

Kapitel 2

Funktionen

Eine grundlegende Aufgabe in Forschung und Praxis besteht darin, den Zusammenhang zwischen verschiedenen Variablen zu beschreiben. Im einfachsten Fall geht es um Abhängigkeiten einer Variablen von einer zweiten Variablen. Das klassische mathematische Instrument zur Erfassung von Abhängigkeiten ist die Funktion. Es ist wichtig, mit einigen Grundfunktionen vertraut zu sein. Dazu gehören jedenfalls die lineare Funktion, die Potenzfunktion, quadratische Polynome, Exponential- und Logarithmusfunktionen sowie die Sinusfunktion. Dieses Kapitel führt in die Theorie und Anwendung der genannten Funktionen ein. Im Besonderen wird im Rahmen von Interpolations- und Regressionsaufgaben auch ein Anwendungsbezug durch Anpassung der Funktionen an vorgegebene Datenpunkte hergestellt. Ein weiterer inhaltlicher Schwerpunkt des Kapitels sind Wahrscheinlichkeits- und Dichtefunktionen. Konkret werden in diesen Abschnitten die Binomialverteilung und die hypergeometrische Verteilung sowie die Normalverteilung betrachtet.

2.1 Lineare Funktionen

2.1.1 Von der Beobachtung zur Funktion

Um zu erkennen, wie eine Variable von einer anderen abhängt, geht man in der Regel von einem Datenmaterial aus, das man durch Beobachtung der beiden Variablen an Untersuchungseinheiten gewonnen hat. Das Datenmaterial fasst man zuerst meist in Form einer **Wertetabelle** zusammen, wie dies beispielhaft in Abb. 2.1 geschehen ist. Die Untersuchungseinheiten stellen hier einen Baum zu aufeinanderfolgenden Zeitpunkten dar, in denen die Höhe gemessen wurde. Die Art der Abhängigkeit der Höhe h vom Alter a wird besonders deutlich, wenn man die einander entsprechenden Variablenwerte als Koordinaten von Punkten auffasst und diese in ein rechtwinkeliges Koordinatensystem einzeichnet. Dabei ist es üblich, wenn h in Abhängigkeit von a betrachtet wird, die a -Werte horizontal und die h -Werte vertikal aufzutragen. Die sich ergebende Punkteverteilung in der (a, h) -Ebene wird als **Streudiagramm** bezeichnet. Es erlaubt eine schnelle Beurteilung, in welcher Weise die betrachteten Variablen zusammenhängen. Oft lässt die Punk-

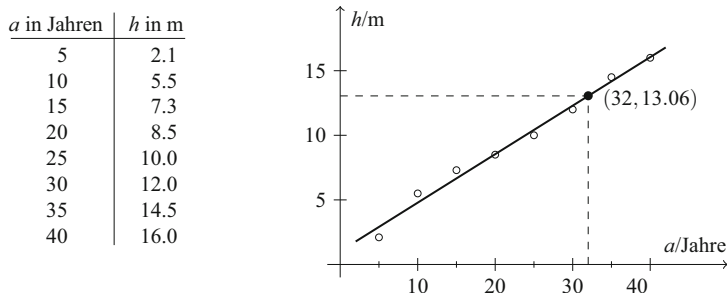


Abb. 2.1 Darstellung der Abhängigkeit der Höhe h (in m) eines Baumes von seinem Alter a (in Jahren) durch eine Wertetabelle und ein Streudiagramm. An die Datenpunkte wurde eine Ausgleichsgerade angepasst, mit der sich für das Alter $a = 32$ Jahre die erwartete Höhe $h = 13.06$ m ergibt

teverteilung im Streudiagramm einen durch eine einfache Kurve beschreibbaren „Trend“ erkennen. In unserem Beispiel handelt es sich offenkundig um einen geradlinigen Trend. Selbstverständlich wird die Punkteverteilung nicht exakt durch die Trendkurve erfasst; vielmehr wird i. Allg. eine „Reststreuung“ der Punkte um die Kurve herum übrigbleiben. Bei dem von uns betrachteten Beispiel geht diese Streuung darauf zurück, dass zwar die horizontal aufgetragenen Variablenwerte vom Beobachter kontrolliert, d. h., fest vorgegeben sind, die vertikal aufgetragene Variable jedoch unregelmäßig auf Grund irgendwelcher Zufallseinflüsse um einen „mittleren Wert“ $\hat{h}$ herum schwankt. Die Trendkurve kann dann so interpretiert werden, dass sie die Abhängigkeit dieser $\hat{h}$ -Werte von a ausdrückt.

Eine einfache Methode, die Trendkurve zu zeichnen, besteht darin, dass man zwischen den Messpunkten „per Augenmaß“ in h -Richtung „ausgleicht“. Dabei trachtet man, die in h -Richtung genommenen Abstände der Messpunkte von der Trendkurve in Summe möglichst klein zu halten. Statt Trendkurve wird meist die Bezeichnung **Ausgleichskurve** verwendet. Rechnerische Verfahren zur Bestimmung von speziellen Ausgleichskurven werden in den folgenden Abschnitten behandelt.

Mit Hilfe der Ausgleichsgeraden können wir im betrachteten Beispiel zu jedem vorgegebenen Wert a (aus einem gewissen Intervall D der Zahlengeraden) einen zugehörigen Durchschnittswert $\hat{h}$ für die Baumhöhe angeben. Über die Ausgleichskurve wird also jedem Wert a aus D (im Beispiel ist D das Intervall $5 \leq a \leq 40$) genau ein Wert $\hat{h}$ zugeordnet. Die allen a -Werten aus dem Intervall D zugeordneten $\hat{h}$ -Werte liegen wieder in einem bestimmten Intervall der reellen Achse, das mit W bezeichnet sei. Die durch die Ausgleichsgerade fixierte Zuordnung von Zahlen aus dem Intervall D zu Zahlen aus dem Intervall W nennt man eine **Ausgleichs- oder Regressionsfunktion**.

Allgemein liegt bei einer Funktion die folgende Situation vor: Gegeben ist eine Variable x , die in einer festgelegten **Definitionsmenge** D variiert. Über eine mit f bezeichnete **Zuordnungsvorschrift** wird jedem Element x von D genau ein „Bild-

element“ y zugeordnet, das einer bestimmten **Wertemenge** W angehört. In diesem Zusammenhang nennt man x auch die „unabhängige“ und y die „abhängige“ Variable. Für die Funktion schreibt man $f : D \rightarrow W$, statt „Bildelement von x “ sagt man auch „Funktionswert $f(x)$ an der Stelle x “. Die Elemente von D und W werden im folgenden stets als reelle Zahlen vorausgesetzt. Die Zuordnungsvorschrift f kann wie im betrachteten Beispiel durch eine Kurve in der (x, y) -Ebene festgelegt sein. Die Kurve nennt man dann den **Graphen** der Funktion. Oft gelingt es auch, die Zuordnung zwischen x und y durch eine Gleichung auszudrücken. Diese Gleichung heißt **Funktionsgleichung** und man schreibt dafür kurz $y = f(x)$.

2.1.2 Geradengleichungen

Im Folgenden wird zuerst die **lineare Funktion** näher betrachtet. Obwohl streng lineare Beziehungen zwischen zwei Variablen in den Biowissenschaften selten sind, spielt die lineare Funktion vor allem auch bei der Approximation von nicht linearen Zusammenhängen eine Rolle. Wir beginnen mit einer Zusammenstellung einiger von der Schule her bekannten Tatsachen und gehen von zwei Größen x und y aus. Diese wollen wir als metrisch voraussetzen. Der Definitionsbereich von x ist ein Intervall D der Zahlengeraden, ebenso der Bereich W , in dem y variiert. Durch die Funktion f wird jedem Wert x aus D ein Wert $y = f(x)$ aus W zugeordnet. Die Funktion f heißt linear, wenn der Funktionsgraph in der (x, y) -Ebene eine Gerade bzw. eine Strecke darstellt. Es seien $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ zwei beliebige auf der Geraden liegende nicht zusammenfallende Punkte. Für eine (nicht zur y -Achse parallele) Gerade ist charakteristisch, dass der aus den Differenzen $\Delta x = x_2 - x_1$ und $\Delta y = y_2 - y_1 = f(x_2) - f(x_1)$ gebildete **Differenzenquotient**

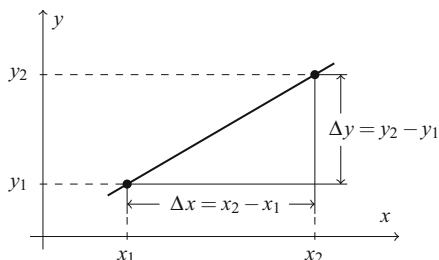
$$k = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \quad (2.1)$$

konstant ist. Man bezeichnet diesen Quotienten als den **Anstieg** der Geraden (vgl. Abb. 2.2). Ist $k > 0$, so steigt die Gerade, für $k < 0$ fällt sie und bei $k = 0$ verläuft die Gerade parallel zur x -Achse. Insbesondere gilt also für irgendeinen, auf der Geraden liegenden Punkt $P(x, y)$ mit $x \neq x_1$ und $x \neq x_2$ die Beziehung $(y - y_1)/(x - x_1) = (y_2 - y_1)/(x_2 - x_1)$. Durch Auflösen nach y folgt daraus die sogenannte **Zwei-Punktform** der Geradengleichung:

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1). \quad (2.2)$$

Von der Zwei-Punktform geht man zweckmäßigerweise bei der Lösung von **Interpolationsaufgaben** aus. Dabei handelt es sich um das folgende Problem: Von einer Funktion g sind an zwei Stellen x_1, x_2 die zugehörigen Funktionswerte $g(x_1), g(x_2)$ bekannt. Gesucht ist ein Näherungswert für den Funktionswert $g(x)$ an einer zwischen x_1 und x_2 liegenden Stelle x . Vom Graphen der Funktion g kennt man

Abb. 2.2 Geradenanstieg
als Differenzenquotient:
 $k = \Delta y / \Delta x$



also 2 Punkte, nämlich $Q_1(x_1, g(x_1))$ und $Q_2(x_2, g(x_2))$. Approximiert man den Graphen von g zwischen Q_1 und Q_2 durch die Strecke von Q_1 nach Q_2 , so spricht man von einer **linearen** Interpolation. Näherungsweise gilt:

$$g(x) \approx g(x_1) + \frac{g(x_2) - g(x_1)}{x_2 - x_1}(x - x_1).$$

Sucht man den Wert der Funktion g an einer Stelle außerhalb des durch x_1 und x_2 begrenzten Intervalls, so spricht man von **Extrapolation**.

Beispiel 2.1

Die CO_2 -Konzentration in der irdischen Atmosphäre belief sich im Jahre 1950 auf 306 ppm und im Jahre 1970 auf 321 ppm (1 ppm CO_2 = 1 cm³ CO_2 pro 1 m³ Luft). Welchen CO_2 -Gehalt erhält man für das Jahr 1965 bei linearer Interpolation, welcher Wert ergibt sich für 1980 bei linearer Extrapolation?

Wir bezeichnen mit $g(x)$ die CO_2 -Konzentration (in ppm) im Jahre x . Aus den Angaben entnimmt man $g(1950) = 306$ und $g(1970) = 321$. Folglich ist die CO_2 -Konzentration im Jahre $x = 1965$ bei linearer Interpolation gegeben durch

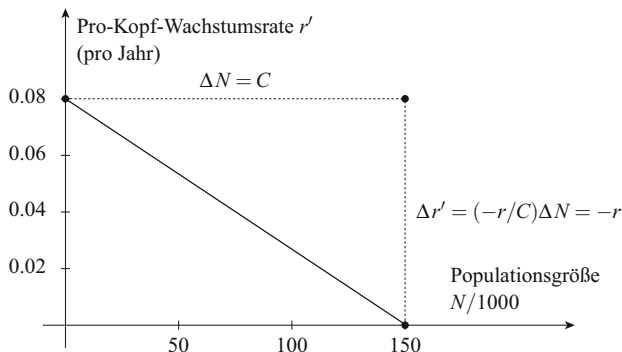
$$306 + \frac{321 - 306}{1970 - 1950}(1965 - 1950) = 317.25.$$

Für das Jahr 1980 ergibt sich bei linearer Extrapolation die CO_2 -Konzentration

$$306 + \frac{321 - 306}{1970 - 1950}(1980 - 1950) = 328.5.$$

```
> # Lösung mit R:
> # Definition der linearen Interpolationsfunktion
> g <- function(x, x1, y1, x2, y2){
+   y1 + (y2 - y1) / (x2 - x1) * (x - x1)
+ }
> # Aufruf der Funktion
> g.65 <- g(1965, 1950, 306, 1970, 321) # Interpolation
> g.80 <- g(1980, 1950, 306, 1970, 321) # Extrapolation
> print(cbind(g.65, g.80), digits=5)
      g.65  g.80
[1,] 317.25 328.5
```

Abb. 2.3 Wachstumsmodell mit linear mit wachsender Populationsgröße fallender pro Kopf-Wachstumsrate (Beispiel 2.2)



Wenn man in der Zwei-Punktform (2.2) für den Differenzenquotienten den Anstieg k schreibt, so erhält man die **Punkt-Richtungsform** $y = y_1 + k(x - x_1)$ der Geradengleichung. Setzt man schließlich $d = y_1 - kx_1$ folgt die **Hauptform** $y = kx + d$. Der Parameter d heißt **y-Achsenabschnitt** und stellt die y -Koordinate des Schnittpunktes der Geraden mit der y -Achse dar. Eine wichtige Grundaufgabe ist es, eine Gerade bei vorgegebenen Werten von k und d in der (x, y) -Ebene darzustellen.

Beispiel 2.2

Nach dem logistischen Wachstumsmodell ist die zeitliche Änderung R der Größe N einer Population durch $R = rN(1 - N/C)$ gegeben; r ist die natürliche Zuwachsrate und $C > 0$ die Kapazität des Lebensraumes. Dies bedeutet, dass die **pro Kopf-Wachstumsrate** $r' = R/N = r(1 - N/C)$ bei positivem r eine linear fallende Funktion der Populationsgröße ist.¹ In der (N, r') -Ebene wird also die Abhängigkeit der pro Kopf-Wachstumsrate von der Populationsgröße durch eine fallende Gerade dargestellt. Diese Gerade ist in Abb. 2.3 für eine hypothetische Population mit $r = 0.08$ p.a. und $C = 150\,000$ Individuen eingezeichnet. Aus der Hauptform $r' = (-r/C)N + r$ kann man die Geradenparameter $k = -r/C = -5.33 \cdot 10^{-7}$ p.a. und $d = r = 0.08$ p.a. ablesen. Neben dem durch den y -Achsenabschnitt d festgelegten Punkt $(0, r)$ erhält man einen zweiten Punkt der Geraden, indem man von $(0, r)$ ausgehend in (positiver) N -Richtung C Einheiten fortschreitet und dabei in (negativer) r' -Richtung r Einheiten abträgt. Man gelangt so zum Punkt $(C, 0)$. Damit ist die Gerade festgelegt.

Für $d = 0$ geht die Hauptform in $y = kx$ über und stellt die Gleichung einer durch den Koordinatenursprung verlaufenden Geraden dar. Sind zwei Größen x und y durch die Gleichung $y = kx$ mit $k \neq 0$ miteinander verbunden, nennt man y **proportional** zu x mit der Proportionalitätskonstanten k . Beispielsweise ist die Masse m eines (homogenen) Körpers proportional zum Volumen V . Man schreibt

¹ Aus dieser Forderung folgt – wie im Abschn. 4.5.2 für eine sich in der Zeit stetig veränderliche Wachstumsgröße gezeigt wird – die logistische Wachstumskurve, die auf den belgischen Mathematiker Pierre-Francois Verhulst (1804–1849) zurückgeht.

dafür kurz $m \sim V$ oder ausführlich $m = \rho V$ mit der Dichte ρ als Proportionalitätskonstanten.

Beispiel 2.3

Ein Tier (Länge L) laufe mit der Geschwindigkeit v eine Steigung hinauf. Hierzu muss es eine zu L^3 und v proportionale Leistung P aufbringen. Die maximale Muskelleistung $P_{\max}$ wird allgemein als zu L^2 proportional angenommen. Wir zeigen, dass die Maximalgeschwindigkeit $v_{\max}$ zu $1/L$ proportional ist (vgl. Smith, 1971).

Aus $P \sim v$ folgt, dass P/v gleich einer von v unabhängigen Größe ist, die aber von L abhängen kann. Für die Abhängigkeit von L ergibt sich aus $P \sim L^3$ (bei konstantem v) die Gleichung $P/v = k_1 L^3$ mit einer bestimmten Proportionalitätskonstanten k_1 . Die Maximalgeschwindigkeit $v_{\max}$ kann daraus durch $v_{\max} = P_{\max}/(k_1 L^3)$ ausgedrückt werden. Da $P_{\max} \sim L^2$, also $P_{\max} = k_2 L^2$ (k_2 bezeichnet wieder eine Proportionalitätskonstante), ergibt sich $v_{\max} = k_2/(k_1 L)$, d. h., $v_{\max} \sim L^{-1}$.

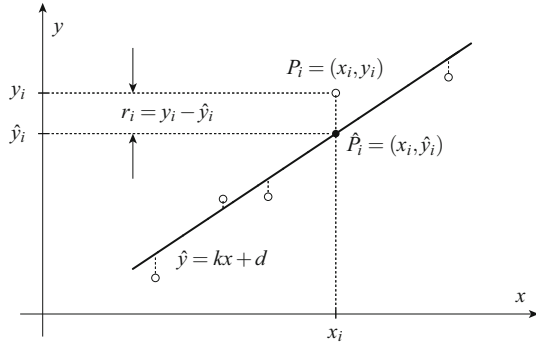
2.1.3 Regressionsgeraden

a) Problemstellung Auch wenn zwischen zwei Größen x und y eine lineare Beziehung besteht, so liegen die als Punkte P_i in einem (x, y) -Koordinatensystem dargestellten Paare (x_i, y_i) von Messwerten i. Allg. nicht auf einer Geraden. Infolge Messfehler werden nämlich die beobachteten Werte um die wahren Werte mehr oder weniger stark streuen. Die bei fehlerfreier Messung linear angeordneten Messpunkte P_i zerfließen so in eine mehr oder weniger verbreiterte „Punktwolke“. Im Folgenden wird gezeigt, wie man eine bestehende lineare Abhängigkeit zwischen zwei Größen x und y aus vorgegebenen x -Werten und (durch Messfehler verfälschten) y -Werten rekonstruieren kann.

Eng mit dieser Problematik verknüpft ist eine andere. In vielen Fällen lässt sich die Abhängigkeit einer Größe von einer anderen überhaupt nicht durch eine Funktion darstellen (auch dann nicht, wenn man alle Messungenauigkeiten ausschalten könnte). Der in Abschn. 2.1 betrachtete Zusammenhang zwischen der Baumhöhe und dem Alter ist von dieser Art. Es ist prinzipiell unmöglich, zu einem vorgegebenen Alter die Höhe eines bestimmten Baumes vorherzusagen. Dies hat seine Ursache darin, dass die Baumhöhe zufallsbedingten (und daher nicht vorhersagbaren) Schwankungen um einen mittleren Wert unterliegt. Wohl drückt das Streudiagramm in Abb. 2.1 eine Abhängigkeit aus, die aber viel „loser“ als bei einer funktionellen Beziehung ist. Passt man eine Gerade an die Punkteverteilung an, so gestattet diese Gerade nicht mehr den Schluss vom Alter auf die Baumhöhe im Einzelfall, sondern nur mehr auf die zu erwartende durchschnittliche Baumhöhe.

Sieht man von der Interpretation des Ergebnisses ab, so geht es – geometrisch ausgedrückt – in den beiden genannten Fällen um folgende Aufgabe: In der (x, y) -Ebene ist eine Gerade „möglichst gut“ an eine bestimmte Anzahl n von vorgegebenen Punkten $P_i = (x_i, y_i)$ ($i = 1, 2, \dots, n$) anzupassen. Die Gerade kann

Abb. 2.4 Zur Methode der kleinsten Quadrate: Die Parameter k und d der an die Datenpunkte anzupassenden Geraden werden aus der Forderung bestimmt, dass die Quadratsumme der Residuen r_i möglichst klein wird



man sich z. B. in der Hauptform $y = kx + d$ angesetzt denken mit zunächst noch unbestimmten Parametern k und d . Für jeden beobachteten x_i -Wert lässt sich dann die y -Koordinate des auf der Geraden über der Stelle x_i liegenden Punktes $\hat{P}_i = (x_i, \hat{y}_i)$ durch $\hat{y}_i = kx_i + d$ ausdrücken (vgl. Abb. 2.4). Im Allgemeinen werden die errechneten Werte $\hat{y}_i$ von den beobachteten y_i -Werten abweichen. Es ist naheliegend, die **Residuen** (Abweichungen) $r_i = y_i - \hat{y}_i$ als Grundlage für die rechnerische Bestimmung der Parameter k und d heranzuziehen. Dies ist allerdings nur dann gerechtfertigt, wenn die Werte der in x -Richtung aufgetragenen Größe als vom Beobachter fixiert betrachtet werden können und nur die in y -Richtung aufgetragene Größe Messfehlern bzw. zufälligen Einflüssen unterliegt und diese Einflüsse voneinander unabhängig wirken.

b) Schätzung der Geradenparameter Um die Geradenparameter zu berechnen, benötigen wir ein Maß für die Güte der Anpassung der Geraden an die Datenpunkte. Die von Carl F. Gauss (1777–1855) vorgeschlagene **Methode der kleinsten Quadrate** verwendet als Gütemaß die Summe

$$Q(k, d) = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - kx_i - d)^2 \quad (2.3)$$

der Quadrate der Residuen $r_i = y_i - \hat{y}_i$. Die Parameter k und d der gesuchten Regressionsgeraden werden aus der Forderung bestimmt, dass Q möglichst klein ist. Man bezeichnet die so bestimmten Parameter als Kleinste Quadrate-Schätzwerte von k bzw. d . Wie in Abschn. 2.7.1 gezeigt wird, nimmt Q für

$$k = \hat{k} = \frac{s_{xy}}{s_x^2} \quad \text{und} \quad d = \hat{d} = \bar{y} - \hat{k}\bar{x} \quad (2.4)$$

den kleinsten Wert

$$S_{rr} = (n-1)s_y^2 \left[1 - \left(\frac{s_{xy}}{s_x s_y} \right)^2 \right] \quad (2.5)$$

an. In den Formeln (2.4) und (2.5) bedeuten $\bar{x}$ und $\bar{y}$ die arithmetischen Mittel der x - bzw. y -Werte,

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{und} \quad s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

die entsprechenden (empirischen) Varianzen und

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (2.6)$$

die (empirische) Kovarianz der beiden Beobachtungsreihen.² Die Gleichung der nach der Methode der kleinsten Quadrate bestimmten **Regressionsgeraden** ist daher durch

$$\hat{y} = \frac{s_{xy}}{s_x^2}(x - \bar{x}) + \bar{y} \quad (2.7)$$

gegeben. Statt y wurde links $\hat{y}$ geschrieben, weil die mit dieser Gleichung erhaltenen Funktionswerte als Schätzwerte für die zu erwartenden y -Werte zu interpretieren sind. Man beachte beim Einzeichnen der Regressionsgeraden in das Streudiagramm, dass jene stets durch das Zentroid $(\bar{x}, \bar{y})$ der Punkteverteilung hindurchgeht. Setzt man $S_{xx} = (n-1)s_x^2$ und $S_{xy} = (n-1)s_{xy}$, kann man den Anstieg der Regressionsgeraden auch durch $\hat{k} = S_{xy}/S_{xx}$ ausdrücken.

c) Bestimmtheitsmaß Die Güte der Anpassung der Regressionsgeraden an die Punkte $P_i = (x_i, y_i)$ des Streudiagramms wird numerisch durch das Bestimmtheitsmaß

$$B = 1 - \frac{S_{rr}}{(n-1)s_y^2} = \left(\frac{s_{xy}}{s_x s_y} \right)^2 \quad (2.8a)$$

beurteilt, das mit dem Quadrat der Produktmomentkorrelation $r_{xy} = s_{xy}/(s_x s_y)$ übereinstimmt. Offensichtlich ist B nicht negativ und stets kleiner oder gleich 1. Im Falle $B = 1$ ist $S_{rr} = 0$, d. h., alle auszugleichenden Punkte (x_i, y_i) liegen auf der Regressionsgeraden. Im anderen Extremfall $B = 0$ ist $S_{rr}/(n-1)$ gleich der Varianz s_y^2 der y_i -Werte; die Regressionsgerade verläuft wegen $s_{xy} = 0$ parallel zur x -Achse und die Variation von y kann durch x überhaupt nicht erklärt werden. Zur Interpretation von zwischen 0 und 1 liegenden Werten des Bestimmtheitsmaßes

² Die Kovarianz ist ein Maß für die gemeinsame Variation zweier verbundener Stichproben $x_1, x_2, \dots, x_n$ und $y_1, y_2, \dots, y_n$. Zwei Stichproben heißen verbunden oder abhängig, wenn x_1 und y_1 die Beobachtungswerte zweier Merkmale an der ersten Untersuchungseinheit darstellen, x_2 und y_2 die entsprechenden Beobachtungswerte an der zweiten Untersuchungseinheit usw. sind.

bringen wir die Formel für B auf die Gestalt $(n-1)s_y^2 = S_{rr} + (n-1)s_y^2 B$, setzen $B = r_{xy}^2$ ein und nehmen folgende Umformungen vor:

$$\begin{aligned} S_{yy} &= (n-1)s_y^2 = S_{rr} + (n-1)s_x^2 \left(\frac{s_{xy}}{s_x^2} \right)^2 \\ &= S_{rr} + \sum_{i=1}^n (x_i - \bar{x})^2 \left(\frac{s_{xy}}{s_x^2} \right)^2 = S_{rr} + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2. \end{aligned}$$

Im letzten Umformungsschritt wurde die Gleichung (2.7) der Regressionsgeraden verwendet. In Analogie zu den Bezeichnungen S_{yy} und S_{rr} führen wir die Abkürzung³

$$S_{\hat{y}\hat{y}} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

für die Quadratsumme der durch die Regressionsgerade erklärten Abweichungen vom Mittelwert $\bar{y}$ ein. Dann lässt sich der Zusammenhang zwischen den drei Quadratsummen kurz durch die Formel $S_{yy} = S_{rr} + S_{\hat{y}\hat{y}}$ ausdrücken. Für das Bestimmtheitsmaß ergibt sich nun die Darstellung:

$$B = 1 - \frac{S_{rr}}{S_{yy}} = \frac{S_{\hat{y}\hat{y}}}{S_{yy}}. \quad (2.8b)$$

Danach ist B als jener Anteil der Gesamtvariation S_{yy} zu deuten, der – über die Regressionsgerade – durch x erklärt werden kann.

