

Fortbildung Mathematik  
Anwendungen der Logarithmus- und Exponentialfunktion

Prof. Dr. Vladimir Shikhman  
Professur für Wirtschaftsmathematik  
Technische Universität Chemnitz  
vladimir.shikhman@mathematik.tu-chemnitz.de

7. März 2024

# Inhaltsverzeichnis

<b>I Grundlagen</b>	<b>3</b>
1 Euler'sche Zahl . . . . .	4
2 Exponentialfunktion . . . . .	6
3 Logarithmusfunktion . . . . .	9
4 Differenzialeigenschaften . . . . .	11
<b>II Anwendungen</b>	<b>12</b>
1 Zinseszinsrechnung . . . . .	13
2 Verlauf einer Epidemie . . . . .	15
3 Signalverarbeitung durch Neurone . . . . .	17
4 Stimulus und Reaktion . . . . .	20
5 Anomalien in den Zahlen . . . . .	23
6 Ausbreitung von Innovationen . . . . .	27
7 Risikoaversion eines Entscheiders . . . . .	31

Teil I

Grundlagen

## 1 Euler'sche Zahl

Wir untersuchen auf Konvergenz die Folge

$$a_n = \left(1 + \frac{1}{n}\right)^n, \quad n \in \mathbb{N}.$$

Dafür wollen wir das *Monotoniekriterium* für Folgen anwenden. Es besagt nämlich, dass jede nach oben beschränkte und monoton wachsende Folge konvergent ist. Dabei heißt eine Folge *nach oben beschränkt*, falls es eine obere Schranke für alle Folgenglieder gibt. Eine Folge heißt *monoton wachsend*, falls das nächste Folgenglied das vorherige nicht unterschreitet. Wir zeigen nun, dass die Folge  $(a_n)_{n \in \mathbb{N}}$  monoton wachsend ist. Dazu vergleichen wir jeweils aufeinanderfolgende Folgenglieder:

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} = \left(\frac{n+2}{n+1}\right)^{n+1} : \left(\frac{n+1}{n}\right)^n = \left(\frac{n+2}{n+1}\right)^{n+1} \cdot \left(\frac{n}{n+1}\right)^n \\ &= \frac{(n+2)^{n+1} \cdot n^n}{(n+1)^{n+1} \cdot (n+1)^n} = \frac{(n+2)^{n+1} \cdot n^{n+1}}{(n+1)^{n+1} \cdot (n+1)^{n+1}} \cdot \frac{n+1}{n} \\ &= \left(\frac{(n+2) \cdot n}{(n+1)^2}\right)^{n+1} \cdot \frac{n+1}{n} = \left(\frac{(n+1+1) \cdot (n+1-1)}{(n+1)^2}\right)^{n+1} \cdot \frac{n+1}{n} \\ &= \left(\frac{(n+1)^2 - 1}{(n+1)^2}\right)^{n+1} \cdot \frac{n+1}{n} = \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \cdot \frac{n+1}{n}. \end{aligned}$$

Um den letzten Term abzuschätzen, benötigen wir die *Ungleichung von Bernoulli*:

$$(1+t)^k \geq 1+k \cdot t \quad \text{für alle } t \geq -1 \text{ und } k \in \mathbb{N}.$$

Zum Nachweis der Bernoullischen Ungleichung ziehen wir die vollständige Induktion heran:

Induktionsanfang:  $(1+t)^1 = 1+1 \cdot t$ .

Induktionsvoraussetzung: Es gelte  $(1+t)^{k-1} \geq 1+(k-1) \cdot t$  für ein  $k \in \mathbb{N}$ .

Induktionsschluss: Wir zeigen  $(1+t)^k \geq 1+k \cdot t$ . Es gilt in der Tat:

$$\begin{aligned} (1+t)^k &= (1+t)^{k-1} \cdot (1+t) \stackrel{\text{Induktions-}}{\geq} \text{voraussetzung} (1+(k-1) \cdot t) \cdot (1+t) \\ &= 1+t+(k-1) \cdot t+(k-1) \cdot t^2 = 1+k \cdot t+(k-1) \cdot t^2 \geq 1+k \cdot t. \end{aligned}$$

Nun setzen wir in der Bernoullischen Ungleichung  $t = -\frac{1}{(n+1)^2} \geq -1$  und  $k = n + 1$ :

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \cdot \frac{n+1}{n} = (1+t)^k \cdot \frac{n+1}{n} \geq (1+k \cdot t) \cdot \frac{n+1}{n} \\ &= \left(1 + (n+1) \cdot \left(-\frac{1}{(n+1)^2}\right)\right) \cdot \frac{n+1}{n} = \left(1 - \frac{1}{n+1}\right) \cdot \frac{n+1}{n} = 1. \end{aligned}$$

