2, die eine Zusammenfassung der 174 verfügbaren Blockzeiten der Flüge von American Airlines von Dallas / Fort Worth (DFW) nach Philadelphia (PHL) im Februar 2006 darstellt.

Tabelle 1.2 Blockzeiten der American Airlines Flüge auf der Strecke DFW-PHL im Februar 2006

[150;160]	[160;170]	[170;180]	[180;190]	[190;200]	[200;210]	[210;220]	[220;230]	[230;240]
7	24	42	54	29	11	5	1	1

Laut Tabelle 1.2 schwankte die Blockzeit der American Airlines Flüge von Dallas nach Philadelphia im Februar 2006 zwischen 150 und 240 Minuten, wobei die meisten Blockzeiten zwischen 180 und 190 Minuten lagen. Laut Flugplan betrug die geplante Blockzeit je nach Verbindung zwischen 180 und 189 Minuten, bei einer Flugstrecke von 1 302 Meilen.

Da man normalerweise Informationen aus einer Grafik wesentlich leichter und schneller erfassen kann, sind die Daten aus Tabelle 1.2 in Abb. 1.3 noch einmal als sogenanntes Histogramm dargestellt. Bei der Erstellung eines Histogramms teilt man in der Regel die Häufigkeiten durch die gesamte Zahl an Beobachtungen und durch die Klassenbreite (die Breite des Intervalls) und stellt das Ergebnis dann als Rechteck über dem entsprechenden Intervall dar. Eine detaillierte Beschreibung der Konstruktion von Histogrammen folgt in Kapitel 2.

Das Histogramm vermittelt einen guten Eindruck von der Verteilung der Blockzeiten. Zusätzlich ist in Abb. 1.3 noch eine Kurve eingezeichnet. Diese Kurve stellt ein stochastisches Modell zur Beschreibung der Blockzeiten dar und kann als Glättung des Histogramms interpretiert werden. Wie dieses stochastische Modell bestimmt worden ist und ob es die Blockzeiten angemessen beschreibt, wird in späteren Kapiteln noch erläutert. An dieser Stelle soll nur festgehalten werden, dass

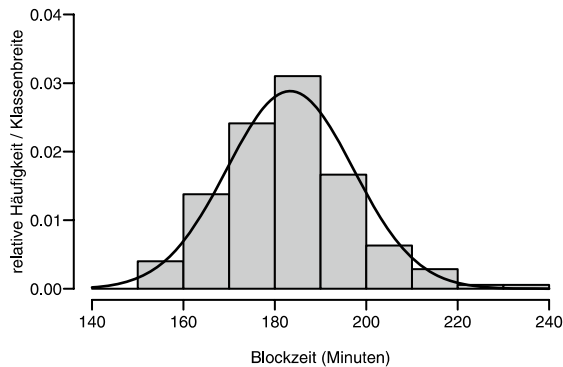


Abb. 1.3 Histogramm der Blockzeiten der American Airlines Flüge auf der Strecke DFW-PHL im Februar 2006 und angepasste Normalverteilung

die Blockzeiten zufälligen Schwankungen unterliegen, wobei sie jedoch gewisse Muster aufweisen (viele Beobachtungen in der Mitte, wenige am Rand), die durch ein stochastisches Modell beschrieben werden können. Mit Hilfe dieses Modells ist es dann möglich, bestimmten Beobachtungen Wahrscheinlichkeiten zuzuordnen. Laut dem Histogramm der Blockzeiten, sowie der eingezeichneten Kurve ist es beispielsweise wahrscheinlicher, dass die tatsächliche Blockzeit zwischen 180 und 190 Minuten beträgt, als dass sie unter 160 Minuten oder über 210 Minuten liegt. Was genau eine Wahrscheinlichkeit ist und wie man sie mit Hilfe eines stochastischen Modells bestimmt, wird ebenfalls in den späteren Kapiteln detailliert erläutert.

Nach Betrachtung der Beispiele der Schwingungsdauer eines Pendels sowie der Blockzeit eines Linienfluges können wir den Unterschied zwischen deterministischen und stochastischen Modellen folgendermaßen beschreiben:

Treten bei den betrachteten Phänomenen *zufällige Schwankungen* auf, so ist damit der Begriff *Wahrscheinlichkeiten* verbunden. Ein **stochastisches Modell** ist zur Beschreibung erforderlich.

Spiele bei den betrachteten Phänomenen zufällige Schwankungen keine Rolle, dann kommt man mit einem **deterministischen Modell** aus.

Hat man es mit Phänomenen zu tun, die durch ein deterministisches Modell beschrieben werden können, dann weiß man genau, wie sich dieses Phänomen unter gegebenen Bedingungen verhält. Der Zusammenhang zwischen den betrachteten Größen ist vollkommen. Keine weiteren Einflüsse als die, die im Modell auftauchen, spielen hinein. Dies ist nicht der Fall, wenn es eine stochastische oder zufällige Variation gibt. Will man ein Phänomen beschreiben, das eine zufällige Variation aufweist, so ist notwendigerweise zu beachten, dass der Zusammenhang zwischen den betrachteten Größen nicht vollkommen ist. Es kommt jetzt darauf an, Begriffe und Maße zu finden, mit denen das Phänomen beschrieben werden kann. Nur dann

kann man, unterstützt durch das Modell, fundierte und angemessene Entscheidungen treffen.

Viele interessante bzw. bedeutende Phänomene sind von Natur aus eher stochastischer als deterministischer Art, z.B. Phänomene, die die Umwelt, das Wetter, das menschliche Verhalten, die Wirtschaft oder Ähnliches betreffen. Man kann z.B. nicht *genau vorhersagen*, wie schnell ein Baum in einem Wald wachsen wird, da sein Wachstum von vielen Faktoren abhängt, wie z.B. der Niederschlagsmenge, die wiederum selbst auch nicht exakt vorhergesagt werden kann. Genauso sind

- das wirtschaftliche Wachstum,
- die Entwicklung der Arbeitslosigkeit,
- die Zahl der zukünftigen Auftragseingänge,
- die Inflationsrate oder
- der morgige Wechsel- oder Aktienkurs

nicht genau vorherzusagen. Man kann auch nicht wissen, wie potenzielle Käufer auf

- eine bestimmte Werbung,
- eine neue Verpackung eines Produkts,
- eine Preisänderung oder
- eine andere Platzierung eines Produkts innerhalb des Geschäftes oder im Regal

reagieren. Man weiß auch nicht im Voraus, wie eine bestimmte Person auf ein gewisses Medikament reagiert.

Solche Phänomene können nur sehr selten durch deterministische Modelle beschrieben werden, da diese Phänomene in der Regel zu komplex sind. Sie werden auf komplizierte Art von vielen unterschiedlichen Faktoren beeinflusst. Häufig kann man jedoch ein stochastisches Modell aufstellen, um damit Wahrscheinlichkeiten anzugeben, dass gewisse Ereignisse eintreten.

So wird man sehr selten in der Lage sein, Aussagen der folgenden Art zu machen: *Dieses Individuum wird positiv auf die Behandlung reagieren.* Mit geeigneten Daten wird es jedoch häufig möglich sein, Aussagen der folgenden Form zu treffen: *Mit einer Wahrscheinlichkeit von 0.9 (bzw. 90 %) wird das Individuum positiv auf die Behandlung reagieren.* Dies soll anhand der folgenden Beispiele stochastischer Probleme und Modelle verdeutlicht werden.

1.2 Beispiele stochastischer Probleme und Modelle

Beispiel 1.3. Aspirin und Herzanfälle

Am 27. Januar 1988 lautete eine Schlagzeile auf der ersten Seite der New York Times *Heart Attack Risk Found to be Cut by Taking Aspirin*. Der zu der Überschrift gehörende Artikel berichtete über die Ergebnisse einer Untersuchung, in der überprüft wurde, ob geringe Dosen Aspirin Herzinfällen bei gesunden Männern mittleren Alters vorbeugen. 22071 Männer waren zufällig in zwei Gruppen aufgeteilt

Tabelle 1.3 Ergebnisse einer Studie zur Wirkung von Aspirin auf Herzanfälle bei Männern mittleren Alters

	Personen	Herzanfälle	Herzanfälle pro 1 000 Personen
Aspirin-Gruppe	11 037	104	9.4
Placebo-Gruppe	11 034	189	17.1

worden. Eine dieser Gruppen, die Behandlungsgruppe, hatte regelmäßig Aspirin erhalten. Die zweite Gruppe, die Placebogruppe, hatte eine Substanz ohne wirksame Inhaltsstoffe erhalten (man spricht dann von einem Placebo).

Es war genau aufgezeichnet worden, wer welche Behandlung erhalten hatte und bei wem im Laufe der Zeit ein Herzanfall aufgetreten war. Dabei wussten weder die Patienten noch die Ärzte, ob es sich bei der verabreichten Substanz um Aspirin oder um ein Placebo handelte. Diese Art von Versuch heißt Doppelblindstudie. Wenn jeder gewusst hätte, mit welcher Substanz er behandelt worden war, so hätte das möglicherweise das Resultat beeinflusst. Tabelle 1.3 enthält eine Zusammenfassung der Ergebnisse der Studie, die in dem Zeitungsartikel veröffentlicht wurden.

Es ist klar, dass das Ergebnis dieses Versuchs nicht durch ein deterministisches Modell beschrieben werden kann. Ob ein Mann mittleren Alters einen Herzanfall erleiden wird, hängt nicht *einzig und allein* davon ab, ob er regelmäßig Aspirin zu sich nimmt oder nicht. Auch andere Faktoren spielen hier eine Rolle. Die Beziehung zwischen der Art der Behandlung und dem Eintreten eines Herzanfalls ist stochastisch. Die Ergebnisse der Studie müssen daher im Kontext eines stochastischen Modells betrachtet werden, auch wenn sie die Vermutung nahe legen, dass die Wahrscheinlichkeit (oder das Risiko), einen Herzanfall zu erleiden, in der Aspirin-Gruppe niedriger ist.

Die genannte Wirkung von Aspirin wurde natürlich weiter untersucht. Bevor man also sofort zur Apotheke läuft, um Aspirin zu kaufen, würde es sich lohnen, Berichte über die vielen Ergebnisse von späteren Untersuchungen zu lesen. Dass man mit der Interpretation statistischer Analysen und stochastischer Modelle sehr vorsichtig sein muss, zeigen auch die folgenden beiden Beispiele.

Beispiel 1.4. Weinkonsum und Herzkrankheiten

Abbildung 1.4a zeigt den Zusammenhang zwischen dem jährlichen Alkoholkonsum in Form von Wein (Liter purer Alkohol / Kopf) und der Todesrate durch Herzkrankheiten (Anzahl Tote / 100 000 Einwohner) für 21 Industrienationen im Jahr 1988, wobei ausgewählten Punkten die entsprechenden Länder zugeordnet sind.²

Die Beobachtungen sowie die an sie angepasste Gerade erwecken den Eindruck, dass sich das Risiko, an einer Herzkrankheit zu sterben, durch erhöhten Weinkonsum deutlich senken lässt. Dies kann man jedoch nicht so einfach sagen, denn aus

² Die Original-Daten stammen aus dem Artikel *Criqui, M.H. and Ringel, B.L. (1994): Does diet or alcohol explain the French paradox? The Lancet 344, December 24/31, 1719-1723*. Die genauen Zahlen wurden allerdings aus dem Artikel *Wine for the Heart: Over All, Risks May Outweigh Benefits*, der am 28. Dezember 1994 auf Seite C10 der New York Times erschien, übernommen.

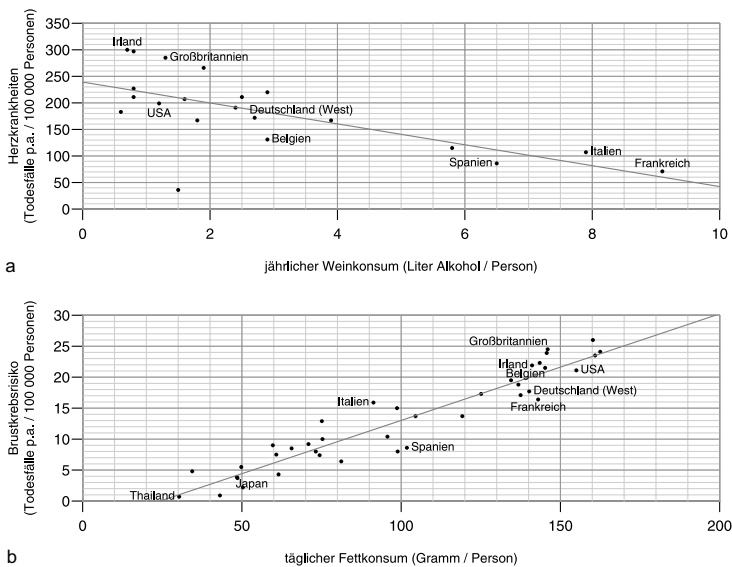


Abb. 1.4 a Weinkonsum und Herzkrankheiten in 21 Industrienationen mit angepasster Gerade.
b Fettkonsum und Brustkrebsrisiko in 39 Nationen mit angepasster Gerade

dem statistischen Zusammenhang folgt nicht automatisch auch ein kausaler Zusammenhang. So kann es beispielsweise völlig andere Gründe dafür geben, dass in Frankreich nur wenige Personen durch Herzkrankheiten sterben, während das Todesrisiko durch Herzkrankheiten in Irland und Großbritannien relativ hoch ist.

Außerdem besteht die Möglichkeit, dass der Weinkonsum neben einer eventuellen positiven Wirkung auf das Risiko von Herzkrankheiten auch deutlich negativen Einfluss auf andere gesundheitliche Aspekte hat. Insofern muss man gerade bei der Auswertung und Interpretation aggregierter Daten sehr vorsichtig sein. Man sollte nicht als Haupteckentnis der ersten Seiten dieses Buches folgern, dass der abendliche Weinkonsum in Verbindung mit der morgendlichen Aspirin-Einnahme nur positive Effekte auf die Gesundheit hat! Ähnlich problematisch ist die Interpretation des folgenden Beispiels.

Beispiel 1.5. Fettkonsum und Brustkrebsrisiko

Abbildung 1.4b stellt den Zusammenhang zwischen dem täglichen Fettkonsum (in Gramm pro Person) und der jährlichen Anzahl durch Brustkrebs verursachter Todesfälle (pro 100 000 Personen) für 39 Nationen Mitte der 1960er Jahre dar.³

In diesem Fall scheint mit zunehmendem Fettkonsum das Brustkrebsrisiko zu steigen. Allerdings muss man auch hier vorsichtig sein. Es könnte sein, dass das

³ Die Original-Daten wurden zuerst als Grafik in dem Artikel Carroll, K.K. (1975): *Experimental Evidence of Dietary Factors and Hormone-Dependent Cancers*. *Cancer Research* 35, 3374-3383, veröffentlicht. Die genauen Zahlen wurden allerdings von der Internetseite http://qrc.depaul.edu/Excel_Files/BreastCancerFatIntake.xls des Quantitative Reasoning Centers der DePaul University heruntergeladen (05. Mai 2008).