Beispiel 2.3

Wir bestimmen mit die Wachstumsdaten in Abb. 2.1 die Parameter k (Anstieg) und d (y -Achsenabschnitt) der Regressionsgeraden sowie den Anteil (in %) der durch das Alter erklärten Variation der Höhenwerte.

Der Stichprobenumfang und die Mittelwerte der beiden Datenreihen sind $n = 8$, $\bar{a} = 180/8 = 22.5$, $\bar{h} = 75.9/8 = 9.4875$. Subtrahiert man von jedem Einzelwert den entsprechenden Mittelwert, erhält man die zentrierten Datenreihen, die in der dritten und vierten Spalte von Tab. 2.1 eingetragen sind. Die weiteren

³ Man beachte, dass das arithmetische Mittel $\bar{\hat{y}}$ der $\hat{y}_i$ -Werte gleich dem arithmetischen Mittel $\bar{y}$ der beobachteten y -Werte ist. Wegen $\hat{y}_i = \bar{y} + k(x_i - \bar{x})$ und $\sum_{i=1}^n (x_i - \bar{x}) = 0$ ist nämlich

$$\begin{aligned} n\bar{\hat{y}} &= \sum_{i=1}^n \hat{y}_i = \sum_{i=1}^n (\bar{y} + k(x_i - \bar{x})) \\ &= n\bar{y} + k \sum_{i=1}^n (x_i - \bar{x}) = n\bar{y}. \end{aligned}$$

Tab. 2.1 Rechenschema zu Beispiel 2.3: Die Tabellenwerte sind exakt bzw. auf 4 Nachkommastellen gerundet; die letzte Zeile enthält die Spaltensummen (auf Grund von Rundungsfehlern können die Werte in der letzten Zeile von der Summe der angeschriebenen Tabellenwerte abweichen)

a_i	h_i	$a_i - \bar{a}$	$h_i - \bar{h}$	$(a_i - \bar{a})^2$	$(h_i - \bar{h})^2$	$(a_i - \bar{a})(h_i - \bar{h})$
5	2.1	-17.5	-7.3875	306.25	54.5752	129.2813
10	5.5	-12.5	-3.9875	156.25	15.9002	49.8438
15	7.3	-7.5	-2.1875	56.25	4.7852	16.4063
20	8.5	-2.5	-0.9875	6.25	0.9752	2.4688
25	10.0	2.5	0.5125	6.25	0.2627	1.2813
30	12.0	7.5	2.5125	56.25	6.3127	18.8438
35	14.5	12.5	5.0125	156.25	25.1252	62.6563
40	16.0	17.5	6.5125	306.25	42.4127	113.9688
180	75.9			1050.00	150.3487	394.7500

Spalten enthalten die Quadrate der zentrierten a_i - bzw. h_i -Werte sowie deren Produkte. Die entsprechenden Spaltensummen sind $S_{aa} = 1050$, $S_{hh} = 150.3487$ bzw. $S_{ah} = 394.75$. Damit ergeben sich die Varianzen $s_a^2 = S_{aa}/(n-1) = 150$ und $s_h^2 = S_{hh}/(n-1) = 21.48$ und die Kovarianz $s_{ah} = S_{ah}/(n-1) = 56.39$. Die Geradenparameter sind $\hat{k} = s_{ah}/s_a^2 = 0.376$ und $\hat{d} = \bar{h} - k\bar{a} = 1.029$. Das Bestimmtheitsmaß ist $B = s_{ah}^2/(s_a^2 s_h^2) = 0.987 = 98.7\%$.

```
> # Lösung mit R:
> a <- seq(from=5, to=40, by=5)
> h <- c(2.1, 5.5, 7.3, 8.5, 10, 12, 14.5, 16)
> aquer <- mean(a); hquer <- mean(h) # arithmetische Mittel
> vara <- var(a); varh <- var(h)      # Varianzen
> covah <- cov(a, h)                 # Kovarianz
> k <- covah/vara; d <- hquer - k*aquer # Geradenparameter
> B <- covah^2/vara/varh # Bestimmtheitsmaß
> print(cbind(k, d, B), digits=4)
      k      d      B
[1,] 0.376 1.029 0.9871
> # Lösung mit der R-Funktion lm()
> linreg <- lm(h ~ a)
> par <- coefficients(linreg)
> d <- par[[1]]; k <- par[[2]]
> B <- summary(linreg)$r.squared
> print(cbind(k, d, B), digits=4)
      k      d      B
[1,] 0.376 1.029 0.9871
```

Die Regressionsgerade wurde bereits in Abb. 2.1 zusammen mit den Originaldaten dargestellt. Sowohl die Zeichnung als auch der nahe bei 1 liegende Wert des Bestimmtheitsmaßes lassen erkennen, dass die Abhängigkeit der (mittleren) Baumhöhe vom Alter innerhalb des betrachteten Altersintervalls recht gut durch die gefundene lineare Regressionsfunktion beschrieben werden kann.

2.1.4 Aufgaben

1. Der Viehbestand y (in Millionen Tieren) stieg in Westeuropa von $y_1 = 82$ im Jahre $t_1 = 1955$ auf $y_2 = 100.6$ im Jahre $t_2 = 1976$.
 - a) Wie lautet die Gleichung der Interpolationsgeraden?
 - b) Welche Stückzahl ergibt sich für das Jahr 2000 bei linearer Extrapolation?
2. Es bezeichne y die Temperatur in Grad Fahrenheit ($^{\circ}\text{F}$) und x die Temperatur in Grad Celsius ($^{\circ}\text{C}$). Der Gefrierpunkt des Wassers ist durch $y = 32^{\circ}\text{F}$ bzw. $y = 0^{\circ}\text{C}$ charakterisiert und der Siedepunkt durch $y = 212^{\circ}\text{F}$ bzw. $y = 100^{\circ}\text{C}$. Man bestimme die Temperatur in $^{\circ}\text{F}$, wenn $x = 36^{\circ}\text{C}$ ist.⁴
3. Bei Seefischen wurden von 1956 bis 1970 weltweit die folgenden Jahresfänge (in Millionen Tonnen) berichtet: 22.7, 22.8, 24.1, 26.8, 29.2, 32.2, 35.6, 36.4, 40.9, 39.6, 43.0, 45.9, 48.7, 47.2, 52.7. Man stelle das pro Jahr durchschnittlich zu erwartende Fangergebnis $\hat{y}$ durch eine lineare Regressionsfunktion in Abhängigkeit von der Zeit t dar (den Nullpunkt der Zeitskala lege man in das Jahr 1950). Welcher Prozentsatz der Gesamtvariation von y kann durch das lineare Modell erklärt werden?
4. Von 1978 bis 1989 wurden in der Donau (km 2060) die folgenden KMnO_4 -Jahresmittelwerte (in mg/l) gemessen: 20, 21, 21, 17.2, 16, 17.4, 18, 18, 18.1, 17.6, 15.2, 12.6. Man bestimme mittels linearer Regression die mittlere Änderung des KMnO_4 -Jahresmittels pro Jahr über den Beobachtungszeitraum.

2.2 Potenzfunktionen

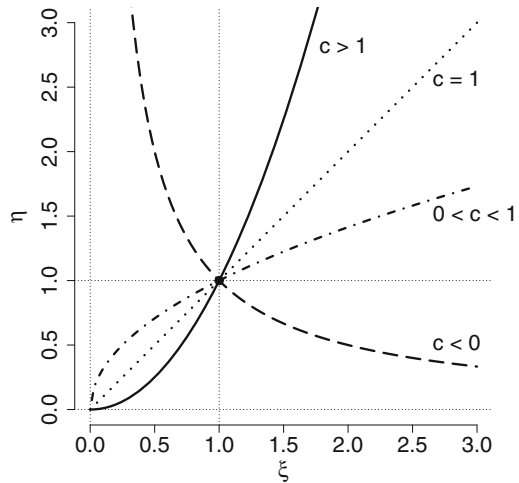
2.2.1 Allometrie

Um die Ausbildung der spezifischen Form eines Organismus quantitativ zu beschreiben, ist es notwendig, die Wachstumsintensitäten in den verschiedenen Dimensionen zu erfassen. Man denke etwa an die Messung der Länge und der Breite von Blättern. Dabei stellt sich die Frage, wie das Längen- und Breitenwachstum miteinander verknüpft sind. Untersuchungen dieser Art gehören in den Bereich der Allometrie. Diese beschäftigt sich mit der Beschreibung des Zusammenhangs zwischen verschiedenen morphometrischen oder physiologischen Größen (wie z. B. Masse, Länge oder Stoffwechselrate). Dabei können sich die Größen sowohl auf Teile eines biologischen Systems als auch auf das Gesamtsystem beziehen.

Wir bezeichnen mit x und y zwei Wachstumsgrößen (z. B. ein Längen- und ein Breitenmaß) eines biologischen Systems. Innerhalb des Zeitintervalls von t bis $t + \Delta t$ möge sich x um Δx und y um Δy verändern. Der Quotient $\Delta x / \Delta t$ stellt die mittlere Änderung pro Zeiteinheit im betrachteten Zeitintervall dar. Ist Δt sehr klein, so ist die mittlere Änderung pro Zeiteinheit geeignet, die Wachs-

⁴Die Temperaturskalen sind nach Daniel G. Fahrenheit (1686–1736, deutscher Physiker) bzw. Anders Celsius (1701–1744, schwedischer Astronom und Physiker) benannt.

Abb. 2.5 Graphen von typischen Potenzfunktionen mit der Gleichung $\eta = \xi^c$ für $\xi > 0$. Alle Graphen verlaufen durch den Punkt $(1, 1)$



tumsintensität in der „Umgebung“ des Zeitpunktes t zu beschreiben: Die mittlere Änderung pro Zeiteinheit geht mit abnehmendem Δt in die Änderung pro Zeiteinheit zum Zeitpunkt t über. Den Quotienten aus der Änderung pro Zeiteinheit und dem jeweiligen Wert der betrachteten Größe nennt man **Wachstumsrate**. So ist $r_x = (\Delta x / \Delta t) / x$ die Wachstumsrate von x (betrachtet in der Umgebung des Zeitpunktes t) und $r_y = (\Delta y / \Delta t) / y$ die Wachstumsrate der Größe y .

Wenn das Verhältnis der Wachstumsraten von y und x gleich einer (von der Zeit unabhängigen) Konstanten c ist, so lässt sich zeigen, dass die Abhängigkeit der Größe y von der Größe x durch die **allometrische Funktion** mit der Gleichung

$$y = bx^c \quad (2.9)$$

beschrieben werden kann, in der b eine positive Konstante ist. Sind x_0 und y_0 zwei zusammengehörende (d. h. zum selben Zeitpunkt gehörende) Werte von x und y , so lässt sich die allometrische Funktion nach Übergang zu den Größen $\xi = x/x_0$ und $\eta = y/y_0$ in der dimensionslosen Form $\eta = \xi^c$ anschreiben. Auf Grund ihrer biologischen Bedeutung sind sowohl x als auch y (daher auch ξ und η) nicht negativ. Die durch die Gleichung $\eta = f(\xi) = \xi^c$ für $\xi > 0$ definierte Funktion heißt **Potenzfunktion** (bei positivem c gehört auch $\xi = 0$ zum Definitionsbereich); sie ist in Abb. 2.5 für verschiedene Werte von c dargestellt. Wenn $c > 0$ ist (positive Allometrie), erhält man **monoton wachsende** Funktionsverläufe, d. h., für beliebige Werte $\xi_1 \geq 0$, $\xi_2 \geq 0$ folgt aus $\xi_2 > \xi_1$ stets $f(\xi_2) \geq f(\xi_1)$. Dagegen ist die Potenzfunktion für $c < 0$ (negative Allometrie) **monoton fallend**; für beliebige $\xi_1 > 0$, $\xi_2 > 0$ folgt nun aus $\xi_2 > \xi_1$ nämlich stets $f(\xi_2) \leq f(\xi_1)$.

Beispiel 2.6

- a) Die Abhängigkeit der Wärmeproduktion W (in kJ pro Tag) vom Körpergewicht M (in kg) wird bei Warmblütern in guter Näherung durch $W = bM^c$ erfasst.

Die Parameter b und c kann man bestimmen, wenn zwei zusammengehörende M - und W -Werte bekannt sind. Gibt man z. B. $M_1 = 500$, $W_2 = 28\,000$ und $M_2 = 0.5$, $W_1 = 175$ vor (diese Angaben entsprechen den Daten eines Pferdes bzw. eines Meerschweinchens), so genügen die Funktionsparameter den Gleichungen $28\,000 = b \cdot 500^c$ und $175 = b \cdot 0.5^c$. Durch Division folgt

$$\frac{28\,000}{175} = \frac{b \cdot 500^c}{b \cdot 0.5^c} = \left(\frac{500}{0.5}\right)^c = 1000^c$$

und schließlich $\ln 160 = \ln 1000^c = c \ln 1000$, d. h. $c = (\ln 160)/(\ln 1000) = 0.7347$ und $b = 175/0.5^c = 291.2$. Für einen Menschen mit 70 kg Gewicht ergibt sich daraus eine Stoffwechselrate von rund 6604 kJ pro Tag.

- b) Allometrische Funktionen treten auch in der Biogeographie auf. So wird zwischen der Fläche A eines Lebensraumes und der Anzahl S der Arten einer taxonomischen Einheit die Arten-Areal-Beziehung $S = bA^c$ angenommen (vgl. z. B. Wilson & Bossert 1973). Kennt man von einem Bezugsareal die Fläche A_0 und die Artenanzahl S_0 , so kann die Arten-Areal-Funktion in der dimensionslosen Form $S = S_0(A/A_0)^c$ angeschrieben werden. Für die Reptilien- und Amphibienfauna auf den Westindischen Inseln gilt beispielsweise $S = S_0(A/A_0)^{0.3}$. Wählt man die Insel Montserrat als Bezugsareal, so ist $A_0 = 106 \text{ km}^2$ und $S_0 = 12$. Damit errechnet man z. B. für Puerto Rico mit $A = 8897 \text{ km}^2$ die Artenanzahl $S = 12(8897/106)^{0.3} \approx 45$.

2.2.2 Linearisierung durch doppelt-logarithmische Transformation

Durch Logarithmieren geht die zwischen den Variablen x und y bestehende allometrische Beziehung $y = bx^c$ mit $b > 0$ in $\log y = \log(bx^c) = \log b + \log x^c = \log b + c \log x$ über (log steht z. B. für lg oder ln). Diese Gleichung bringt zum Ausdruck, dass zwischen den durch logarithmische Transformation aus den Originalvariablen x , y gewonnenen Variablen $\eta = \log y$, $\xi = \log x$ die lineare Abhängigkeit $\eta = c\xi + d$ mit $d = \log b$ besteht. Man spricht von einer **doppelt-logarithmischen** Variablentransformation (kurz log/log-Transformation), durch die allometrische Zusammenhänge in lineare übergeführt werden können.

Die doppelt-logarithmische Transformation ermöglicht also die Darstellung eines bekannten allometrischen Zusammenhanges mit Hilfe der Geradengleichung. In der Praxis sind die Funktionsparameter b und c meist unbekannt und müssen im Zuge einer Regressionsanalyse ermittelt werden. Auch dabei leistet die log/log-Transformation gute Dienste; nach Übergang zu den Hilfsvariablen $\xi = \log x$ und $\eta = \log y$ wird die Kurvenanpassung auf die in Abschn. 2.1.3 behandelte Ermittlung einer Regressionsgeraden zurückgeführt.

Wir nehmen an, dass zwischen den Größen x und y der allometrische Zusammenhang $y = bx^c$ bestehe. Ferner möge von x und y ein Datenmaterial in Form

einer Wertetabelle aus insgesamt n zusammengehörenden x_i - und y_i -Werten vorliegen (der Index i dient zur Unterscheidung der n Wertepaare). Um von einem vorgegebenen x -Wert auf den (durchschnittlichen) y -Wert $\hat{y}$ schließen zu können, müssen die Funktionsparameter b und c geschätzt werden. Dabei geht man folgendermaßen vor:

- Aus gegebenen x_i - und y_i -Werten werden durch Logarithmieren die Werte ξ_i bzw. η_i der Hilfsvariablen $\xi = \log x$ bzw. $\eta = \log y$ ermittelt.
- In der (ξ, η) -Ebene wird die Regressionsgerade $\hat{\eta} = \hat{k}\xi + \hat{d}$ bestimmt.
- Aus den Schätzwerten $\hat{k}$ und $\hat{d}$ folgen für die gesuchten Parameter c und b die Schätzwerte $\hat{c} = \hat{k}$ und $\hat{b} = a^{\hat{d}}$ (a ist die Basis des verwendeten Logarithmen-systems). Zur Beurteilung der Güte der Anpassung kann wieder das (mit den Originaldaten aus $B = 1 - S_{rr}/S_{yy}$ errechnete) Bestimmtheitsmaß verwendet werden.

Beispiel 2.7

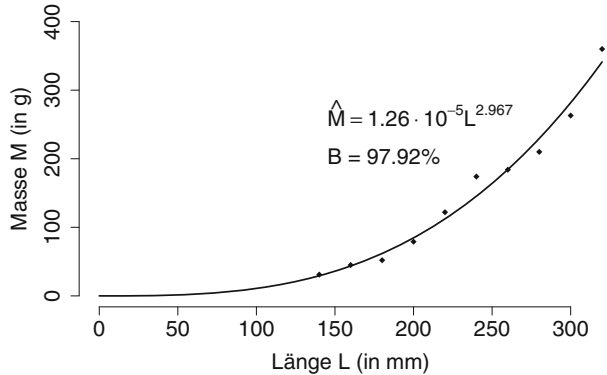
Von 10 Bachforellen (*Salmo trutta forma fario*) wurden die Masse M (in g) und die Länge L (in mm) bestimmt. Die Darstellung der in der ersten und zweiten Spalte von Tab. 2.2 angegebenen Wertepaare als Punkte in der (L, M) -Ebene lässt eine Abhängigkeit nach der allometrischen Formel $M = bL^c$ vermuten. Es sollen die Funktionsparameter b und c geschätzt und die Güte der Regressionsfunktion mit Hilfe des Bestimmtheitsmaßes beurteilt werden.

- Zuerst werden die Werte der Hilfsvariablen mit den Transformationsgleichungen $\xi = \ln L$, $\eta = \ln M$ berechnet. Diese Werte sind in Tab. 2.2 angeführt. Zusätzlich enthält die Tabelle alle für die lineare Regression und die Ermittlung des Bestimmtheitsmaßes notwendigen Zwischenergebnisse. Der Stichprobenumfang ist $n = 10$.
- Für die lineare Regressionsanalyse in der (ξ, η) -Ebene werden die arithmetischen Mittel $\bar{\xi} = 5.405$ und $\bar{\eta} = 4.754$, die Varianz $s_{\xi}^2 = 0.0760$ der ξ -Werte

Tab. 2.2 Daten und Rechenschema zu Beispiel 2.7. Die ξ -Spalte und η -Spalte enthalten die natürlichen Logarithmen der Originalvariablen L bzw. M

L	M	ξ	η	$(\xi - \bar{\xi})^2$	$(\xi - \bar{\xi})(\eta - \bar{\eta})$	$\hat{M} = bL^c$	$(M - \hat{M})^2$	$(M - \bar{M})^2$
140	31	4.94	3.43	0.2147	0.6119	29.37	2.67	14 641
160	45	5.08	3.81	0.1088	0.3126	43.64	1.85	11 449
180	52	5.19	3.95	0.0450	0.1703	61.89	97.88	10 000
200	79	5.30	4.37	0.0114	0.0411	84.60	31.42	5329
220	122	5.39	4.80	0.0001	-0.0006	112.25	95.01	900
240	174	5.48	5.16	0.0057	0.0306	145.31	822.92	484
260	184	5.56	5.21	0.0242	0.0717	184.26	0.07	1024
280	210	5.63	5.35	0.0528	0.1362	229.57	383.06	3364
300	263	5.70	5.57	0.0893	0.2443	281.72	350.30	12 321
320	360	5.77	5.89	0.1320	0.4112	341.17	354.73	43 264
Σ		54.05	47.54	0.6840	2.0292		2139.90	102 776

Abb. 2.6 Anpassung einer allometrischen Funktion an die Datenpunkte von Beispiel 2.7. Der Funktionsgleichung $\hat{M} = 1.26 \cdot 10^{-5} L^{2.967}$ liegt die lineare Regressionsgleichung $\hat{\eta} = 2.967\xi - 11.2807$ mit $\xi = \ln L$ und $\eta = \ln M$ zugrunde



sowie die Kovarianz $s_{\xi\eta} = 0.2255$ benötigt. Die genannten Statistiken können mit den entsprechenden Spaltensummen aus Tab. 2.2 berechnet werden. Als Anstieg der Regressionsgeraden ergibt sich $\hat{k} = s_{\xi\eta}/s_{\xi}^2 = 2.967$. Damit berechnet man die Gleichung $\hat{\eta} = k(\xi - \bar{\xi}) + \bar{\eta} = 2.967\xi - 11.2807$ der Regressionsgeraden.

- c) Die Gleichung zur Bestimmung der zu vorgegebenen Werten von L zu erwartenden (durchschnittlichen) Werte $\hat{M}$ von M lautet

$$\hat{M} = e^{\hat{\eta}} = e^{2.967\xi - 11.2807} = (e^{\ln L})^{2.967} e^{-11.2807} = 1.26 \cdot 10^{-5} L^{2.967}.$$

Damit wurden die Werte $\hat{M}_i$ in der siebten Spalte von Tab. 2.2 bestimmt. Zur visuellen Beurteilung der Anpassungsgüte der berechneten allometrischen Funktion stellt man diese in dem mit den Originalvariablen gezeichneten Streudiagramm dar (vgl. Abb. 2.6). Die numerische Beschreibung der Anpassungsgüte erfolgt mit dem Bestimmtheitsmaß, das nun mit der Formel $B = 1 - S_{rr}/S_{MM}$ zu berechnen ist. In dieser Formel bezeichnen

$$S_{rr} = \sum_{i=1}^n (M_i - \hat{M}_i)^2 \quad \text{und} \quad S_{MM} = \sum_{i=1}^n (M_i - \bar{M})^2$$

die Summe der Quadrate der Abweichungen (Residuen) der beobachteten M -Werte von den geschätzten bzw. die Summe der Quadrate der Abweichungen der beobachteten M -Werte vom Mittelwert $\bar{M}$. Diese Summen werden mit Hilfe der beiden letzten Spalten in Tab. 2.2 bestimmt. Man erhält mit $\bar{M} = 152$, $S_{rr} = 2139.90$ und $S_{MM} = 102776$ schließlich $B = 0.9792 = 97.92\%$.

```
> # Lösung mit R:
> L <- c(140, 160, 180, 200, 220, 240, 260, 280, 300, 320)
> M <- c(31, 45, 52, 79, 122, 174, 184, 210, 263, 360)
> n <- length(L)
> # log/log-Transformation und lineare Regression
> xi <- log(L); eta <- log(M)
> linreg <- lm(eta ~ xi)
```

```

> par <- coefficients(linreg)
> k <- c <- par[[2]]; d <- par[[1]]; b <- exp(d)
> print(cbind(k, c, d, b), digits=4)
      k      c      d      b
[1,] 2.967 2.967 -11.28 1.261e-05
> # Bestimmtheitsmaß
> Mexp <- exp(k*xi+d) # erwartete M_Werte
> Srr <- sum((M-Mexp)^2); SMM <- (n-1)*var(M); B=1-Srr/SMM
> print(cbind(Srr, SMM, B), digits=4)
      Srr      SMM      B
[1,] 2140 102776 0.9792
> # Abb. 2.6 (Streudiagramm mit Regressionsfunktion)
> par(mai=c(1.2, 1.2, 0.2, 0.2), cex.axis=1.3, cex.lab=1.3)
> plot(L, M, type="p", col="black", xlab="Länge L (in mm)",
+      ylab="Masse M (in g)", pch=18, frame.plot=F,
+      xlim=c(0, 320), ylim=c(0, 400), lwd=2)
> curve(b*x^c, lty=1, lwd=2, ad=T)
> text(140, 260, pos=4, cex=1.3,
+      expression(hat(M)=1.26 %.% 10^{ -5}*L^{ 2.967}))
> text(140, 200, expression("B = 97.92%"), pos=4, cex=1.3)

```

Das eben behandelte Beispiel stellt eine **nicht-lineare** Regressionsaufgabe dar, die durch eine log/log-Transformation in eine lineare übergeführt werden konnte. In diesem Zusammenhang beachte man: Für die mit Hilfe von linearisierenden Transformationen bestimmte Ausgleichsfunktion ist i. Allg. die mit den Originalvariablen gebildete Summe der Abweichungsquadrate nicht minimal. Diese Tatsache ist für die Praxis jedoch bedeutungslos, solange – wie im Beispiel 2.7 – S_{rr}/S_{MM} klein bleibt.

2.2.3 Gebrochen lineare Funktionen

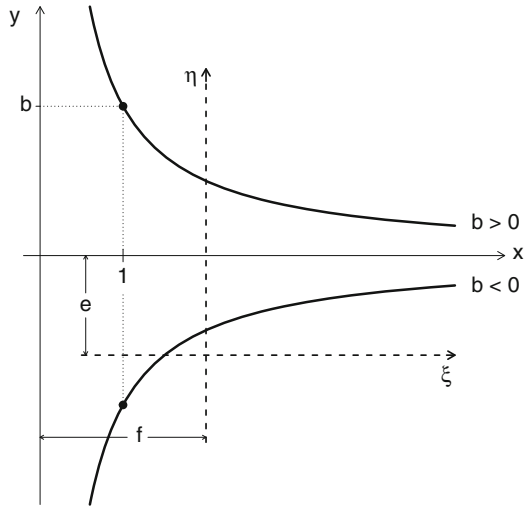
Besondere Beachtung verdient auch der Sonderfall $c = -1$ der allometrischen Funktion. Besteht zwischen zwei Größen x und y der Zusammenhang

$$y = bx^{-1} = \frac{b}{x}, \quad (2.10)$$

so bezeichnet man die beiden Größen als **indirekt proportional**. Im Gegensatz zur bisherigen Verabredung möge nun b eine positive oder negative Konstante sein. Je nach dem Vorzeichen von b sind dann die in Abb. 2.7 gezeichneten Funktionsgraphen zu unterscheiden, wobei wir uns auf positive x -Werte beschränken. Wenn $b > 0$ ist, fallen mit wachsendem x die Funktionswerte monoton gegen den „Grenzwert“ null; ist $b < 0$, so steigen die Funktionswerte monoton gegen null. Auf Grund dieses Verhaltens können die dargestellten Kurven als einfache Modelle zur Beschreibung monotoner Wachstumsvorgänge verwendet werden. Zu diesem Zweck werden zuerst noch zwei Koordinatentransformationen vorgenommen.