Für alle  $n \in \mathbb{N}$  gilt also  $a_{n+1} \geq a_n$ . Eine ähnliche Rechnung liefert die Beschränktheit von  $(a_n)_{n \in \mathbb{N}}$ . Nach dem Monotoniekriterium folgt insgesamt, dass die Folge  $(a_n)_{n \in \mathbb{N}}$  konvergiert. Ihr Grenzwert wird als *Euler'sche Zahl* bezeichnet:

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Das ist eine der wichtigsten Konstanten in der Mathematik. Sie wurde nach Leonhard Euler benannt, der sie in seinem Werk „Mechanica sive motus scientia analytice exposita“ aus dem Jahr 1736 mit dem Buchstaben e bezeichnete. Die Euler'sche Zahl lässt sich mit Hilfe dieser Definition approximieren, z. B.

$$e \approx \left(1 + \frac{1}{100000}\right)^{100000} = 2,7182\dots$$

Die Euler'sche Zahl ist irrational, d. h. als Bruch zweier ganzer Zahlen nicht darstellbar. Sie ist auch transzendent, d. h. sie lässt sich auch nicht als Nullstelle eines Polynoms mit rationalen Koeffizienten darstellen.

## 2 Exponentialfunktion

Man definiert die (*natürliche*) *Exponentialfunktion* als die unendliche Summe

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \dots + \frac{x^k}{k!} + \dots,$$

wobei die Fakultät  $k!$  das Produkt aller natürlichen Zahlen von 1 bis  $k$  angibt:

$$k! = 1 \cdot 2 \cdot \dots \cdot (k-1) \cdot k.$$

Wir untersuchen diese Reihe auf Konvergenz. Eine Reihe heißt *konvergent*, falls die zugehörige Folge ihrer Partialsummen konvergiert. Ihr Wert entspricht in diesem Fall dem reellen Grenzwert der Partialsummen. Eine hinreichende Bedingungen für die Konvergenz der Reihen ist das sogenannte Quotientenkriterium. Das *Quotientenkriterium* besagt, dass für die Konvergenz das betragsmäßige Verhältnis der aufeinanderfolgenden Reihenglieder entscheidend ist. Ist der Grenzwert dieses Verhältnisses kleiner als 1, so liegt Konvergenz vor. Es stellt sich heraus, dass die oben definierte Reihe nach dem Quotientenkriterium für jedes  $x \in \mathbb{R}$  konvergiert:

$$\left| \frac{x^k}{k!} \right| : \left| \frac{x^{k-1}}{(k-1)!} \right| = \frac{|x|^k}{k!} \cdot \frac{(k-1)!}{|x|^{k-1}} = \frac{|x|}{k} \rightarrow 0 < 1 \quad \text{für } k \rightarrow \infty.$$

Mit Hilfe der Definition weisen wir folgende Eigenschaften der Exponentialfunktion nach:

- $e^0 = 1$ . Es gilt:

$$e^0 = 1 + 0 + \frac{0^2}{2} + \dots + \frac{0^k}{k!} + \dots = 1.$$

- $e^{x+y} = e^x \cdot e^y$ . Aufgrund des Binomischen Lehrsatzes gilt:

$$\begin{aligned} e^x \cdot e^y &= \left( 1 + x + \frac{x^2}{2} + \dots + \frac{x^k}{k!} + \dots \right) \cdot \left( 1 + y + \frac{y^2}{2} + \dots + \frac{y^k}{k!} + \dots \right) \\ &= 1 + (x + y) + \left( \frac{x^2}{2} + x \cdot y + \frac{y^2}{2} \right) + \dots + \left( \sum_{\ell=0}^k \frac{x^\ell}{\ell!} \cdot \frac{y^{k-\ell}}{(k-\ell)!} \right) + \dots \\ &= 1 + (x + y) + \frac{1}{2} (x + y)^2 + \dots + \frac{1}{k!} \left( \sum_{\ell=0}^k \frac{k!}{\ell!(k-\ell)!} x^\ell \cdot y^{k-\ell} \right) + \dots \\ &= 1 + (x + y) + \frac{1}{2} (x + y)^2 + \dots + \frac{1}{k!} (x + y)^k + \dots = e^{x+y}. \end{aligned}$$