Brustkrebsrisiko von völlig anderen Faktoren abhängt oder dass der Fettkonsum mit anderen Faktoren einhergeht, die ursächlich für ein erhöhtes Brustkrebsrisiko sind. Wenn man im Internet nach anderen Studien sucht, die den Zusammenhang zwischen Krebsrisiko und Fettkonsum untersuchen, findet man sehr unterschiedliche Ergebnisse, so dass der kausale Zusammenhang mit der obigen Grafik sicher nicht bewiesen ist.

Nachdem die letzten Beispiele eher in der Medizin angesiedelt waren, möchten wir uns im Folgenden wieder den Wirtschaftswissenschaften zuwenden und noch einige Anwendungsbeispiele statistischer Methoden und stochastischer Modelle mit wirtschaftswissenschaftlichem Hintergrund vorstellen. Stochastische Modelle spielen, wie in vielen anderen wirtschaftswissenschaftlichen Bereichen, auch in der betrieblichen Finanzwirtschaft eine wichtige Rolle. Jeder hat sicherlich schon von Begriffen wie *Risikomanagement*, *Volatilität* oder *Value at Risk* gehört oder gelesen. Ein Aspekt, der diese Begriffe verbindet, ist das stochastische Verhalten der Wertpapierkurse, die durch entsprechende Modelle abgebildet werden können. Nicht geeignet sind in diesem Fall deterministische Modelle. Die Kurse an allen Aktienmärkten der Welt unterliegen einer Vielzahl von Einflüssen. Da eine deterministische Erfassung und Auswertung aller Einflüsse unmöglich ist, müssen die Kursschwankungen durch ein stochastisches Modell beschrieben werden. Im Folgenden werden einige Möglichkeiten der statistischen Analyse von Aktienkursen vorgestellt.

Beispiel 1.6. Entwicklung von Aktienkursen

In Abb. 1.5 ist die Entwicklung des Deutschen Aktienindex (DAX) sowie des Aktienkurses der Deutsche Bank Aktie von Anfang 2006 bis Ende 2007 dargestellt.⁴ Ein Ausschnitt der Kurse ist in Tabelle 1.4 gegeben.

Tabelle 1.4 Schlusskurse des DAX und der Deutsche Bank Aktie in den Jahren 2006 und 2007

Datum	DAX	Deutsche Bank	Datum	DAX	Deutsche Bank
02. Jan 06	5 449.98	81.93	12. Dez 07	8 076.12	91.16
03. Jan 06	5 460.68	81.74	13. Dez 07	7 928.31	88.75
04. Jan 06	5 523.62	83.47	14. Dez 07	7 948.36	89.15
05. Jan 06	5 516.53	83.50	17. Dez 07	7 825.44	87.79
06. Jan 06	5 536.32	84.24	18. Dez 07	7 850.74	87.73
09. Jan 06	5 537.11	84.55	19. Dez 07	7 837.32	87.45
10. Jan 06	5 494.71	84.70	20. Dez 07	7 869.19	87.15
11. Jan 06	5 532.89	86.71	21. Dez 07	8 002.67	87.87
12. Jan 06	5 542.13	86.78	27. Dez 07	8 038.60	89.14
13. Jan 06	5 483.09	85.64	28. Dez 07	8 067.32	89.40
⋮	⋮	⋮			

⁴ Historische Kurse findet man auf vielen Seiten im Internet. Die hier verwendeten Daten wurden am 28. April 2008 auf der Seite <http://de.finance.yahoo.com/> abgefragt und beziehen sich auf das elektronische Handelssystem XETRA.

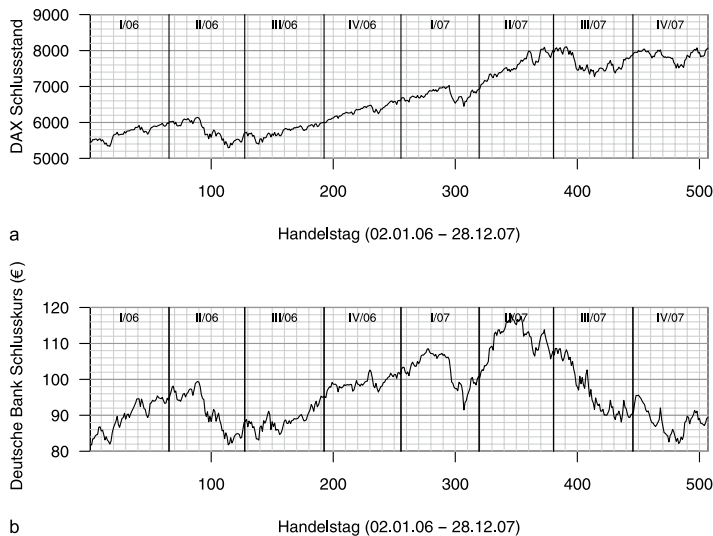


Abb. 1.5 Zeitliche Entwicklung des DAX und des Aktienkurses der Deutschen Bank in den Jahren 2006 und 2007

Der DAX ist der wichtigste deutsche Aktienindex. Ein Aktienindex stellt einen Maßstab dar, der die generelle Entwicklungsrichtung des Aktienmarktes beschreibt und an dem die Entwicklung einzelner Aktien gemessen werden kann. Der DAX wird als gewichteter Durchschnitt der Kurse der bedeutendsten deutschen Aktien berechnet. Er wurde zum 01.01.1988 mit einem Basiswert von 1000 Punkten eingeführt; der aktuelle Wert des DAX beschreibt somit die Entwicklung des deutschen Aktienmarktes im Vergleich zu diesem Datum. In die Berechnung des DAX gehen 30 wichtige deutsche Aktienwerte aus den Technologie- und klassischen Branchen ein, wobei als Auswahlkriterien für die Aufnahme der Börsenumsatz und die Marktkapitalisierung (Produkt aus der Anzahl frei verfügbarer Aktien und dem Aktienkurs) dienen. Detailliertere Informationen zum DAX folgen in Abschnitt 13.2.3.

Die Aktie der Deutschen Bank gehört zu den 30 Aktien, die aktuell in die Berechnung des DAX eingehen; am 28.12.2007 betrug ihr Gewicht bei der Berechnung 5.63 %. Da der Aktienkurs der Deutschen Bank den Stand des DAX direkt beeinflusst, ist es nicht überraschend, dass sich der Aktienkurs der Deutschen Bank und der DAX im Betrachtungszeitraum ähnlich entwickelt haben. Man sieht jedoch vor allem in der rechten Hälfte der jeweiligen Grafiken, dass es auch Phasen gibt, in denen sich die beiden Kurse gegensätzlich entwickeln. Dies ist in diesem Fall vermutlich darauf zurückzuführen, dass die Deutsche Bank vom Beginn der Finanzkrise im Jahr 2007 deutlich stärker betroffen war als der Aktienmarkt allgemein.

Einen weiteren Einblick in ein Detail der Kursentwicklung der Deutsche Bank Aktie vermittelt Abb. 1.6, in der die (nicht in Tabelle 1.4 angegebenen) Eröffnungskurse der Deutsche Bank Aktie des oben genannten Zeitraums gegen die Schlusskurse des jeweiligen Vortags abgetragen sind.

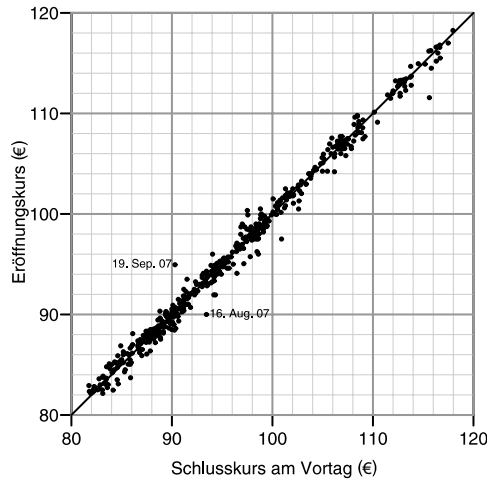


Abb. 1.6 Eröffnungskurs der Deutsche Bank Aktie und Schlusskurs am Vortag in den Jahren 2006 und 2007

Wie man erkennt, liegen die meisten Punkte sehr dicht an der eingezeichneten Winkelhalbierenden, d.h. fast immer sind Schlussstand und nachfolgender Eröffnungsstand nahezu identisch. In den Formulierungen *fast immer* und *nahezu* spiegelt sich der Zufall wider. *Nahezu* bedeutet, dass der Eröffnungskurs mal ein wenig über, mal ein wenig unter dem Schlusskurs des Vortages liegt. *Fast immer* bedeutet das, dass es auch Fälle geben kann, in denen der grundsätzliche Zusammenhang zwischen Eröffnungskurs und vorherigem Schlusskurs nicht gilt.

In Abb. 1.6 sind zwei Punkte besonders gekennzeichnet, bei denen der Eröffnungskurs besonders stark vom Schlusskurs des Vortages abweicht. Am 16. August 2007 lag der Eröffnungskurs der Deutsche Bank Aktie fast 4 % unter dem Schlusskurs des Vortages, während er am 19. September 2007 rund 5 % oberhalb des vorangegangenen Schlusskurses lag.

Im Nachhinein ist es manchmal möglich, die Ursache solcher „Störungen“ zu bestimmen, beispielsweise Ankündigungen eines Unternehmens, politische Entwicklungen oder auch technische Fortschritte (z.B. in der Biotechnologie). So lag die Ursache für den niedrigen Eröffnungskurs am 16. August 2007 vermutlich in den starken Kursverlusten an den Börsen in den USA und Asien während der Nacht zuvor, und der Eröffnungskurs am 19. September 2007 war wohl eine Folge der Ankündigung einer Zinssenkung durch die US-Notenbank am Vorabend.⁵

Das Problem aus Anlegersicht ist, dass derartige Geschehnisse nicht im Voraus bekannt sind bzw. sich deren Auswirkungen auf die Kursentwicklung nicht exakt abschätzen lassen. Trägt man den Eröffnungsstand gegen den Schlussstand des vor-

⁵ Vergleiche den Artikel *DAX kommt unter die Räder* auf der Seite <http://www.sueddeutsche.de/finanzen/artikel/494/128284/article.html> sowie den Artikel *DAX startet nach Zinssenkung mit Kursprung* unter http://www.focus.de/finanzen/boerse/aktien/boerse-am-morgen_aid_133236.html (Download am 29. April 2008).

letzten Tages ab (Abb. 1.7a), sind die Punkte sogar noch weiter von der Winkelhalbierenden entfernt als in Abb. 1.6. Noch stärker sind die Abweichungen, wenn drei weitere Tage zwischen Eröffnungs- und Schlusskurs liegen (Abb. 1.7b).

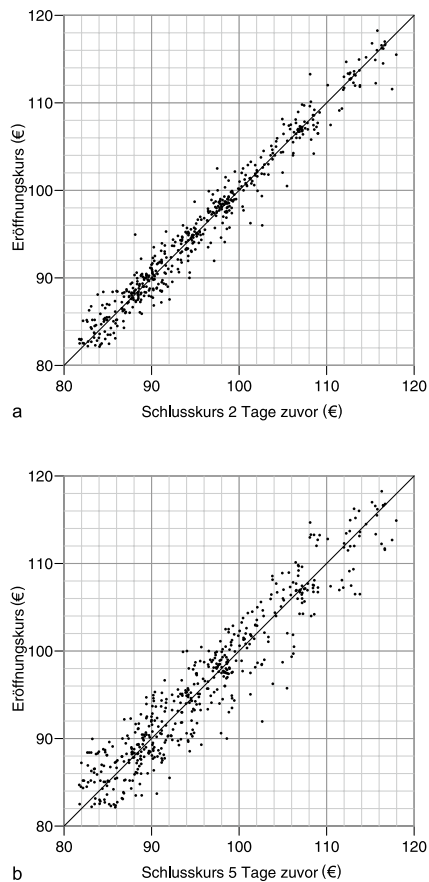


Abb. 1.7 Eröffnungskurs der Deutsche Bank Aktie in den Jahren 2006 und 2007 und Schlusskurs **a** zwei Tage und **b** fünf Tage zuvor

Die Abbildungen 1.6 und 1.7 verdeutlichen, dass bei der Verwendung des Schlusskurses der Aktie an einem bestimmten Tag zur Vorhersage des Eröffnungskurses an einem zukünftigen Tag die Unsicherheit wächst, je weiter man in die Zukunft prognostiziert. Während Abb. 1.6 einen nahezu deterministischen Zusammenhang zeigt, sind die Zusammenhänge in Abb. 1.7 stochastischer Natur. Je weiter eine Prognose in die Zukunft weist, desto länger ist der Zeitraum, in dem verschiedenste Ereignisse die Kurse in eine zum Zeitpunkt der Prognose unvorhersehbare Richtung lenken. Dieses Verhältnis zwischen Zukunftszeitraum und Unsicherheit stellt (neben vielen anderen Phänomenen) ein Phänomen dar, das mit Hilfe statistischer Methoden analysiert werden kann.

Eine andere Möglichkeit, den Aktienkurs der Deutsche Bank Aktie zu betrachten, stellen die täglichen Renditen dar, d.h. die täglichen Veränderungen in Prozent. Am 03.01.2006 beispielsweise betrug laut Tabelle 1.4 der Schlusskurs der Deutsche Bank Aktie 81.74 €; für den nächsten Tag wurde er mit 83.47 € angegeben. Für die prozentuale Veränderung, also die tägliche Rendite, ergibt sich demnach:

$$\text{einfache Rendite} = 100 \cdot \frac{(83.47 - 81.74)}{81.74} \approx 2.12 \%$$

Der Kurs der Aktie ist also am 04.01.2006 um 2.12 % gestiegen.

Alternativ zu den soeben betrachteten Renditen werden in der Statistik oft die sogenannten *kontinuierlichen Renditen* berechnet. Diese stellen den Logarithmus des Verhältnisses aus dem Schlusskurs am betrachteten Tag und dem des Vortages, multipliziert mit 100, dar:

$$\text{kontinuierliche Rendite} = 100 \cdot \log \left(\frac{83.47}{81.74} \right) \approx 2.09 \%$$

Dieser Wert liegt sehr nah bei der einfachen Rendite. Es kann gezeigt werden, dass dies immer gilt, wenn die Renditen hinreichend klein sind.