An der Stelle $x = 0$ ist der Funktionswert nicht definiert. Nähert man sich von positiven x -Werten kommend dem Nullpunkt, so geht der Funktionswert ge-

Abb. 2.7 Funktionsgraphen von $y = b/x$ für $b > 0$ und $b < 0$. Durch Übergang auf die Koordinaten $\xi = x - f$ und $\eta = y + e$ erhält man die Funktionsgleichung $\eta = b/(\xi + f) + e$



gen $+\infty$ oder $-\infty$, je nachdem ob $b > 0$ oder $b < 0$ gilt. Die Stelle $x = 0$ wird daher eine **Unendlichkeitsstelle** (oder auch **Polstelle**) genannt. Da reale Größen keine beliebig großen Werte annehmen, wollen wir die Unendlichkeitsstelle aus unserem Betrachtungsbereich ausschließen. Dazu verschieben wir die y -Achse um f Einheiten nach rechts. Zusätzlich verschieben wir die x -Achse um einen bestimmten Betrag e parallel zu sich nach unten; wir erreichen dadurch, dass sich die Kurven mit wachsendem x nicht dem Grenzwert null, sondern monoton dem Wert e nähern. Den Verschiebungen der Achsen wird in Abb. 2.7 durch die neuen Koordinaten $\eta = y + e$ und $\xi = x - f$ Rechnung getragen; die entsprechenden Koordinatenachsen sind strichliert eingezeichnet. In den neuen Koordinaten nimmt die Gleichung $y = b/x$ die Gestalt

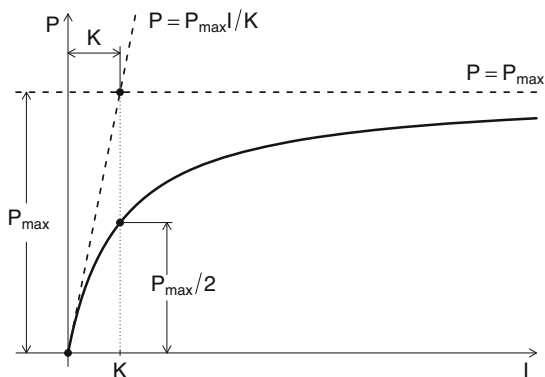
$$\eta = \frac{b}{\xi + f} + e \quad (2.11)$$

an. Die durch (2.11) dargestellten Funktionen werden als gebrochen linear bezeichnet. Sie besitzen an der Stelle $\xi = 0$ den Funktionswert $\eta_0 = b/f + e$ und nähern sich mit wachsendem ξ monoton dem Grenzwert e . Bei Beschränkung auf $\xi \geq 0$ sind die Funktionsgraphen in der (ξ, η) -Ebene Teile von **Hyperbeln** mit den achsenparallelen Asymptoten $\eta = e$ und $\xi = -f$.

Durch Spezialisierung der Parameter e , f und b lassen sich aus (2.11) für die Praxis wichtige Funktionsgleichungen gewinnen. Wir bringen zunächst ein Beispiel für eine nach dem Gesetz $\eta = b/(\xi + f)$ von einem Anfangswert $\eta_0 = b/f$ mit wachsendem ξ monoton gegen null abnehmende Zustandsgröße η . Im zweiten Beispiel geht es um die Funktionsgleichung (2.11) mit der Spezialisierung $b = -fe$. Damit geht (2.11) über in:

$$\eta = \frac{b}{\xi + f} + e = -\frac{fe}{\xi + f} + e = \frac{e(-f + \xi + f)}{\xi + f} = \frac{e\xi}{\xi + f}.$$

Abb. 2.8 Graphische Darstellung der Funktion P mit der Gleichung $P(I) = P_{\max}/(I + K)$ für $I \geq 0$. Der Funktionsgraph ist Teil eines Hyperbelastes



Durch diese Gleichung wird in der (ξ, η) -Ebene – bei Beschränkung auf den ersten Quadranten – ein Teil eines Hyperbelastes definiert, der im Nullpunkt beginnt und sich mehr und mehr der Asymptote $\eta = e$ nähert.

Beispiel 2.8

Bei einer bimolekularen Reaktion des Typs $A + B \xrightarrow{k} P$ möge die Konzentration von A stets gleich der von B sein. Die Konzentration c der reagierenden Moleküle nimmt dann im Verlaufe der Zeit t gemäß $c = c_0/(c_0 k t + 1)$ ab. Dabei ist c_0 die Anfangskonzentration von A und B und k die Reaktionskonstante. Diese Formel geht aus (2.11) hervor, wenn dort $e = 0$, $f = 1/(c_0 k)$ und $b = c_0 f = 1/k$ gesetzt wird.

Beispiel 2.9

Die Abhängigkeit der Photosyntheserate P von der Lichtintensität I kann durch die Funktionsgleichung

$$P = P(I) = \frac{P_{\max} I}{I + K}$$

mit positiven Konstanten $P_{\max}$ und K erfasst werden. Durch Vergleich mit $\eta = e\xi/(\xi + f)$ sieht man, dass es sich wieder um eine gebrochene lineare Funktion mit den Verschiebungsparametern $e = P_{\max}$ und $f = K$ sowie dem Anfangswert $\eta_0 = 0$ handelt. Die Photosyntheserate nähert sich mit wachsender Lichtintensität I dem Sättigungswert $P_{\max}$.

Um das Verhalten der Funktion in der Umgebung der Stelle $I = 0$ zu studieren, kann man im Nenner der Formel $P = P_{\max} I/(I + K)$ die Größe I gegenüber K vernachlässigen. Für kleine Werte von I gilt also näherungsweise $P \approx (P_{\max}/K)I$. Mit gegen null strebendem I nähert sich der Graph der Funktion P der Nullpunktsgerechten mit dem Anstieg $P_{\max}/K$; diese Gerade stellt die Tangente des Funktionsgraphen von P im Nullpunkt dar. Mit Hilfe dieser Tangente, der Sättigungsgeraden $P = P_{\max}$ und des Kurvenpunktes $(K, P_{\max}/2)$ lässt sich der Funktionsgraph von P schnell skizzieren (vgl. Abb. 2.8).

Man beachte, dass sich die beiden betrachteten Funktionstypen durch eine **Reziproktransformation** linearisieren lassen. So geht die Funktionsgleichung $y = b/(x + f)$ durch die Transformation $\eta = 1/y$ und $\xi = x$ in die Geradengleichung $\eta = \xi/b + f/b$ über. Zur Linearisierung von $y = ex/(x + f)$ gehe man zu den neuen Koordinaten $\eta = 1/y$ und $\xi = 1/x$ über, womit man in der (ξ, η) -Ebene die Geradengleichung $\eta = f\xi/e + 1/e$ erhält. Damit lässt sich die Anpassung von gebrochen linearen Funktionen an vorgegebene Datenpunkte auf ein lineares Ausgleichsproblem zurückführen.

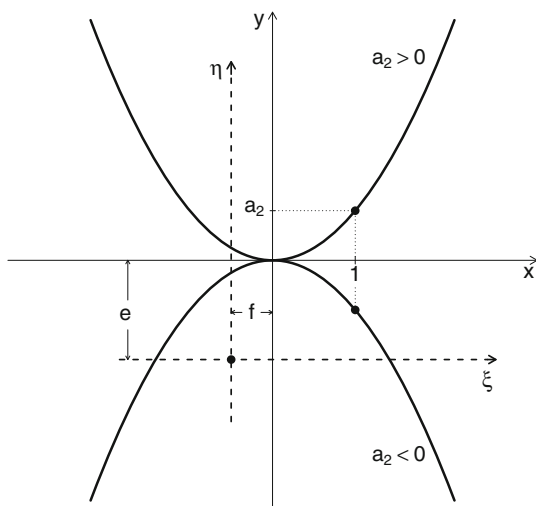
2.2.4 Quadratische Polynome

Für $c = 2$ geht die Gleichung $y = bx^c$ der allometrischen Funktion in $y = bx^2$ über. Diese Gleichung ist ein Sonderfall der Funktionsgleichung

$$y = Q(x) = a_2x^2 + a_1x + a_0 \quad (2.12)$$

für das allgemeine **quadratische Polynom** Q mit beliebigen reellen Koeffizienten $a_2 \neq 0$, a_1 , a_0 und der gesamten x -Achse als Definitionsbereich. Um die Eigenschaften dieser Funktion an Hand des Funktionsgraphen studieren zu können, gehen wir von der „verkürzten“ Funktion mit der Gleichung $y = a_2x^2$ aus. Durch diese Gleichung wird in der (x, y) -Ebene bekanntlich eine **Parabel** dargestellt (vgl. Abb. 2.9). Der **Scheitel** S liegt im Nullpunkt, die **Scheiteltangente** fällt mit der x -Achse und die **Parabelachse** mit der y -Achse zusammen. Bei positivem a_2 ist die Parabel „nach oben offen“ und stellt insgesamt eine **nach unten konvexe** Kurve dar; bei negativem a_2 ist die Parabel „nach unten offen“ und ist eine **nach oben konvexe** Kurve.

Abb. 2.9 Funktionsgraphen von $y = a_2x^2$ für $a_2 > 0$ und $a_2 < 0$. Durch Übergang auf die Koordinaten $\xi = x + f$, $\eta = y + e$ ergibt sich das Polynom $\eta = a_2\xi^2 + a_1\xi + a_0$ mit $a_1 = -2fa_2$ und $a_0 = a_2f^2 + e$



Durch Parallelverschiebung des Koordinatensystems verändert sich eine dargestellte Kurve nicht, wohl aber muss die Kurvengleichung modifiziert werden. Wir nehmen nun an, dass die x -Achse parallel zu sich um e Einheiten nach unten und die y -Achse parallel zu sich um f Einheiten nach links verschoben wird. Die neuen Achsen – wir bezeichnen sie als ξ - bzw. η -Achse – sind in Abb. 2.9 strichliert eingezeichnet. Die neuen Koordinaten hängen mit den alten über die Transformationsgleichungen $\xi = x + f$ und $\eta = y + e$ zusammen. Die Funktionsgleichung $y = a_2x^2$ lautet in den Koordinaten

$$\eta = a_2(\xi - f)^2 + e = a_2(\xi^2 - 2f\xi + f^2) + e = a_2\xi^2 - 2fa_2\xi + a_2f^2 + e. \quad (2.13)$$

Setzt man $a_1 = -2fa_2$ und $a_0 = a_2f^2 + e$, so erhält man wieder die allgemeine Gleichung für das quadratische Polynom. Der Graph eines jeden quadratischen Polynoms stellt daher in der (x, y) -Ebene eine Parabel dar mit zur x -Achse paralleler Scheiteltangente und dem Scheitel $S(f, e)$; die Koordinaten des Scheitels stimmen also mit den Verschiebungsparametern f und e überein, die sich ihrerseits wieder über die Formeln $f = -a_1/(2a_2)$, $e = a_0 - a_2f^2$ aus den Koeffizienten des Polynoms berechnen lassen.

Für die Darstellung des Funktionsgraphen von Q ist es ferner zweckmäßig, allfällige Schnittpunkte der Kurve mit der ξ -Achse zu bestimmen. Die ξ -Koordinaten der Schnittpunkte werden als reelle **Nullstellen** des Polynoms bezeichnet und sind als Lösungen der quadratischen Gleichung $Q(x) = 0$ zu berechnen. Reelle Nullstellen gibt es nur dann, wenn die Diskriminante $D = a_1^2 - 4a_0a_2$ nicht negativ ist. Wegen

$$D = 4a_2 \left(\frac{a_1^2}{4a_2} - a_0 \right) = 4a_2(f^2a_2 - a_0) = -4a_2e$$

ist das aber genau dann der Fall, wenn entweder $a_2 > 0$ und $e \leq 0$ ist oder $a_2 < 0$ und $e \geq 0$. Diese Bedingungen sind geometrisch unmittelbar einsichtig. So bedeutet die erste, dass die durch die Gleichung dargestellte Parabel nach oben offen sein muss und der Scheitel unter oder auf der ξ -Achse liegen muss. Offensichtlich gibt es dann zwei Schnittpunkte oder einen Berührungspunkt. Analog kann die zweite Bedingung interpretiert werden. Da die Schnittpunkte symmetrisch zur Parabelachse liegen, muss die Scheitelkoordinate f gleich dem arithmetischen Mittel der ξ -Koordinaten der Schnittpunkte sein.

Beispiel 2.10

Nach dem Populationsmodell in Beispiel 1.17 ist die zeitliche Änderung R' der Größe N einer Population durch $R' = rN(1 - N/C) - EN$ gegeben. Nach dieser Gleichung ist R' ein quadratisches Polynom in N mit den Koeffizienten $a_2 = -r/C$, $a_1 = r - E$ und $a_0 = 0$. Der Funktionsgraph stellt daher in der (N, R') -Ebene eine nach unten offene Parabel dar mit den Scheitelkoordinaten

$$N_S = \frac{C(r - E)}{2r} \quad \text{und} \quad R_S = \frac{C(r - E)^2}{4r}.$$

Die beiden Nullstellen $N_1 = 0$, $N_2 = (1 - E/r)C$ wurden bereits in Beispiel 1.17 berechnet.

2.2.5 Aufgaben

1. An 5 Exemplaren der Wandermuschel *Dreissena polymorpha pallas* wurden (nach 5 Altersklassen aufgegliedert) die Massen (in g) und die Längen (in mm) bestimmt und die Klassenmittelwerte L bzw. M berechnet. Die so erhaltenen Zahlenpaare (L, M) sind: (7.56, 0.055), (11.92, 0.213), (16.40, 0.564), (24.83, 1.894) und (29.03, 3.012). Man stelle die Abhängigkeit der Masse M von der Länge L durch eine allometrische Funktion dar und beurteile die Güte der Anpassung mit dem Bestimmtheitsmaß.⁵
2. Wir betrachten die enzymatische Reaktion $E + S \rightleftharpoons ES \rightarrow E + P$, in der E das Enzym, S das Substrat, ES das Enzymsubstrat und P das Ergebnis der Umwandlung bezeichnen. Die Reaktionskonstanten seien k_1 (für die Reaktion $E + S \rightarrow ES$), k_2 (für $E + S \leftarrow ES$) und k_3 (für $ES \rightarrow E + P$). Die Konzentrationen $[S]$ und $[P]$ mögen konstant gehalten werden; ebenso konstant sei die Gesamtmenge des Enzyms, d. h. $[E] + [ES] = E_0$. Im Gleichgewicht gilt $k_1[E][S] = (k_2 + k_3)[ES]$, also $k_1(E_0 - [ES])[S] = (k_2 + k_3)[ES]$. Daraus folgt die Michaelis-Menten-Gleichung⁶

$$v = k_3[ES] = v_m \frac{[S]}{[S] + K}$$

für die Reaktionsgeschwindigkeit mit den Konstanten $v_m = k_3 E_0 > 0$ und $K = (k_2 + k_3)/k_1 > 0$.

- a) Welchem Wert nähert sich v mit wachsender Substratkonzentration?
 - b) Für kleine Werte von $[S]$ nimmt v nahezu linear mit $[S]$ zu. Wie lautet die Gleichung dieser linearen Approximationsfunktion?
 - c) Man skizziere den durch die Funktionsgleichung $v = v([S])$ bestimmten Kurvenverlauf für $[S] \geq 0$.
3. Die Abhängigkeit der Photosyntheserate P (in % eines Standards) von der Temperatur t (in °C) wurde durch die quadratische Funktion $P = 45 + 7t - 0.25t^2$ dargestellt. Man bestimme jene Temperatur, bei der P den maximalen Wert besitzt.
 4. Im Hardy-Weinbergschen Gleichgewicht bestehen (an einem autosomalen Genort mit 2 Allelen) zwischen den Genotyphäufigkeiten D , H und R einerseits und den Genhäufigkeiten p und $q = 1 - p$ andererseits die Beziehungen $D = p^2$, $H = 2pq$, $R = q^2$. Man stelle die Abhängigkeit der Genotyphäufigkeiten von der Genhäufigkeit p ($0 \leq p \leq 1$) graphisch dar. Um welche Kurven handelt es sich bei den Funktionsgraphen?

⁵ Vgl. Schulz, N.: Die Wandermuschel im Keutschacher See. Carinthia II, 170/90, 549 (1980).

⁶ Leonor Michaelis (1875–1949) arbeitete als Biochemiker in Deutschland und den USA, Gemeinsam mit der kanadischen Medizinerin Maud L. Menten (1879–1960) schuf er die Grundlagen zum Verständnis von enzymatischen Reaktionen.

2.3 Exponential- und Logarithmusfunktionen

2.3.1 Exponentialfunktionen

Der Zusammenhang zweier Wachstumsgrößen x und y wird durch die allometrische Funktion beschrieben, wenn dem Wachstum folgende Gesetzmäßigkeit zugrunde liegt: Mit einer kleinen Änderung Δx von x ist eine kleine Änderung Δy von y derart verbunden, dass $\Delta y/\Delta x$ stets proportional zum Quotienten y/x der Wachstumsgrößen ist.

Eine große Anzahl von verschiedenartigen Phänomenen kann auf einen anderen grundlegenden Wachstumsmechanismus zurückgeführt werden. So wächst die Größe y einer Population im Anfangsstadium oft so, dass die Zunahme Δy innerhalb einer kleinen Zeitspanne Δt proportional zum Produkt aus der momentanen Populationsgröße y und dem Zeitintervall Δt ist. Ein anderes Beispiel ist der radioaktive Zerfall: Hier ist die Abnahme $-\Delta m$ der Masse einer radioaktiven Substanz in einem kleinen Zeitintervall Δt proportional zum Produkt $m\Delta t$. Schließlich gilt auch für die Verdünnung einer Substanz, etwa einer bestimmten Pharmakonzentration c im Blut, dass die Abnahme $-\Delta c$ in einer kleinen Zeitspanne Δt proportional zu $c\Delta t$ ist. Den angeführten Beispielen ist gemeinsam, dass die auf die Zeiteinheit bezogene Änderung $\Delta y/\Delta t$ der jeweiligen Bestandsgröße (im Folgenden mit y bezeichnet) bei kleinem Δt proportional zum Momentanwert der Bestandsgröße ist; für hinreichend kleine Zeitspannen Δt gilt also $\Delta y/\Delta t \sim y$ oder, wenn die Proportionalitätskonstante mit r bezeichnet wird,

$$\frac{\Delta y}{\Delta t} = ry.$$

Je nach Anwendungsfall wird $r = (\Delta y/\Delta t)/y$ als Wachstumsrate, Zerfallskonstante, Ausscheiderate oder ähnlich bezeichnet. Die sich auf der Grundlage des **Prinzips bestandsproportionaler Veränderungen** ergebende Zeitabhängigkeit kann – wie im Abschn. 4.5.2 gezeigt wird – durch die Gleichung

$$y = y_0 e^{rt} \quad (2.14)$$

der **allgemeinen Exponentialfunktion** ausgedrückt werden, in der y_0 den Wert der Bestandsgröße zum Zeitpunkt $t = 0$ darstellt.

Wir betrachten zuerst den Sonderfall der sogenannten **natürlichen Exponentialfunktion** mit der Gleichung $y = y(t) = e^t$ ($e = 2.71828\dots$). Diese Funktion ist für jeden Wert von t definiert, stets positiv und streng monoton wachsend.⁷ Der Funktionsgraph verläuft in der (t, y) -Ebene durch den Punkt $(0, 1)$ und nähert sich mit gegen $-\infty$ gehendem t asymptotisch der t -Achse. Ersetzt man in der Funktionsgleichung y durch y/y_0 , so erhält man die Gleichung $y = y_0 e^t$ einer Exponentialkurve, die durch den Punkt $(0, y_0)$ verläuft. Ist $y_0 = c > 0$, so ist der Kurvenverlauf ähnlich dem der natürlichen Exponentialfunktion; den Kurvenverlauf für $y_0 = -c < 0$ erhält man aus dem für $y_0 = c > 0$ durch Spiegelung an

⁷ Für beliebige Werte t_1, t_2 gilt: Aus $t_2 > t_1$ folgt stets $y(t_2) > y(t_1)$.

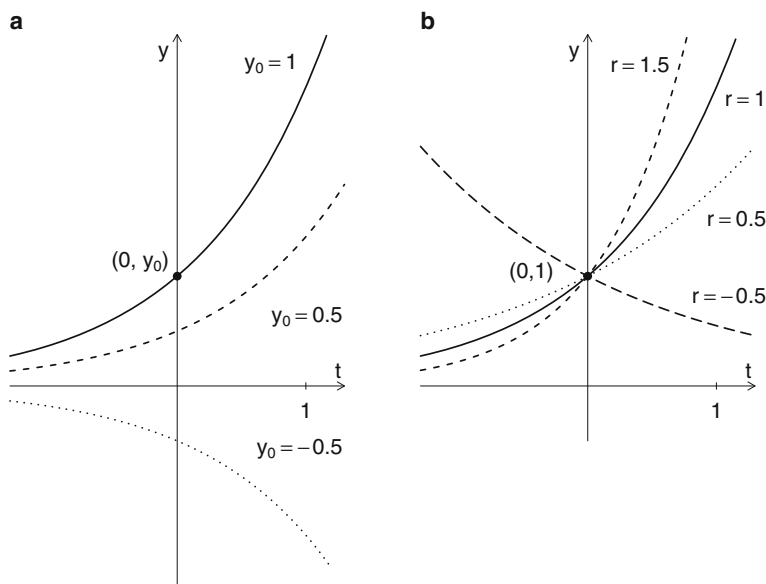


Abb. 2.10 Typische Kurvenverläufe von $y(t) = y_0 e^t$ für $y_0 = -0.5, 0.5, 1$ (a) und von $y/y_0 = e^t$ für $r = -0.5, 0.5, 1, 1.5$ (b)

der t -Achse (vgl. Abb. 2.10). Ersetzt man schließlich noch t durch rt , so folgt die Gleichung $y = y_0 e^{rt}$ der allgemeinen Exponentialfunktion. Verglichen mit dem Sonderfall $r = 1$ ergeben sich für $r > 1$ steiler und für $0 < r < 1$ flacher verlaufende Funktionsgraphen. Der Funktionsgraph bei negativem $r = -r_0 < 0$ folgt aus dem für $r = r_0 > 0$ durch Spiegelung an der y -Achse. Die Gleichung $y = y_0 e^{rt}$ ist also geeignet, bei positivem r bestandsproportionale Wachstumumsprozesse (Geburtenprozesse) wiederzugeben, bei negativem r bestandsproportionale Zerfallsprozesse (Sterbeprozesse).

Beispiel 2.11

- a) Die Einwohnerzahl von Mexico-City betrug 5.5 Millionen im Jahre 1960 und stieg bis zum Jahr 1975 auf 8.1 Millionen an. Unter der Annahme eines exponentiellen Wachstums soll die Einwohnerzahl y (in Millionen) als Funktion der Zeit dargestellt werden.

Wenn der Nullpunkt der Zeitskala in das Jahr 1960 verlegt wird, so ist $y = 5.5$ für $t = 0$ und $y = 8.1$ für $t = 15$. Zu bestimmen sind die Parameter y_0 und r der Wachstumsfunktion $y = y_0 e^{rt}$. Setzt man $t = 0$, folgt sofort $y_0 = y(0) = 5.5$; für $t = 15$ gilt $8.1 = 5.5 e^{15r}$, woraus man durch Logarithmieren $\ln 8.1 = \ln 5.5 + \ln e^{15r} = \ln 5.5 + 15r$ erhält. Somit ist die Wachstumsrate $r = (\ln 8.1 - \ln 5.5)/15 = 2.58\%$ pro Jahr.

Wir bestimmen auch noch die *Verdopplungszeit*, also jene Zeitspanne t' , nach der sich die Einwohnerzahl verdoppelt. Aus der Bedingung $y = 2y_0 = y_0 e^{rt'}$

folgt $t' = (\ln 2)/r = 26.9$ Jahre. Setzt man $r = (\ln 2)/t'$ in die Wachstumsgleichung ein, so geht diese über in die Gleichung

$$y = y_0 \left(e^{(\ln 2)/t'} \right)^t = y_0 \left(2^{1/t'} \right)^t = y_0 1.026^t$$

einer Exponentialfunktion mit der Basis $a = 1.026$. Umgekehrt kann jede mit einer Basis $a > 0$ vorgegebene Exponentialfunktion $y = y_0 a^t$ auf die Gestalt $y = y_0 e^{rt}$ gebracht werden, wobei $r = \ln a$ zu setzen ist.

- b) Vom Kaliumisotop K^{42} , das zur radioaktiven Markierung verwendet wird, zerfallen innerhalb einer Stunde $h = 5.39\%$ der zu Beginn vorhandenen Menge. Zur Bestimmung der *Zerfallskonstanten* λ betrachten wir den durch das *radioaktive Zerfallsgesetz* $m = m_0 e^{-\lambda t}$ beschriebenen Zerfall im Zeitraum von $t_0 = 0$ bis $t_1 = 1$ (Zeitangaben in Stunden). Dabei bezeichnen m_0 die Ausgangsmenge der radioaktiven Substanz und m die nach t Zeiteinheiten noch nicht zerfallene Menge. Von $t_0 = 0$ bis $t_1 = 1$ zerfällt die Substanzmenge

$$m_0 - m_0 e^{-\lambda} = m_0 (1 - e^{-\lambda})$$

und diese Menge soll nach der Angabe hm_0 sein, d. h. $1 - e^{-\lambda} = h$. Somit ist $\lambda = -\ln(1 - h) = 0.05541$ pro Stunde. Die *Halbwertszeit* $t_{1/2}$, also jene Zeitdauer, innerhalb der die Hälfte der Ausgangsmenge zerfällt, ergibt sich schließlich aus der Formel $t_{1/2} = (\ln 2)/\lambda = 12.51$ Stunden.