- $e^x \neq 0$ ,  $e^{-x} = \frac{1}{e^x}$ . Es gilt:

$$1 = e^0 = e^{x-x} = e^x \cdot e^{-x}.$$

- $e^x > 0$ . Für  $x \geq 0$  gilt:

$$e^x = 1 + x + \frac{x^2}{2} + \dots + \frac{x^k}{k!} + \dots > 0.$$

Für  $x < 0$  folgt daraus:

$$e^x = \frac{1}{e^{-x}} > 0.$$

- $e^x < e^y$  für  $x < y$ . Speziell für  $0 \leq x < y$  gilt:

$$e^x = 1 + x + \frac{x^2}{2} + \dots + \frac{x^k}{k!} + \dots < 1 + y + \frac{y^2}{2} + \dots + \frac{y^k}{k!} + \dots = e^y.$$

Der allgemeine Fall  $x < y$  folgt daraus:

$$e^y = e^{y-x+x} = e^{y-x} \cdot e^x > e^0 \cdot e^x = 1 \cdot e^x = e^x.$$

Der Graph der Exponentialfunktion ist in Abbildung I.1 dargestellt.

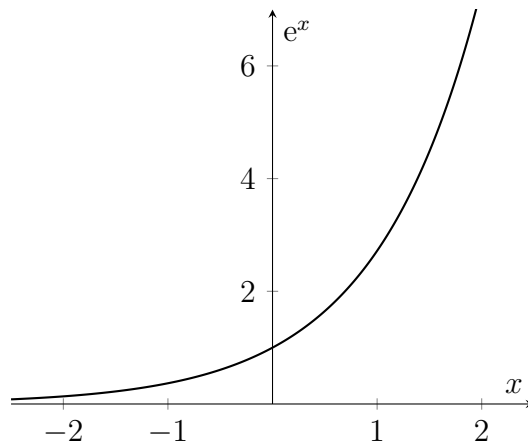


Abbildung I.1: Graph der Exponentialfunktion  $e^x$

## Zusammenhang mit der Euler'schen Zahl

Es stellt sich heraus, dass es sich bei dem Wert  $e^1$ , definiert über den Reihenansatz, um die *Euler'sche Zahl*  $e$ , definiert über den Folgenansatz, handelt:

$$\underbrace{\sum_{k=0}^{\infty} \frac{1}{k!}}_{=e^1} = \underbrace{\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n}_{=e}.$$

## 2. Exponentialfunktion

---

Nach dem Sandwich-Lemma reicht es, die  $n$ -te Partialsumme zwischen zwei Folgen einzuschließen und dann  $n \rightarrow \infty$  gegen Unendlich laufen zu lassen:

$$\underbrace{\left(1 + \frac{1}{n}\right)^n}_{\rightarrow e} \leq \sum_{k=0}^n \frac{1}{k!} \leq e, \quad n \in \mathbb{N}.$$

Das für die Konvergenz hinreichende *Sandwich-Lemma* wird oft *Satz von den zwei Polizisten* genannt. Ist ein von zwei „Polizisten“  $\left(1 + \frac{1}{n}\right)^n$  und  $e$  verfolgter „Verbrecher“  $\sum_{k=0}^n \frac{1}{k!}$  zwischen ihnen eingeschlossen, so wird er über kurz oder lang gestellt. Für den Nachweis der obigen Ungleichung brauchen wir die folgende Identität:

$$\begin{aligned} \binom{n}{k} \cdot \frac{1}{n^k} &= \frac{n!}{k!(n-k)!} \cdot \frac{1}{n^k} = \frac{(n-k)!(n-k+1) \cdot \dots \cdot n}{k!(n-k)!} \cdot \frac{1}{n^k} \\ &= \frac{1}{k!} \cdot \frac{(n-k+1) \cdot \dots \cdot n}{n^k} = \frac{1}{k!} \cdot \prod_{\ell=0}^{k-1} \left(\frac{n-\ell}{n}\right) = \frac{1}{k!} \cdot \prod_{\ell=0}^{k-1} \left(1 - \frac{\ell}{n}\right). \end{aligned}$$

Wir schätzen nun mit Hilfe des Binomischen Lehrsatzes ab:

$$\left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} \cdot \frac{1}{n^k} \cdot 1^{n-k} = \sum_{k=0}^n \frac{1}{k!} \cdot \underbrace{\prod_{\ell=0}^{k-1} \left(1 - \frac{\ell}{n}\right)}_{\leq 1} \leq \sum_{k=0}^n \frac{1}{k!}.$$

Damit ist die linke Seite der obigen Ungleichung gültig. Für  $m \geq n$  haben wir ferner:

$$\left(1 + \frac{1}{m}\right)^m = \sum_{k=0}^m \binom{m}{k} \cdot \frac{1}{m^k} \geq \sum_{k=0}^n \binom{m}{k} \cdot \frac{1}{m^k} = \sum_{k=0}^n \frac{1}{k!} \cdot \prod_{\ell=0}^{k-1} \left(1 - \frac{\ell}{m}\right).$$