Abbildung 1.8 zeigt ein Histogramm der (kontinuierlichen) Renditen der Deutsche Bank Aktie für den betrachteten Zeitraum. Man erkennt, dass die Renditen in den Jahren 2006 und 2007 im Bereich von -6% bis 6% lagen und eine steigende Häufigkeit für Werte festzustellen ist, die näher bei Null liegen.

Wie bei der Blockzeit in Abb. 1.3 stellt die in Abb. 1.8 enthaltene Kurve ein stochastisches Modell dar. Sie kann wiederum als Glättung des Histogramms betrachtet werden. Diese Kurve gibt einen Hinweis darauf, wie die Renditen in dem betrachteten Zeitraum verteilt sind. Wie bereits erwähnt, wird die Konstruktion solcher Modelle und ihre Überprüfung in späteren Kapiteln detailliert erläutert. Zu diesem Zeitpunkt bleibt festzuhalten:

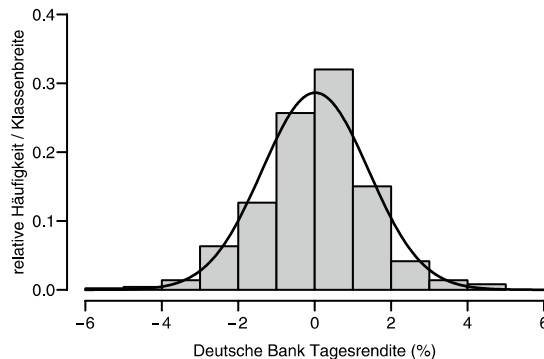


Abb. 1.8 Kontinuierliche Tagesrenditen der Deutsche Bank Aktie in den Jahren 2006 und 2007 mit angepasster Normalverteilung

- Die Renditen sind nicht deterministisch. Sie variieren zufällig.
- Die Renditen zeigen bestimmte Muster, z.B. sind Werte nahe Null am häufigsten, und die Häufigkeit nimmt ab, je weiter die Renditen von Null entfernt sind.

Obwohl die Renditen zufällig variieren, kann man sehen, dass gewisse Aussagen über ihr Verhalten dennoch möglich sind. Wie brauchbare Aussagen aus vorliegenden Daten gebildet und interpretiert werden können, ist eines der zentralen Themen dieses Buches.

Auch die Beziehung zwischen der Entwicklung der Deutsche Bank Aktie und der Entwicklung des DAX kann mit Hilfe statistischer Methoden untersucht werden. Weiter oben wurde bereits erwähnt, dass ein Aktienindex als Maßstab dient, an dem die Kursentwicklung einzelner Aktien gemessen werden kann. Eines der bekanntesten Modelle der Finanzwirtschaft ist das Capital Asset Pricing Model (CAPM). Im CAPM werden die Renditen einzelner Aktien mit der sogenannten Marktrendite verglichen, die zum Beispiel durch die Entwicklung eines Aktienindex gemessen werden kann. In Abb. 1.9 sind die (kontinuierlichen) täglichen Renditen der Deutsche Bank Aktie gegen die entsprechenden Renditen des DAX abgetragen.

Es ist gut zu erkennen, dass es einen Zusammenhang zwischen den jeweiligen Renditen gibt, allerdings weicht dieser Zusammenhang tendenziell von der Winkelhalbierenden ab. Die Renditen der Deutsche Bank Aktie scheinen deutlicher zu schwanken als die des DAX; während die Renditen des DAX nur selten außerhalb des Intervalls $[-2\%; 2\%]$ liegen, kommen bei den Renditen der Deutsche Bank Aktie wesentlich häufiger extremere Werte vor.

Dies wird auch durch die angepasste Gerade verdeutlicht, die für eine gegebene DAX-Rendite eine etwas extremere Rendite für die Deutsche Bank Aktie vorhersagt. In Kapitel 11 wird im Rahmen der Regressionsanalyse noch einmal auf die Analyse des Zusammenhangs zwischen Deutsche Bank Rendite und DAX-Rendite,

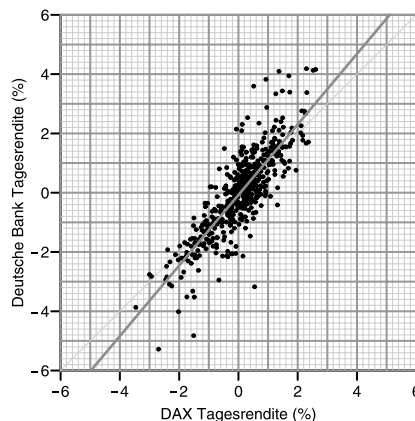


Abb. 1.9 Kontinuierliche Renditen der Deutsche Bank Aktie und des DAX in den Jahren 2006 und 2007 und angepasste Gerade

sowie die Berechnung und Interpretation der Geraden zur Beschreibung des Zusammenhangs eingegangen.

Die Kurse bzw. Börsenindizes der Aktienmärkte sind nur ein Beispiel für stochastische Phänomene in der Finanzwelt. Es gibt viele andere Risiken/Unsicherheiten, beispielsweise Risiken, die sich aus Veränderungen bei Währungspreisen oder Zinsen ergeben. Banken und andere Kreditinstitute sollten zum Beispiel das Risiko von Kreditausfällen bei der Durchführung ihrer Geschäfte berücksichtigen.

Ein weiterer Bereich, in dem stochastische Modelle eine bedeutende Rolle spielen, ist der Versicherungsbereich. Das Wesentliche an einem Versicherungsgeschäft ist die Kompensation eines Schadens im Schadensfall. Daher müssen Versicherungen mögliche Risiken und damit verbundene Schäden sorgfältig quantifizieren. Wie sonst könnte eine Versicherung ihre Versicherungsprämien kalkulieren? Beispielsweise müssen Versicherungen abschätzen, mit welcher Wahrscheinlichkeit eine Person erkrankt (Krankenversicherung) oder mit welcher Wahrscheinlichkeit ein Autofahrer einen Unfall verursacht (Kfz-Haftpflichtversicherung). Sogar für Weltraumfahrzeuge müssen derartige Überlegungen angestellt werden, wenn sich eine Versicherung in diesem Zweig behaupten will.

Besonders große Probleme ergeben sich für Versicherungen und Rückversicherer aus Naturkatastrophen, deren Anzahl in den letzten Jahren zugenommen zu haben scheint. Daher ist es für Versicherungen von großem Interesse, das Risiko des Auftretens solcher Katastrophen mit Hilfe statistischer Methoden abzuschätzen.

Beispiel 1.7. Erdbeben und Tsunamis

Abbildung 1.10 zeigt Histogramme der Zeit zwischen den weltweit im Zeitraum 1982–2002 registrierten Tsunamis (oben), sowie der Zeit zwischen den weltweit im Zeitraum 1973–2007 registrierten Erdbeben der Stärke 7.0 oder größer (unten), jeweils in Tagen.⁶

Es fällt auf, dass beide Grafiken eine sehr ähnliche Struktur haben. Sowohl zwischen Tsunamis als auch zwischen starken Erdbeben vergehen sehr häufig nur wenige Tage, während längere Zeiträume zwischen jeweils zwei aufeinander folgenden Ereignissen nur sehr selten vorkommen.

Wie bei den bereits in den vorherigen Beispielen betrachteten Histogrammen, wurde auch hier versucht, die Daten jeweils mit Hilfe eines stochastischen Modells zu „glätten“. In Kapitel 6 wird die zu Grunde liegende Modellfamilie näher vorgestellt. Dabei wird auch aufgezeigt, dass es weitere Phänomene gibt, deren Histogramme diesen typischen exponentiell fallenden Verlauf aufweisen.

In direktem Zusammenhang mit der Zeit, die zwischen zwei Erdbeben (oder zwei Tsunamis) vergeht, steht die Anzahl in einem bestimmten Zeitraum auftreten der Erdbeben (oder Tsunamis). In Abb. 1.11 ist die Anzahl monatlich beobachteter Erdbeben der Stärke 7.0 dargestellt. Dieser Abbildung liegen dieselben Daten zu Grunde wie dem obigen Histogramm der Zeit zwischen zwei Erdbeben.

⁶ Die Tsunami-Daten wurden aus dem 2003 erschienen Heft 1 des Jahrgangs 21 der Zeitschrift *Science of Tsunami Hazards* entnommen. Die Erdbeben-Daten werden vom National Earthquake Information Center des U.S. Geological Survey auf der Internetseite http://neic.usgs.gov/neis/epic/epic_global.html zur Verfügung gestellt (Download am 25.04.2008).

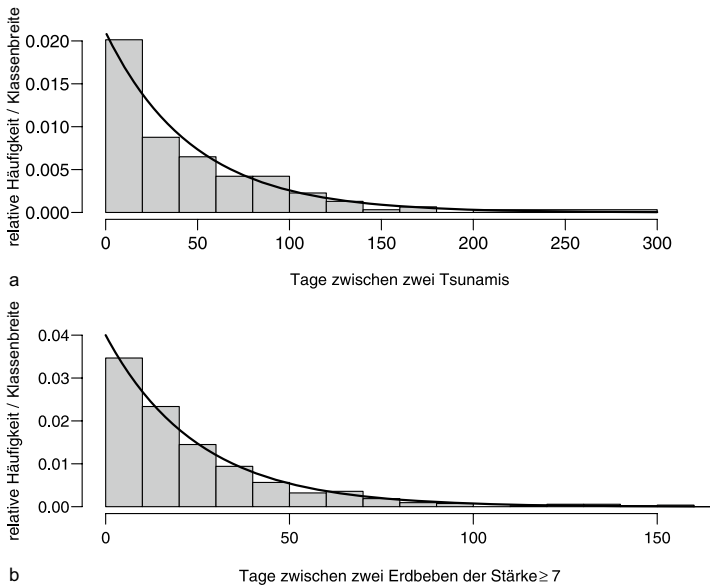


Abb. 1.10 **a** Tage zwischen den weltweit im Zeitraum 1982–2002 beobachteten Tsunamis und angepasste Exponentialverteilung. **b** Tage zwischen den weltweit im Zeitraum 1973–2007 beobachteten Erdbeben der Stärke 7.0 oder größer und angepasste Exponentialverteilung

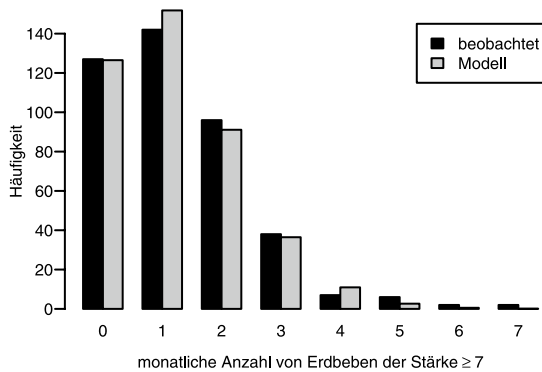


Abb. 1.11 Anzahl monatlicher Erdbeben der Stärke 7.0 oder größer im Zeitraum 1973–2007 und angepasste Poissonverteilung

Insgesamt wurden von Anfang 1973 bis Ende 2007 weltweit 532 Erdbeben registriert, die eine Stärke von 7.0 oder größer hatten. Für die jeweilige Anzahl an Erdbeben pro Monat ist die entsprechende beobachtete Häufigkeit als schwarze Säule eingetragen. Beispielsweise gab es 127 Monate, in denen kein starkes Erdbeben auftrat. Man nennt diesen Abbildungstyp Säulen- oder Balkendiagramm.

Die grauen Balken, die jeweils an die schwarzen Balken grenzen, repräsentieren ein stochastisches Modell (eine Poissonverteilung), das an die Daten angepasst wurde und die Anzahl monatlicher Erdbeben beschreiben soll. Wie ein solches Modell

ermittelt wird und ob es die Daten angemessen beschreibt, wird in späteren Kapiteln noch behandelt. Dann wird man auch verstehen, welchen unmittelbaren Zusammenhang es zwischen dem Modell für die Zeit zwischen zwei Erdbeben und dem Modell für die Anzahl von Erdbeben gibt.

Es gäbe noch viele weitere Beispiele aus den Geowissenschaften (oder auch Ingenieurwissenschaften), die auch wirtschaftswissenschaftliche Relevanz haben. Will man zum Beispiel Brücken oder Staudämme bauen, so muss man etwas über das Auftreten extremer Belastungen, d.h. abnormal hoher Wasserstände wissen. Man baut so, dass Belastungen bis zu einer bestimmten Grenze ausgehalten werden können. Es ist unmittelbar einleuchtend, dass hohe Belastungen, also hohe Wasserstände, zufällig auftreten. Statistische Methoden können daher helfen, das Risiko und die Höhe extremer Belastungen zu quantifizieren.

Statistische Methoden werden auch eingesetzt, um abzuschätzen, ob ein Goldvorkommen (oder auch eine Ölquelle) wirtschaftlich ausbeutbar ist. Man untersucht beispielsweise den Goldgehalt vorhandener Erzproben und zieht dann Rückschlüsse auf den gesamten Goldgehalt einer potenziellen Mine.

Auch im Marketing spielen statistische Methoden eine große Rolle. Beispielsweise wird im Marketing mit Hilfe von Marktforschungsstudien abgeschätzt, auf welchen Märkten noch Chancen bestehen, wie sich bestimmte Werbemaßnahmen auswirken oder auf welche Produkteigenschaften Konsumenten wie reagieren. Ein Produktmerkmal, dass bei sehr vielen Produkten einen großen Einfluss auf den Absatz hat, ist der Preis, wie das folgende Beispiel verdeutlicht.

Beispiel 1.8. Preis und Absatz von Traubensaft

In diesem Beispiel wird ein Datensatz betrachtet, der Informationen über die wöchentlichen Verkaufszahlen einer bestimmten Traubensaft-Sorte in einem bestimmten Supermarkt im Großraum Chicago im Zeitraum zwischen Juli 1992 und Juni 1996 in Abhängigkeit vom Verkaufspreis enthält.⁷ Der Zusammenhang zwischen Verkaufspreis und Verkaufsmenge ist in Abb. 1.12 dargestellt.

Die zusätzlich eingezeichnete Gerade ist die sogenannte Preis-Absatz-Funktion, die den prinzipiellen Zusammenhang zwischen dem Verkaufspreis und der Absatzmenge beschreibt. Es ist offensichtlich, dass die tatsächlichen Verkaufszahlen stark um die Preis-Absatz-Funktion schwanken, allerdings kann mit Hilfe der Geraden in etwa die durchschnittliche wöchentliche Verkaufsmenge für einen vorgegebenen Preis abgelesen werden. Wie man eine solche Preis-Absatz-Funktion bestimmen kann, wird im Rahmen der Regressionsanalyse in Kapitel 11 näher erläutert. Die Regressionsanalyse findet auch im folgenden Beispiel Anwendung, das vielleicht für einige Leser auch von privatem Interesse ist.