- c) Nach dem Lambert-Beerschen⁸ Gesetz nimmt die Intensität I von Licht bei Durchgang durch ein Medium der Dicke x von anfangs I_0 auf $I = I_0 e^{-\mu x}$ ab. Die positive Konstante μ wird Absorptionskoeffizient genannt. Man weiß, dass in reinem Meerwasser sich die Lichtintensität pro Meter Meerestiefe um den Faktor $1/4$ verringert. Daraus ergibt sich mit Hilfe von $0.25 I_0 = I_0 e^{-\mu}$ die Absorptionskonstante $\mu = 1.386 \text{ m}^{-1}$. Damit findet man, dass in einer Tiefe von 5 m unter dem Meeresspiegel die Lichtintensität nur mehr rund 0.1% der auf die Oberfläche auftreffenden Tageslichtintensität beträgt.

2.3.2 Linearisierung durch einfach-logarithmische Transformation

In diesem Abschnitt setzen wir Exponentialfunktionen des Typs $y = y_0 e^{rt}$ mit $y_0 > 0$ und $r \neq 0$ voraus. In der (t, y) -Ebene ist der Funktionsgraph je nach dem Vorzeichen von r eine monoton wachsende oder fallende Exponentialkurve. Trägt man auf der vertikalen Achse $\eta = \ln y$ statt y auf, so erhält man statt der Exponentialkurve eine steigende oder fallende Gerade.⁹ Die Gerade hat in der (t, η) -Ebene die

⁸ Benannt nach Johann H. Lambert (1728–1777, elsässischer Mathematiker und Physiker) und August Beer (1825–1863, deutscher Physiker).

⁹ Im Gegensatz zur doppelt-logarithmischen Transformation wird hier also nur der Logarithmus der abhängigen Variablen gebildet. Dieser Umstand wird durch die Bezeichnung „einfach-logarithmische Transformation“ (kurz log-Transformation) zum Ausdruck gebracht.

Gleichung $\eta = \ln y_0 + rt$, der Geradenanstieg ist $k = r$ und der η -Achsenabschnitt $d = \ln y_0$. Exponentielle Abhängigkeiten können also in linearisierter Form in der (t, η) -Ebene wiedergegeben werden. Diese Tatsache kann – wie im Folgenden gezeigt wird – auch bei Anpassung einer Exponentialkurve an vorgegebene Datenpunkte genutzt werden.

Es seien t und y zwei Größen, von denen insgesamt n Wertepaare (t_i, y_i) bestimmt wurden. Der Einfluss der Variablen t auf die Größe y soll mit Hilfe einer Regressionsfunktion beschrieben werden. Es ist also t die unabhängige Variable (oder Einflussgröße) und y die abhängige Variable (oder Zielgröße). Auf Grund von Voruntersuchungen oder theoretischen Überlegungen wird eine exponentielle Abhängigkeit des Typs $y = y_0 e^{rt}$ vermutet. Wenn diese Vermutung zu Recht besteht, dann müssen in der (t, η) -Ebene die Punkte (t_i, η_i) mit $\eta_i = \ln y_i$ im Großen und Ganzen eine lineare Anordnung aufweisen. In diesem Fall ist es sinnvoll, an die Datenpunkte (t_i, η_i) eine Gerade anzupassen. Die Gerade setzen wir in der Hauptform $\eta = kt + d$ mit zunächst unbestimmten Parametern k und d an. Damit liegt in der (t, η) -Ebene ein lineares Regressionsproblem vor, das mit den Methoden von Abschn. 2.1.3 gelöst werden kann. Aus den bekannten Schätzwerten $\hat{k}$ und $\hat{d}$ ergeben sich dann die Schätzwerte $\hat{r} = \hat{k}$ und $\hat{y}_0 = e^{\hat{d}}$ für die Parameter der gesuchten Exponentialfunktion. Diese ermöglicht es, von einem vorgegebenen t -Wert auf den entsprechenden mittleren Wert $\hat{y} = \hat{y}_0 e^{\hat{r}t}$ der Zielgröße y hoch zu schätzen.

Es folgt ein Zahlenbeispiel für ein Problem mit exponentieller Regressionsfunktion. Durch eine log-Transformation der Zielgröße wird die nicht lineare Regressionsaufgabe auf eine lineare zurückgeführt.

Beispiel 2.12

Die Bevölkerungsgröße y der U.S.A. zeigt zwischen den Jahren 1790 und 1890 einen angenähert exponentiellen Wachstumsverlauf nach der Formel $y = y_0 e^{rt}$. Die Parameter y_0 und r sollen durch Anpassung der Exponentialkurve an die in Tab. 2.3 angegebenen Datenpunkte (t_i, y_i) bestimmt werden.¹⁰ Zur Vereinfachung der Rechnung ist der Nullpunkt der Zeitskala in das Jahr 1790 verlegt.

Zur Lösung der gestellten Aufgabe wird zuerst eine logarithmische Transformation der y -Werte vorgenommen und dann die Regressionsgerade $y' = \hat{k}t + \hat{d}$ in der (t, y') -Ebene bestimmt. Der Stichprobenumfang n ist 11; die Mittelwerte der t - und y -Werte sind $\bar{t} = 550/11 = 50$ und $\bar{y} = 263.62/11 = 23.97$. Mit Hilfe der Summen der η -, $(t - \bar{t})^2$ - und $(t - \bar{t})(\eta - \bar{\eta})$ -Spalte in Tab. 2.3 erhält man den Mittelwert $\bar{\eta} = 31.01/11 = 2.82$, die Varianz $s_t^2 = 11\,000/10 = 1100$ und die Kovarianz $s_{t\eta} = 309.18/10 = 30.92$. Daraus folgen die Schätzwerte $\hat{k} = s_{t\eta}/s_t^2 = 0.02811$ und $\hat{d} = 2.82 - 0.02811 \cdot 50 = 1.414$ für die Parameter k und d der Regressionsgeraden. Somit ist $\hat{r} = \hat{k} = 0.02811$ ein Schätzwert für die Wachstumsrate und $\hat{y}_0 = e^{\hat{d}} = 4.11$ die mit der Regressionsfunktion geschätzte Bevölkerungsgröße zum Zeitpunkt $t = 0$ (Kalenderjahr 1790). Die durch $\hat{r}$ und $\hat{y}_0$ festgelegte

¹⁰ Historical Statistics on the United States. US Department of Commerce, Bureau of the Census (1970, <http://www.census.gov/history/pdf/histstats-colonial-1970.pdf>).

Tab. 2.3 Daten und Rechenschema zu Beispiel 2.12. Die zweite und dritte Spalte enthalten die Werte der Variablen t (Zeit in Jahren nach 1790) bzw. y (US-Bevölkerung in Millionen), in der dritten Spalte stehen die logarithmierten y -Werte. In den weiteren Spalten werden die Größen bereitgestellt, die für die lineare Regression bzw. die Berechnung des Bestimmtheitsmaßes gebraucht werden. Die Spaltensummen in der letzten Zeile können auf Grund von Rundungsfehlern von den Summen der angeschriebenen Tabellenwerte abweichen

i	t_i	y_i	η_i	$(t_i - \bar{t})^2$	$(t_i - \bar{t})(\eta_i - \bar{\eta})$	$\hat{y}_i = y_0 e^{k t_i}$	$(y_i - \hat{y})^2$	$(y_i - \bar{y})^2$
1	0	3.93	1.37	2500	72.52	4.11	0.033	401.42
2	10	5.31	1.67	1600	45.98	5.45	0.018	348.03
3	20	7.24	1.98	900	25.18	7.21	0.001	279.74
4	30	9.64	2.27	400	11.06	9.55	0.007	205.22
5	40	12.87	2.55	100	2.64	12.65	0.047	123.11
6	50	17.07	2.84	0	0.00	16.76	0.095	47.55
7	60	23.19	3.14	100	3.25	22.20	0.978	0.60
8	70	31.44	3.45	400	12.58	29.41	4.135	55.87
9	80	39.82	3.68	900	25.96	38.95	0.756	251.37
10	90	50.16	3.92	1600	43.85	51.59	2.051	686.15
11	100	62.95	4.14	2500	66.16	68.34	29.017	1519.79
Σ	550	263.62	31.01	11 000	309.18		37.138	3918.85

Exponentialkurve mit der Gleichung

$$\hat{y} = \hat{y}_0 e^{\hat{k}t} = 4.11e^{0.02811t}$$

ist in Abb. 2.11 dargestellt. Man erkennt durch Vergleich mit den ebenfalls eingezeichneten, den Beobachtungswerten entsprechenden Punkten, dass die Bevölkerungsentwicklung in guter Näherung tatsächlich exponentiell verlaufen ist. Numerisch kann die Güte der Anpassung in der (t, y) -Ebene wieder mit dem Bestimmtheitsmaß $B = 1 - S_{rr}/S_{yy}$ mit $S_{yy} = (n-1)s_y^2 = 3918.85$ und $S_{rr} = 37\,138$ beurteilt werden. Man erhält für B den nahe bei 1 liegenden Wert 0.991.

```
> # Lösung mit R
> t <- seq(from=0, to=100, by=10)
> y <- c( 3.93, 5.31, 7.24, 9.64, 12.87, 17.07,
+       23.19, 31.44, 39.82, 50.16, 62.95)
> n <- length(t)
> # log-Transformation und lineare Regression
> eta <- log(y); linreg <- lm(eta ~ t)
> par <- coefficients(linreg); k <- par[[2]]; d <- par[[1]]
> print(cbind(k, d), digits=4)
      k      d
[1,] 0.02811 1.414
> # Bestimmtheitsmaß
> yexp <- exp(k*t+d) # erwartete y-Werte
> Srr <- sum((y-yexp)^2); Sy <- (n-1)*var(y); B=1-Srr/Sy
> print(cbind(Srr, Sy, B), digits=4)
      Srr   Sy   B
[1,] 37.14 3919 0.9905
> # Abb. 2.11
```

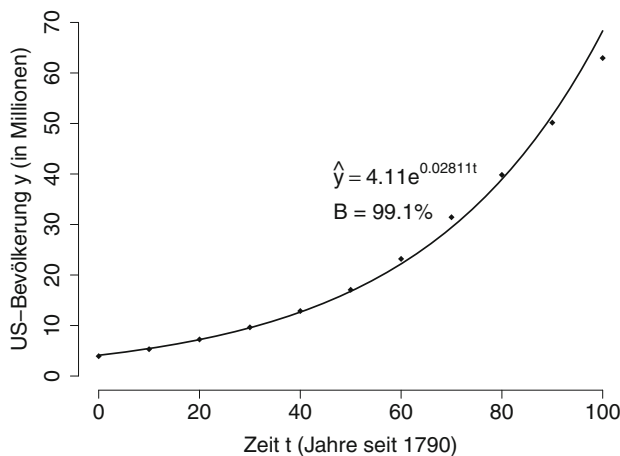


Abb. 2.11 Anpassung einer Exponentialfunktion an die Datenpunkte von Beispiel 2.12. Die Ausgleichsfunktion $\hat{y} = 4.11e^{0.02811t}$ wurde durch lineare Regression von $\eta = \ln y$ auf t bestimmt

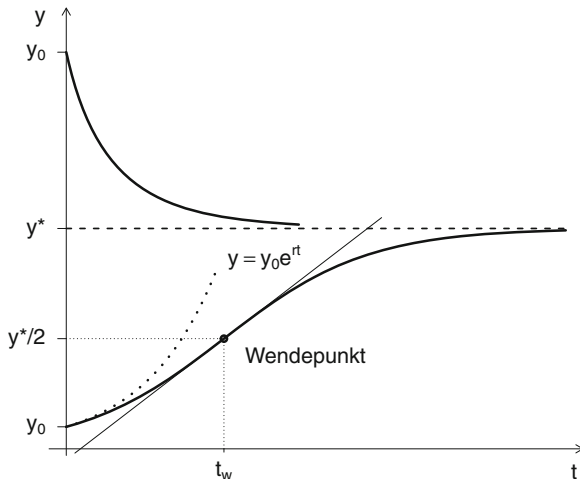
```
> par(mai=c(1.2, 1.2, 0.1, 0.1), cex.axis=1.2, cex.lab=1.2)
> # Streudiagramm mit Originalvariablen
> plot(t, y, type="p", col="black",
+      xlab="Zeit t (Jahre seit 1790)", pch=18, frame.plot=F,
+      ylab="US-Bevölkerung y (in Millionen)",
+      xlim=c(0, 100), ylim=c(0, 70), lwd=2)
> curve(exp(k*x+d), lty=1, lwd=2, ad=T)
> text(45, exp(k*80+d), pos=4, cex=1.2,
+      expression(hat(y)==4.11*e^{0.02811*t}))
> text(45, 32, expression("B = 99.1%"), pos=4, cex=1.2)
```

2.3.3 Logistisches Wachstum

Stellt man das Wachstum eines Organismus oder einer Population in Abhängigkeit von der Zeit graphisch dar, so erhält man eine *Wachstumskurve*. Die Erfahrung zeigt, dass Wachstumskurven oft einen *S-förmigen* Verlauf besitzen: Die das Wachstum beschreibende Größe y nimmt in der Anfangsphase „exponentiell“ mit einer angenähert konstanten Rate r zu und strebt dann aber als Folge einer abnehmenden Wachstumsrate einem Sättigungswert y^* zu. Man ist versucht, den Wachstumsverlauf in der Anfangsphase durch eine Exponentialfunktion mit der Gleichung $y = y_0 e^{rt}$ (mit $y_0 < y^*/2$ und $r > 0$) und in der Sättigungsphase durch eine exponentielle Annäherung an den Sättigungswert zu approximieren (vgl. Abschn. 2.7.3). Eine gemeinsame Darstellung des Zeitverhaltens in beiden Phasen ermöglicht die **logistische Funktion** mit der Gleichung

$$y = \frac{y^*}{1 + \left(\frac{y^*}{y_0} - 1\right)e^{-rt}}. \quad (2.15a)$$

Abb. 2.12 Typische Kurvenverläufe der logistischen Funktion mit den positiven Parametern y_0 , y^* und r . Für $y_0 < y^*$ sind die Funktionen monoton wachsend, für $y_0 > y^*$ monoton fallend



Die in Abb. 2.12 gezeichneten Funktionsgraphen werden **logistische Kurven** genannt. Für Anfangswerte $y_0 < y^*/2$ ist die logistische Kurve „S-förmig“ und lässt drei Abschnitte erkennen: Eine exponentielle Anfangsphase, eine „lineare“ Wachstumsphase um den „Wendepunkt“ $(t_w, y^*/2)$ der Kurve mit $t_w = \ln(y^*/y_0 - 1)/r$ und eine Sättigungsphase mit exponentieller Annäherung an die Grenzgerade $y = y^*$. Führt man an Stelle des Anfangswertes y_0 die Wendepunktkoordinate t_w und an Stelle der Wachstumsrate r ihren Kehrwert $\gamma = 1/r$ als Parameter ein, erhält man die Gleichung der logistischen Funktion in der Form

$$y = \frac{y^*}{1 + e^{-(t-t_w)/\gamma}}. \quad (2.15b)$$

Mit den Gleichungen $y_0 = y^*/(1 + e^{rt_w})$ und $r = 1/\gamma$ kann man bei bekanntem t_w und γ wieder die ursprünglichen Parameter bestimmen.¹¹

Die logistische Funktion wird verwendet, um S-förmige Wachstumsverläufe zu modellieren. Wir legen im Folgenden die logistische Gleichung (2.15a) zugrunde. Die Beobachtung einer Größe y zu n aufeinanderfolgenden Zeitpunkten $t_1, t_2, \dots, t_n$ liefert die Zeitreihe $y_1, y_2, \dots, y_n$. Trägt man die Wertepaare (t_i, y_i) als Punkte in die (t, y) -Ebene ein, erhält man eine punktweise Darstellung der Wachstumskurve. Die Anpassung der logistischen Kurve an diese Punkte stellt ein nicht-lineares Ausgleichsproblem dar. Die Rückführung auf ein lineares Problem ist möglich, wenn der Sättigungswert y^* bekannt ist. Dazu formen wir (2.15a) zunächst in $y^*/y - 1 = (y^*/y_0 - 1)e^{-rt}$ um. Durch Logarithmieren erhält man

¹¹ Die Parametrisierung (2.15b) wird z.B. in der R-Funktion `SSlogis()` verwendet, mit der Startwerte für die Parameter y^* , t_w und γ im Rahmen einer nicht-linearen Regression mit der R-Funktion `nls()` berechnet werden.

daraus

$$\ln\left(\frac{y^*}{y} - 1\right) = \ln\left(\frac{y^*}{y_0} - 1\right) + \ln e^{-rt} = \ln\left(\frac{y^*}{y_0} - 1\right) - rt. \quad (2.16)$$

Zwischen der neuen Variablen $\eta = \ln(y^*/y - 1)$ und t besteht also eine lineare Abhängigkeit, die wieder in der gewohnten Form $\eta = kt + d$ mit $k = -r$ und $d = \ln(y^*/y_0 - 1)$ angeschrieben werden kann. Folglich hat man in der (t, η) -Ebene eine Gerade an die Punkte (t_i, η_i) mit $\eta_i = \ln(y^*/y_i - 1)$ anzupassen. Die Lösung dieser linearen Regressionsaufgabe liefert die Schätzwerte $\hat{k}$ und $\hat{d}$ für die Geradenparameter k bzw. d . Wegen $r = -k$ folgt daraus unmittelbar auch ein Schätzwert für r , nämlich $\hat{r} = -\hat{k}$; um einen Schätzwert $\hat{y}_0$ für y_0 zu erhalten, muss $\hat{d} = \ln(y^*/\hat{y}_0 - 1)$ nach $\hat{y}_0$ aufgelöst werden. Die Rechnung ergibt zunächst $e^{\hat{d}} = y^*/\hat{y}_0 - 1$ und schließlich $\hat{y}_0 = y^*/(e^{\hat{d}} + 1)$. Damit sind alle Parameter der logistischen Funktion bestimmt und man kann die zu den vorgegebenen t_i -Werten gehörenden Schätzwerte $\hat{y}_i$ für y berechnen. Mit Hilfe der Summe S_{rr} der Quadrate der Residuen $(y_i - \hat{y}_i)$ sowie der Summe S_{yy} der Quadrate der Abstände der y_i -Werte von ihrem Mittelwert $\bar{y}$ errechnet man das Bestimmtheitsmaß aus $B = 1 - S_{rr}/S_{yy}$.

Im Allgemeinen ist der Sättigungswert y^* unbekannt. Um die Linearisierung gemäß (2.16) anwenden zu können, wählt man für y^* zunächst einen plausiblen Näherungswert. Für diesen werden sodann die Parameter r und y_0 wie oben dargestellt geschätzt und die Restquadratsumme S_{rr} berechnet. Es ist durchaus möglich, dass durch eine andere Wahl von y^* eine bessere Anpassung, d. h. ein kleineres S_{rr} , erreicht wird. Ob dem so ist, kann man erkennen, indem man die Rechnung nochmals mit einem vergrößerten y^* -Wert durchführt. Ist der neue Wert von S_{rr} größer als der zuerst erhaltene, wird man es auch noch mit einem verkleinerten y^* -Wert versuchen. Ist dagegen S_{rr} kleiner geworden, so kann man vielleicht mit einem noch größeren y^* -Wert eine weitere Optimierung erreichen. Dieses schrittweise Vorgehen zur Bestimmung einer optimalen Anpassung wird durch das folgende Beispiel veranschaulicht.

Beispiel 2.13

Das Wachstum einer Hefekultur kann mit der logistischen Funktion modelliert werden. Wir greifen auf die Daten eines Experimentes zurück, die in der zweiten und dritten Spalte von Tab. 2.4 angeschrieben sind. Hier stellen die t_i -Werte die Messzeitpunkte (in Stunden) dar. Der Nullpunkt der Zeitskala fällt mit dem ersten Messzeitpunkt zusammen. Die m_i -Werte sind die zu den Beobachtungszeitpunkten gemessenen Massen m_i (in mg). Die entsprechenden Datenpunkte (t_i, m_i) sind in Abb. 2.13 eingezeichnet. Man erkennt, dass die Masse m im Verlaufe der Zeit einem Grenzwert m^* zustrebt, der bei ca. 500 mg liegt. Wir wollen das Wachstum durch die logistische Gleichung

$$m = \frac{m^*}{1 + \left(\frac{m^*}{m_0} - 1\right)e^{-rt}} \quad (2.17)$$

mit geeignet gewählten Konstanten m^* , m_0 und r beschreiben.

Tab. 2.4 Daten und Rechenschema zum Wachstum einer Hefekultur (Beispiel 2.13). In der zweiten und dritten Spalte sind die Messzeiten t_i (in Stunden) bzw. die Massen m_i (in mg) angeführt. Die vierte Spalte enthält die Werte der Hilfsvariablen $\eta_i = \ln(m^*/m_i - 1)$, die mit dem angenommenen Sättigungswert $m^* = \hat{m}^* = 500$ berechnet wurden. Die Spaltensummen in der letzten Zeile können auf Grund von Rundungsfehlern von den Summen der angeschriebenen Tabellenwerte abweichen

i	t_i	m_i	η_i	$(t_i - \bar{t})^2$	$(t_i - \bar{t})(\eta_i - \bar{\eta})$	$\hat{m}_i$	$(m_i - \hat{m})^2$	$(m_i - \bar{m})^2$
1	0	7.5	4.184	64	-34.500	8.27	0.587	62 945
2	2	26.0	2.903	36	-18.186	22.99	9.042	54 005
3	4	53.0	2.132	16	-9.041	60.72	59.559	42 185
4	6	140.0	0.944	4	-2.145	141.92	3.690	14 016
5	8	276.0	-0.209	0	0.000	265.97	100.571	310
6	10	394.0	-1.313	4	-2.370	382.60	130.018	18 390
7	12	453.0	-2.266	16	-8.551	451.67	1.780	37 873
8	14	484.0	-3.409	36	-19.689	482.01	3.954	50 900
9	16	492.0	-4.119	64	-31.929	493.58	2.484	54 574
Σ	72	2325.5	-1.152	240	-126.41		311.686	335 199

Unabhängig von der Festlegung des Sättigungswertes m^* können berechnet werden: Das arithmetische Mittel $\bar{t} = 72/9 = 8$ der Messzeiten, die mittlere Masse $\bar{m} = 2325.5/9 = 258.4$ sowie die Varianz $s_t^2 = 240/8 = 30$. Die Summe der Abweichungsquadrate $(m_i - \bar{m})^2$ der beobachteten Massen von ihrem Mittelwert, die als Maß für die Gesamtvariation in die Berechnungsformel für das Bestimmtheitsmaß eingeht, ist $S_{mm} = 335.199$.

Für den Sättigungswert nehmen wir zuerst den Schätzwert $m^* = \hat{m}^* = 500$. Die Bestimmung der Regressionsparameter m_0 und r erfolgt mit Hilfe von Tab. 2.4. Zuerst werden aus den beobachteten m -Werten die logarithmisch transformierten Werte $\eta_i = \ln(\hat{m}^*/m_i - 1)$ und die Produkte $(t_i - \bar{t})(\eta_i - \bar{\eta})$ gebildet; diese sind in der 4. und 6. Spalte angeschrieben. Wir notieren das arithmetische Mittel $\bar{\eta}_i = -1.152/9 = -0.128$ der η_i -Reihe und die Kovarianz $s_{t\eta} = -126.41/8 = -15.80$ sowie die Regressionsparameter

$$\hat{k} = \frac{s_{t\eta}}{s_t^2} = -0.5267, \quad \hat{d} = \bar{\eta} - \hat{k}\bar{t} = 4.086,$$

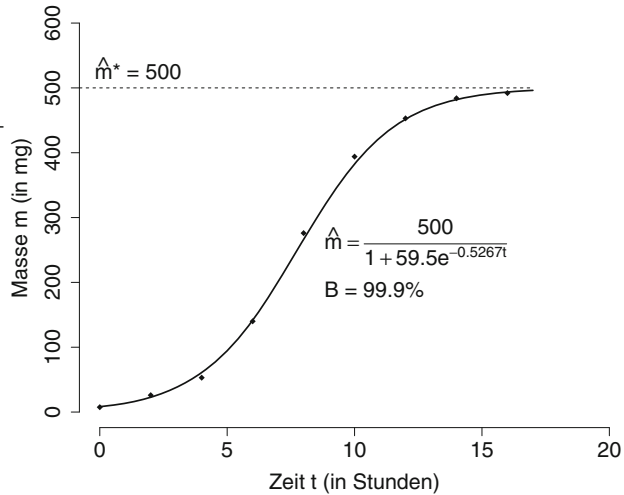
$$\hat{r} = -\hat{k} = 0.5267, \quad \hat{m}_0 = \frac{\hat{m}^*}{e^{\hat{d}} + 1} = 8.266.$$

Damit lautet die Gleichung der angepassten logistischen Funktion

$$\hat{m} = \frac{\hat{m}^*}{1 + \left(\frac{\hat{m}^*}{\hat{m}_0} - 1\right)e^{-\hat{r}t}} = \frac{500}{1 + 59.5e^{-0.5267t}}.$$

Diese Funktion ist in Abb. 2.13 dargestellt. Setzt man rechts die t_i -Werte ein, so folgen die mit der Regressionsfunktion hochgeschätzten Massen $\hat{m}_i$, die in der 7. Spalte von Tab. 2.4 aufgelistet sind. Schließlich enthält die vorletzte Spalte die Quadrate der Residuen $(m_i - \hat{m}_i)$ und deren Summe $S_{rr} = 311.686$. Der letzten Spalte

Abb. 2.13 Anpassung der logistischen Funktion (2.17) an die Datenpunkte von Beispiel 2.13. Mit dem (angenommenen) Sättigungswert $\hat{m}^* = 500$ wurden die Parameter r und y_0 durch eine lineare Regression von $\eta = \ln(\hat{m}^*/m - 1)$ auf t bestimmt



entnimmt man die Quadratsumme $S_{mm} = 335.199$ und berechnet mit den Summen das Bestimmtheitsmaß $B = 1 - S_{rr}/S_{mm} = 99.9\%$.

```
> # Lösung mit R (Linearisierungsmethode):
> t <- seq(from=0, to=16, by=2)
> m <- c(7.5, 26, 53, 140, 276, 394, 453, 484, 492)
> n <- length(m)
> # Linearisierung und lineare Regression
> ms <- 500; eta <- log(ms/m-1)
> linreg1 <- lm(eta ~ t); par <- coefficients(linreg1)
> d <- par[[1]]; k <- par[[2]] # Geradenparameter
> r <- -k; m0 <- ms/(exp(d)+1) # Parameter der logist. Funkt.
> # Bestimmtheitsmaß
> mexp <- ms/(1+(ms/m0-1)*exp(-r*t)); Srr <- sum((m-mexp)^2)
> Smm <- (n-1)*var(m); B <- 1-Srr/Smm
> print(cbind(k, d, ms, m0, r, Srr, B), digits=4)
      k      d  ms  m0  r   Srr    B
[1,] -0.5267 4.086 500 8.266 0.5267 311.7 0.9991
> # Abb. 2.13
> par(cex.axis=1.2, cex.lab=1.2)
> plot(t, m, type="p", xlab="Zeit t (in Stunden)",
+      ylab="Masse m (in mg)", pch=18, frame.plot=F,
+      xlim=c(0, 20), ylim=c(0, 600), lwd=2)
> segments(-2, ms, 17, ms, lty=2)
> text(1.5, ms, expression(hat(m)*" = 500"), pos=3, cex=1.2)
> curve(ms/(1+(ms/m0-1)*exp(-r*x)), xlim=c(0,17), lwd=2, ad=T)
> text(8.5, 250, pos=4, cex=1.2,
+      expression(hat(m)==frac(500, 1+59.5*e^{ -0.5267*t })))
> text(8.5, 185, expression("B = 99.9%"), pos=4, cex=1.2)
```

Der nahe bei 1 liegende Wert von B drückt bereits eine sehr gute Anpassung der logistischen Kurve aus. Wir untersuchen noch, wie sich die Restquadratsumme S_{rr} ändert, wenn der Sättigungswert m^* variiert wird. Für $\hat{m}^* = 501$ steigt S_{rr}

auf 436.3 an. Verringert man andererseits $\hat{m}^*$, so erhält man $S_{rr} = 211.3$ für $\hat{m}^* = 499$, $S_{rr} = 144.1$ für $\hat{m}^* = 498$, $S_{rr} = 124.3$ für $\hat{m}^* = 497$. Eine weitere Verringerung führt wieder zu einem Anstieg von S_{rr} . Bei Beschränkung auf ganzzahlige Sättigungswerte führt die Linearisierungsmethode auf die optimalen Parameter $\hat{m}^* = 497$, $\hat{m}_0 = 7.57$ und $\hat{r} = 0.5484$.