Man geht hier zum Grenzwert  $m \rightarrow \infty$  über und erhält die rechte Seite der obigen Ungleichung:

$$e = \lim_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^m \geq \lim_{m \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} \cdot \underbrace{\prod_{\ell=0}^{k-1} \left(1 - \frac{\ell}{m}\right)}_{\rightarrow 1} = \sum_{k=0}^n \frac{1}{k!}.$$



### 3 Logarithmusfunktion

Die Exponentialfunktion  $e : \mathbb{R} \rightarrow (0, \infty)$  wächst monoton und hat einen stetigen lückenlosen Verlauf. Daher ist sie bijektiv und lässt sich somit umkehren. Jede positive Zahl  $u > 0$  wird von einem eindeutigen Urbild  $x$  erreicht, sodass  $e^x = u$  gilt. Die so konstruierte Umkehrfunktion  $\ln : (0, \infty) \rightarrow \mathbb{R}$  heißt der (*natürliche*) *Logarithmus*, d. h.

$$x = \ln u \quad \Leftrightarrow \quad e^x = u.$$

Der Graph der Logarithmusfunktion ist in Abbildung I.2 dargestellt.

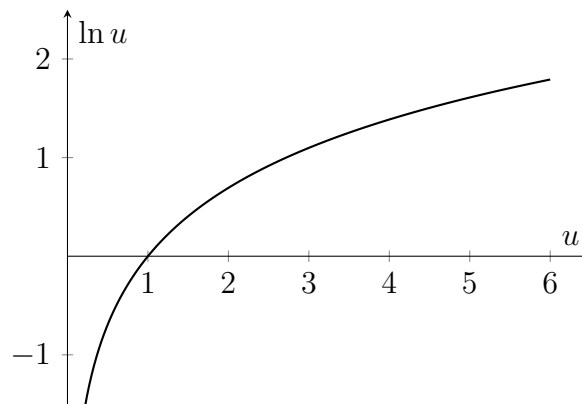


Abbildung I.2: Graph der Logarithmusfunktion  $\ln u$

Folgende Eigenschaften der Logarithmusfunktion sind schnell hergeleitet:

- $\ln 1 = 0$ ,  $\ln e = 1$ . Es gilt:

$$\ln 1 = \ln(e^0) = 0, \quad \ln e = \ln(e^1) = 1.$$

- $\ln(u \cdot v) = \ln u + \ln v$ . Es gilt:

$$e^{\ln u + \ln v} = e^{\ln u} \cdot e^{\ln v} = u \cdot v = e^{\ln(u \cdot v)}.$$

- $\ln \frac{1}{u} = -\ln u$ . Es gilt:

$$0 = \ln 1 = \ln\left(u \cdot \frac{1}{u}\right) = \ln u + \ln \frac{1}{u}.$$

- $\ln u < \ln v$  für  $u < v$ . Wegen der Monotonie der Exponentialfunktion gilt:

$$e^{\ln u} = u < v = e^{\ln v} \quad \Leftrightarrow \quad \ln u < \ln v.$$

### 3. Logarithmusfunktion

---

#### Allgemeine Basis

Mit Hilfe der natürlichen Exponential- und Logarithmusfunktion kann man auch allgemeine Exponential- bzw. Logarithmusfunktionen zur Basis  $a > 0$  definieren:

$$a^x = e^{x \cdot \ln a}, \quad \log_a u = \frac{\ln u}{\ln a}.$$

Allgemeine Exponential- und Logarithmusfunktionen sind zueinander invers:

$$a^x = u \quad \Leftrightarrow \quad x = \log_a u.$$

Es gelten für sie folgende Rechenregeln, die leicht nachzuprüfen sind:

- $(a^x)^y = a^{x \cdot y}$ ,  $a^{x+y} = a^x \cdot a^y$ ,  $a^x \cdot b^x = (a \cdot b)^x$ ,  $\left(\frac{1}{a}\right)^x = a^{-x}$ .
- $\log_a \left(\frac{u}{v}\right) = \log_a u - \log_a v$ ,  $\log_a(u \cdot v) = \log_a u + \log_a v$ ,  $\log_a u^v = v \cdot \log_a u$ .

## 4 Differenzialeigenschaften

Nun können wir die Ableitungen der beiden uns interessierenden Funktionen ausrechnen.