⁷ Der hier verwendete Datensatz stellt einen kleinen Ausschnitt der Scanner-Daten dar, die in der *Dominick's Database* des Kilts Centers for Marketing der Chicago Graduate School of Business auf der Internetseite <http://research.chicagogsb.edu/marketing/databases/index.aspx> zur Verfügung gestellt werden (Stand 29.08.2008). Aus diesen Scanner-Daten wurden ein Supermarkt und ein Produkt ausgewählt. Außerdem wurden die Daten auf diejenigen Wochen beschränkt, in denen es keine besonderen Promotions-Aktivitäten für die gewählte Traubensaft-Sorte gab, damit man einen möglichst unverfälschten Eindruck vom Einfluss des Preises erhält.

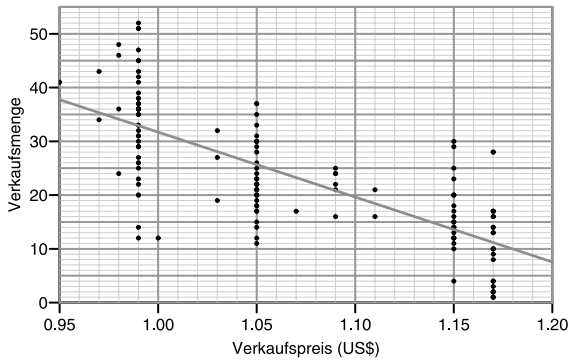


Abb. 1.12 Preis und Absatz eines Traubensaft-Produktes sowie angepasste Gerade

Beispiel 1.9. Verkaufspreis bei Online-Auktionen

Wir betrachten einen Datensatz, der Beobachtungen aller Auktionen neuer Handys vom Modell Nokia 8310 auf der Online-Plattform *www.ricardo.ch* von Oktober 2001 bis Januar 2002 enthält.⁸ Dabei werden nur diejenigen Auktionen berücksichtigt, bei denen genau ein neues Handy angeboten wurde und die erfolgreich, d.h. mit einem Verkauf, abgeschlossen wurden.

Abbildung 1.13a stellt die Entwicklung des Maximalgebotes (in Schweizer Franken, CHF) im Lauf der Zeit (beginnend mit dem Tag der ersten beobachteten Auktion) dar. Es wird deutlich, dass das Maximalgebot tendenziell mit der Zeit sinkt. Das bedeutet, dass die höchsten Preise im Durchschnitt im Oktober 2001 erzielt wurden, als das Handy-Modell gerade auf den Markt kam; anschließend gingen die Preise degressiv zurück.

Diese Aussage wird durch die eingezeichnete Kurve unterstützt, die wie bei vorangegangenen ähnlichen Beispielen, wieder ein stochastisches Modell darstellt, das den Zusammenhang zwischen Zeit und Maximalgebot beschreibt. Diese Kurve wurde unter Ausschluss des „Ausreißers“ mit einem Maximalgebot von 800 CHF ermittelt; mit Einschluss dieser Beobachtung sähe ihr Verlauf auffallend anders aus.

Abbildung 1.13b verdeutlicht einen anderen Aspekt der Daten. Sie stellt das Maximalgebot in Abhängigkeit vom Wochentag mit sogenannten Boxplots dar. 50% der Maximalgebote für einen gegebenen Wochentag liegen innerhalb der jeweiligen Box; außerdem kennzeichnet der Querstrich in der Box den Gebotspreis, der die Beobachtungen dieses Wochentages so teilt, dass 50% der Maximalgebote kleiner oder gleich diesem Wert und die anderen 50% dementsprechend größer oder gleich diesem Wert sind. Außerdem sagt die Breite der Boxplots etwas über die Anzahl der Beobachtungen aus, d.h. je breiter die Box, desto mehr Auktionen endeten am

⁸ Die Daten wurden in dem Artikel *Diekmann, A. und Wyder, D. (2002): Vertrauen und Reputationseffekte bei Internet-Auktionen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 54(4), 674-693* im Hinblick auf Reputationseffekte analysiert und stehen auf der Internetseite <http://www.socio.ethz.ch/research/datafiles/> der Professur für Soziologie der ETH Zürich zur Verfügung (Download am 25.04.2008).

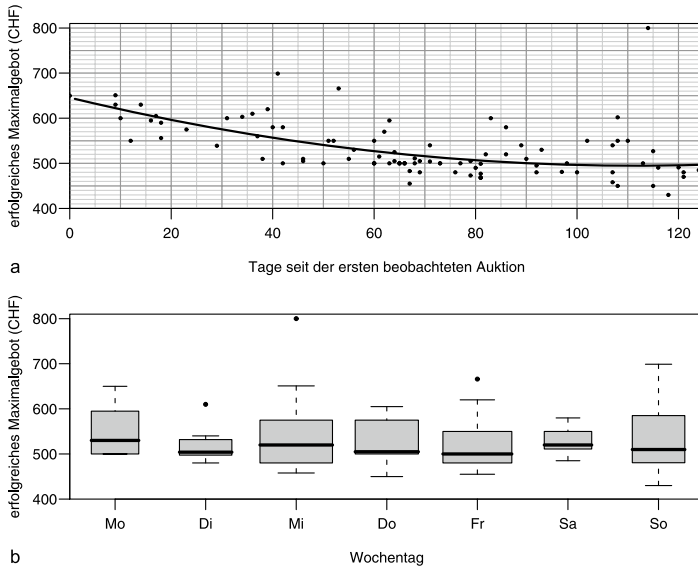


Abb. 1.13 Maximalgebot bei Online-Auktionen des Handy-Modells Nokia 8310 auf www.ricardo.ch von Oktober 2001 bis Januar 2002 in Abhängigkeit von **a** der Zeit seit dem ersten Angebot und **b** vom Wochentag

entsprechenden Wochentag. Die genaue Konstruktion solcher Boxplots wird in Kapitel 2 näher erläutert.

Durch die Darstellung wird beispielsweise deutlich, dass das Maximalgebot dienstags tendenziell niedriger war als montags, oder dass das Maximalgebot sonntags stärker schwankte als samstags. Es kann jedoch auch sein, dass diese Eindrücke durch rein zufällige Schwankungen entstanden sind und nicht „nachweisbar“ auf den Einfluss des Wochentages zurückgeführt werden können. In Kapitel 12 wird eine Methode vorgestellt, mit der der Einfluss des Wochentages auf das Maximalgebot überprüft werden kann.

Beispiel 1.10. Anrufe in einem Call-Center

Heutzutage sind interne oder externe Call-Center, die Anrufe von Kunden entgegen nehmen und Fragen beantworten oder Aufträge zur Bearbeitung weiterleiten, weit verbreitet. Dabei kann einerseits ein langes oder häufiges Verbleiben des Kunden in einer Warteschleife, bis ein freier Call-Center-Mitarbeiter zur Verfügung steht, ein Ärgernis darstellen, dass die Zufriedenheit des Kunden mit der Service-Qualität des entsprechenden Unternehmens stark negativ beeinflusst. Andererseits verursacht jedoch jeder weitere Mitarbeiter in einem Call-Center neue Kosten. Aus diesem Grund müssen die personellen Ressourcen in einem Call-Center möglichst effizient geplant werden. Zu diesem Zweck können als Planungsgrundlage historische Anrufrdaten mit Hilfe statistischer Methoden ausgewertet werden.

In diesem Beispiel werden Daten untersucht, die Informationen zu den Anrufen im Call-Center einer anonymen israelischen Bank im Januar 1999 enthalten.⁹ Dabei wird zunächst nur die Anzahl eintreffender Anrufe pro 5-Minuten-Intervall am Mittwoch, 20.01.1999, zwischen 10 und 17 Uhr betrachtet. Dieser Tag wurde zufällig ausgewählt, das Zeitintervall dagegen bewusst, da die Anrufe in diesem Intervall relativ gleichmäßig auf die einzelnen Stunden verteilt waren, während in den Stunden vor und nach dem betrachteten Intervall deutlich weniger Anrufe eintrafen. Insgesamt wurden in dem gegebenen Intervall 711 Anrufe registriert. Die Häufigkeiten der beobachteten Anrufrufen pro 5-Minuten-Intervalle sind in Tabelle 1.5 gegeben. In Abb. 1.14a sind diese Häufigkeiten als Säulendiagramm dargestellt. Für die jeweilige Anzahl an Anrufen pro 5-Minuten-Intervall ist die entsprechende beobachtete Häufigkeit als schwarze Säule eingetragen, während die grauen Säulen ein stochastisches Modell repräsentieren, das an die Daten angepasst wurde (eine Poissonverteilung).

Neben der Anzahl eintreffender Anrufe ist auch die Dauer der einzelnen Telefonate entscheidend für die Anzahl benötigter Call-Center-Mitarbeiter. Abbildung 1.14b zeigt daher ein Histogramm der Anrufdauer (in Sekunden) derjenigen 590 Anrufe, die tatsächlich von einem Call-Center-Mitarbeiter bedient wurden (die anderen Anrufer sind von einem Computer bedient worden oder haben vorher bereits aufgelegt; außerdem wurde ein extrem hoher Wert von der weiteren Analyse ausgeschlossen). Die Häufigkeiten sind in Tabelle 1.6 enthalten.

Wenn man die Tabelle mit dem Histogramm vergleicht, fällt auf, dass das Histogramm bei 800 Sekunden abgeschnitten ist, obwohl 14 Anrufe zwischen 800 und 1 400 Sekunden gedauert haben (der längste beobachtete Anruf hatte eine Länge von 1 354 Sekunden). Diese extrem langen Anrufe wären in dem Histogramm kaum zu erkennen gewesen, so dass das Histogramm so beschränkt wurde, dass der linke Teil des Histogramms etwas größer dargestellt werden kann.

Für praktische Anwendungen (sowie für die Auswahl eines potenziellen stochastischen Modells) können diese extremen Beobachtungen allerdings von großer Bedeutung sein. Eventuell ist es wichtig zu wissen, wie oft extrem lange Anrufe auftreten. Dann kann wie folgt vorgegangen werden, um das durchschnittliche Auftreten langer Anrufdauern zu bestimmen. Aus Tabelle 1.6 kann man ablesen, dass 14

Tabelle 1.5 Anzahl der Anrufe pro 5-Minuten-Intervall im Call-Center einer israelischen Bank am Mittwoch, 20.01.1999, zwischen 10 und 17 Uhr

Anzahl Anrufe	4	5	6	7	8	9	10	11	12	13	14	15	16	Summe
Häufigkeit	6	8	11	13	11	3	9	8	6	4	3	1	1	711

⁹ Die Daten werden (ebenso wie weitere Daten für das gesamte Jahr 1999) von Prof. Avishai Mandelbaum vom Technion — Israel Institute of Technology auf der Internetseite <http://iew3.technion.ac.il/serveng/callcenterdata/index.html> zur Verfügung gestellt (Download am 25.04.2008) und umfassen unter anderem den Zeitpunkt, an dem ein Anruf im Call-Center eintrifft, sowie den Zeitpunkt, zu dem der Anruf beendet wurde.

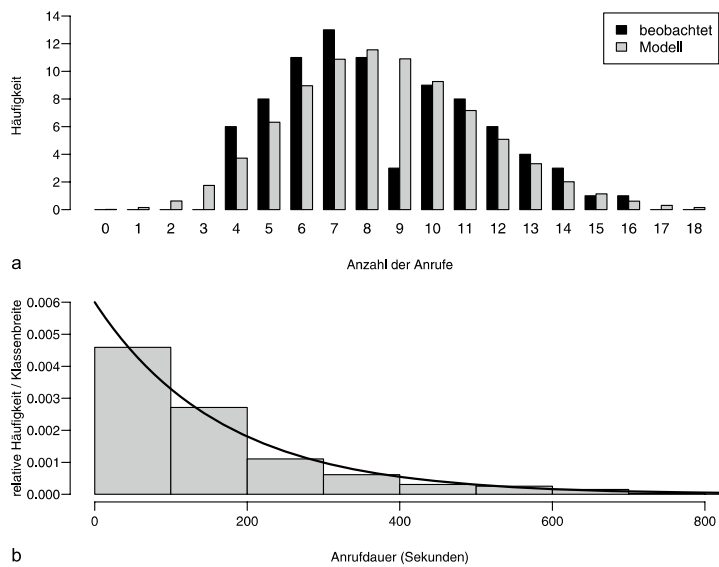


Abb. 1.14 Anrufe im Call-Center einer israelischen Bank am Mittwoch, 20.01.1999, zwischen 10 und 17 Uhr. **a** Anzahl der Anrufe pro 5-Minuten-Intervall und Poissonverteilung. **b** Dauer der angenommenen Anrufe und Exponentialverteilung

Tabelle 1.6 Dauer der von einem Mitarbeiter im Call-Center einer israelischen Bank am Mittwoch, 20.01.1999, zwischen 10 und 17 Uhr bedienten Anrufe (Sekunden)

[0;100]	(100;200]	(200;300]	(300;400]	(400;500]	(500;600]	(600;700]	(700;800]	(800;1 400]
271	160	65	36	18	15	9	2	14

von 590 Anrufen eine Dauer von mehr als 800 Sekunden hatten. Wenn man davon ausgeht, dass der beobachtete Zeitraum typisch für das Call-Center war, kann man sagen, dass im Durchschnitt $14/590 \cdot 100\% \approx 2.4\%$ aller Anrufe extrem lang dauern. Dann würde man erwarten, dass im Schnitt ungefähr jeder 42. Anruf extrem lang ist.

Bei der Interpretation einer solchen Aussage muss man allerdings vorsichtig sein, denn sie bedeutet nicht, dass regelmäßig jeder 42. Anruf extrem lang dauert. Direkt nach einem langen Anruf kann man *nicht* sagen: *Jetzt haben wir 40 Anrufe lang Ruhe*. Es handelt sich um eine Aussage über zufällige Ereignisse. Man kann nicht sagen, wann gewisse genau bestimmte Ereignisse eintreten werden, sondern nur die *Wahrscheinlichkeiten* angeben, mit denen sie auftreten.