Wenn man die Summe der Residuenquadrate $(m_i - \hat{m}_i)^2$ direkt mit der logistischen Gleichung (2.17) minimiert, so erhält man die optimale Restquadratsumme $S_{rr} = 78.52$ und die Regressionsparameter $\hat{m}^* = 495.5$, $\hat{m}_0 = 6.568$ und $\hat{r} = 0.5654$. Die nicht-lineare Optimierung ist allerdings wesentlich aufwendiger und praktisch nur mit einer entsprechenden Software zu bewältigen. Die folgenden Zeilen geben die Lösung mit der R-Funktion `nls()` wieder.

```
> # Nichtlineare Optimierung mit automatischer Startwertsuche
> t <- seq(from=0, to=16, by=2)
> m <- c(7.5, 26, 53, 140, 276, 394, 453, 484, 492)
> n <- length(t)
> nlreg <- nls(m ~ SSlogis(t, ms, tw, gamma))
> npar <- coefficients(nlreg)
> ms <- npar[[1]]; tw <- npar[[2]]; gamma <- npar[[3]]
> r <- 1/gamma; m0 <- ms/(1+exp(r*tw))
> mexp <- ms/(1+(ms/m0-1)*exp(-r*t))
> Srr <- sum((m-mexp)^2); Smm <- (n-1)*var(m); B <- 1-Srr/Smm
> print(cbind(ms, m0, r, Srr, B), digits=4)
      ms    m0      r    Srr      B
[1,] 495.5 6.568 0.5654 78.52 0.9998
```

2.3.4 Logarithmusfunktionen

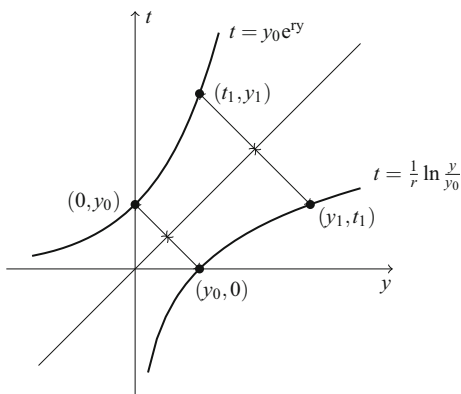
Durch die Gleichung $y = y_0 e^{rt}$ ($r \neq 0, y_0 > 0$) der in Abschn. 2.3.1 betrachteten Exponentialfunktion wird jedem reellen Wert von t bei positivem y_0 ein positiver Funktionswert y zugeordnet. Diese Zuordnung kann umgekehrt werden. Jedem positivem Wert von y entspricht genau eine reelle Zahl t . Indem man $y = y_0 e^{rt}$ logarithmiert, findet man für t die Darstellung:

$$t = t(y) = \frac{1}{r} \ln \frac{y}{y_0}.$$

Die durch diese Gleichung für reelle $y > 0$ definierte Funktion heißt **allgemeine Logarithmusfunktion**. Den Funktionsgraphen in der (y, t) -Ebene erhält man aus der in der (t, y) -Ebene dargestellten Exponentialfunktion, wenn man die t - mit der y -Achse vertauscht und die Exponentialkurve an der Geraden $y = t$ spiegelt (vgl. Abb. 2.14). Ist $y_0 = r = 1$, so spricht man speziell von der **natürlichen Logarithmusfunktion**. Durch geeignete Wahl von $r > 1$ kann man von der Basis e auf Logarithmen mit einer anderen Basis umsteigen. Setzt man beispielsweise $r = \ln 10$, so erhält man $t = \ln(y/y_0)/\ln 10 = \lg(y/y_0)$.

Die Logarithmusfunktion – wir schreiben ihre Gleichung in der Form $t = \ln(y/y_0)$ mit $r = 1$ an – besitzt folgende wichtige Eigenschaft: Wird von einem

Abb. 2.14 Zusammenhang zwischen Logarithmus- und Exponentialkurven



beliebigen Ausgangswert $y_1 > y_0$ ausgehend y_1 um einen bestimmten Faktor $f > 1$ auf $y_2 = f y_1$ vergrößert, so nimmt die abhängige Variable t stets um denselben (von y_1 unabhängigen) Betrag

$$t(y_2) - t(y_1) = \ln \frac{y_2}{y_0} - \ln \frac{y_1}{y_0} = \ln \frac{f y_1}{y_0} - \ln \frac{y_1}{y_0} = \ln f$$

zu. Das bedeutet, dass gleiche prozentuelle Änderungen einer Variablen y stets dieselbe (absolute) Änderung der von y logarithmisch abhängigen Variablen t bewirken.

Beispiel 2.14

- a) Aus der Erfahrung weiß man, dass gleichen prozentuellen Änderungen der Intensität I eines physikalischen Reizes wenigstens für mittlere Werte von I gleich große Änderungen in der Stärke R der Reizempfindung entsprechen. Dieser Umstand legt es nahe, für die Reizempfindung eine Skala zu benutzen, die mit der „Reizskala“ durch die logarithmische Transformation $R = c \lg(I/I_0)$ verknüpft ist (Weber-Fechnersche Formel). Dabei bedeuten c eine positive Skalenkonstante und I_0 die Reizschwelle, der der Nullpunkt der „Empfindungsskala“ zugeordnet ist.

Für die Wahrnehmung von Lautstärken ist z. B. $I_0 = 10^{-12} \text{ W/m}^2$ (Sinuston, 2 kHz). Man erhält die Lautstärke in dB (Dezibel), wenn $c = 10$ gesetzt wird. Da die Schallintensität proportional zum Quadrat des Schalldrucks p (in Pa) ist, gilt für die Lautstärke $L = 10 \lg(I/I_0) = 20 \lg(p/p_0)$. Hier ist p_0 der gerade noch hörbare Schalldruck, der der Referenz-Intensität I_0 entspricht. Man beachte, dass den äquidistanten Werten 0, 10, 20, ... der dB-Skala die Schallintensitäten $I_0, 10I_0, 100I_0, \dots$ entsprechen; jede Verzehnfachung der Schallintensität führt also zu einer Erhöhung der empfundenen Lautstärke um jeweils 10 dB.

- b) Eine ähnliche Situation wie bei der Bewertung von durch physikalische Reize ausgelösten Empfindungen liegt in der Pharmakologie bei der Untersuchung von

dosisabhängigen Effekten vor. Im allgemeinen kann erst ab einer bestimmten minimalen Dosis D_0 (der sogenannten Schwellendosis) ein Effekt E nachgewiesen werden. Der Effekt nimmt dann keineswegs linear mit wachsender Dosis D zu. Stellt man den Effekt (in % des maximalen Effekts) in Abhängigkeit von der logarithmisch aufgetragenen Dosis graphisch dar, so erhält man meist eine S -förmige *Dosis-Wirkungs-Kurve*, die in ihrem „mittleren Teil“ durch eine Gerade angenähert werden kann; das ist der Bereich, in dem die Weber-Fechnersche Formel $E = c \lg(D/D_0)$ gültig ist, d.h., gleiche prozentuelle Dosissteigerungen führen zu gleichen Zunahmen des Effekts.

2.3.5 Aufgaben

1. Zur Untersuchung der Elimination eines Pharmakons aus dem Plasma wurde das Pharmakon mit einer (im Vergleich zur Eliminationsdauer langlebigen) radioaktiven Substanz markiert und die zeitliche Abnahme der Aktivität A (in Mikrocurie) nach erfolgter intravenöser Injektion beobachtet. Es ergaben sich die folgenden Aktivitäten: 7.5 (20 Minuten nach Injektion), 6.4 (nach 40 Minuten), 4.3 (nach 80 Minuten), 2.4 (nach 120 Minuten), 1.6 (nach 160 Minuten) und 0.5 (nach 240 Minuten). Wie groß ist die Halbwertszeit $t_{1/2}$ für die Elimination des Medikaments, wenn die ausgeschiedene Pharmakonmenge (und damit auch die Aktivität) nach einer Exponentialfunktion abnimmt?
2. In einem Experiment zur Untersuchung des Wachstums einer *Drosophila*-Population (eingeschlossen in einer Flasche mit begrenzter Nahrungsbeigabe) wurden die folgenden Populationswerte y gezählt: 6 (1 Tag nach Beginn des Versuchs), 21 (nach 3 Tagen), 67 (nach 5 Tagen), 163 (nach 7 Tagen), 256 (nach 9 Tagen) und 319 (nach 11 Tagen). Man beschreibe den Wachstumsverlauf mit Hilfe der logistischen Funktion.
3. Es sei p die Wahrscheinlichkeit eines Ereignisses E (z. B. Tod auf Grund einer Erkrankung) und x eine metrischen Einflussgröße (z. B. Übergewicht). Wenn p für kleine Werte von x gering ist und mit wachsendem x gegen den Wert 1 strebt, kann zur Darstellung der Abhängigkeit der Wahrscheinlichkeit p von x die logistische Gleichung (2.15a) verwendet werden. Mit $y = p$, $t = x$, $y^* = 1$ und $0 < y_0 < 1$ ergibt sich aus (2.15a):

$$p = \frac{1}{1 + \left(\frac{1}{y_0} - 1\right)e^{-rx}}.$$

- a) Man zeige, dass $\ln \frac{p}{1-p}$ ($0 < p < 1$) linear von x abhängt.
- b) Die Abbildung $p \rightarrow \text{logit}(p) = \ln \frac{p}{1-p}$ ($0 < p < 1$) wird als Logit-Funktion bezeichnet. Man zeichne den Graphen der Logit-Funktion.

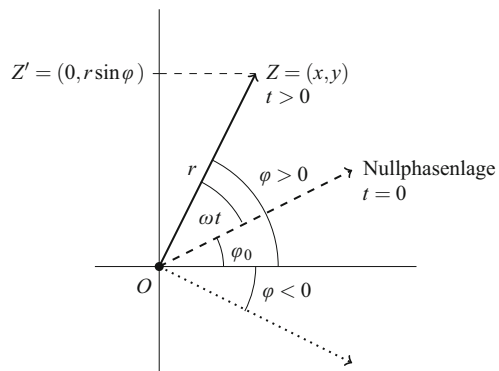
2.4 Sinusförmige Veränderungen

2.4.1 Die allgemeine Sinusfunktion

Die wichtigsten mathematischen Bausteine zur Beschreibung von periodischen Vorgängen (z. B. biologischen Rhythmen) sind die Sinus- und die Kosinusfunktion. In Abschn. 1.3 wurde der Sinus und Kosinus eines spitzen Winkels als das Verhältnis spezieller Seiten (nämlich von Gegenkathete zu Hypotenuse bzw. Ankathete zu Hypotenuse) im rechtwinkligen Dreieck eingeführt. Eine Verallgemeinerung auf beliebige Winkel ist auf der Basis des **Zeigermodells** möglich, das einen einfachen periodischen Vorgang wiedergibt (vgl. Abb. 2.15):

Ein im Nullpunkt O der (x, y) -Ebene fixierter **Zeiger** der Länge r rotiere im entgegengesetzten Uhrzeigersinn. Zum Zeitpunkt $t = 0$ schließe der Zeiger mit der positiven x -Achse den Winkel φ_0 ein; diesen Winkel bezeichnet man als den **Nullphasenwinkel** des Zeigers. Der Zeiger möge gleichförmig rotieren, d. h., in gleichen Zeitintervallen wird immer der gleiche „Winkelweg“ zurückgelegt; der pro Zeiteinheit zurückgelegte (und im Bogenmaß dargestellte) Winkelweg wird als **Winkelgeschwindigkeit** ω des Zeigers bezeichnet. Multipliziert man ω mit t , so erhält man den in t Zeiteinheiten überstrichenen Winkel. Der vom Zeiger in der Zeit von 0 (Nullphasenlage) bis t überstrichene Winkel ist also durch ωt gegeben. Jede Phase der Zeigerbewegung kann somit eindeutig durch den von der positiven x -Achse genommenen **Phasenwinkel** $\varphi = \omega t + \varphi_0$ beschrieben werden. Zur Nullphasenlage (Zeitpunkt $t = 0$) gehört der Phasenwinkel $\varphi = \varphi_0$; zum Zeitpunkt $t = T = 2\pi/\omega$ ist $\varphi = \omega T + \varphi_0 = 2\pi + \varphi_0$, d. h., der Zeiger hat von der Nullphasenlage aus gerechnet genau einen vollen Umlauf getan. Wieder T Zeiteinheiten später ist ein weiterer Umlauf beendet usw. Mit wachsendem t nimmt daher der Phasenwinkel beliebig große Zahlenwerte an. Auch vor der Nullphasenlage, also für negative Werte von t , kann die Zeigerbewegung mit dem Phasenwinkel erfasst werden, wenn wir vereinbaren, dass bei negativem Phasenwinkel der Betrag des Winkels von der positiven x -Achse weg im Uhrzeigersinn aufgetragen wird.

Abb. 2.15 Zeigerdarstellung der allgemeinen Sinusfunktion mit der Gleichung $y = r \sin \varphi = r \sin(\omega t + \varphi_0)$; die Zeigerlänge r ist die Amplitude, ω die Kreisfrequenz und φ_0 der Nullphasenwinkel



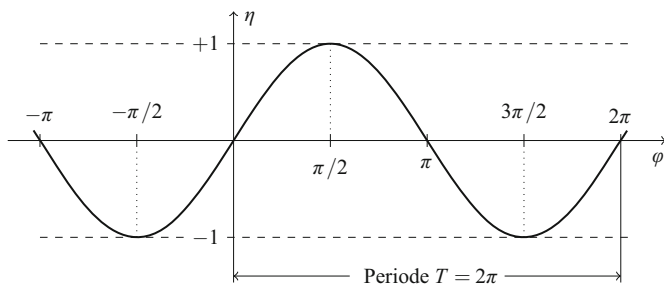


Abb. 2.16 Graph der Sinusfunktion $\varphi \rightarrow \eta$ mit $\eta(\varphi) = \sin \varphi$ (Sinuslinie)

Der bei einem vollen Umlauf zurückgelegte Winkelweg ist 2π . Die dazu benötigte Zeitdauer wird als die **Periode** T bezeichnet. Zwischen der Winkelgeschwindigkeit und der Periode besteht der Zusammenhang $\omega = 2\pi/T$. Die Bezeichnung Periode bringt zum Ausdruck, dass der rotierende Zeiger in allen Zeitpunkten $t_1, t_2, \dots$, die voneinander T Einheiten entfernt liegen, dieselbe Lage in der (x, y) -Ebene einnimmt; denn aus $t_2 - t_1 = T$ folgt z. B. wegen $\varphi_2 - \varphi_1 = \omega(t_2 - t_1)$, dass der vom Zeitpunkt t_1 bis zum Zeitpunkt t_2 überstrichene Winkelweg gleich $\omega T = 2\pi$ ist, was einem vollen Umlauf entspricht. Der Kehrwert f von T gibt die Anzahl der Perioden pro Zeiteinheit an und wird als **Frequenz** bezeichnet. Mit der Frequenz kann die Winkelgeschwindigkeit in der Form $\omega = 2\pi f$ dargestellt werden und wird daher auch als **Kreisfrequenz** bezeichnet.

Wenn ein Zeiger der Länge r gleichförmig rotiert, durchläuft die Zeigerspitze Z mit konstanter Geschwindigkeit einen Kreis mit dem Radius r . Bei Projektion auf die y -Achse geht die Kreisbewegung in eine **Schwingung** über, d. h., die Projektion Z' der Zeigerspitze vollführt auf der y -Achse eine auf- und abgehende Bewegung. Zur Beschreibung der Bewegung von Z' genügt es anzugeben, wie die y -Koordinate der Zeigerspitze vom Phasenwinkel φ und der Zeit t abhängt. An Hand des in Abb. 2.15 eingezeichneten rechtwinkligen Dreiecks OZZ' erkennt man, dass $\sin \varphi = y/r$ ist. Diese Interpretation ermöglicht es, die zunächst nur für spitze Winkel gültige Berechnungsvorschrift von Sinuswerten auf beliebige reelle Winkel zu erweitern.

Als **Sinus** wird jene Funktion definiert, die einem vorgegebenen Phasenwinkel φ den im Zeigermodell eindeutig bestimmten Wert $\sin \varphi = y/r$ zuordnet. Man prüft leicht nach, dass z. B. $\sin 0 = \sin 2\pi = 0$ ist oder $\sin(\pi/2) = 1$ und allgemein $-1 \leq \sin \varphi \leq 1$ gilt. Trägt man über der φ -Achse zu jedem φ -Wert den entsprechenden Sinuswert $\eta = y/r = \sin \varphi$ auf, so erhält man die in Abb. 2.16 dargestellte **Sinuslinie**, die zwischen dem Minimalwert -1 und dem Maximalwert $+1$ hin- und her pendelt. Da bei jedem Zeigerumlauf dieselben y -Werte erzeugt werden, genügt es, sich bei der Darstellung der Sinuslinie auf ein Winkelintervall der Länge 2π zu beschränken, z. B. auf das Intervall von 0 bis 2π .

Man spricht von einer sinusförmigen Schwingung, wenn sich eine Größe in Abhängigkeit von der Zeit so wie die betrachtete y -Koordinate der Spitze eines gleichförmig rotierenden Zeigers verhält. Die Zeitabhängigkeit wird durch die **allgemeine Sinusfunktion** mit der Gleichung

$$y = r \sin(\omega t + \varphi_0) \quad (2.18)$$

beschrieben. In diesem Zusammenhang heißt r die **Schwingungsamplitude**, ω die **Kreisfrequenz** und φ_0 wie bisher der Nullphasenwinkel.

Beispiel 2.15

Der Luftstrom Q (bewegte Luftmenge pro Zeiteinheit) beim Aus- und Einatmen ändert sich in Abhängigkeit von der Zeit näherungsweise nach einer Sinusfunktion. Die positive „Sinushalbwellen“ möge den Luftstrom beim Ausatmen und die negative den Luftstrom beim Einatmen wiedergeben. Für eine ruhende Person betrage die Dauer eines Atmungszyklus 5 s (Sekunden) und der maximale Luftstrom sei 0.5 l/s (Liter pro Sekunde). Wir bestimmen, wie Q von der Zeit t abhängt, wenn am Anfang ($t = 0$) die Lungen leer sind.

Unmittelbar aus den Angaben folgt die Amplitude $r = 0.5$ l/s und wegen $T = 5$ (in s) die Kreisfrequenz $\omega = 2\pi/T = 0.4\pi$ (in s^{-1}). Wenn die Lungen am Anfang leer sind, stehen wir am Beginn der Einatmungsphase, die durch eine negative Sinushalbwellen beschrieben wird. Somit ist $Q = 0.5 \sin(0.4\pi t + \pi)$ in l/s.

Bei der graphischen Darstellung der allgemeinen Sinusfunktionen wird horizontal meist ωt anstelle von t aufgetragen. Es erweist sich als zweckmäßig, zuerst die zur ωt -Achse parallelen Geraden $y = +r$ bzw. $y = -r$ einzuzichnen, zwischen denen die Sinuslinie schwankt. Aus dem Zeigermodell ist unmittelbar zu erkennen, dass an die Phasenlage $\omega t = -\varphi_0$ eine positive Halbwellen der Sinuslinie anschließt. Man zeichnet also von der Stelle $-\varphi_0$ auf der ωt -Achse ausgehend eine positive Sinushalbwellen (bis zur Stelle $-\varphi_0 + \pi$), an die – nach beiden Seiten – abwechselnd negative und positive Halbwellen anzufügen sind. Für eine genauere Darstellung wird man sich zusätzlich an einigen Stellen Kreisfunktionswerte berechnen und die entsprechenden Punkte in das $(\omega t, y)$ -Diagramm eintragen.

Beispiel 2.16

Der (langjährige) monatliche Mittelwert M der Lufttemperatur schwankt im Verlaufe eines Jahres in guter Näherung sinusförmig um das Jahresmittel J . Für Wien ist beispielsweise $J = 9.83^\circ\text{C}$ und die monatlichen Schwankungen $A = M - J$ (in Kelvin) können durch $A = 10.45 \sin(2\pi t/12 - 1.86)$ dargestellt werden; dabei ist t die Zeit in Monaten ($t = 0.5$ bezeichnet die Mitte des Monats Januar, $t = 1.5$ die Mitte des Monats Februar usw.).

In Abb. 2.17 ist horizontal $x = \omega t = 2\pi t/12$ aufgetragen. Die eingezeichneten Skalenstriche kennzeichnen die jeweiligen Monatsmitten. Aus $2\pi t/12 - 1.86 =$

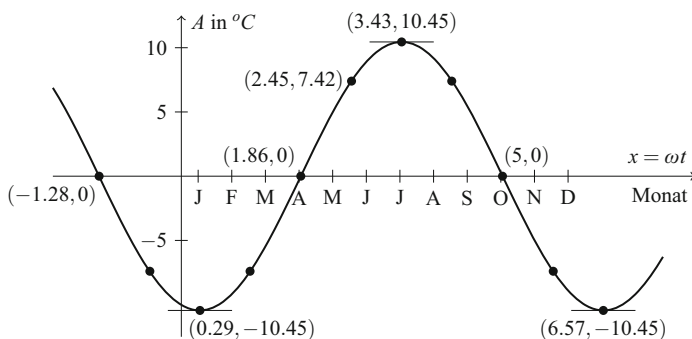


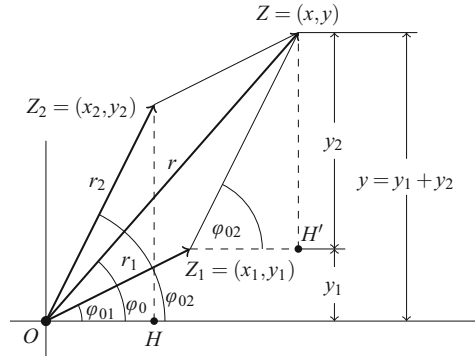
Abb. 2.17 Verlauf des monatlichen Temperaturmittels M (in $^{\circ}\text{C}$, Wien) um den Jahresmittelwert $J = 9.83$ als allgemeine Sinusfunktion: $M - J = 10.45 \sin(\pi/6t - 1.86)$

$x - 1.86 = 0$ folgt die Nullstelle 1.86; zwischen den Skalenwerten 1.86 und $1.86 + \pi = 5.00$ liegt eine positive Halbwelle der Sinusfunktion, die für $x = (1.86 + 5.00)/2 = 3.43$ den Maximalwert 10.45 annimmt. An die positive Halbwelle anschließend liegt zwischen 5.00 und $5.00 + \pi = 8.14$ bzw. $1.86 - \pi = -1.28$ und 1.86 jeweils eine negative Halbwelle (die zugehörigen Minimalwerte werden für $(5.00 + 8.14)/2 = 6.57$ bzw. $(-1.28 + 1.86)/2 = 0.29$ angenommen). Um die Funktion genauer zeichnen zu können, empfiehlt es sich, noch weitere Funktionswerte zu berechnen, etwa für die Mitten zwischen den Nullstellen und den Extremstellen. So findet man z. B. für $x = (1.86 + 3.43)/2 = 2.65$ den Funktionswert $10.45 \sin(2.65 - 1.86) = 7.42$.

In Beispiel 2.16 wurden die Abweichungen des Monatsmittels M vom Jahresmittel J der Lufttemperatur durch die Sinusfunktion $M - J = 10.45 \sin(2\pi t/12 - 1.86)$ erfasst. Es folgt, dass der monatliche Mittelwert der Lufttemperatur durch $M = J + 10.45 \sin(2\pi t/12 - 1.86)$ gegeben ist. Sinusförmige Schwankungen einer Größe y um einen festen Wert y_0 können also durch die Funktionsgleichung $y = y_0 + r \sin(\omega t + \varphi_0)$ dargestellt werden.