- Exponentialfunktion:

$$\begin{aligned}
 (e^x)' &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right)' \\
 &\stackrel{\text{Summenregel}}{=} (1)' + (x)' + \left(\frac{x^2}{2!}\right)' + \left(\frac{x^3}{3!}\right)' + \dots \\
 &\stackrel{\text{Faktorregel}}{=} 0 + 1 + \frac{1}{2!} (x^2)' + \frac{1}{3!} (x^3)' + \dots \\
 &\stackrel{\text{Produktregel}}{=} 1 + \frac{2x}{1! \cdot 2} + \frac{3x^2}{2! \cdot 3} + \dots = \\
 &= 1 + x + \frac{x^2}{2!} + \dots = e^x.
 \end{aligned}$$

- Logarithmusfunktion:

$$e^{\ln x} = x \quad \Rightarrow \quad (e^{\ln x})' = x' \stackrel{\text{Kettenregel}}{\Rightarrow} e^{\ln x} \cdot \ln' x = 1 \quad \Rightarrow \quad \ln' x = \frac{1}{x}.$$

Da die Ableitungen der Exponential- und Logarithmusfunktion auf ihren jeweiligen Definitionsbereichen positiv sind, d. h.

$$(e^x)' = e^x > 0 \text{ für } x \in \mathbb{R}, \quad \ln' x = \frac{1}{x} > 0 \text{ für } x > 0,$$

sind die beiden Funktionen streng monoton steigend. Sie haben also keine Extremstellen. Wir untersuchen auch ihre zweiten Ableitungen:

$$(e^x)'' = e^x > 0 \text{ für } x \in \mathbb{R}, \quad \ln'' x = \left(\frac{1}{x}\right)' = -\frac{1}{x^2} < 0 \text{ für } x > 0.$$

Daraus folgt, dass die Exponentialfunktion konvex und die Logarithmusfunktion konkav ist. Die Konvexität bzw. Konkavität einer Funktion besagt, dass ihr Graph nach unten bzw. nach oben gekrümmt ist. Außerdem weist die Exponential- und Logarithmusfunktion keine Wendepunkte auf.

Teil II

Anwendungen

# 1 Zinseszinsrechnung

Die Vergabe von Krediten ist eine der wichtigsten Aufgaben von Banken und Sparkassen. Heutzutage geschieht dies auf der Basis einer Verzinsung: Man zahlt nicht nur den aufgenommenen Kredit zurück, sondern auch die darauf entfallenden Zinsen zu einem im Voraus vereinbarten Zinssatz. Grundlegend für das Finanzwesen ist deshalb die folgende Fragestellung:

Auf welche Art und Weise werden Zinsen berechnet?

Seit jeher war es in diesem Kontext höchst umstritten, ob man überhaupt Zinsen berechnen darf. In den Worten des Hl. Thomas von Aquin:

„Zins für geliehenes Geld nehmen [...] an sich ungerecht, weil etwas verkauft wird, was nicht besteht. [...] Nach der gleichen Überlegung begeht eine Ungerechtigkeit, wer Wein oder Getreide herleiht und dafür eine doppelte Vergütung verlangt, auf der einen Seite die Wiedererstattung der gleichen Sache, auf der anderen aber einen Preis für den Gebrauch, den man Zins nennt.“

Martin Luther relativierte dieses Verbot:

„Wenn nu das geschicht an ubertretung des geystlichen gesetzes / das man auff's hundert 4, 5, 6 Gulden gibt / lest sichs tragen / Doch soll allzeyt Gottis forschet sorgfelig seyn.“

Noch umstrittener war die Frage, ob man auf Zinsen ebenfalls weitere Zinsen berechnen darf. Gottfried Wilhelm Leibniz meinte anfangs:

„Aber ist vielleicht der zwischenzeitliche Zins nicht nach einfachen Zinsen, sondern nach dem Zinseszins zu berechnen? Ich denke, dass eine Berechnung nach dem Zinseszins nicht gebilligt werden kann. [...] die Angelegenheiten der meisten Menschen an den meisten Orten verhalten sich nicht so, dass sie aus den Zinsen des ersten Jahres wieder andere Zinsen nach demselben Verhältnis erwerben können“.

Nach einem darauf folgenden Studium verwarf Leibniz seine Bedenken gegen den Zinseszins und leitete die Zinseszinsformel her. Die Zinseszinsformel wurde von ihm in einer juristisch-ökonomischen Abhandlung ausgearbeitet, welche in der Leipziger Zeitschrift „Acta Eruditorum“ 1682 veröffentlicht wurde.