Es sollen aber auch noch einmal die nicht so extremen Beobachtungen betrachtet werden, die im Histogramm der Anrufdauer in Abb. 1.14 dargestellt sind. Man erkennt deutlich, dass fast die Hälfte aller Anrufe zwischen 0 und 100 Sekunden gedauert hat. Darüber hinaus nimmt die Anzahl beobachteter Anrufe mit zunehmender Anrufdauer ab. Auch hier wurde versucht, ein stochastisches Modell zu

finden, das die beobachteten Daten angemessen beschreibt; es wurde wiederum als Kurve über das Histogramm gezeichnet (hier eine Exponentialverteilung). Mit Hilfe dieses Modells könnte man jetzt noch genauere Aussagen über das Auftreten extrem langer oder auch extrem kurzer Anrufe treffen, als es mit den Häufigkeiten der Fall war (Voraussetzung ist allerdings, dass das Modell die Daten angemessen beschreibt).

An dieser Stelle könnte aber noch auffallen, dass die Grafiken in Abb. 1.14 eine gewisse Ähnlichkeit mit den Erdbeben-Grafiken in den Abbildungen 1.10 und 1.11 aufweisen. Dies ist kein Zufall, denn wie sich später noch zeigen wird, liegt in beiden Fällen vermutlich ein ähnlicher stochastischer Prozess zu Grunde.

Bisher haben wir Beispiele mit eher betriebswirtschaftlichem Bezug betrachtet. Es gibt jedoch auch viele volkswirtschaftliche Anwendungsmöglichkeiten statistischer Methoden. Die Disziplin, die sich mit der Anwendung statistischer Methoden auf volkswirtschaftliche Fragestellungen beschäftigt, hat sogar einen eigenen Namen: **Ökonometrie**. Ökonometriker untersuchen beispielsweise, ob es einen Zusammenhang zwischen der Inflationsrate und der Arbeitslosenquote gibt (wie er von der sogenannten Phillips-Kurve unterstellt wird) oder ob sich die Hartz-Gesetze tatsächlich positiv auf den Arbeitsmarkt ausgewirkt haben.

Beispiel 1.11. Entwicklung der Arbeitslosenquote in Deutschland

Es ist sicherlich weithin bekannt, dass die Arbeitslosenquote in Deutschland starken Saison-Schwankungen unterliegt. So ist die Arbeitslosenquote im Winter in der Regel deutlich höher als im Sommer, weil beispielsweise viele Arbeitnehmer im Baugewerbe vorübergehend entlassen werden. Abbildung 1.15 verdeutlicht dies durch die Darstellung der zeitlichen Entwicklung der monatlichen Arbeitslosenquote in Deutschland (in % aller Erwerbspersonen) von Anfang 2002 bis Ende 2007.¹⁰

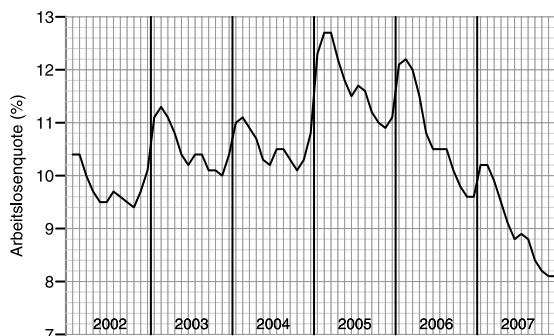


Abb. 1.15 Zeitliche Entwicklung der monatlichen Arbeitslosenquote in Deutschland von Januar 2002 bis Dezember 2007

¹⁰ Die Arbeitslosenquoten wurden am 25. April 2008 auf der Internetseite <https://www-genesis.destatis.de/genesis/online/logon> des statistischen Informationssystems GENESIS-Online des Statistischen Bundesamtes abgefragt.

Wenn man die Entwicklung einer Größe mit der Zeit betrachtet, spricht man auch von einer *Zeitreihe*. Die Zeitreihe der Arbeitslosenquote zeigt das erwartete saisonale Muster. Darüber hinaus scheint sie aber auch einen Trend aufzuweisen, der bis 2005 langsam ansteigt und danach wieder relativ stark abnimmt. Im Rahmen der Zeitreihenanalyse (Kapitel 13) werden einfache Methoden vorgestellt, mit denen man solche Zeitreihen untersuchen kann.

An dieser Stelle soll abschließend noch ein Beispiel aus der Qualitätskontrolle betrachtet werden, das im weiteren Verlauf des Buches mehrfach verwendet wird.

Beispiel 1.12. Brenndauer von Glühlampen

Tabelle 1.7 gibt die Brenndauer (in Stunden) von 30 Glühlampen (40 Watt) an, die im Rahmen einer Qualitätskontrolle untersucht wurden.¹¹ Eine Einteilung der Brenndauern in Intervalle ist Tabelle 1.8 zu entnehmen. In Abb. 1.16 sind die Daten grafisch dargestellt. Die Striche am unteren Rand der Grafik kennzeichnen, bei welchen Werten die beobachteten Brenndauern lagen. Die glatte Kurve stellt wieder ein stochastisches Modell dar (eine Normalverteilung), das an die Daten angepasst wurde. Wie bereits mehrfach erwähnt, werden im Laufe dieses Buches Methoden vorgestellt, mit deren Hilfe man Modelle anpassen und interpretieren kann.

Tabelle 1.7 Brenndauer von 30 Glühlampen (Stunden)

699	756	814	827	863	889	924	956	1 003	1 028
1 049	1 055	1 058	1 061	1 063	1 068	1 085	1 134	1 160	1 178
1 197	1 204	1 222	1 252	1 255	1 262	1 303	1 310	1 550	1 562

Tabelle 1.8 Brenndauer von 30 Glühlampen (Stunden), gruppiert

[600;800]	(800;1 000]	(1 000;1 200]	(1 200;1 400]	(1 400;1 600]
2	6	13	7	2

Es soll nun die folgende Frage betrachtet werden: *Wie groß ist die Brenndauer einer Glühlampe dieser Art?* Es gibt keine einfache Antwort auf diese Frage. Diese Frage lässt sich nämlich nicht durch die Nennung einer einzigen Zahl beantworten. Jede Glühlampe hat eine andere bzw. zufällige Brenndauer. Man hat es wieder mit einer stochastischen Situation zu tun. Die Frage kann nur mit Aussagen über Wahrscheinlichkeiten beantwortet werden.

In diesem Buch geht es darum, einen Eindruck von der statistischen Analyse zu bekommen, die man etwa wie folgt zusammenfassen kann:

Eine statistische Analyse besteht in der Regel aus dem Suchen, Anpassen, Überprüfen und Interpretieren stochastischer Modelle.

¹¹ Die Original-Daten stammen aus dem Artikel *Davis, D.J. (1952): An Analysis of Some Failure Data. Journal of the American Statistical Association 47, 113-150*. Von den in diesem Artikel angegebenen Brenndauern von insgesamt 417 Glühlampen wurden hier 30 zufällig ausgewählt.

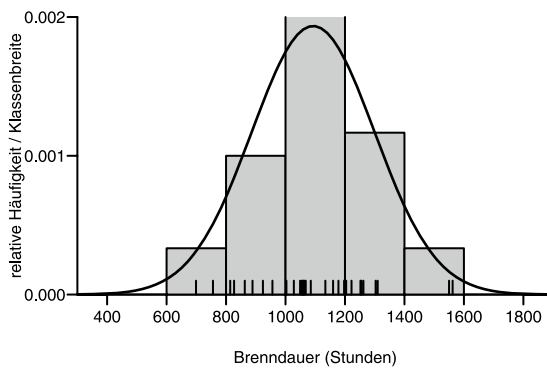


Abb. 1.16 Brenndauer von Glühbirnen und Normalverteilung

In diesem Kernsatz sind die Inhalte dieses Buches zusammengefasst. Das Ziel dieses Buches ist, zu beschreiben, wie man bei der Durchführung der einzelnen Schritte vorgeht. Dabei gibt es ein eingeführtes Vokabular, um über stochastische Modelle und Begriffe, wie Wahrscheinlichkeit objektiv zu diskutieren. Dieses Buch wird das Grundvokabular vermitteln, das nötig ist, um stochastische Phänomene objektiv zu beschreiben. Von großer Bedeutung sind Grundbegriffe wie *Zufall*, *Wahrscheinlichkeit*, *Variation* und *Schätzung*. Sie sind nicht nur Meilensteine auf dem Weg, kompliziertere Methoden zu verstehen, sie eröffnen auch die Möglichkeit, Aspekte unseres Lebens und unserer Umwelt besser zu verstehen. Viele dieser Ideen werden mit mathematischen Symbolen dargestellt.

Die oben dargestellten Beispiele (mit Ausnahme des Pendels) haben gemeinsam, dass sie mit dem Begriff *Unbestimmtheit* oder *Unsicherheit* zu tun haben. Das englische Wort **uncertainty** (oder auch das deutsche Wort *Ungewissheit*) drückt das am besten aus. Im täglichen Leben hat man viel mit Ungewissheit zu tun. Sehr viele Entscheidungen, die man jeden Tag trifft, enthalten Ungewissheit (oder Unbestimmtheit). Dies können triviale Entscheidungen sein, wie z.B. die Frage, ob man einen Regenschirm mitnehmen soll für den Fall, dass es regnet, oder die Frage, ob es sich lohnt, für eine bestimmte Vorlesung sehr früh aufzustehen. Aber es gibt auch wichtigere Entscheidungen, die unter der Bedingung eines ungewissen Ausgangs getroffen werden müssen, z.B. in den Bereichen Gesundheit, Umwelt, Politik, Wirtschaft oder Wissenschaft. Entscheidungen, die unter der Bedingung eines ungewissen Ausgangs gefällt werden, sind sehr zahlreich. Einige Beispiele für Fragestellungen, die entschieden werden müssen, sind die folgenden:

- Ist eine gewisse Maßnahme zum Umweltschutz effektiv oder nicht?
- Ist eine neue Medizin verträglich genug, um freigegeben zu werden?
- Sollte man in ein bestimmtes Projekt investieren oder nicht?
- Wie wird der Markt auf eine gewisse Produktänderung reagieren?

Es ist wohl nicht nötig, weiter zu begründen, dass es sinnvoll ist, über den Begriff *uncertainty* nachzudenken, insbesondere über das Verstehen und die Messbarkeit von Unsicherheit. Die Hauptsache ist, dass man beginnt, kritisch über Fragestellun-

gen nachzudenken, die unter der Bedingung der Ungewissheit entschieden werden. Diese Begriffe sollten nach der Lektüre des Buches zum intellektuellen Handwerkszeug gehören. Man sollte dann auch in der Lage sein, nicht nur wissenschaftliche Problemstellungen, sondern auch einfache reale Fragestellungen unter Ungewissheit mit Hilfe statistischer Methoden systematisch zu analysieren. Einige Beispiele sind die folgenden:

- Ist es für einen rational denkenden Menschen sinnvoll, Lotto zu spielen?
- Wie groß ist die Chance, bei der Auslosung von Tickets für eine Fußball-Weltmeisterschaft Karten für mindestens ein Spiel zu erhalten?
- Lohnt sich an Stelle des Kaufs einer regulären, 12 Monate gültigen BahnCard 25 der Kauf einer Aktions-BahnCard 25, deren Gültigkeitsdauer vom Erfolg der deutschen Fußball-Nationalmannschaft abhängt, wie z.B. bei der Weltmeister-BahnCard zur WM 2006 oder der Fan-BahnCard zur EM 2008?
- Rentiert sich die Investition in eine Anleihe, deren Rückzahlungshöhe auf eine bestimmte Weise an die Entwicklung des Aktienindex Dow Jones EURO STOXX 50 gekoppelt ist?

Bevor in den folgenden Abschnitten dieses Kapitels noch einige wichtige statistische Grundbegriffe eingeführt werden, sind noch einige kurze Bemerkungen zu dem Beispiel eines Pendels notwendig. Das Pendel, das in *Beispiel 1.1* betrachtet wurde, ist kein *reales* Pendel, sondern ein *mathematisches*. Es ist eine Idealisierung, da von idealierten Bedingungen ausgegangen wird, beispielsweise, dass sich das Pendel im Vakuum befindet. Ein reales Pendel jedoch ist kein perfekter Gegenstand. Es bewegt sich nicht exakt nach der Formel des deterministischen Modells. Angenommen man misst für verschiedene Pendel-Längen die genaue Zeit, die das Pendel braucht, um einmal hin und zurück zu schwingen. Dann ist unwahrscheinlich, dass die Beobachtungen exakt auf der Kurve liegen, die in Abb. 1.1 die Schwingungsdauer des Pendels beschreibt. Sie würden vom Charakter her vermutlich eher wie in Abb. 1.17 aussehen.

Ebenso wie die Dauer eines Linienfluges nicht nur von der Flugstrecke abhängt, wird die Schwingungsdauer des realen Pendels außer von der Länge noch durch an-

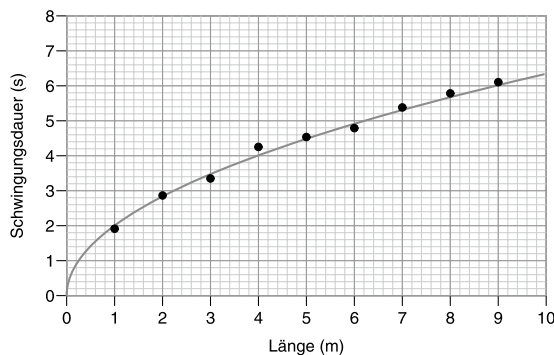


Abb. 1.17 Beobachtungen für die Schwingungsdauer realer Pendel

dere Faktoren beeinflusst; man stelle sich als extremes Beispiel ein Pendel vor, das auf einer Wiese im Wind schwingt. Das bedeutet, dass auch die Beziehung zwischen der Länge und der Schwingungsdauer eines realen Pendels im Grunde stochastisch ist. Trotzdem ist das deterministische Modell für praktische Zwecke wie die Zeitmessung genau genug.

1.3 Grundgesamtheit und Stichprobe

Die Begriffe *Grundgesamtheit* und *Stichprobe* sollen anhand des Glühbirnen-Beispiels (*Beispiel 1.12*) eingeführt werden. In größerem Zusammenhang gesehen kommt Glühbirnen nur eine relativ geringe Bedeutung zu. Es besteht kein dringendes persönliches Interesse an Glühbirnen oder ähnlichen geringwertigen Gegenständen des privaten Bedarfs. Dieses Beispiel wird verwendet, weil es sehr einfach zu verstehen ist und gleichzeitig viel Grundlegendes an ihm erklärt werden kann.

Angenommen, die Informationen über die Glühbirnen-Brenndauer aus den Tabellen 1.7 und 1.8 bzw. der Abb. 1.16 sind nicht verfügbar, und man möchte in einem Geschäft eine Glühbirne kaufen. Dann ist es nicht unerheblich zu wissen, wie lange diese Glühbirne leuchten wird. Es stellt sich also die einfache Frage: *Wie lange wird diese Glühbirne brennen?* Die Antwort auf diese Frage ist gar nicht so einfach. Zwei mögliche Antworten sind die folgenden:

- Es ist nicht möglich, die Frage für eine bestimmte Glühbirne zu beantworten, da alle Glühbirnen eine unterschiedliche Lebensdauer haben.
- Die Beantwortung kann eindeutig erfolgen, aber erst dann, wenn die Glühbirne durchgebrannt ist.