Aus der allgemeinen Sinusfunktion folgt die in Abb. 2.16 dargestellte Sinusfunktion, wenn man die Amplitude gleich 1 und den Nullphasenwinkel gleich 0 setzt. Dieser Festlegung entspricht im Zeigermodell der sogenannte **Sinuszeiger** (Zeiger der Länge 1, der zum Zeitpunkt $t = 0$ in die positive x -Richtung weist). Ein anderer wichtiger Sonderfall ergibt sich, wenn man die Amplitude wieder 1, aber den Nullphasenwinkel gleich $\pi/2$ wählt. Der entsprechende Zeiger eilt dem Sinuszeiger stets um den Winkel $\pi/2$ voraus und wird als **Kosinuszeiger** bezeichnet. Durch Projektion der Zeigerspitze auf die y -Achse erhält man nunmehr eine Schwingung, die durch die **Kosinusfunktion** mit der Gleichung $y = \cos \omega t = \sin(\omega t + \pi/2)$ beschrieben wird. Damit kann auch die **Tangensfunktion** über die Vorschrift $\tan \omega t = \sin \omega t / \cos \omega t$ auf beliebige Winkel ωt , für die $\cos \omega t \neq 0$ ist, erweitert werden.

Abb. 2.18 Geometrische Addition von Zeigern nach der Parallelogrammmethode



2.4.2 Überlagerung von Sinusschwingungen

Wir betrachten die durch die Gleichungen

$$y_1 = r_1 \sin(\omega t + \varphi_{01}) \quad \text{und} \quad y_2 = r_2 \sin(\omega t + \varphi_{02})$$

beschriebenen Schwingungen derselben Kreisfrequenz ω . Wenn man die Schwingungen überlagert, also die Summe $y = y_1 + y_2$ bildet, so erhält man wieder eine sinusförmige Schwingung mit der Kreisfrequenz ω . Dies erkennt man schnell mit Hilfe des Zeigermodells. Den beiden zu überlagernden Schwingungskomponenten entsprechen zwei mit der Winkelgeschwindigkeit ω rotierende Zeiger, deren Längen durch r_1 bzw. r_2 und deren Nullphasenwinkel durch φ_{01} bzw. φ_{02} gegeben sind. Die beiden Zeiger setzen wir nach der **Parallelogrammmethode**, so wie in Abb. 2.18 vorgezeigt, zu einem resultierenden Zeiger zusammen. Dabei ist zu beachten, dass die Dreiecke OHZ_2 und $Z_1H'Z$ kongruent (deckungsgleich) sind. Man erkennt, dass die Spitze Z des resultierenden Zeigers in jedem Zeitpunkt gerade die Summe $y = y_1 + y_2$ als y -Koordinate besitzt. Es folgt aus dieser Interpretation, dass sich $y = y_1 + y_2$ als eine allgemeine Sinusfunktion in der Form $y = r \sin(\omega t + \varphi_0)$ darstellen lässt mit von r_1, r_2, φ_{01} sowie φ_{02} abhängigen Parametern r und φ_0 .

Wichtig für die Praxis ist der Sonderfall, dass $y_1 = r_1 \sin \omega t$ und $y_2 = r_2 \cos \omega t$ zu addieren sind. Durch geometrische Addition der y_1 und y_2 entsprechenden Zeiger findet man für die Länge r des resultierenden Zeigers $y = y_1 + y_2$

$$r = \sqrt{r_1^2 + r_2^2}. \quad (2.19a)$$

Der Nullphasenwinkel von y ist die zwischen $-\pi/2$ und $\pi/2$ liegende Lösung

$$\varphi_0 = \arctan\left(\frac{r_2}{r_1}\right) \quad (2.19b)$$

der Gleichung $\tan \varphi_0 = r_2/r_1$. Damit kann die Summe $y = r_1 \sin \omega t + r_2 \cos \omega t$ in die allgemeine Sinusfunktion $r \sin(\omega t + \varphi_0)$ umgerechnet werden. Umgekehrt kann man eine allgemeine Sinusfunktion der Gestalt $r \sin(\omega t + \varphi_0)$ in die Summe $r_1 \sin \omega t + r_2 \cos \omega t$ mit $r_1 = r \cos \varphi_0$ und $r_2 = r \sin \varphi_0$ entwickeln.

2.4.3 Kurvenanpassung bei periodischen Daten

Für viele Zwecke ist es ausreichend, einen in der Zeit t periodischen Vorgang (mit der bekannten Periodendauer T) näherungsweise durch eine sinusförmige Schwingung (mit der Kreisfrequenz $\omega = 2\pi/T$) um einen stationären Wert a_0 wiederzugeben. Die gesuchte Näherungsfunktion setzen wir in der Form

$$y(t) = a_0 + r_1 \sin(\omega t + \varphi_{01}) = a_0 + a_1 \cos \omega t + b_1 \sin \omega t \quad (2.20a)$$

mit zunächst unbestimmten Konstanten a_0 , r_1 und φ_{01} bzw. a_0 , a_1 und b_1 an. Von dem zu beschreibenden Vorgang mögen in einer Periode $n > 3$ Beobachtungswerte $w_1, w_2, \dots, w_n$ bekannt sein. Die entsprechenden Beobachtungszeitpunkte (man spricht auch von Stützstellen) seien $t_1, t_2, \dots, t_n$. Aus den Beobachtungsdaten sind die Parameter a_0 , a_1 und b_1 der Näherungsfunktion zu bestimmen. Dazu verlangen wir wie bei den in den vorangehenden Abschnitten betrachteten Ausgleichsproblemen, dass die über alle Beobachtungswerte erstreckte Summe $Q(a_0, a_1, b_1)$ der Quadrate der Abweichungen $w_i - y(t_i)$ minimal wird (Methode der kleinsten Quadrate). Die Rechnung vereinfacht sich, wenn die t -Werte äquidistant liegen, was wir im folgenden voraussetzen wollen. Die Minimierung der Quadratsumme $Q(a_0, a_1, b_1)$ führt auf die Formeln

$$a_0 = \frac{1}{n} \sum_{i=1}^n w_i, \quad a_1 = \frac{2}{n} \sum_{i=1}^n w_i \cos \omega t_i, \quad \text{und} \quad b_1 = \frac{2}{n} \sum_{i=1}^n w_i \sin \omega t_i. \quad (2.20b)$$

Damit kann auch die Amplitude r_1 und der Nullphasenwinkel φ_{01} der allgemeinen Sinusfunktion berechnet werden, mit der der vorgegebene periodische Vorgang approximiert wird. Im Sonderfall $b_1 = 0$ gilt $r_1 = |a_1|$ und $\varphi_{01} = +\pi/2$ für $a_1 > 0$ bzw. $\varphi_{01} = -\pi/2$ für $a_1 < 0$. Im Fall $b_1 \neq 0$ gelten die Formeln:

$$r_1 = \sqrt{a_1^2 + b_1^2}, \quad \varphi_{01} = \begin{cases} \arctan(a_1/b_1) & \text{für } b_1 > 0 \\ \arctan(a_1/b_1) - \pi & \text{für } b_1 < 0. \end{cases} \quad (2.20c)$$

Beispiel 2.17

Die langjährigen Monatsmittelwerte (Januar bis Dezember) der Lufttemperatur (in °C) von Wien sind: -1, 0, 5, 10, 15, 18, 20, 19, 16, 10, 5 und 1. Wir verwenden die Näherungsfunktion $y = a_0 + a_1 \cos \omega t + b_1 \sin \omega t$ mit $\omega = 2\pi/12 = \pi/6 =$

0.5236, um den Verlauf der mittleren Monatstemperatur zu approximieren. Mit der Variablen t wird die Zeit in Monaten ab Jahresbeginn gemessen; die angegebenen Temperaturwerte $w_1 = -1, w_2 = 0, \dots, w_{12} = 1$ ordnen wir den Monatsmitten $t_1 = 0.5, t_2 = 1.5, \dots, t_{12} = 11.5$ zu. Mit den Formeln (2.20b) erhält man

$$\begin{aligned} a_0 &= \frac{1}{12}(-1 + 0 + 5 + \dots) = 9.83, \\ a_1 &= \frac{2}{12}(-1 \cdot \cos(0.5\pi/6) + \dots + 1 \cdot \cos(11.5\pi/6)) = -10.01, \\ b_1 &= \frac{2}{12}(-1 \cdot \sin(0.5\pi/6) + \dots + 1 \cdot \sin(11.5\pi/6)) = -3.00 \end{aligned}$$

und weiter mit den Formeln (2.20c)

$$\begin{aligned} r_1 &= \sqrt{(-10.01)^2 + (-3)^2} = 10.45, \\ \varphi_{01} &= \arctan \frac{-10.01}{-3} - \pi = -1.86. \end{aligned}$$

Die Approximationsfunktion für den Verlauf des monatlichen Temperaturmittels lautet daher

$$\begin{aligned} y(t) &= 9.83 - 10.01 \cos(\pi t/6) - 3.00 \sin(\pi t/6) \\ &= 9.83 + 10.45 \sin(\pi t/6 - 1.86). \end{aligned}$$

Die Funktion $y(t) - 9.83 = 10.45 \sin(\pi t/6 - 1.86)$ wurde bereits in Abb. 2.17 graphisch dargestellt. Um die Güte der Approximation beurteilen zu können, sind in Tab. 2.5 neben den Beobachtungszeitpunkten t_i und den beobachteten Temperaturwerten w_i die Quadrate der zentrierten Temperaturen $w_i - a_0$, die mit der Ausgleichsfunktion geschätzten Temperaturmittel $y(t_i)$ sowie die Quadrate der Residuen $w_i - y(t_i)$ aufgelistet. Als Summe der Quadrate der Residuen entnimmt man der Tabelle $S_{rr} = 2.352$; die Summe S_{ww} der Quadrate der Abweichungen der beobachteten w_i -Werte von ihrem Mittelwert a_0 ist $S_{ww} = 657.67$. Somit ergibt sich das Bestimmtheitsmaß $B = 1 - S_{rr}/S_{ww} = 99.6\%$.

```
> # Lösung mit R:
> t <- seq(from=0.5, to=11.5, by=1)
> w <- c(-1, 0, 5, 10, 15, 18, 20, 19, 16, 10, 5, 1)
> n <- length(w); omega <- 2*pi/n
> a0 <- mean(w); a1 <- 2*sum(w %c% cos(omega*t))/n
> b1 <- 2*sum(w %c% sin(omega*t))/n
> r1 <- sqrt(a1*a1+b1*b1)
> phi01 <- atan(a1/b1); if (b1<0) {phi01 <- phi01-pi}
> yexp <- a0+r1*sin(omega*t+phi01)
> res2 <- (w-yexp)^2; Srr <- sum(res2)
> Sww <- (n-1)*var(w); B <- 1-Srr/Sww
> print(cbind(a0, a1, b1, r1, phi01, B), digits=4)
      a0      a1      b1      r1  phi01      B
[1,] 9.833 -10.01 -3.004 10.45 -1.862 0.9964
```

Tab. 2.5 Rechenschema zu Beispiel 2.17. Die t_i - und w_i -Spalten enthalten die Beobachtungszeitpunkte (Monatsmitten) und die monatlichen Temperaturmittel. Die weiteren Spalten werden zur Berechnung des Bestimmtheitsmaßes verwendet; die entsprechenden Tabellenwerte und Spaltensummen sind gerundet

i	t_i	w_i	$(w_i - a_0^2)$	$y(t_i)$	$(w_i - y(t_i))^2$
1	0.5	-1	117.36	-0.61	0.150
2	1.5	0	96.69	0.63	0.398
3	2.5	5	23.36	4.34	0.434
4	3.5	10	0.03	9.52	0.228
5	4.5	15	26.69	14.79	0.045
6	5.5	18	66.69	18.72	0.525
7	6.5	20	103.36	20.28	0.078
8	7.5	19	84.03	19.04	0.001
9	8.5	16	38.03	15.33	0.455
10	9.5	10	0.03	10.14	0.021
11	10.5	5	23.36	4.88	0.014
12	11.5	1	78.03	0.94	0.003
Σ		118	657.67		2.354

Das Funktion (2.20a) liefert bei zwar periodischen, aber nicht sinusförmigen Oszillationen um einen stationären Wert nur eine mehr oder weniger grobe Approximation in Form der sogenannten *dominanten Schwingungskomponente* $r_1 \sin(2\pi f t + \varphi_{01})$ mit der *Grundfrequenz* $f = 1/T$. Bessere Approximationen können erreicht werden, wenn man Komponenten hinzufügt, die mit den Frequenzen $2f, 3f, \dots, pf$ ($p < (n-1)/2$) schwingen. Auf diese Weise erhält man das **trigonometrische Polynom**

$$a_0 + r_1 \sin(\omega t + \varphi_{01}) + \dots + r_p \sin(p\omega t + \varphi_{0p}).$$

Die Anpassung dieser Funktion an periodische Daten kann wieder nach der Methode der kleinsten Quadrate vorgenommen werden (vgl. z. B. Hämmerlein & Hoffmann 1994).

2.4.4 Aufgaben

- Der Verlauf der langjährigen Monatsmittelwerte y (Jänner bis Dezember) der Lufttemperatur (in °C) kann für Sydney näherungsweise durch $y = 17.6 + 4.5 \cos(0.524t) + 1.7 \sin(0.524t)$ wiedergegeben werden. In dieser Formel sind die Monate der Reihe nach durch die t -Werte 0.5, 1.5 usw. erfasst.
 - Man fasse die Terme $4.5 \cos(0.524t)$ und $1.7 \sin(0.524t)$ in einer allgemeinen Sinusfunktion zusammen.
 - In welchem Monat ist die mittlere Lufttemperatur am größten?

2. Im Jahre 1989 wurden in der Donau (bei Ybbs-Persenbeug) die folgenden, monatlichen Temperaturwerte y_i gemessen (jeweils zur Monatsmitte, Angaben in °C): 4.5, 4.0, 6.7, 10.5, 12.6, 15.1, 17.6, 16.0, 15.5, 10.4, 7.3, 1.4. Man approximiere die Abhängigkeit der Wassertemperatur T von der Zeit t durch ein trigonometrisches Polynom der Gestalt (2.20a).

Hinweis: Man erfasse die Monate Jänner bis Dezember durch die t -Werte 0.5, 1.5 usw.

2.5 Wahrscheinlichkeitsfunktionen

2.5.1 Diskrete Zufallsvariable

Zur Einführung betrachten wir ein einfaches Zufallsexperiment. Eine Münze mit den Ausgängen K (Kopf) und Z (Zahl) wird zweimal geworfen. Wir stellen die Frage, wie oft dabei ein K auftritt. Mit dieser Frage haben wir implizit eine Variable eingeführt, nämlich die Anzahl X der Versuche mit dem Ausgang K . Die Ergebnismenge des Zufallsexperimentes ist $\Omega = \{(K, K), (K, A), (A, K), (A, A)\}$, wobei das Element (K, A) die Ausgänge K und A beim ersten bzw. zweiten Werfen bedeutet usw. Offensichtlich kann X die Werte 0, 1 oder 2 annehmen. Welchen Wert X annimmt, kann nicht vorhergesagt werden, der Wert von X kommt zufällig zustande. Man bezeichnet X daher als eine Zufallsvariable. Man beachte, dass durch die Zufallsvariable jedem Element von Ω eine reelle Zahl zugeordnet wird und umgekehrt durch die Vorgabe eines möglichen Wertes von X ein Element oder mehrere Elemente von Ω bestimmt werden.

In Verallgemeinerung des betrachteten Experimentes sei nun Ω die Ergebnismenge eines Zufallsexperimentes mit endlich vielen oder abzählbar unendlich vielen Ausgängen. Jedem Ausgang ω des Zufallsexperimentes möge durch die Funktion X eine reelle Zahl $X(\omega)$ zugeordnet werden. Man nennt X eine **diskrete Zufallsvariable**¹² und die Werte, die X bei jeder Ausführung des Zufallsexperimentes annimmt, Realisierungen von X . Ist x eine Realisierung von X , so bildet die durch $E_x = \{\omega | X(\omega) = x\}$ bestimmte Menge eine Teilmenge (d. h. ein Ereignis) von Ω . Mit dem auf Ω definierten Wahrscheinlichkeitsmaß P wird die Wahrscheinlichkeit dieses Ereignisses auch kurz durch $P(E_x) = P(X = x)$ dargestellt.

Die Variation der diskreten Zufallsvariablen X wird durch Angabe der Menge M der (höchstens abzählbar unendlich vielen) möglichen Werte $x_1, x_2, \dots$ und durch Angabe der Wahrscheinlichkeiten $p_i = P(X = x_i)$ für $x_i \in M$ beschrieben. Die Funktion $f : M \rightarrow [0, 1]$ mit $f(x_i) = p_i = P(X = x_i)$ für $x = x_i \in M$ und $f(x) = 0$ für $x \notin M$ heißt **Wahrscheinlichkeitsfunktion** von X . Man beachte, dass $p_1 + p_2 + \dots = 1$ ist. Zur grafischen Darstellung der Wahrscheinlichkeits-

¹² Zufallsvariable werden meist mit lateinischen Großbuchstaben (z. B. X) bezeichnet, die Werte mit den entsprechenden (indizierten) Kleinbuchstaben (z. B. $x_1, x_2, \dots$). Im vorhin betrachteten Beispiel hat X die Bedeutung eines Zählmerkmals.

funktion werden über der X -Achse an den Stellen x_i ($i = 1, 2, \dots$) Strecken der Länge p_i aufgetragen, wodurch ein aus einer Folge von Strecken bestehendes **Stabdiagramm** entsteht. Die Summe der Längen aller Strecken ist 1. Denkt man sich die Strecken mit Masse belegt, so bringt das Stabdiagramm bildlich zum Ausdruck, wie die gesamte „Wahrscheinlichkeitsmasse“ (nämlich 1) einer diskreten Zufallsvariablen auf der reellen Achse verteilt ist. Der Vergleich mit einer Masseverteilung ist auch bei der Interpretation von Verteilungskennwerten (z. B. des Mittelwerts) nützlich.

Beispiel 2.18

Ein einfaches Beispiel einer diskreten Zufallsvariablen ist die Bernoulli-Variable X , mit der Bernoulli-Experimente beschrieben werden. Das sind Zufallsexperimente mit zwei Ausgängen ω_1 und ω_2 , die mit gewissen Wahrscheinlichkeiten p bzw. $1 - p$ eintreten. Ordnet man X den Wert 1 zu, wenn das Ereignis $\{\omega_1\}$ eintritt, und andernfalls den Wert 0, so erhält man die sogenannte Zweipunktverteilung mit den durch $P(X = 1) = p$ und $P(X = 0) = 1 - p$ gegebenen Wahrscheinlichkeiten. Das Ausspielen eines Würfels kann z. B. als Bernoulli-Experiment gedeutet werden, wenn man unter ω_1 den Ausgang „Sechser“ versteht und unter ω_2 die übrigen Ausgänge zusammenfasst.

2.5.2 Binomialverteilung

Man betrachte ein Bernoulli-Experiment, das n -mal ($n \geq 1$) ausgeführt wird. Bei jeder einzelnen Ausführung des Experiments möge der Ausgang E mit der Wahrscheinlichkeit p eintreten.¹³ Den Versuchsausführungen ordnen wir die Zufallsvariablen $X_1, X_2, \dots, X_n$ zu; jede dieser Zufallsvariablen möge den Wert 1 annehmen, wenn die entsprechende Ausführung den Ausgang E zeigt, andernfalls den Wert 0. Die Summe $X = X_1 + X_2 + \dots + X_n$ bedeutet dann die Anzahl der Versuchsausführungen mit dem Ausgang E .

Wie man leicht erkennt, kann die Zufallsvariable X einen Wert aus der Menge $M = \{0, 1, \dots, n\}$ annehmen. Es gilt: Der Wert x aus M wird mit der Wahrscheinlichkeit

$$P(X = x) = B_{n,p}(x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad (2.21)$$

angenommen. Die Funktion $f : \mathbb{R} \rightarrow [0, 1]$ mit $f(x) = B_{n,p}(x)$ für $x \in M$ und $f(x) = 0$ für $x \notin M$ wird als Wahrscheinlichkeitsfunktion der **Binomialver-**

¹³ Dabei wird vorausgesetzt, dass die Versuchsausführungen voneinander unabhängig sind, d. h. das Ergebnis einer Ausführung hängt nicht von den Ergebnissen der vorangehenden Ausführungen ab. Das Ergebnis E wird oft als „Erfolg“ und entsprechend p als Erfolgswahrscheinlichkeit bezeichnet.

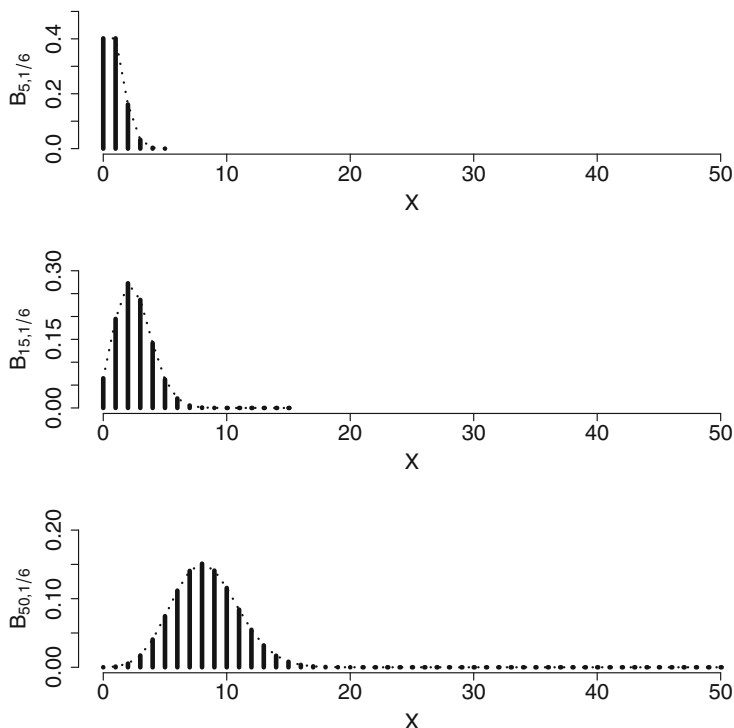


Abb. 2.19 Verteilung der Anzahl X der Sechser beim n -maligen Ausspielen eines Würfels. X ist binomialverteilt mit den Parametern n und $p = 1/6$. Die Verteilungen sind für $n = 5, 15$ und 50 durch Stabdiagramme dargestellt. Man erkennt, dass mit wachsendem n die Form der Verteilung zunehmend symmetrisch wird

teilung mit den Parametern n und p bezeichnet.¹⁴ Für eine mit den Parametern n und p binomialverteilte Zufallsvariable X schreibt man kurz $X \sim B_{n,p}$. Abb. 2.19 zeigt als Beispiel einer binomialverteilten Zufallsvariablen die Verteilung der Anzahl der Sechser beim n -maligen Ausspielen eines Würfels für verschiedene Werte von n . Die Binomialverteilung mit den Parametern $n = 1$ und p wird speziell als **Bernoulli-Verteilung** bezeichnet (dieser Sonderfall entspricht der 1-maligen Ausführung eines Bernoulli-Experimentes).

Mit der Binomialverteilung werden Zufallsexperimente vom Typ „Auswählen mit Zurücklegen“ modelliert. Bei diesem Modell eines Zufallsexperimentes stel-

¹⁴ Die Bezeichnung weist darauf hin, dass die Binomialwahrscheinlichkeiten (2.21) in der Entwicklung des Binoms $(q + p)^n$ mit $q = 1 - p$ auftreten. So erhält man z. B. für $n = 3$ die Entwicklung $(q + p)^3 = 1 \cdot q^3 p^0 + 3 \cdot q^2 p^1 + 3 \cdot q^1 p^2 + 1 \cdot q^0 p^3 = B_{3,p}(0) + B_{3,p}(1) + B_{3,p}(2) + B_{3,p}(3) = (1 - p + p)^3 = 1$. Man beachte, dass allgemein die Summe $B_{n,p}(0) + B_{n,p}(1) + \dots + B_{n,p}(n)$ der Binomialwahrscheinlichkeiten stets gleich eins ist. Dies folgt unmittelbar aus dem binomischen Lehrsatz.

le man sich z. B. eine „Urne“ mit N Elementen vor, von denen eine bestimmte Zahl a eine weiße Farbe und die restlichen $N - a$ eine rote Farbe aufweisen. Die Wahrscheinlichkeit, ein weißes Element auszuwählen, ist $p = a/N$. Legt man jedes ausgewählte Element wieder zurück, bleibt die Wahrscheinlichkeit für die Auswahl eines weißen Elementes unverändert gleich p . Die Wahrscheinlichkeit, bei zufälliger Auswahl von n Elementen x weiße und $n - x$ rote in einer bestimmten Reihenfolge auszuwählen, ergibt sich mit Hilfe der Multiplikationsregel für unabhängige Ereignisse als $p^x(1 - p)^{n-x}$. Da es nun $\binom{n}{x}$ verschiedene Möglichkeiten gibt, die x weißen und $n - x$ roten Elemente anzuordnen, ist das Ereignis $A =$ „Auswahl von x weißen Elementen bei n -maligem Auswählen mit Zurücklegen“ die Vereinigung der $\binom{n}{x}$ disjunkten Ereignisse, ein bestimmtes Anordnungsmuster der weißen und roten Elemente (z. B. zuerst x weiße gefolgt von $n - x$ roten) zu erhalten. Somit ist $P(A) = B_{n,p}(x)$.

Der Funktionsterm $B_{n,p}(x)$ gibt die Wahrscheinlichkeit an, dass eine mit den Parametern n und p binomialverteilte Zufallsvariable X genau den Wert x ($x = 0, 1, \dots, n$) annimmt. Häufig tritt die Wahrscheinlichkeit $P(X \leq x)$ auf, dass X einen Wert kleiner oder gleich x annimmt. Für diese Wahrscheinlichkeit – wir bezeichnen sie kurz mit $F_B(x)$ – erhält man mit Hilfe der Additionsregel

$$\begin{aligned} F_B(x) &= P(X \leq x) = P(X = 0) + P(X = 1) + \dots + P(X = x) \\ &= B_{n,p}(0) + B_{n,p}(1) + \dots + B_{n,p}(x). \end{aligned} \quad (2.22)$$

Man bezeichnet die durch (2.22) in Verbindung mit $F_B(x) = 0$ für $x \notin \{0, 1, \dots, n\}$ definierte Funktion $F_B : \mathbb{R} \rightarrow [0, 1]$ als **Verteilungsfunktion** der Binomialverteilung.