Wir leiten die Zinseszinsformel für die unterjährige Verzinsung her. Dafür nehmen wir an, dass der Betrag von  $K$  Euro als Kredit aufgenommen wird und zwar zum Jahreszinssatz von  $i$ %. In einem Jahr soll der Schuldner also  $K \cdot (1 + i)$  Euro zurückzahlen. Der Gläubiger kommt auf die Idee, stattdessen die halben Zinsen von  $\frac{i}{2}$ % halbjährlich dem Schuldner in Rechnung zu stellen:

$$\underbrace{K \cdot \left(1 + \frac{i}{2}\right)}_{\substack{\text{Verzinsung von} \\ K \text{ Euro} \\ \text{im ersten Halbjahr}}} \cdot \underbrace{\left(1 + \frac{i}{2}\right)}_{\substack{\text{Verzinsung von} \\ K \cdot \left(1 + \frac{i}{2}\right) \text{ Euro} \\ \text{im zweiten Halbjahr}}} = K \cdot \left(1 + \frac{i}{2}\right)^2.$$

## 1. Zinseszinsrechnung

---

Im zweiten Halbjahr wird der Zinseszins angewandt und der Schuldner muss am Ende des Jahres mehr als früher zurückzahlen. Tatsächlich gilt:

$$\underbrace{K \cdot \left(1 + \frac{i}{2}\right)^2}_{\text{Schuld mit Zinseszins}} = K \cdot \left(1 + i + \frac{i^2}{4}\right) = \underbrace{K \cdot (1 + i)}_{\text{Schuld ohne Zinseszins}} + K \cdot \frac{i^2}{4}.$$

Der Zinseszins bringt dem Gläubiger zusätzliche  $K \cdot \frac{i^2}{4}$  Euro ein. Man setzt diesen Gedanken fort und unterteilt das Jahr in  $n$  gleiche Zeitabschnitte (etwa Quartale, Monate, Tage usw.). Im  $n$ -ten Anteil des Jahres verzinst man also zum Zinssatz  $\frac{i}{n}$ . Die *Zinseszinsformel* für die unterjährig Verzinsung des Kredits  $K$  lautet:

$$K_n = K \cdot \underbrace{\left(1 + \frac{i}{n}\right) \cdot \left(1 + \frac{i}{n}\right) \cdot \dots \cdot \left(1 + \frac{i}{n}\right)}_{n\text{-mal}} = K \cdot \left(1 + \frac{i}{n}\right)^n.$$

Es ist entscheidend, ob die Folge  $(K_n)_{n \in \mathbb{N}}$  einen Grenzwert besitzt. Ansonsten könnte der Gläubiger den Schuldner lediglich über den Zinseszins ausplündern, indem man das Jahr in immer mehrere Zeitabschnitte für  $n \rightarrow \infty$  unterteilt. Die Zinseszinsfolge wollen wir im Weiteren auf Konvergenz untersuchen.

Dafür wird in der folgenden Rechnung  $m = \frac{n}{i}$  gesetzt und die Definition der Euler'schen Zahl benutzt:

$$\begin{aligned} \lim_{n \rightarrow \infty} K_n &= \lim_{n \rightarrow \infty} K \cdot \left(1 + \frac{i}{n}\right)^n = K \cdot \lim_{n \rightarrow \infty} \left(1 + \frac{1}{\frac{n}{i}}\right)^{\frac{n}{i} \cdot i} \\ &= K \cdot \lim_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^{m \cdot i} = K \cdot \left(\lim_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^m\right)^i = K \cdot e^i. \end{aligned}$$

Das ist die Formel der sogenannten *stetigen Verzinsung*, die im Finanz- und Versicherungswesen eine wichtige Rolle spielt. Beispielsweise liefert die stetige Verzinsung von  $K = 10\,000$  Euro zum Zinssatz  $i = 10\%$  den Betrag

$$10\,000 \cdot e^{0,1} \approx 11\,051,7.$$

Beim Vergleich mit dem einmal jährlich verzinsten Betrag

$$10\,000 \cdot (1 + 0,1) = 11\,000$$

ist eine Differenz von immerhin knapp 52 Euro festzustellen.

## 2 Verlauf einer Epidemie

Eine Epidemie, auch Seuche genannt, ist ein zeitlich und örtlich begrenztes vermehrtes Auftreten von Krankheitsfällen einheitlicher Ursache innerhalb einer menschlichen Population und entspricht damit einem großen Ausbruch einer Krankheit. In der Epidemiologie wird von einer Epidemie gesprochen, wenn die Anzahl an neuen Erkrankungsfällen (Inzidenzen) über einen gewissen Zeitraum in einer bestimmten Region zunimmt. Nach der Geschwindigkeit der Zunahme der Erkrankungsfälle werden Explosiv- und Tardivepidemie unterschieden. Bei der Explosiv-epidemie handelt es sich um eine Epidemie mit schlagartigem Anstieg der Erkrankungszahlen. Tardivepidemie (vom lateinische tardus/ardivus „langsam“, „verspätet“) heißt eine Epidemie mit langsam, aber stetig ansteigenden Erkrankungszahlen. Bei einer Länder und Kontinente übergreifenden Ausbreitung wird von einer Pandemie gesprochen. Als eine Endemie wird demgegenüber das andauernd gehäufte Auftreten einer Krankheit in einer umschriebenen Population bezeichnet. Hierbei bleibt die Inzidenz annähernd gleich, ist aber gegenüber nichtendemischen Gebieten erhöht. Die folgende Fragestellung aus dem Bereich der Epidemieforschung ist diesbezüglich relevant:

Nach welchem Muster verbreitet sich eine Epidemie?