Keine der beiden Antworten ist von besonderem Nutzen. Die erste sagt nicht das, was man eigentlich wissen möchte, und auf die zweite Antwort kann man nicht warten. Man hat es, wie bereits mehrfach erwähnt, mit einer stochastischen Situation zu tun. Alle Individuen (Glühbirnen) sind verschieden. Jede Glühbirne hat ihre eigene Lebensdauer. Die Frage hat keine einfache Antwort wie etwa *1 000 Stunden*.

Eine Möglichkeit, in einer solchen Situation weiterzukommen, ist die folgende: man kann andere *ähnliche* Glühbirnen testen und beobachten, wie lange diese brennen. Dadurch erhält man indirekt Informationen über die mögliche Brenndauer der zu kaufenden Birne. Das entscheidende Wort ist hier *ähnlich*. Was damit gemeint ist, wird im Folgenden genauer betrachtet.

Zunächst ist zu entscheiden, welche Glühbirnen derjenigen ähnlich sind, die man kaufen möchte. Es bringt nichts, eine andere Sorte zu untersuchen. Das bedeutet, man muss eine *Menge von Glühbirnen* bestimmen, aus der man einige zum Testen auswählt. Eine solche Menge wird in der Statistik als *Grundgesamtheit* bezeichnet.

Die **Grundgesamtheit** ist die Menge der Objekte, Personen oder anderer Dinge, über die man Informationen gewinnen möchte.

Überraschenderweise ist es in vielen Fällen keineswegs trivial, die zur Beantwortung einer bestimmten Frage geeignete Grundgesamtheit einzugrenzen. Welche Menge von Glühbirnen sollte man als Grundgesamtheit in diesem Beispiel verwenden:

- Alle Glühbirnen dieses Typs, die jemals hergestellt wurden,
- nur diejenigen Glühbirnen, die in einem bestimmten Jahr produziert wurden oder
- diejenigen, die in einer bestimmten Produktionsperiode angefertigt wurden?

Diese Frage soll zunächst zurückgestellt werden. Für den Moment wird angenommen, dass die Grundgesamtheit durch alle Glühbirnen der gewünschten Sorte im ausgesuchten Geschäft beschrieben wird und dass diese Grundgesamtheit genau 100 Glühbirnen umfasst. Diese Grundgesamtheit ist schematisch in Abb. 1.18 dargestellt.

Nun steht man vor der nächsten Frage: *Wie viele Glühbirnen* soll man untersuchen? In der Regel betrachtet man aus Kostengründen oder anderen Gründen nur eine Teilmenge der relevanten Grundgesamtheit und spricht dann von einer **Stichprobe**. Da man nicht annehmen kann, dass die Brenndauer aller Glühbirnen gleich lang ist, macht es dabei keinen Sinn, nur die Brenndauer einer einzigen Glühbirne zu beobachten. Im Allgemeinen ist es unvernünftig, nur ein Mitglied einer Grundgesamtheit zu untersuchen, es sei denn, man hat es mit einem deterministischen Phänomen zu tun. Andererseits kann man nicht alle 100 Glühbirnen testen. Dann bleiben keine mehr übrig.

Es gibt allerdings Situationen, in denen es — zumindest prinzipiell — sinnvoll bzw. möglich ist, alle Mitglieder einer Grundgesamtheit zu untersuchen, z.B. wenn die Grundgesamtheit durch die Menge aller Autos in Deutschland gegeben ist und man den Schadstoffausstoß messen möchte. Dadurch würden natürlich enorme Kosten entstehen. Ein weiteres Beispiel, in dem man im Prinzip die ganze Grundgesamtheit untersuchen könnte, findet sich im Bereich des Prüfungswesens. Aber auch

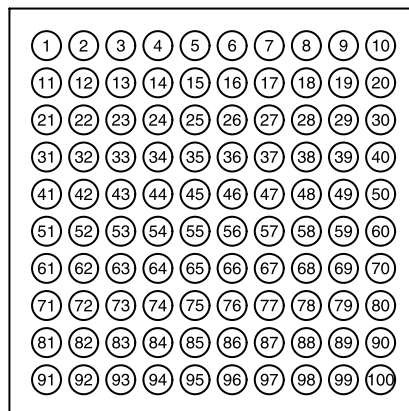


Abb. 1.18 Schematische Darstellung der Grundgesamtheit von 100 Glühbirnen

Jahresabschlussprüfungen in Unternehmen werden aus Zeit- und Kostengründen auf Stichprobenbasis durchgeführt.

Die Festlegung der **Stichprobengröße**, also der genauen Anzahl der zu untersuchenden Objekte, wird von vielen Faktoren abhängen, unter anderem davon

- wieviel es *kostet*, ein Mitglied der Grundgesamtheit zu untersuchen und
- mit welcher *Genauigkeit* man die Grundgesamtheit beschreiben möchte.

Die Frage, wie viele Objekte aus der Grundgesamtheit man untersuchen sollte, ist in der Tat sehr schwierig, und wird zunächst zurückgestellt. Stattdessen folgen zunächst einige Anmerkungen zum nächsten Problem, der **Stichprobenziehung**.

Sicherlich ist in der Zwischenzeit klar geworden, dass die ursprüngliche, so einfach erscheinende Frage: *Wie lange wird diese Glühbirne brennen?* sich als eine sehr komplizierte Frage erweist, wobei das nicht heißen soll, dass einfache Dinge nur aus Spaß an der Theorie kompliziert gemacht werden. Man muss solche Überlegungen auch anstellen, um bedeutendere Fragen solcher Art zu beantworten, und die Konsequenzen der Antworten in anderen Situationen (Kernkraftwerke, neue Medikamente, Investitionen) sind wesentlich gravierender als im Falle einer Glühbirne.

Angenommen, man hat sich für 30 Glühbirnen als eine gut zu untersuchende Anzahl von Glühbirnen entschieden und möchte also eine Stichprobe von 30 Glühbirnen nehmen und ihre Brenndauer beobachten. Dann stellt sich die Frage, welche der 100 Glühbirnen im Supermarkt-Regal man für die Stichprobe auswählen soll:

- Soll man sie von vorne wegnehmen oder vielleicht nicht? Vielleicht gibt es in dem Supermarkt die Praxis, die älteren Glühbirnen, deren Brenndauer vielleicht geringer ist, ganz nach vorne in das Regal zu legen. Andererseits könnten die alten auch länger halten.
- Soll man 15 von vorne und 15 von hinten nehmen? Vielleicht gibt es auch noch andere Gründe für systematische Verfälschungen, von denen man nichts weiß, z.B. dass 5 alte hinten im Regal liegen, während vorne 95 neue liegen. Dann hätte man in der Stichprobe 5 alte und 25 neue Glühbirnen. Auch das würde einen verfälschten Eindruck von der Grundgesamtheit geben.

Eine Möglichkeit, das Problem zu lösen, ist die Glühbirnen *zufällig* auszuwählen. Man könnte z.B. jeder Glühbirne in der Grundgesamtheit eine Nummer (1 bis 100) zuordnen, diese Zahlen auf kleine Zettel schreiben, die Zettel gründlich mischen und dann 30 Zettel ziehen. Dieses Verfahren ist unter dem Namen einfache Zufallsauswahl bekannt. Die Stichprobe, die man erhält, nennt man einfache **Zufallsstichprobe**. Der große Vorteil einer zufälligen Auswahl ist, dass systematische Fehler vermieden werden. Oder umgekehrt formuliert, Zufallsstichproben sind *in der Regel repräsentativ*, weil bei einfachen Zufallsstichproben alle Mitglieder der Grundgesamtheit die gleiche Chance haben, in die Stichprobe zu kommen. In Abb. 1.19 ist ein mögliches Ergebnis der Zufallsauswahl dargestellt.

Es gibt jedoch keine Methode, die *garantiert*, dass die gewählte Stichprobe die Grundgesamtheit korrekt repräsentiert. Die Zufallsauswahl ist ein Versuch, systema-

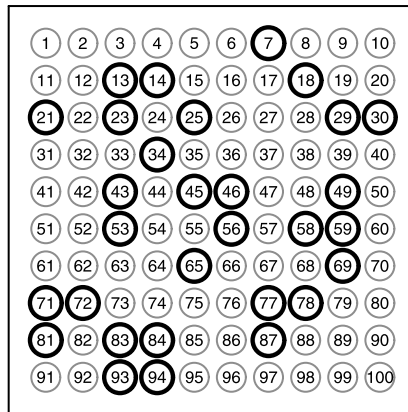


Abb. 1.19 Schematische Darstellung der Grundgesamtheit von 100 Glühlampen und einer Stichprobe von 30 Glühlampen

tische Fehler zu vermeiden. Sie liefert jedoch nicht immer repräsentative Stichproben. Angenommen, 70 der 100 Glühlampen taugen nichts und die anderen 30 haben eine lange Brenndauer. Auch bei einer zufälligen Auswahl könnte man, wenn man absolutes Pech hat, gerade die 30 guten erwischen. Das kann passieren, ist aber bei einer zufälligen Auswahl sehr unwahrscheinlich. Zufällige Stichproben tendieren also dazu, repräsentativ zu sein.

Zusammenfassend sind die Schritte, die wir bisher durchlaufen haben, die folgenden. Man möchte die Frage *Wie lange hält die Glühlampe?* beantworten. Man schafft das nicht, ohne die Glühlampe durchzubrennen. Deshalb betrachtet man andere Glühlampen aus der Grundgesamtheit und beobachtet, wie lange diese brennen. Im Idealfall würde man alle anderen außer der gewählten Glühlampe untersuchen. Das würde die maximal mögliche Information liefern. Jedoch wäre es zu teuer, alle zu untersuchen. Deshalb untersucht man nur 30 Glühlampen. Man möchte, dass diese 30 Glühlampen einen „fairen“ Eindruck über die Lebensdauer der Glühlampen in der Grundgesamtheit vermitteln, d.h. man möchte, dass die Glühlampen in der Stichprobe *repräsentativ* für alle Glühlampen der Grundgesamtheit sind. Zusammenfassend ist von Bedeutung, aus *welcher Grundgesamtheit wie viele* und *welche* Elemente in die Stichprobe sollen.

Die Prinzipien, die an diesem simplen Beispiel aufgezeigt werden können, gelten auch für Fragen, die von größerer Bedeutung sind, wie z.B.:

- Wann fällt der Motor eines Flugzeugs aus, das man gerade besteigen will?
- Wann wird es einen Störfall im Kernkraftwerk nahe eines Wohnortes geben?

In allen Bereichen des Lebens gibt es ähnliche Fragen. Weitere Beispiele sind:

- Wie lange braucht ein bestimmter Patient, um gesund zu werden?
- Wie viele Autos, Schuhe, Brötchen oder Bananen werden wir heute verkaufen?
- Wie wird der Goldpreis in einem Jahr sein?
- Wie lange wird es dauern, bis Tiere dieser Art ausgestorben sind?

Im folgenden Beispiel soll das Problem der systematischen Verfälschung noch einmal aufgegriffen und genauer betrachtet werden. Es geht also wieder um die Frage, welche Elemente der Grundgesamtheit in die Stichprobe gelangen sollen.

Beispiel 1.13. Lebensmittelausgaben Göttinger Studenten

Angenommen, man möchte die Ausgaben der Göttinger Studenten für Lebensmittel in der letzten Woche schätzen. Zunächst muss man klären, wen man zur Gruppe der Göttinger Studenten (Grundgesamtheit) zählt. Werden Studenten mit einbezogen, die in der letzten Woche gar nicht in Göttingen waren oder die bei ihren Eltern leben?

Im Folgenden soll angenommen werden, dass man sich darüber geeinigt hat, über welche Gruppe von Studenten man spricht, dass also eine klar abgegrenzte Grundgesamtheit definiert wurde. Außerdem sollen bei den folgenden Erklärungen rein technische Schwierigkeiten vernachlässigt werden, wie z.B. die Frage, wie man erfährt, wieviel ein bestimmter Student in der letzten Woche ausgegeben hat. Es soll einfach davon ausgegangen werden, dass alle Studenten genau wissen, wieviel sie in der letzten Woche für Lebensmittel ausgegeben haben und dass sie ehrlich antworten. Schließlich soll noch angenommen werden, dass ausreichend Mittel zur Verfügung stehen, um die Befragung von 50 Studenten zu finanzieren; d.h. auch der Stichprobenumfang ist festgelegt.

Man wird nicht sehr viel herausfinden, wenn die ausgewählte Stichprobe nicht repräsentativ für die Grundgesamtheit der Göttinger Studenten ist. Man würde höchstens etwas über die Ausgaben der 50 Studenten aus der Stichprobe erfahren, nichts jedoch über die Grundgesamtheit; d.h. die Ergebnisse sind nicht generalisierbar, wenn man nicht annehmen kann, dass die Stichprobe repräsentativ ist.

Es gibt viele Möglichkeiten, Stichproben auszuwählen, die *nicht* repräsentativ sind: man könnte z.B. die ersten 50 Studenten aus der Göttinger Mensa-Schlange auswählen. Dann wären allerdings diejenigen Studenten, die nicht in der Mensa essen, mit Sicherheit nicht in der Stichprobe vertreten. Das würde bedeuten, dass man diejenigen Studenten nicht erfasst,

- die es sich nicht leisten können, in der Mensa zu essen,
- die lieber in Restaurants essen oder
- die gar nicht zu Mittag essen usw.

Da davon auszugehen ist, dass diese Gruppen andere Lebensmittelausgaben haben, ist es keine gute Idee, die ersten 50 Studenten der Mensa-Schlange zu wählen. Die Stichprobe ist mit Blick auf die Grundgesamtheit verfälscht.

Man könnte sich ein komplizierteres System überlegen, das nicht zu einer so unrepräsentativen Stichprobe führt. Wenn man sich eine gute Methode überlegen möchte, um eine Stichprobe von 50 Göttinger Studenten auszuwählen, wird man merken, dass es nicht einfach ist, eine Methode zu finden, die praktisch durchführbar und darüber hinaus nicht systematisch verfälscht ist.

Mit einer systematischen Verfälschung oder Verzerrung ist gemeint, dass eine Gruppe ganz übersehen wird oder nicht im richtigen Verhältnis (im Vergleich zur

Grundgesamtheit) vertreten ist. Der wichtigste Punkt bei einer *zufälligen Auswahl* ist, dass jeder Student die *gleiche Chance* hat, ausgewählt zu werden.