Beispiel 2.19

Bei einem Herstellungsverfahren ist die Wahrscheinlichkeit, dass die Durchmesser D einer Injektionsnadel die obere Toleranzgrenze T_o überschreitet oder die untere Toleranzgrenze T_u unterschreitet jeweils gleich 2,5 %. Wir bestimmen die Wahrscheinlichkeit P , dass von $n = 10$ hergestellten Nadeln höchstens eine außerhalb des Toleranzbereiches liegt.

Für die Wahrscheinlichkeit, dass entweder $D > T_o$ oder $D < T_u$ gilt, ergibt sich mit der Additionsregel für disjunkte Ereignisse $p = 5\%$. Es folgt, dass in einer Stichprobe mit dem Umfang $n = 10$ die Anzahl X der Nadeln mit $D > T_o$ oder $D < T_u$ binomialverteilt ist mit den Parametern $n = 10$ und $p = 0.05$. Gesucht ist die Wahrscheinlichkeit P des Ereignisses, dass X kleiner oder gleich 1, also entweder 0 oder 1 ist. Wegen

$$\begin{aligned} P(X = 0) &= B_{10,0.05}(0) = \binom{10}{0} 0.05^0 0.95^{10} = 0.95^{10} = 0.599, \\ P(X = 1) &= B_{10,0.05}(1) = \binom{10}{1} 0.05^1 0.95^9 = 10 \cdot 0.05 \cdot 0.95^9 = 0.315 \end{aligned}$$

ergibt sich $P = P(X = 0) + P(X = 1) = 0.914$.

```

> # Lösung mit R:
> # Berechnung mit der Wahrscheinlichkeitsfunktion dbinom()
> P <- dbinom(0,10,0.05)+dbinom(1,10,0.05)
> print(P, digits=3)
[1] 0.914
> # Berechnung mit der Verteilungsfunktion pbinom()
> P <- pbinom(1,10,0.05); print(P, digits=3)
[1] 0.914

```

2.5.3 Hypergeometrische Verteilung

Mit der hypergeometrischen Verteilung werden Zufallsexperimente vom Typ „**Auswählen ohne Zurücklegen**“ beschrieben. Darunter ist folgendes Ziehungs-schemata zu verstehen: Es sei M eine Menge von N Elementen, von denen $a < N$ vom Typ A (diese bilden die Teilmenge M_A) und $b = N - a$ vom Typ B (diese bilden die Teilmenge M_B) sind. Die Teilmengen M_A und M_B sind disjunkt und ergeben vereinigt die Menge M . Aus der Menge M werden aufs Geratewohl (und ohne Zurücklegen) n ($n \leq N$) Elemente gezogen. Wir fragen nach der Anzahl X der gezogenen Elemente, die vom Typ A sind, und berechnen die Wahrscheinlichkeit $P(X = x)$, dass sich unter den gezogenen Elementen genau x Elemente vom Typ A (und folglich $n - x$ Elemente vom Typ B) befinden. Der Auswahlvorgang ist schematisch in Abb. 2.20 dargestellt. Die Werte von X sind offensichtlich ganzzahlig; wenn $N - a \geq n$ ist, ist der kleinste mögliche Wert von X gleich null, andernfalls $n - (N - a)$. Wenn $a \geq n$ ist, ist der größte mögliche Wert von X gleich n , andernfalls a . Die möglichen Werte von X liegen also im Intervall $[x_{\min}, x_{\max}]$ mit $x_{\min} = \max(0, n - (N - a))$ und $x_{\max} = \min(n, a)$.

Zur Berechnung der Wahrscheinlichkeit $P(X = x)$ verwenden wir die Laplace-Formel (1.14). Jedes Element der Ergebnismenge Ω des betrachteten Zufallsexperiments (Zufallsziehung von n Elementen aus M ohne Zurücklegen) ist eine n -Kombination der N Elemente aus M . Die Anzahl dieser Kombinationen ist $C(n, N) = \binom{N}{n}$. Das Ereignis E , genau x Elemente des Typs A und $n - x$ Elemente des Typs B (ohne Berücksichtigung der Reihenfolge) zu ziehen, tritt ein, wenn

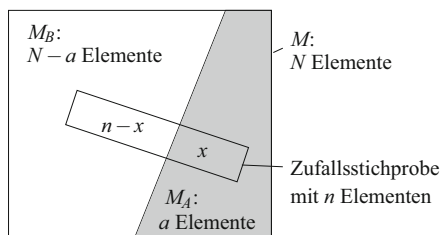
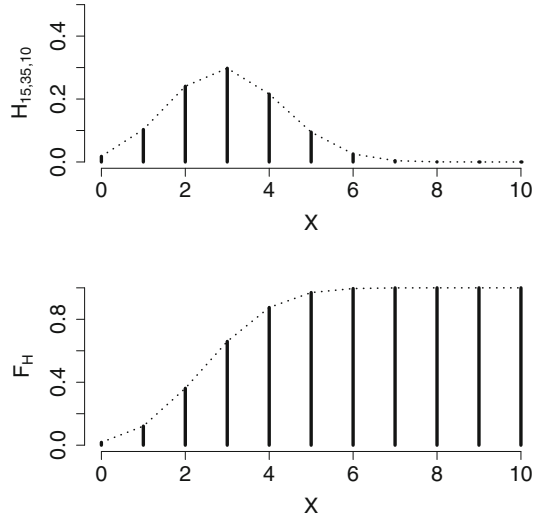


Abb. 2.20 Ziehungs-schemata der hypergeometrischen Verteilung: Aus der Menge M mit a Elementen des Typs A (Teilmenge M_A) und $N - a$ Elementen des Typs B (Teilmenge M_B) werden $n \leq N$ Elemente ohne Zurücklegen gezogen

Abb. 2.21 Wahrscheinlichkeitsfunktion $H_{15,35,10}$ der hypergeometrischen Verteilung mit den Parametern $a = 15$, $N - a = 35$ und $n = 10$ (obere Grafik) und die entsprechende Verteilungsfunktion F_H (untere Grafik)



in der gezogenen Zufallsstichprobe genau x Elemente aus M_A und $n - x$ Elemente aus M_B sind. Nun gibt es $C(x, a) = \binom{a}{x}$ verschiedenen Möglichkeiten, x Elemente aus M_A (ohne Berücksichtigung der Reihenfolge) auszuwählen. Analog gibt es $C(n - x, N - a) = \binom{N-a}{n-x}$ Möglichkeiten, $n - x$ Elemente aus M_B auszuwählen. Somit ist die Anzahl der Elemente in E gleich dem Produkt $C(x, a)C(n - x, N - a)$ und die gesuchte Wahrscheinlichkeit gleich¹⁵

$$P(X = x) = H_{a,N-a,n}(x) = \frac{C(x, a)C(n - x, N - a)}{C(n, N)} = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}. \quad (2.23)$$

Wir fassen zusammen: Die Zufallsvariable X (Anzahl der Elemente des Typs A beim Ziehen von n Elementen ohne Zurücklegen) nimmt jeden (ganzzahligen) Wert x aus dem Intervall $[x_{\min}, x_{\max}]$ mit der Wahrscheinlichkeit (2.23) an. Man bezeichnet X als **hypergeometrisch verteilt** mit den Parametern a , $N - a$ und n und schreibt dafür kurz $X \sim H_{a,N-a,n}$. Analog zur Binomialverteilung ist die Verteilungsfunktion der $H_{a,N-a,n}$ -Verteilung an der Stelle x gleich der Wahrscheinlichkeit

$$F_H(x) = P(X \leq x) = H_{a,N-a,n}(0) + H_{a,N-a,n}(1) + \cdots + H_{a,N-a,n}(x), \quad (2.24)$$

dass X einen Wert kleiner oder gleich x annimmt. Die Wahrscheinlichkeitsfunktion und die Verteilungsfunktion einer hypergeometrisch verteilten Zufallsvariablen sind in Abb. 2.21 beispielhaft dargestellt.

¹⁵ Die Sonderfälle $x = 0$ und $x = 1$ erfordern eine eigene Überlegung. Für $x = 0$ enthält die gezogene Zufallsstichprobe nur Elemente des Typs B ; es gibt $C(n, N - a) = \binom{N-a}{n}$ verschiedene Möglichkeiten, n Elemente aus M_B zu ziehen. Für $x = 1$ gibt es in der Zufallsstichprobe ein Element aus M_A und $n - 1$ Elemente aus M_B ; das eine Element aus M_A kann eines der a Elemente von M_A sein, die $n - 1$ Elemente aus M_B können durch $C(n - 1, N - a) = \binom{N-a}{n-1}$ verschiedene Ziehungen realisiert werden. Die Formel (2.23) behält in beiden Fällen ihre Gültigkeit, wenn $C(0, a) = \binom{a}{0} = 1$ und $C(1, a) = \binom{a}{1} = a$ vereinbart wird.

Beispiel 2.20

Ein Produkt wird vom Hersteller in Packungen von 40 Stück ausgeliefert. Im Rahmen der Eingangskontrolle wird nach folgendem Plan geprüft: Es werden 5 Stück aus der gelieferten Packung zufällig ausgewählt (ohne Zurücklegen) und auf Fehler überprüft. Ist höchstens ein Stück fehlerhaft, wird die Packung angenommen, andernfalls zurückgeschickt. Wir bestimmen die Wahrscheinlichkeit, dass bei diesem Prüfplan die Packung zurückgewiesen wird, wenn sie 2 defekte Stücke enthält.

Es sei X die Anzahl der fehlerhaften Einheiten in der Prüfstichprobe mit dem Umfang $n = 5$. Die Packung umfasst $N = 40$ Stück. Unter der Annahme von $a = 2$ fehlerhaften Einheiten in der Packung soll die Wahrscheinlichkeit $P(X > 1) = 1 - P(X \leq 1)$ berechnet werden. Auf Grund des Ziehungsschemas ist X hypergeometrisch verteilt mit den Parametern $a = 2$, $N - a = 38$ und $n = 5$. Es ist $P(X \leq 1) = P(X = 0) + P(X = 1) = H_{2,38,5}(0) + H_{2,38,5}(1)$. Wegen

$$H_{2,38,5}(0) = \frac{\binom{2}{0} \binom{38}{5}}{\binom{40}{5}} = \frac{1 \cdot 38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 / 5!}{40 \cdot 39 \cdot 38 \cdot 37 \cdot 36 / 5!} = 0.7628,$$

$$H_{2,38,5}(1) = \frac{\binom{2}{1} \binom{38}{4}}{\binom{40}{5}} = \frac{2 \cdot 38 \cdot 37 \cdot 36 \cdot 35 / 4!}{40 \cdot 39 \cdot 38 \cdot 37 \cdot 36 / 5!} = 0.2244$$

ergibt sich $P(X > 1) = 1 - (0.7628 + 0.2244) = 0.0128$.

```
> # Lösung mit R:
> # Berechnung mit der Wahrscheinlichkeitsfunktion dhyper()
> # Pab = P(X>1) = 1-P(X<=1) = 1-P(X=0)-P(X=1)
> Pab <- 1-dhyper(0,2,38,5)-dhyper(1,2,38,5)
> print(Pab, digits=3)
[1] 0.0128
> # Berechnung mit der Verteilungsfunktion phyper()
> Pab <- 1-phyper(1,2,38,5); print(Pab, digits=3)
[1] 0.0128
```

2.5.4 Kennzahlen der Wahrscheinlichkeitsfunktion

Mit der Wahrscheinlichkeitsfunktion f wird die Zufallsvariation einer diskreten Zufallsvariablen X vollständig beschrieben. Eine Kurzbeschreibung der Zufallsvariation erhält man, wenn man die Wahrscheinlichkeitsfunktion hinsichtlich ihrer „Lage“ auf der horizontalen Achse (Merkmalsachse) und hinsichtlich ihrer „Form“ durch geeignete Kennzahlen beschreibt. Das wichtigste Lagemaß ist der **Mittelwert**

$$\mu_X = x_1 f(x_1) + x_2 f(x_2) + \dots \quad (2.25)$$

von X . Zu seiner Berechnung wird jeder Wert x_i von X mit der entsprechenden Wahrscheinlichkeit $f(x_i) = P(X = x_i)$ multipliziert und die erhaltenen Produkte aufsummiert. Statt Mittelwert von X sagt man auch **Erwartungswert** und

schreibt dafür $E(X)$. Die Lage des Mittelwerts kann durch ein mechanisches Modell veranschaulicht werden: Man denke sich auf der X -Achse an jeder Stelle x_i ($i = 1, 2, \dots$) die Masse $f(x_i)$ angebracht und die X -Achse wie ein Waagebalken um einen Punkt ξ drehbar gelagert. Der Balken ist im Gleichgewicht, wenn $\xi = \mu_X$ ist.

Die klassische Kennzahl zur Beschreibung der Form einer Wahrscheinlichkeitsfunktion f ist die **Varianz** bzw. deren Quadratwurzel, die man als **Standardabweichung** bezeichnet. Es sei X wieder eine diskrete Zufallsvariable und $x_1, x_2, \dots$ die möglichen Werte von X . Dann ist die Varianz von X durch

$$\sigma_X^2 = (x_1 - \mu_X)^2 f(x_1) + (x_2 - \mu_X)^2 f(x_2) + \dots \quad (2.26a)$$

definiert. Quadriert man die Binome in (2.26a) aus, so ergibt sich eine weitere Formel für die Varianz, nämlich

$$\begin{aligned} \sigma_X^2 &= (x_1^2 - 2x_1\mu_X + \mu_X^2)f(x_1) + (x_2^2 - 2x_2\mu_X + \mu_X^2)f(x_2) + \dots \\ &= x_1^2 f(x_1) + x_2^2 f(x_2) + \dots - 2\mu_X(x_1 f(x_1) + x_2 f(x_2) + \dots) \\ &\quad + \mu_X^2(f(x_1) + f(x_2) + \dots) \\ &= x_1^2 f(x_1) + x_2^2 f(x_2) + \dots - \mu_X^2. \end{aligned} \quad (2.26b)$$

Die Varianz (bzw. Standardabweichung) ist groß, wenn der Graph von f einen flachen, langgestreckten Verlauf um den Mittelwert μ_X besitzt, und kleiner, wenn f nur in einem kleinen Intervall um μ_X von null verschiedene Werte annimmt. Wie in Abschn. 2.7.4 gezeigt wird, gilt für eine mit den Parametern n und p **binomialverteilte** Zufallsvariable:

$$\mu_X = np \quad \text{und} \quad \sigma_X^2 = np(1-p) = \mu_X \left(1 - \frac{\mu_X}{n}\right)$$

Ist X **hypergeometrisch** verteilt mit den Parametern a , $N - a$ und n , sind der Mittelwert durch $\mu_X = na/N$ und die Varianz durch $\sigma_X^2 = \mu_X(1 - \mu_X/n)(N - n)/(N - 1)$ gegeben.

2.5.5 Aufgaben

1. Bei einem medizinischen Eingriff ist das Risiko einer bestimmten Komplikation gleich 0.5 %. Der Eingriff wird bei 100 Personen vorgenommen. Wie groß ist die Wahrscheinlichkeit, dass höchstens 2-mal eine Komplikation auftritt.
2. Mit welcher Wahrscheinlichkeit nimmt eine mit den Parametern $n = 15$ und $p = 0.5$ binomialverteilte Zufallsvariable einen Wert im einfachen Streubereich um den Mittelwert an, d. h. einen Wert aus dem Intervall $[\mu_X - \sigma_X, \mu_X + \sigma_X]$?
3. In einem Betrieb mit 30 Mitarbeitern wurden 20 prophylaktisch gegen Grippe geimpft. Insgesamt erkrankten 6 Mitarbeiter an Grippe. Unter der Annahme,

dass das Erkrankungsrisiko mit und ohne Impfung gleich groß ist, bestimme man die Wahrscheinlichkeit, dass 2 der geimpften Mitarbeiter an Grippe erkranken.

2.6 Dichtefunktionen

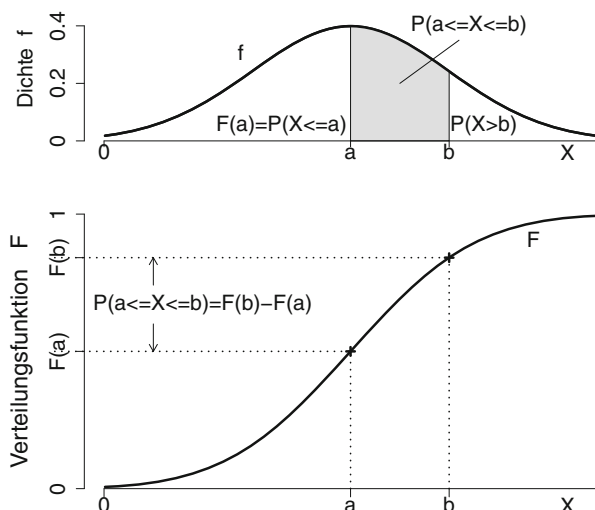
2.6.1 Stetige Zufallsvariable

In diesem Abschnitt geht es um metrische Merkmale, auch Messmerkmale genannt. Bei einem Messmerkmal kann die Maßzahl des Messergebnisses irgendein Wert aus einem Intervall der reellen Zahlen sein. Es folgt, dass ein Messmerkmal grundsätzlich überabzählbar viele Werte annehmen kann. Die beim wiederholten Messen einer Größe X (z. B. der Spaltöffnungsgröße auf den Blättern einer Pflanze) beobachtbare Variation der Messergebnisse zeigt an, dass der Vorgang, der die Messwerte erzeugt, grundsätzlich als Zufallsexperiment zu deuten ist. Dabei sind zwei verschiedene Situationen zu unterscheiden. Es kann die Messgröße zwar einen festen „wahren“ Wert besitzen, durch den Messvorgang wird dem wahren Wert aber ein Messfehler überlagert, der zu regellos um den wahren Wert schwankenden Messergebnissen führt; in diesem Fall ist das Zufallsexperiment mit dem Messvorgang verbunden. Oder jeder Wert der Messgröße selbst ist das Ergebnis eines Zufallsexperimentes (wie z. B. die konkrete Realisierung einer Spaltöffnung bestimmter Länge).

Die Zufallsvariation eines diskreten Merkmals wurde mit Hilfe der Wahrscheinlichkeitsfunktion beschrieben, die jedem Merkmalswert die Wahrscheinlichkeit zuordnet, mit der das Merkmal diesen Wert annimmt. Bei einem Messmerkmal mit überabzählbar vielen möglichen Werten muss ein anderer methodischer Ansatz zur Beschreibung der Merkmalsvariation gewählt werden. An die Stelle der Wahrscheinlichkeitsfunktion tritt nun die **Wahrscheinlichkeitsdichte** oder kurz Dichte. Das ist eine für alle reellen x definierte reellwertige und nicht-negative Funktion f mit folgender Eigenschaft: Für beliebige reelle Zahlen $a, b > a$ ist die Wahrscheinlichkeit $P(a \leq X \leq b)$ dafür, dass das X einen Wert im Intervall $[a, b]$ annimmt, gleich der Maßzahl der unter dem Graphen von f liegenden Fläche von der Stelle $x = a$ bis zur Stelle $x = b$ (vgl. Abb. 2.22, obere Grafik).¹⁶ Wegen $P(-\infty < X < +\infty) = 1$ muss von f im Besonderen verlangt werden, dass die gesamte unter dem Graphen von f liegende Fläche die Maßzahl eins hat. Wir bezeichnen ein Messmerkmal X , dessen Variation durch eine Wahrschein-

¹⁶ Wählt man als Intervall speziell $a = x$ und $b = x + \Delta x$ mit sehr kleinem $\Delta x > 0$, hat man näherungsweise $P(x \leq X \leq x + \Delta x) \approx f(x)\Delta x$. Daraus erkennt man, dass der Funktionswert $f(x)$ die „Dimension“ einer „Wahrscheinlichkeit pro Länge“ hat. Auf Längen, Flächen oder Volumina bezogene Größen werden in der Physik als Dichten bezeichnet. Diese Terminologie wurde in den Bezeichnungen „Wahrscheinlichkeitsdichte“ und „Dichtekurve“ (für den Graphen der Wahrscheinlichkeitsdichte) übernommen.

Abb. 2.22 Wahrscheinlichkeitsdichte (*obere Grafik*) und Verteilungsfunktion (*untere Grafik*) einer stetigen Zufallsvariablen. Dargestellt ist die Verteilung einer mit dem Mittelwert $\mu = 2.5$ und der Standardabweichung $\sigma = 1$ normalverteilten Zufallsvariablen X



lichkeitsdichte mit den genannten Eigenschaften beschrieben wird, als eine **stetige Zufallsvariable**.

In der unteren Grafik von Abb. 2.22 ist die **Verteilungsfunktion** F einer stetigen Variablen X beispielhaft dargestellt. Die Verteilungsfunktion ordnet jeder reellen Zahl x die Wahrscheinlichkeit $F(x) = P(X \leq x)$, dass X höchstens gleich x ist, als Funktionswert zu. Dieser kann geometrisch als „Unterschreitungsfläche“, d. h. als Maßzahl der Fläche unter der Dichtekurve bis zur Stelle x interpretiert werden.

Beispiel 2.21

Für eine über dem Intervall $[0, 1]$ gleich verteilte stetige Zufallsvariable X ist die Wahrscheinlichkeitsdichte durch

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{für } x < a \text{ oder } b > 1 \end{cases}$$

gegeben. Die dadurch definierte Verteilung heißt **stetige Gleichverteilung** (oder Rechteckverteilung) über dem Intervall $[a, b]$. Wie man leicht nachrechnet, ist die von der Dichtekurve und der reellen Achse eingeschlossene Fläche gleich 1. Die Verteilungsfunktion F von X ist eine stetige Funktion mit $F(x) = 0$ für $x < a$, $F(x) = (x - a)/(b - a)$ für $a \leq x \leq b$ und $F(x) = 1$ für $x > b$. Ist z. B. $a = 0$ und $b = 2$, bestimmt man mit Hilfe von F die Wahrscheinlichkeit, dass X kleiner oder gleich 0.5 ist, aus $P(X \leq 0.5) = F(0.5) = 0.5/2 = 0.25$.

Im Allg. ist die Berechnung von Funktionswerten der Verteilungsfunktion viel komplizierter als im vorangehenden Beispiel. Glücklicherweise gibt es für in der Praxis bedeutsame Verteilungsfunktionen statistische Tabellen, mit denen man

Funktionswerte zu vorgegebenen Argumenten bestimmen kann. Meist bedient man sich heute aber entsprechender Funktionen, die in statistischen Softwareprodukten zur Verfügung gestellt werden. Wenn man die Funktionswerte der Verteilungsfunktion F einer stetigen Zufallsvariablen X bestimmen kann, lassen sich damit schnell die Wahrscheinlichkeiten berechnen, dass X kleiner oder größer als ein vorgegebener Wert a ist oder zwischen zwei vorgegebenen Werten $a, b > a$ liegt. Es ist nämlich

$$\begin{aligned}P(X < a) &= P(X \leq a) = F(a), \\P(X > a) &= 1 - P(X \leq a) = 1 - F(a), \\P(a < X < b) &= F(b) - F(a).\end{aligned}$$

2.6.2 Normalverteilung

Wohl das bekannteste Verteilungsmodell für eine stetige Zufallsvariable ist die **Standardnormalverteilung**, die in vielen Bereichen der Statistik auftritt. Es ist üblich, eine standardnormalverteilte Zufallsvariable durch den Großbuchstaben Z , ihre Dichtefunktion durch φ und die Verteilungsfunktion durch Φ zu bezeichnen. Für jedes reelle z kann der Wert der Dichtefunktion aus der Funktionsgleichung

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (2.27)$$

berechnet werden. Der Graph der Dichtefunktion ist in Abb. 2.23 dargestellt. Man erkennt den symmetrischen Verlauf der Dichtekurve um $z = 0$. Die Dichtefunktion nimmt an dieser Stelle ihr Maximum $\varphi(0) = 1/\sqrt{2\pi} = 0.399$ an und nähert sich mit wachsendem Abstand vom Nullpunkt der Z -Achse. Dabei durchläuft sie an der Stelle $z = -1$ einen Wendepunkt und ebenso an der Stelle $z = +1$. Eine markante Eigenschaft der Standardnormalverteilung ist also die Symmetrie um $z = 0$. Aus der Symmetrie folgt, dass der **Mittelwert** μ_Z von Z gleich null ist. Der Mittelwert μ_Z kann nämlich als horizontale Schwerpunktskoordinate gedeutet werden: Wenn man die Z -Achse mit variabler Masse so belegt, dass $\varphi(z)$ die Massendichte an der Stelle z ist, liegt der Schwerpunkt dieser Massenverteilung genau an der Stelle $\mu_Z = 0$. Statt vom Mittelwert μ_Z spricht man auch vom Erwartungswert $E(Z)$.

Die **Varianz** σ_Z^2 von Z ist als Erwartungswert des Quadrats der Abweichung $Z - \mu_Z$ der Zufallsvariablen Z vom Mittelwert $\mu_Z = 0$ definiert. Wie später noch gezeigt wird, ist $\sigma_Z^2 = 1$. Daher ist auch die **Standardabweichung** $\sigma_Z = \sqrt{\sigma_Z^2} = 1$. Das durch die Wendepunktsabszissen begrenzte Intervall $[-1, 1]$ stellt somit den einfachen Streubereich $[\mu_Z - \sigma_Z, \mu_Z + \sigma_Z]$ um den Mittelwert dar. Die Standardnormalverteilung wird auch als $N(0, 1)$ -Verteilung bezeichnet; dabei steht der Buchstabe N für „Normalverteilung“ und die zwischen den Klammern stehen-

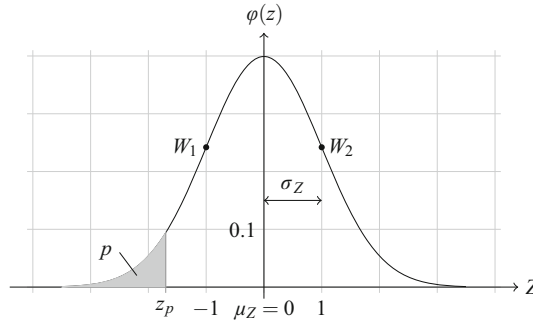


Abb. 2.23 Dichtefunktion φ der standardnormalverteilten Zufallsvariablen Z . Wegen der Symmetrie der Dichtekurve um $Z = 0$ ist der Mittelwert $\mu_Z = 0$. Die Standardabweichung $\sigma_Z = 1$ kann geometrisch als Abstand des Wendepunktes W_1 bzw. W_2 von der Symmetrieachse gedeutet werden. Das Quantil z_p ist jener Wert von Z , der mit der Wahrscheinlichkeit p unterschritten wird.

den Parameter repräsentieren den Mittelwert $\mu_Z = 0$ bzw. die Varianz $\sigma_Z^2 = 1$. Für eine standardnormalverteilte Zufallsvariable schreibt man kurz $Z \sim N(0, 1)$.