Setzen wir voraus, dass  $A_0 > 0$  die Anzahl der aktuell Infizierten ist. Aus Erfahrung zeichnet sich das folgende Infizierungsschema ab:

- Neuinfizierte werden nur von zuletzt Infizierten angesteckt.
- Jeder Kranke kann  $x$  potenzielle Neuinfizierte anstecken.
- Das Ansteckungsrisiko sinkt linear mit der Zeit.

Die letzte Eigenschaft bedeutet, dass die Ansteckungskraft der Infizierten mit fortschreitender Zeit abnimmt. Die Population wird zum einen mit der Zeit widerstandsfähiger. Zum anderen mutiert das Virus und passt sich dahingehend an die Menschen an, dass diese bloß erkranken und nicht sterben. Um den Verlauf der Epidemie zu modellieren, bezeichnen wir die Anzahl der Neuinfizierten zum Zeitpunkt  $k$  mit  $A_k$ . Deren Entwicklung sieht also wie folgt aus:

$$A_k = \frac{x}{k} \cdot A_{k-1}, \quad k \in \mathbb{N}.$$

Die Anzahl der Neuinfizierten zum Zeitpunkt  $k$  ist proportional zur Anzahl der Neuinfizierten zum Zeitpunkt  $k - 1$ . Dabei zeugt der abklingende Proportionalitätsfaktor  $\frac{1}{k}$  von immer kleiner werdenden Ansteckungsrisiken. Aus der obigen Dynamik leiten wir her:

$$A_k = \frac{x}{k} \cdot A_{k-1} = \frac{x}{k} \cdot \frac{x}{k-1} \cdot A_{k-2} = \dots = \frac{\overbrace{x \cdot x \cdot \dots \cdot x}^{k\text{-mal}}}{k \cdot (k-1) \cdot \dots \cdot 1} \cdot A_0 = \frac{x^k}{k!} \cdot A_0.$$

Die Gesamtanzahl so infizierter Menschen lässt sich als eine unendliche Summe darstellen:

$$A_0 + A_1 + A_2 + \dots + A_k + \dots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \cdot A_0 = e^x \cdot A_0.$$

## 2. Verlauf einer Epidemie

---

Wenn exemplarisch mit nur einem potenziellen Neuinfizierten gerechnet wird, steigt die Anzahl der Infizierten trotzdem um knapp das Dreifache deren Anfangszahl:

$$\sum_{k=0}^{\infty} \frac{1}{k!} \cdot A_0 = e \cdot A_0 \approx 2,7182 \cdot A_0.$$

Im Unterschied zu einer Krankheit, welche innerhalb einer Population fortwährend mit etwa gleicher Fallzahl auftritt, verbreitet sich eine Epidemie mit einer größeren Reproduktionsrate. Dies bedeutet bei einer Infektionskrankheit, dass die Zahl an Infizierten zunimmt und die Zahl an Neuinfektionen ansteigt. Für die Ausbreitung bedeutsam ist die Rate, mit der durch Kontakt mit Infizierten neue Infizierte auftreten. Sie entspricht zu Beginn der Basisreproduktionszahl. Anfangs erhöht sich die Zahl an neuen Infektionsfällen pro Zeitintervall im Vergleich zum vorigen um einen ungefähr gleichen Anteil und wächst exponentiell. Der Anstieg neuer Infektionsfälle in absoluten Zahlen fällt daher zunächst eher gering aus und wächst mit fortschreitendem Geschehen stärker an.



### 3 Signalverarbeitung durch Neurone

*Neuronen* sind spezielle Zellen im Nervensystem, die Informationen an andere Nerven-, Muskel- oder Drüsenzellen übertragen. Typischerweise hat ein Neuron einen Zellkörper, einen Axon und Dendriten, siehe Abbildung II.1. Der Zellkörper enthält den Zellkern, der von Zytoplasma umgeben ist. Dendriten erstrecken sich vom Zellkörper des Neurons und empfangen Nachrichten von anderen Neuronen. Synapsen sind die Kontaktstellen, an denen ein Neuron mit einem anderen kommuniziert. Die Dendriten sind mit Synapsen bedeckt, die sich an den Enden der Axone anderer Neuronen bilden. Wenn Neuronen Nachrichten senden, übertragen sie elektrische Impulse entlang ihrer Axone.