Um eine zufällige Auswahl zu treffen, könnte man tatsächlich so vorgehen wie im Beispiel der Glühbirnen, d.h. man schreibt die Namen aller Göttinger Studenten auf jeweils einen Zettel, durchmischt diese Zettel und zieht dann 50 heraus. Eine einfachere Möglichkeit wäre, sich die Matrikelnummern im Studentensekretariat ausdrucken zu lassen und diese für die Ziehung zu verwenden. Diese oder ähnliche mechanische Vorgehensweisen wären hier noch durchführbar, weil es nicht allzu viele Göttinger Studenten gibt.

Es gibt jedoch Situationen, in denen man mit den beschriebenen Vorgehensweisen nicht weiterkommt. Betrachtet man noch einmal *Beispiel 1.3*, Aspirin und Herzanfälle, dann ist nicht klar, wie groß die Grundgesamtheit bei dieser Studie war. Die Stichprobe allein bestand schon aus 22072 Männern mittleren Alters. Man würde ziemlich viel Aufwand betreiben, um die zugehörige Grundgesamtheit auf Papier zu bringen. Natürlich führt man die Auswahl in solchen Situationen mit Hilfe eines Computers durch, der so programmiert wird, dass er zufällig aus einer Liste auswählt. Auch für kleine Stichproben ist es nicht nötig, mechanische Geräte wie Zettel oder den Apparat zur Ziehung der Lottozahlen im Fernsehen zu benutzen. Stattdessen kann man Statistik-Programme verwenden, um Zufallszahlen wie in dem Glühbirnen-Beispiel zu ziehen.

In der Praxis wird es allerdings oft, insbesondere bei räumlich sehr ausgedehnten Studien, zu kostspielig sein, eine einfache Zufallsstichprobe zu ziehen. Angenommen, man möchte etwas über die Wirkung von saurem Regen auf Birken in Europa oder auch nur in Niedersachsen erfahren. Man kann sich vorstellen, wie viel es kosten würde, eine Liste aller Bäume zu erstellen, um damit eine einfache Zufallsstichprobe aus dieser Grundgesamtheit zu ziehen.

Man muss in solchen Fällen andere Methoden verwenden, die weniger kostspielig sind. Es gibt eine ganze Theorie, die **Stichprobentheorie**, über die unterschiedlichen Möglichkeiten, Stichproben zu ziehen, ohne dabei einen systematischen Fehler zu machen oder ihn zumindest zu kontrollieren. Im Folgenden soll kurz auf solche Möglichkeiten eingegangen werden.

Bei **geschichteten Zufallsstichproben** (stratified sampling) ist man in der Lage die Grundgesamtheit in Gruppen (strata) zu unterteilen. Angenommen, man wäre daran interessiert, das Einkommen eines Landes zu schätzen, in dem viele Einwohner arm und wenige reich sind. Der Stichprobenmittelwert wird dann stark davon abhängen, wie viele reiche Personen in der Stichprobe enthalten sind. Man kann sich leicht vorstellen, dass es bei einer relativ kleinen Stichprobe nicht unwahrscheinlich ist, eine Stichprobe zu erhalten, die nicht repräsentativ ist, d.h. das richtige Verhältnis von arm und reich widerspiegelt. Anders ausgedrückt, würde man eine sehr große Zufallsstichprobe benötigen, um repräsentative Ergebnisse zu erhalten. Bei den geschichteten Zufallsstichproben wird die Grundgesamtheit zunächst in Gruppen (Arme und Reiche) unterteilt. Anschließend werden Personen aus den verschiedenen Gruppen zufällig für die Stichprobe ausgewählt. Bei der Berechnung des mittleren Einkommens muss dann das (bekannte) Verhältnis der Gruppen in der Grundgesamtheit berücksichtigt werden. Es ist somit mög-

lich, mit einer relativ kleinen Stichprobe ein repräsentatives Ergebnis zu erhalten.

Bei der Ziehung von **Klumpenstichproben** (cluster sampling) werden ebenfalls Gruppen gebildet. Diese Gruppen unterscheiden sich jedoch nicht, sondern stellen kleine Abbilder der Grundgesamtheit dar. Bei der bereits erwähnten Untersuchung aller Birken in Niedersachsen z.B. müsste man für eine Zufallsstichprobe quer durch das Land reisen, um die Stichprobe zu untersuchen. Bildet man statt dessen Cluster, beispielsweise einzelne Wälder, und wählt dann zufällig einige Cluster und anschließend einzelne Bäume aus den Clustern aus, so können die Kosten einer repräsentativen Erhebung gesenkt werden. Man muss nur noch die entsprechenden Cluster anfahren.

Die **bewussten Auswahlverfahren** (quota sampling) legen bestimmte Quoten in der Stichprobe von vornherein fest. Wenn z.B. bei der Untersuchung der Lebensmittelausgaben Göttinger Studenten bekannt ist, dass 40 % der Göttinger Studenten weiblich sind, wäre es möglich, in einer Stichprobe der Größe 50 genau 20 Frauen und 30 Männer aufzunehmen. Man erreicht Repräsentativität bezüglich des Geschlechts. Ebenso könnte mit weiteren Merkmalen vorgegangen werden, wie beispielsweise Fachrichtung oder Semesterzahl. Problematisch ist, dass man nicht in der Lage ist, alle relevanten Merkmale zu kontrollieren, so dass durch die Vermeidung einer zufälligen Auswahl die Gefahr besteht, genau das Gegenteil, nämlich eine unrepräsentative Stichprobe zu erhalten.

Bei statistischen Untersuchungen muss man sich immer darüber im Klaren sein, wie eine Stichprobe erzeugt worden ist und welche Konsequenzen dies haben könnte. Die Verfahren wurden hier kurz beschrieben, um zu zeigen, dass es Möglichkeiten gibt, wenn die einfache Zufallsauswahl ungeeignet erscheint. Im Folgenden soll jedoch immer mit Zufallsstichproben gearbeitet werden.

Unabhängig von den beschriebenen Verfahren der Stichprobentheorie ist eine Möglichkeit, die Kosten gering zu halten, ein kleine und leicht zugängliche Grundgesamtheit zu verwenden. Jedoch ist bei der Wahl der Grundgesamtheit zu beachten, dass die Grundgesamtheit, aus der man die Stichprobe zieht, die *Allgemeingültigkeit* der Ergebnisse und Schlussfolgerungen bestimmt. Diese gelten jeweils nur für die Grundgesamtheit, aus der die Stichprobe gezogen wurde. Wie dies zu verstehen ist, wird noch einmal an dem bereits bekannten Beispiel von Aspirin und Herzanfällen (*Beispiel 1.3*) erläutert.

Angenommen, man wollte eine ähnliche Studie in Deutschland wiederholen, um zu sehen, ob die Ergebnisse auch hier gültig sind. Das Risiko eines Herzinfarkts unter deutschen Männern mittleren Alters könnte ja anders sein als das unter amerikanischen Männern mittleren Alters. Beispielsweise unterscheidet sich die Nahrung in den beiden Ländern, ebenso Sport- und Rauchgewohnheiten usw. Und auch wenn sich die Risiken für einen Herzinfarkt in den beiden Ländern nicht unterscheiden, könnte es sein, dass die Verminderung des Risikos, die aus der Einnahme von Aspirin resultiert, sich hier von der unterscheidet, die für die amerikanische Grundgesamtheit ermittelt wurde.

Die Originalstudie war auf gesunde Männer mittleren Alters beschränkt. Das umfasst (auch in Deutschland) sehr viele Personen. Man könnte diese Gruppe ein-

schränken, indem man nur gesunde Männer mittleren Alters aus Niedersachsen betrachtet. Wenn diese Gruppe immer noch zu groß ist, könnte man sie weiter einschränken auf gesunde, 40-jährige Männer aus Niedersachsen. Die Grundgesamtheit wird noch einmal kleiner und überschaubarer. Wenn man die Grundgesamtheit immer weiter eingrenzt, wird man aber am Ende nur sehr eingeschränkte Schlussfolgerungen ziehen können. Beispielsweise sind Aussagen der folgenden Art von sehr geringem Interesse: *Eine Studie über gesunde, 40-jährige männliche, ledige Linkshänder aus Niedersachsen ergab, dass Aspirin das Risiko, einen Herzanfall zu erleiden, reduziert.*

Begründeterweise würden die meisten diese Nachricht nur mit einem Achselzucken zur Kenntnis nehmen, da sie nicht Teil der Grundgesamtheit sind. Es muss daher noch einmal betont werden, dass die Grundgesamtheit, aus der man die Stichprobe zieht, die *Allgemeingültigkeit* der Ergebnisse und Schlussfolgerungen bestimmt. Wenn man nur eine sehr spezielle Grundgesamtheit untersucht, sind die Ergebnisse also von begrenzter Verwendbarkeit.

Das bedeutet jedoch nicht, dass es immer besser ist, eine große Grundgesamtheit zu haben (auch wenn der Kostenaspekt einmal vernachlässigt wird). Für einen gesunden, 40-jährigen, männlichen, ledigen Linkshänder aus Niedersachsen wäre die oben genannte hypothetische Studie einer allgemeineren vorzuziehen. Für Personen, die diesem Linkshänder sehr *ähnlich* sind, wären die Ergebnisse wesentlich spezifischer und folglich (bei sonst gleichen Bedingungen) auch genauer.

Bisher wurde davon ausgegangen, dass es nur eine Frage des Aufwands ist, eine adäquate Stichprobe zu ziehen (beispielsweise aus allen Göttinger Studenten). In einigen Situationen ist es jedoch gar nicht möglich, eine Zufallsstichprobe zu ziehen. Selbst wenn Kosten keine Rolle spielen, muss man mit gegebenen Daten auskommen. Ein solcher Fall liegt zum Beispiel bei der Blockzeit eines Linienfluges (*Beispiel 1.2*) vor. Daten wie die Blockzeit aller American Airlines Flüge von Dallas / Fort Worth (DFW) nach Philadelphia (PHL) im Februar 2006 sind eine wichtige Grundlage für die Erstellung von Flugplänen, denn um einen möglichst effizienten Flugplan zu erstellen, müssen die Fluggesellschaften wissen, wie groß die Blockzeiten für einzelne Strecken sind.

Dabei interessiert weniger, wie groß die Blockzeiten im Februar 2006 oder in den letzten 5 Jahren waren, sondern wie groß sie in der Zukunft sein werden, z.B. im nächsten Jahr. Die verfügbare Stichprobe umfasst aber nur die Blockzeiten vergangener Flüge der Grundgesamtheit. Diese Stichprobe ist natürlich keine zufällige Stichprobe aus der Grundgesamtheit. Es ist nicht möglich, eine zufällige Stichprobe aus der Grundgesamtheit (alle Blockzeiten vom Erstflug einer Strecke bis zum letzten jemals durchgeführten Flug) zu ziehen. Man benutzt die zur Verfügung stehenden Informationen, da es keine weiteren Informationen gibt. In solchen Situationen muss man *annehmen*, dass die Blockzeiten in der Stichprobe repräsentativ für die Grundgesamtheit sind. Vorher muss man sich jedoch sehr genau überlegen, ob es irgendeinen Grund gibt, der gegen eine solche Annahme spricht (vielleicht eine extrem lange Blockzeit, die daraus resultierte, dass sich ein inzwischen entlassener Pilot verfliegen hatte) oder ob sich die äußeren Rahmenbedingungen seit der Erhebung der Stichprobe geändert haben (z.B. eine neue Flugroute).

Man kann auch geeignete Korrekturen (vielleicht Ausschluss extrem kurzer oder extrem langer Blockzeiten) vornehmen, um diese Quelle von Unsicherheit auszugleichen. Eine andere Möglichkeit ist, dass der Flugplan so große Sicherheitspuffer enthält, dass auch noch extremere Blockzeiten als die bisher beobachteten ohne größere Verspätungen aufgefangen werden könnten. Dies würde natürlich wiederum die Produktivität der Flotte reduzieren.

Zusammenfassend lauten die wichtigsten Erkenntnisse dieses Abschnitts:

- In der Regel werden Zufallsstichproben gezogen, z.B. mit Hilfe von Computern.
- Weitere Methoden werden in der Stichprobentheorie behandelt.
- Es gibt Situationen, in denen man keine Zufallsstichprobe ziehen kann.

1.4 Zufallsvariablen

Im Folgenden soll eines der wichtigsten statistischen Konzepte, das der *Zufallsvariablen*, vermittelt werden. Hierzu soll wieder die Brenndauer einer hypothetischen Glühbirne betrachtet werden (*Beispiel 1.12*). Es wird nun angenommen, dass eine zufällige Stichprobe von 30 Glühbirnen gezogen wurde. Der nächste Schritt ist nun, die Glühbirnen in der Stichprobe zu untersuchen, d.h. die Brenndauern der 30 Glühbirnen zu beobachten und die Ergebnisse festzuhalten. Wir wollen auch davon ausgehen, dass dies bereits geschehen ist und die 30 beobachteten Zahlen für die Brenndauern genau die Werte aus Tabelle 1.7 sind. Das dazugehörige Histogramm wurde bereits in Abb. 1.16 dargestellt.

Allgemein geben die Beobachtungen nur Aufschluss über die Brenndauern in der Stichprobe. Tatsächlich ist man jedoch an der Brenndauer der Glühbirnen in der Grundgesamtheit interessiert. Um die Brenndauer in der Grundgesamtheit zu schätzen, benutzt man ein Modell, in diesem Fall die glatte Kurve in Abb. 1.16. Die Kurve ist das Modell für die Brenndauer in der Grundgesamtheit, das man aus den Stichprobendaten erhält. In den folgenden Kapiteln wird ausführlich beschrieben, wie man solche Modelle berechnet. Im Moment soll nur betont werden, dass diese Kurve als *geglättete* Version des Histogramms aufgefasst werden kann.

Genau wie das Histogramm sagt auch die Kurve, wo die Punkte konzentriert sind. So kann man z.B. aus der Kurve abschätzen, dass es in der Grundgesamtheit mehr Glühbirnen mit einer Brenndauer in der Nähe von 1000 Stunden als mit einer Brenndauer in der Nähe von 500 Stunden gibt.

Die Kurve heißt **Dichtefunktion**, da sie etwas über die Dichte der Punkte sagt. Abbildung 1.20 zeigt die Dichtefunktion ohne das Histogramm. Es handelt sich übrigens um eine so genannte Normalverteilung, die im weiteren Verlauf dieses Buches und in der Statistik allgemein eine bedeutende Rolle spielt und zum Beispiel auch auf dem ehemaligen 10 DM-Schein abgebildet war.