Neben dem Mittelwert und der Standardabweichung sind der **Median** und der **Quartilabstand** weitere Maßzahlen zur Kennzeichnung der zentralen Lage bzw. der Streuung einer stetigen Zufallsvariablen. Der Median $z_{0.5}$ ist derjenige Wert von Z , der mit 50 %-iger Wahrscheinlichkeit unter- bzw. überschritten wird, d. h., für $z_{0.5}$ gilt: $P(Z \leq z_{0.5}) = P(Z \geq z_{0.5}) = 0.5$. Die an der Stelle $z_{0.5}$ der Z -Achse errichtete Ordinate teilt die Fläche „unter“ der Dichtekurve in zwei Hälften. Für die Standardnormalverteilung ist daher $z_{0.5} = 0$.

Der Median ist ein Spezialfall eines allgemeineren Lagemaßes, des sogenannten **p -Quantils** z_p mit $0 < p < 1$; z_p ist der durch die Forderung $P(Z \leq z_p) = \Phi(z_p) = p$ festgelegte Wert von Z , also jener Wert, der mit der Wahrscheinlichkeit p unterschritten wird. Die Definition des Quantils z_p ist in Abb. 2.23 veranschaulicht. Das 25 %-Quantil $z_{0.25}$ und das 75 %-Quantil $z_{0.75}$ heißen auch das untere bzw. das obere **Quartil** von Z . Die Differenz $z_{0.75} - z_{0.25}$ aus dem oberen und unteren Quartil ist der **Quartilabstand** IQR (inter quartile range).

Beispiel 2.22

Es sei Z eine $N(0, 1)$ -verteilte Zufallsvariable.

- Wir berechnen die Wahrscheinlichkeit P , dass Z einen Wert im einfachen Streubereich $[-1, 1]$ um den Mittelwert annimmt. Dazu schreiben wir $P = P(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1)$. Wegen der Symmetrie der Dichtefunktion ist $\Phi(-1) = P(Z \leq -1) = P(Z > 1) = 1 - P(Z \leq 1) = 1 - \Phi(1)$. Daher ergibt sich für die gesuchte Wahrscheinlichkeit $P = 2\Phi(1) - 1 = 68.3\%$; der Wert $\Phi(1) = 0.8413$ wurde mit der R-Funktion `pnorm()` bestimmt.
- Um das 95 %-Quantil $z_{0.95}$ von Z zu bestimmen, hat man die Gleichung $P(Z \leq z_{0.95}) = \Phi(z_{0.95}) = 0.95$ zu lösen. Die Lösung findet man mit Hilfe der Quan-

tilsfunktion Φ^{-1} , die z. B. in R durch die Funktion `qnorm()` bereitgestellt wird. Damit findet man $z_{0.95} = \Phi^{-1}(0.95) = 1.645$.

```
> # Lösung mit R:
> options(digits=4)
> # a) Berechnung von  $P(-1 \leq Z \leq 1)$ 
> P <- pnorm(1) - pnorm(-1); P
[1] 0.6827
> # b) Berechnung des 95%-Quantils  $z_{0.95}$ 
> z_0.95 <- qnorm(0.95); z_0.95
[1] 1.645
```

Eine Zufallsvariable X heißt **normalverteilt** mit dem Mittelwert μ_X und der Varianz σ_X^2 , wenn die Variable $Z = (X - \mu_X)/\sigma_X$ standardnormalverteilt ist. Zur Dichtefunktion f von X gelangt man auf folgende Weise: Es seien z_1 und $z_2 > z_1$ zwei Werte von Z und $x_1 = \mu_X + \sigma_X z_1$ bzw. $x_2 = \mu_X + \sigma_X z_2$ die entsprechenden Werte von X . Dann gilt

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(\mu_X + \sigma_X z_1 \leq X \leq \mu_X + \sigma_X z_2) \\ &= P\left(z_1 \leq \frac{X - \mu_X}{\sigma_X} \leq z_2\right) = P(z_1 \leq Z \leq z_2). \end{aligned} \quad (2.28)$$

Wenn $\Delta x = x_2 - x_1 = \sigma_X(z_2 - z_1)$ sehr klein ist, können die Wahrscheinlichkeiten $P(x_1 \leq X \leq x_2)$ und $P(z_1 \leq Z \leq z_2)$ näherungsweise durch $f(x_1)\Delta x$ bzw. $\varphi(z_1)(z_2 - z_1) = \varphi(z_1)\Delta x/\sigma_X$ ersetzt werden. Damit folgt aus (2.28) im Grenzfall $\Delta x \rightarrow 0$ für jedes reelle x_1

$$f(x_1) = \frac{1}{\sigma_X} \varphi\left(\frac{x_1 - \mu_X}{\sigma_X}\right).$$

An die Stelle der Dichtefunktion (2.27) tritt daher die Gleichung

$$f(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{(x - \mu_X)^2 / (2\sigma_X^2)} \quad (2.29)$$

für die Dichtefunktion der mit dem Mittelwert μ_X und der Varianz σ_X^2 normalverteilten Variablen X . In Abb. 2.24 sind die Dichtefunktionen von zwei normalverteilten Zufallsvariablen gemeinsam mit der Dichte der Standardnormalverteilung als Bezugskurve dargestellt.¹⁷

Der Wert $F(x) = P(X \leq x)$ der Verteilungsfunktion von X an der Stelle x ist wegen

$$\begin{aligned} F(x) &= P(X \leq x) = P\left(\frac{X - \mu_X}{\sigma_X} \leq \frac{x - \mu_X}{\sigma_X}\right) \\ &= P\left(Z \leq \frac{x - \mu_X}{\sigma_X}\right) = \Phi\left(\frac{x - \mu_X}{\sigma_X}\right) \end{aligned}$$

¹⁷ Für „ X ist normalverteilt mit dem Mittelwert μ_X und der Varianz σ_X^2 “ schreibt man auch kurz $X \sim N(\mu_X, \sigma_X^2)$.

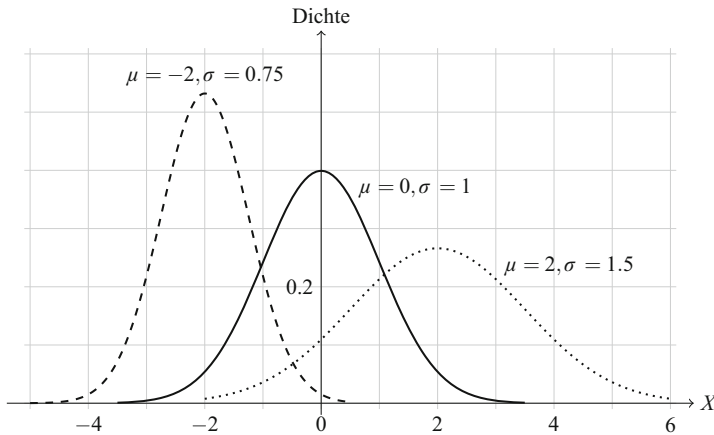


Abb. 2.24 Dichtekurven von normalverteilten Zufallsvariablen mit verschiedenen Mittelwerten und Standardabweichungen. Als Referenzkurve ist auch die Dichte der Standardnormalverteilung eingezeichnet

gleich dem Wert $\Phi(z)$ der Verteilungsfunktion von $Z = (X - \mu_X)/\sigma_X$ an der Stelle $z = (x - \mu_X)/\sigma_X$. Gibt man für die Verteilungsfunktion den Wert p vor, ergibt sich einerseits aus $F(x) = p$ die Lösung $x = x_p$ und andererseits aus $\Phi((x - \mu_X)/\sigma_X) = p$ die Lösung $(x - \mu_X)/\sigma_X = z_p$. Somit ist das p -Quantil x_p von X und das p -Quantil z_p von Z durch die Beziehung

$$x_p = \mu_X + \sigma_X z_p \quad (2.30)$$

miteinander verknüpft. Dieser Zusammenhang bildet die Grundlage für sogenannte Normal- $Q-Q$ -Plots, mit denen man auf graphischem Wege überprüfen kann, ob die Werte einer Stichprobe Realisierungen einer normalverteilten Zufallsvariablen sind. Setzt man in (2.30) speziell $p = 0.5$, erhält man $x_{0.5} = \mu_X + \sigma_X z_{0.5} = \mu_X$, d. h. der Median $x_{0.5}$ von X fällt mit dem Mittelwert μ_X zusammen.

Beispiel 2.23

- a) Es sei $X \sim N(\mu, \sigma^2)$ mit $\mu = 15$ und $\sigma^2 = 16$. Wir berechnen die Wahrscheinlichkeiten $P(X < 10)$, $P(X > 10)$, $P(10 \leq X \leq 20)$ sowie das Quantil $x_{0.95}$. Bezeichnet F die Verteilungsfunktion von X , so gilt $P(X < 10) = F(10)$, $P(X > 10) = 1 - F(10)$ und $P(10 \leq X \leq 20) = F(20) - F(10)$. Das 95 %-Quantil von X ist die Lösung der Gleichung $P(X \leq x) = 0.95$, die mit der Quantilsfunktion F^{-1} durch $x_{0.95} = F^{-1}(0.95)$ ausgedrückt werden kann. Zur numerischen Berechnung werden im Folgenden die R-Funktionen `pnorm()` und `qnorm()` verwendet.

```

> # Lösung mit R:
> mu <- 15; sigma <- sqrt(16)
> P1 <- pnorm(10,mu,sigma)      # P1=P(X<10)
> P2 <- 1-pnorm(10,mu,sigma)    # P2=P(X>10)
> P3 <- pnorm(20,mu,sigma)-P1   # P3=P(10<=X<=20)
> x0.95 <- qnorm(0.95,mu,sigma) # x0.95=95%-Quantil
> print(cbind(P1, P2, P3, x0.95), digits=4)
      P1      P2      P3 x0.95
[1,] 0.1056 0.8944 0.7887 21.58

```

- b) Es sei X eine $N(\mu, \sigma^2)$ -verteilte Zufallsvariable. Wir berechnen die Wahrscheinlichkeit, dass X einen extremen Wert annimmt. Dabei möge ein Wert als extrem gelten, wenn er kleiner als $x_{0.25} - 1.5IQR$ oder größer als $x_{0.75} + 1.5IQR$ ist. Hier ist $IQR = x_{0.75} - x_{0.25}$ der Abstand der Quartile $x_{0.25}$ und $x_{0.75}$. Wegen $x_{0.75} = \mu + \sigma z_{0.75}$, $x_{0.25} = \mu + \sigma z_{0.25} = \mu - \sigma z_{0.75}$ und $IQR = \sigma(z_{0.75} - z_{0.25}) = 2\sigma z_{0.75}$ ist $P_1 = P(X < x_{0.25} - 1.5IQR) = F(\mu - \sigma z_{0.75} - 3\sigma z_{0.75}) = \Phi(-4z_{0.75})$. Aus Symmetriegründen ist $P_2 = P(X > x_{0.75} + 1.5IQR) = P_1$, so dass sich für die gesuchte Wahrscheinlichkeit $P_1 + P_2 = 2\Phi(-4z_{0.75}) = 0.7\%$ ergibt. Die numerische Berechnung wurde mit den folgenden R-Anweisungen durchgeführt:

```

> # Lösung mit R:
> z0.75 <- qnorm(0.75); P <- 2*pnorm(-4*z0.75)
> print(cbind(z0.75, P), digits=4)
      z0.75      P
[1,] 0.6745 0.006977

```

2.6.3 Aufgaben

- Die Masse (in mg) einer Wirksubstanz W in einem Präparat sei normalverteilt mit dem Mittelwert $\mu = 10$ und der Standardabweichung $\sigma = 0.5$.
 - Mit welcher Wahrscheinlichkeit nimmt W einen Wert zwischen 9 und 11 an?
 - Man bestimme ferner das 25 %- und das 75 %-Quantil von W .
- Man berechne die Wahrscheinlichkeit, dass eine $N(\mu, \sigma^2)$ -verteilte Zufallsvariable X einen Wert außerhalb des 4-fachen Streubereiches um den Mittelwert annimmt, d. h. dass X kleiner als $\mu - 4\sigma$ oder größer als $\mu + 4\sigma$ ist.

2.7 Ergänzungen

2.7.1 Berechnung der Parameter der Regressionsgeraden

In Abschn. 2.1.3 wurden für die Parameter k und d der Regressionsgeraden $y = kx + d$ die Schätzwerte $\hat{k} = s_{xy}/s_x^2$ bzw. $\hat{d} = \bar{y} - \hat{k}\bar{x}$ angeführt. Diese Schätzwerte ergeben sich, wenn man die Summe (2.3) der Quadrate der Residuen minimiert. Eine elementare Lösung der Optimierungsaufgabe ist möglich, wenn man die Stichprobenmittel $\bar{x}$ und $\bar{y}$ einführt und die Umformung

$$\begin{aligned} Q &= \sum_{i=1}^n (y_i - \bar{y} - k(x_i - \bar{x}) + \bar{y} - k\bar{x} - d)^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 + k^2 \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{y} - k\bar{x} - d)^2 - 2k \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ &\quad + 2(\bar{y} - k\bar{x} - d) \sum_{i=1}^n (y_i - \bar{y}) - 2k(\bar{y} - k\bar{x} - d) \sum_{i=1}^n (x_i - \bar{x}) \end{aligned}$$

vornimmt. Die Summanden in der letzten Zeile sind wegen $\sum_{i=1}^n (x_i - \bar{x}) = 0$ und $\sum_{i=1}^n (y_i - \bar{y}) = 0$ gleich null. Die verbleibenden Summen können mit Hilfe der Definitionsgleichungen für die Varianzen s_x^2 , s_y^2 und die Kovarianz s_{xy} durch

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})^2 &= (n-1)s_x^2, \quad \sum_{i=1}^n (y_i - \bar{y})^2 = (n-1)s_y^2 \quad \text{und} \\ \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= (n-1)s_{xy} \end{aligned}$$

ersetzt werden. Man erhält so nach Umstellen der Summanden

$$\begin{aligned} Q &= k^2(n-1)s_x^2 - 2k(n-1)s_{xy} + n(\bar{y} - k\bar{x} - d)^2 + (n-1)s_y^2 \\ &= (n-1)s_x^2 \left(k - \frac{s_{xy}}{s_x^2} \right)^2 + n(\bar{y} - k\bar{x} - d)^2 + (n-1)s_y^2 \left(1 - \frac{s_{xy}^2}{s_x^2 s_y^2} \right). \end{aligned}$$

Daraus kann unmittelbar abgelesen werden, dass Q den kleinsten Wert

$$S_{rr} = (n-1)s_y^2 \left(1 - \left(\frac{s_{xy}}{s_x s_y} \right)^2 \right)$$

annimmt, wenn $k = \hat{k} = s_{xy}/s_x^2$ und $d = \hat{d} = \bar{y} - \hat{k}\bar{x}$ gesetzt wird.

2.7.2 Regressionsgerade durch den Nullpunkt

Bei der Bestimmung der Regressionsgeraden wurde in Abschn. 2.1.3 vorausgesetzt, dass beide Geradenparameter geschätzt werden müssen. Es gibt aber Abhängigkeiten, bei denen die Regressionsgerade ausdrücklich eine **Nullpunktsgerade** sein muss und als solche in der Form $y = kx$ anzusetzen ist. Damit reduziert sich (2.3) auf die nur mehr von k abhängige Quadratsumme

$$Q_0(k) = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - kx_i)^2. \quad (2.31)$$

Die Forderung $Q_0(k) = \min!$ führt auf die Formel

$$\hat{k}_0 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \quad (2.32)$$

für den Anstieg der Nullpunktsgereaden. Man erhält (2.32) aus (2.4), wenn dort $\bar{x} = \bar{y} = 0$ gesetzt wird. Der Nullpunkt übernimmt also die Rolle des Zentroids der Punkteverteilung, durch den die Regressionsgerade hindurchgeht.

Zu zeigen ist, dass die Quadratsumme $Q_0(k) = \sum (y_i - kx_i)^2$ an der Stelle $k = \hat{k}_0 = \sum x_i y_i / \sum x_i^2$ ein eindeutig bestimmtes Minimum besitzt.¹⁸ Dazu wird die Umformung

$$\begin{aligned} Q_0(k) &= \sum (y_i - kx_i)^2 = \sum y_i^2 - 2k \sum x_i y_i + k^2 \sum x_i^2 \\ &= \sum x_i^2 \left(k^2 - 2k \frac{\sum x_i y_i}{\sum x_i^2} \right) + \sum y_i^2 \\ &= \sum x_i^2 \left(k - \frac{\sum x_i y_i}{\sum x_i^2} \right)^2 - \sum x_i^2 \left(\frac{\sum x_i y_i}{\sum x_i^2} \right)^2 + \sum y_i^2 \\ &= \sum x_i^2 \left(k - \frac{\sum x_i y_i}{\sum x_i^2} \right)^2 + \sum y_i^2 \left(1 - \frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2} \right) \end{aligned}$$

vorgenommen. Setzt man hier $k = \hat{k}_0$, sieht man unmittelbar, dass $Q_0(k) > Q_0(\hat{k}_0)$ für alle $k \neq \hat{k}_0$ gilt.

Beispiel 2.24

Beim freien Fall hängt die Falldauer T über die Beziehung $T = \sqrt{2H/g}$ von der Fallhöhe H ab. Die Konstante g ist die Erdbeschleunigung. Zur Bestimmung von g wurde für verschiedene Fallhöhen (in m) die Falldauer (in s) bestimmt. Es ergaben sich die folgenden Messwertpaare (H, T) : (1, 0.47), (2, 0.68), (3, 0.74), (4, 0.82),

¹⁸ Der Summationsindex läuft jeweils von $i = 1$ bis $i = n$.

(5, 1.09). Wir bestimmen einen Schätzwert für g mit Hilfe einer linearen Regression durch den Nullpunkt.

Die Variable T ist proportional zu $x = \sqrt{H}$, die Proportionalitätskonstante $k_0 = \sqrt{2/g}$ ist der Anstieg einer Nullpunktsgersten. Mit den Summen

$$\sum_i^5 x_i^2 = 1 + 2 + 3 + 4 + 5 = 15,$$

$$\sum_{i=1}^5 x_i T_i = 1 \cdot 0.47 + \sqrt{2} \cdot 0.68 + \sqrt{3} \cdot 0.74 + 2 \cdot 0.82 + \sqrt{5} \cdot 1.09 = 6.791$$

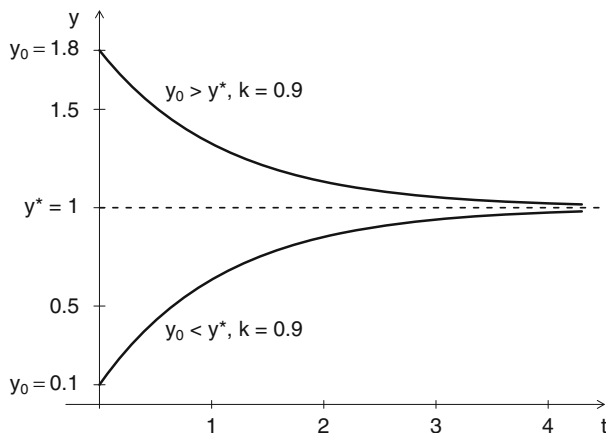
ergibt sich aus (2.32) der Schätzwert $\hat{k}_0 = 0.4527$ für den Anstieg und weiter der Schätzwert $\hat{g} = 2/\hat{k}_0^2 = 9.759$ für g .

```
> # Lösung mit R:
> H <- 1:5; x <- sqrt(H)
> T <- c(0.47, 0.68, 0.74, 0.82, 1.09)
> # Berechnung des Anstiegs mit (2.32)
> sumxT <- sum(x*T); sumx2 <- sum(x*x)
> k0 <- sumxT/sumx2; g <- 2/k0^2
> print(cbind(sumxT, sumx2, k0, g), digits=4)
      sumxT sumx2      k0      g
[1,] 6.791    15 0.4527 9.759
> # Berechnung des Anstiegs mit der R-Funktion lm()
> mod <- lm(T ~ 0+x)
> k0 <- mod$coefficients[[1]]; g <- 2/k0^2
> print(cbind(k0, g), digits=4)
      k0      g
[1,] 0.4527 9.759
```

2.7.3 Exponentielle Annäherung an einen Gleichgewichtswert

Es sei y eine Variable, die den Zustand eines Systems beschreibt, z. B. die Temperatur eines Körpers, die Höhe eines Baumes oder die Konzentration eines gelösten Stoffes in einer Zelle. Das betrachtete System wird im allgemeinen ein Teil eines umfassenderen Systems sein, das bei Änderungen im Teilsystem auf dieses rückwirkt. Als Folge vorhandener Rückkopplungen kann das Teilsystem veranlasst werden, einen stationären Zustand $y = y^*$ einzunehmen. Man bezeichnet y^* als einen Gleichgewichtswert. Stimmt y nicht mit y^* überein, so setzt eine „Rückbewegung“ zum Gleichgewichtszustand ein. Wir nehmen an, dass sich die Geschwindigkeit, mit der y dem Wert y^* zustrebt, proportional zur Abweichung $(y - y^*)$ der Zustandsgröße vom Gleichgewichtswert vermindert (mit einer gewissen Proportionalitätskonstanten $k > 0$). Dann wird die Annäherung der Größe y an

Abb. 2.25 Exponentielle Annäherung der Zustandsgröße y an den Gleichgewichtswert y^* nach dem Zeitgesetz $y = y^* + (y_0 - y^*)e^{-kt}$



den Wert y^* in Abhängigkeit von der Zeit t durch die Formel

$$y = y^* + (y_0 - y^*)e^{-kt} \quad (2.33)$$

wiedergegeben. Je nachdem, ob der Ausgangswert y_0 größer oder kleiner als y^* ist, fällt bzw. steigt y mit wachsendem t monoton gegen den Wert y^* . Die Funktionsgraphen stellen in y -Richtung verschobene Exponentialkurven dar (vgl. Abb. 2.25). Es folgen einige Anwendungen, bei denen die beschriebene exponentielle Annäherung an einen Gleichgewichtswert beobachtet werden kann.

Beispiel 2.25

- Ein Körper mit der Anfangstemperatur T_0 befindet sich in einer Umgebung mit der konstanten Temperatur $T_U < T_0$. Offensichtlich wird die Körpertemperatur T im Verlaufe der Zeit t auf die Umgebungstemperatur absinken. Nach dem Newtonschen Abkühlungsgesetz erfolgt dies exponentiell mit einer für den Wärmeaustausch typischen Konstanten $k > 0$ nach der Formel $T = T_U + (T_0 - T_U)e^{-kt}$.
- Eine Zelle befinde sich in einer Flüssigkeit, die einen gelösten Stoff mit der konstanten Konzentration c_F enthält. Die Konzentration c des gelösten Stoffes in der Zelle sei am Anfang c_0 . Als Folge des Konzentrationsunterschiedes wird ein Diffusionsprozess einsetzen, der eine Angleichung der Zellkonzentration an die Umgebungskonzentration bewirkt. Der Ausgleichsvorgang erfolgt nach dem Gesetz $c = c_F + (c_0 - c_F)e^{-kt}$. Dabei ist $k > 0$ eine Konstante, die die Durchlässigkeit der Zellmembran sowie die Zelloberfläche und das Zellvolumen erfasst.

2.7.4 Mittelwert und Varianz einer $B_{n,p}$ -verteilten Zufallsvariablen

Wendet man (2.25) auf eine mit den Parametern n und p binomialverteilte Zufallsvariable X an, erhält man den Mittelwert:

$$\begin{aligned}
 \mu_X &= \sum_{x=0}^n x B_{n,p}(x) = \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} \\
 &= \sum_{x=1}^n x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{k=0}^{n-1} (k+1) \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1} \\
 &= \sum_{k=0}^{n-1} \frac{(k+1)n(n-1)\dots(n-k-1+1)}{(k+1)!} p p^k (1-p)^{n-1-k} \\
 &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np[(1-p) + p]^{n-1} = np
 \end{aligned}$$

Hier wurde in der letzten Zeile die Summe mit dem binomischen Lehrsatz umgeformt. Zur Berechnung der Varianz verwenden wir Formel (2.26b) und erhalten:

$$\begin{aligned}
 \sigma_X^2 &= \sum_{x=0}^n x^2 \binom{n}{x} p^x (1-p)^{n-x} - \mu_X^2 \\
 &= \sum_{x=0}^n [x(x-1) + x] \binom{n}{x} p^x (1-p)^{n-x} - \mu_X^2 \\
 &= \sum_{x=0}^n x(x-1) \binom{n}{x} p^x (1-p)^{n-x} + \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} - \mu_X^2 \\
 &= \sum_{x=2}^n x(x-1) \frac{n(n-1)(n-2)\dots(n-x+1)}{x!} p^x (1-p)^{n-x} + \mu_X - \mu_X^2 \\
 &= n(n-1)p^2 \sum_{x=2}^n \frac{(n-2)\dots(n-x+1)}{(x-2)!} p^{x-2} (1-p)^{n-2-x+2} + \mu_X - \mu_X^2 \\
 &= n(n-1)p^2 \sum_{k=0}^{n-2} \frac{(n-2)\dots(n-2-k+1)}{k!} p^k (1-p)^{n-2-k} + \mu_X - \mu_X^2 \\
 &= n(n-1)p^2[(1-p) + p]^{n-2} + \mu_X - \mu_X^2 = \mu_X \left(1 - \frac{\mu_X}{n}\right)
 \end{aligned}$$

Mathematische Methoden in den Biowissenschaften

Eine Einführung mit R

Timischl, W.

2016, XIV, 456 S. 110 Abb.,

ISBN: 978-3-662-48952-9