Charakteristisch für die Dichtefunktion ist die Tatsache, dass die Fläche zwischen ihr und der x -Achse immer genau 1 beträgt. Die Dichtefunktion ermöglicht es, den Anteil der Brenndauern in der Grundgesamtheit, der in ein bestimmtes Intervall

fällt, zu schätzen. Angenommen, man möchte schätzen, wie viele Glühbirnen der Grundgesamtheit eine Brenndauer zwischen 1000 und 1500 Stunden haben. Dies kann berechnet werden durch die Größe der Fläche unter der Kurve zwischen den beiden Punkten (oder besser vertikalen Linien) 1000 und 1500. In Abb. 1.21 ist diese Fläche gekennzeichnet.

Berechnet man die Größe der Fläche, so erhält man den Wert 0.65, d.h. die Fläche entspricht 65 % der Gesamtfläche. Daher schätzt man, dass 65% der Glühbirnen der Grundgesamtheit eine Brenndauer zwischen 1000 und 1500 haben. Da alle diese Glühbirnen derjenigen ähnlich sind, die man kaufen möchte, kann man das Resultat auch so interpretieren: *Mit einer Wahrscheinlichkeit von 65% wird die gekaufte Glühbirne zwischen 1000 und 1500 Stunden brennen.*

Die Dichtefunktion in den Abbildungen 1.20 und 1.21 stellt somit die Antwort auf die ursprüngliche Frage: *Wie lange wird diese Glühbirne brennen?* dar. Zumindest ist dies die konkreteste Antwort, die man überhaupt auf diese Frage bekommen kann. Man weiß immer noch nicht genau, wie lange die gekaufte Glühbirne brennen wird, aber man kann mit Hilfe der Dichtefunktion Wahrscheinlichkeiten für alle in Frage kommenden Möglichkeiten geben.

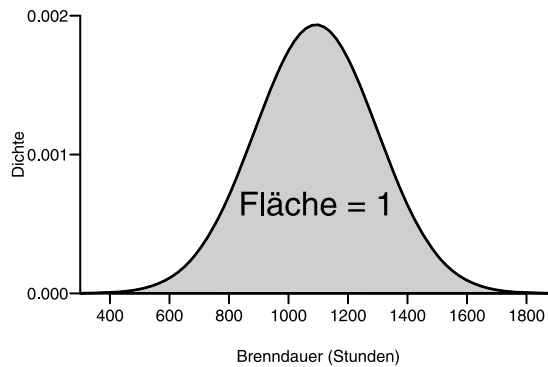


Abb. 1.20 Dichtefunktion als Modell für die Brenndauer von Glühbirnen

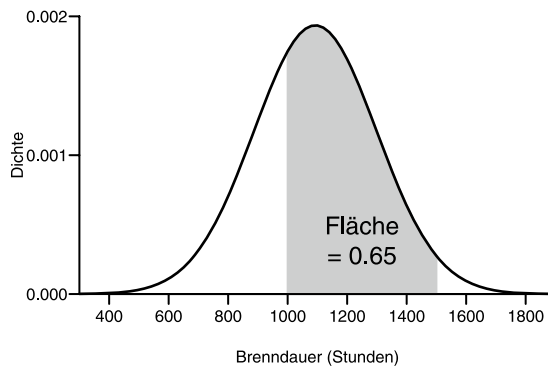


Abb. 1.21 Dichtefunktion als Modell für die Brenndauer von Glühbirnen und Wahrscheinlichkeit für eine Brenndauer zwischen 1000 und 1500 Stunden

Die Antwort auf die gestellte Frage ist daher nicht eine einzige Zahl, wie etwa 1 200 *Stunden*. Sie ist etwas anderes. Tatsächlich ist sie eines der faszinierendsten Objekte in der Mathematik. Man nennt sie eine **Zufallsvariable**. In späteren Kapiteln wird noch genau definiert, was unter einer Zufallsvariablen zu verstehen ist.

Im Rahmen dieses Buchs werden Zufallsvariablen mit Großbuchstaben bezeichnet, wie z.B. X, Y, Z . Damit sollen sie von gewöhnlichen Variablen unterschieden werden. In dem Glühbirnen-Beispiel könnte man also sagen, X sei die Brenndauer der gekauften Glühbirne. Im Folgenden werden noch einmal die Eigenschaften von X betont.

Eine Zufallsvariable X ist kein einzelner Wert. X hat einen ganzen *Bereich* möglicher Werte.

Es sieht so aus, als ob die Brenndauer der Glühbirne irgendwo zwischen 0 und 1 800 Stunden liegt. Wenn man sehr viel Glück hat, brennt sie länger. Diese Möglichkeit ist nicht auszuschließen. Vielleicht hält sie 5000 Stunden, vielleicht auch noch länger. Daher sollen als Wertebereich von X alle Zahlen betrachten, die größer als 0 sind.

Das Verhalten einer Zufallsvariablen X kann durch Wahrscheinlichkeiten beschrieben werden.

Es ist im Moment noch nicht wichtig, alle Details über die Kurve zu behalten, die als Dichtefunktion bezeichnet wurde. Es ist nur wichtig, zu verstehen, dass es in stochastischen Situationen nur möglich ist, Wahrscheinlichkeiten für gewisse Dinge anzugeben. Das bedeutet, Entscheidungen beruhen auf Wahrscheinlichkeiten, nicht auf Gewissheiten. Im weiteren Verlauf des Buches werden Dichtefunktionen noch eine große Rolle spielen und im Detail erklärt.

Angenommen, die gekaufte Glühbirne soll mindestens 1 600 Stunden glühen. Die Wahrscheinlichkeit ist durch die Fläche unterhalb der Dichtefunktion rechts von 1 600 gegeben. Diese Fläche ist in Abb. 1.22a gekennzeichnet. Sie ist gleich 0.01, d.h. es gibt eine Wahrscheinlichkeit von 1 zu 100 bzw. 1%, dass die gekaufte Glühbirne länger als 1 600 Stunden brennen wird.

Wenn man also tatsächlich eine Glühbirne braucht, die länger als 1 600 Stunden brennt, sollte man eine andere Sorte wählen und mit der ganzen Prozedur von vorne beginnen. Andererseits ist man vielleicht schon zufrieden, wenn die Glühbirne länger als 800 Stunden brennt. Ein Blick auf Abb. 1.22b zeigt, dass dies ziemlich sicher ist. Die Fläche, die man jetzt berechnen muss, ist diejenige unter der Dichtefunktion rechts von 800. Diese Fläche beträgt 0.92. Die Chance, dass der Wunsch, eine Glühbirne mit einer größeren Brenndauer als 800 Stunden zu erhalten, erfüllt wird, ist also 92%.

Es wurde gezeigt, dass die Antwort auf die ursprüngliche Frage *Wie lange wird diese Glühbirne brennen?* keine Zahl ist, sondern etwas, das gerade Zufallsvariable genannt wurde. Wenn man also nach der Brenndauer einer Glühbirne gefragt wird,

kann man als Antwort keine Zahl geben, sondern ein Bild wie das in Abb. 1.20. Damit besteht die Antwort auf die Frage aus einer ganzen Reihe von Möglichkeiten, und es wurden Informationen aus der Stichprobe genutzt, um die Wahrscheinlichkeiten dieser Möglichkeiten zu schätzen.

Die Geschichte mit der Glühbirne ist allerdings noch nicht zu Ende. Man nimmt die Glühbirne schließlich mit nach Hause, benutzt sie und eines Tages geht sie kaputt, z.B. nach 1452 Stunden. Zu diesem speziellen Zeitpunkt geschieht etwas sehr Bedeutendes. Alle Ungewissheit, über die so lange gesprochen wurde, verschwindet in diesem Moment. Die Brenndauer einer Glühbirne wird plötzlich zu einer gewöhnlichen Zahl: 1452 Stunden.

Die Frage nach der Brenndauer der Glühbirne begann mit einer komplizierten Sache, die Zufallsvariable genannt wurde und die nur durch Wahrscheinlichkeiten beschrieben werden kann. Nachdem die Glühbirne durchgebrannt ist, ist die Antwort auf die Frage eine einfache Zahl. Die tatsächliche Brenndauer der Glühbirne wurde beobachtet. In der Statistik drückt man das folgendermaßen aus: man hat jetzt eine **Realisation** der Zufallsvariablen, d.h. sie hat jetzt einen konkreten Wert angenommen, den Wert 1452 Stunden.

Damit hat die ursprüngliche Frage zwei verschiedene Antworten:

- *Bevor* die Glühbirne kaputt geht, kann die Antwort nur durch das Nennen möglicher Werte und ihrer Wahrscheinlichkeiten gegeben werden.
- *Nachdem* die Glühbirne kaputt ist, wird die Antwort zu einer gewöhnlichen Zahl.

Die Unbestimmtheit ist nach dem Durchbrennen der Glühbirne verschwunden. Der wichtige Punkt ist natürlich, dass man sich entscheiden muss, die Glühbirne zu kau-

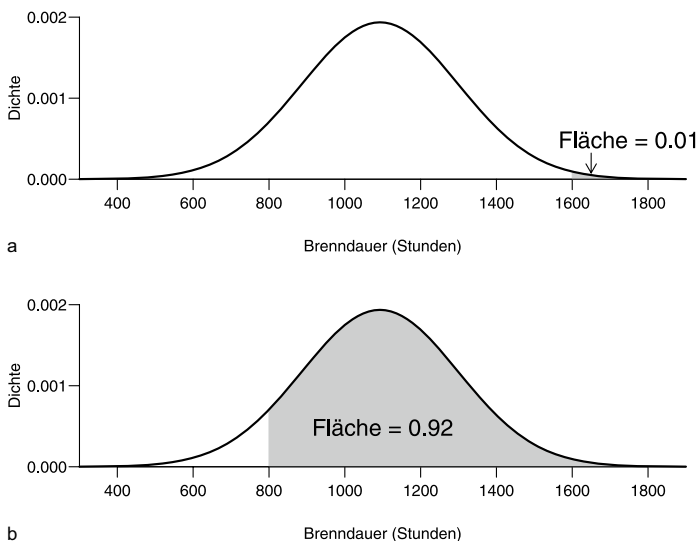


Abb. 1.22 Dichtefunktion als Modell für die Brenndauer von Glühbirnen und Wahrscheinlichkeit für eine Brenndauer von **a** mehr als 1600 Stunden bzw. **b** von mehr als 800 Stunden

fen, bevor man ihre Brenndauer beobachten kann und nicht nachher. Dasselbe passiert bei unzähligen anderen Entscheidungen, die man jeden Tag treffen muss.

Beispiel 1.14. Investition von 1 000 €

Angenommen, man hätte 1 000 € zur Verfügung und möchte diese für ein Jahr anlegen. Zum einen besteht die Möglichkeit, das Geld zu einem festen Zinssatz, z.B. 5%, auf ein Bankkonto zu legen. Wenn man das Risiko ignoriert, dass die Bank in Konkurs geht, ist dies eine sichere Anlage (und selbst dann wäre das Geld bei einer deutschen Bank abgesichert). Man weiß, dass man nach einem Jahr 1 050 € haben wird. Dies ist ein deterministischer Zusammenhang.

Als Alternative könnte man das Geld auf dem Aktienmarkt anlegen. In diesem Fall ist die Auszahlung, die am Ende des Jahres zufließt, im Voraus nicht bekannt. Sie ist eine Zufallsvariable. Angenommen, man hat die Dichtefunktion dieser Zufallsvariablen durch die Kurve in Abb. 1.23 geschätzt.

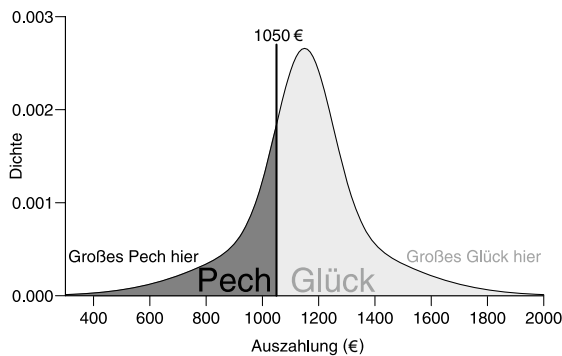


Abb. 1.23 Dichtefunktion für die Auszahlung am Ende des Anlagezeitraums

Der Wert 1 050 € ist durch den senkrechten Strich dargestellt und zeigt, was man bei der Anlage bei der Bank garantiert bekommt. Das Bild verdeutlicht, dass gute Chancen bestehen, mehr Geld zu bekommen, wenn man in Aktien investiert. Jedoch gibt es auch eine positive Wahrscheinlichkeit, dass man nach einem Jahr weniger als 1 050 € erhält, evtl. sogar noch weniger als die eingesetzten 1 000 €.

Dieses Beispiel kann man auch anhand der realen Daten aus *Beispiel 1.6* betrachten, in dem die Entwicklung Deutsche Bank Aktienkurses untersucht wurde. Hätte man am 02.01.2006 1 000 € in Aktien der Deutschen Bank zu einem Schlusskurs von 81.93 € investiert (also 12.21 Aktien gekauft), dann wäre das investierte Kapital bis zum 29.12.2006 auf 1 236.91 € gestiegen, da an diesem Tag der (nicht in Tabelle 1.4 angegebene) Schlusskurs bei 101.34 € lag. Hätte man dagegen erst am 02.01.2007 zugeschlagen und $1000/102.89 \approx 9.72$ Aktien erstanden, dann wäre der Aktienbesitz am 28.12.2007 bei einem Schlusskurs von 89.40 € gerade noch 868.89 € Wert gewesen (dabei wurde natürlich vernachlässigt, dass man nur ganze Stücke kaufen kann und bei Kauf und Verkauf noch Transaktionskosten anfallen).

Informationen dieser Art werden verwendet, um ein Modell wie in Abb. 1.23 zu erstellen, das einen bei der Anlageentscheidung unterstützt. Es kann aber auch im folgenden Jahr alles ganz anders kommen. Erst nach Ablauf des Jahres wird man wissen, was man hätte tun sollen. Die Zufallsvariable wird dann eine Realisation sein und damit einen konkreten Wert haben. Die Statistik kann zwar nicht sagen, welche Entscheidung die bessere ist, sie kann aber Methoden anbieten, mit deren Hilfe man das Risiko, das mit den verschiedenen Möglichkeiten verbunden ist, abwägen kann. Diese Methoden werden z.B. verwendet, um den sogenannten *Value-at-Risk* zu berechnen, der später noch näher erläutert wird.

Statistik für Bachelor- und Masterstudenten
Eine Einführung für Wirtschafts- und
Sozialwissenschaftler

Zucchini, W.; Schlegel, A.; Nenadic, O.; Sperlich, S.
2009, XII, 453 S., Softcover
ISBN: 978-3-540-88986-1