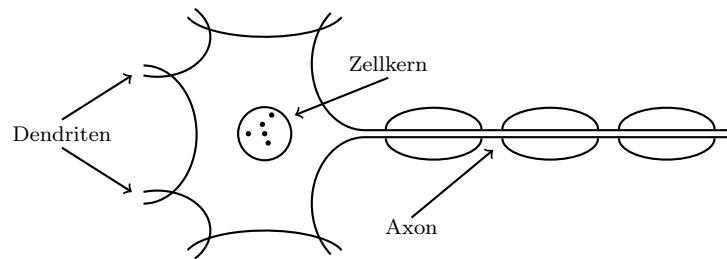


Abbildung II.1: Neuron

Lassen Sie uns das mathematische Modell eines Neurons präsentieren, siehe Abbildung II.2. Dafür sollen die Eingaben durch  $m-1$  Dendriten erfolgen, wie durch die Zahlen  $x_1, \dots, x_{m-1} \in \mathbb{R}$  gegeben. Abhängig von der Bedeutung eines bestimmten Dendriten wird diese Information im Zellkern verarbeitet:

$$w_0 + x_1 \cdot w_1 + \dots + x_{m-1} \cdot w_{m-1},$$

wobei  $w_1, \dots, w_{m-1} \in \mathbb{R}$  einige unbekannte Gewichte sind und  $w_0 \in \mathbb{R}$  die Rolle eines Bias spielt. Lassen Sie uns diese lineare Kombination kurz als  $x^T \cdot w$  schreiben, wobei die Vektoren der Eingaben und Gewichte durch

$$x = (1, x_1, \dots, x_{m-1})^T, \quad w = (w_0, w_1, \dots, w_{m-1})^T.$$

bezeichnet werden. Dann leitet das Axon das aggregierte Signal an andere Neuronen weiter durch

$$y = f(x^T \cdot w),$$

wobei  $y \in \mathbb{R}$  die Ausgabe ist und  $f: \mathbb{R} \rightarrow \mathbb{R}$  eine passend gewählte *Aktivierungsfunktion* ist. Diese Gleichung kann als *verallgemeinerte lineare Regression* interpretiert werden.

Um konkret zu werden, möchten wir, dass ein Neuron zu Klassifizierungszwecken trainiert wird. Dafür besteht der Datensatz einerseits aus  $n$  gegebenen Eingabevektoren:

$$x_i = (1, (x_i)_1, \dots, (x_i)_{m-1})^T, \quad i = 1, \dots, n.$$

Andererseits werden sie in zwei Klassen  $C_{yes}$  und  $C_{no}$  unterteilt. Entsprechend sind die Datenproben durch binäre Ausgaben gekennzeichnet, d.h. für alle  $i = 1, \dots, n$  gilt:

$$y_i = \begin{cases} 1, & \text{if } i \in C_{yes}, \\ 0, & \text{if } i \in C_{no}. \end{cases}$$

### 3. Signalverarbeitung durch Neurone

---

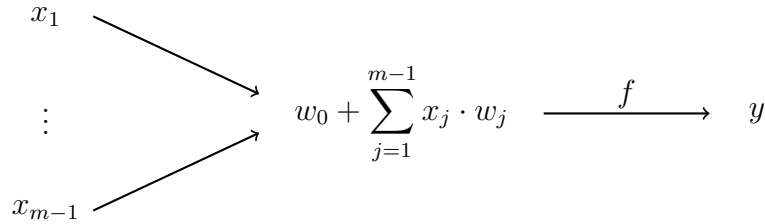


Abbildung II.2: Mathematisches Modell eines Neurons

Basierend auf dem Datensatz  $(x_i, y_i) \in \mathbb{R}^m \times \{0, 1\}$ ,  $i = 1, \dots, n$ , müssen wir die Gewichte  $w$  trainieren. Danach kann die Klassifizierung einfach durch Verwendung der *Schwellenwert-Aktivierungsfunktion* durchgeführt werden:

$$f_T(z) = \begin{cases} 1, & \text{if } z \geq 0, \\ 0, & \text{else.} \end{cases}$$

Tatsächlich kennzeichnen wir einen Neuling  $x$  einfach durch Einstellung:

$$y = f_T(x^T \cdot w).$$

Eine andere Möglichkeit wäre es, die Schwellenwert-Aktivierungsfunktion  $f_T$  durch die *Sigmoid-Aktivierungsfunktion* zu approximieren, siehe Abbildung II.3:

$$f_S(z) = \frac{1}{1 + e^{-z}}.$$

Der Wert  $f_S(z)$  kann als die Wahrscheinlichkeit interpretiert werden, dass  $z$  nichtnegativ ist. Durch Verwendung der Sigmoid-Aktivierungsfunktion können wir einen Neuling  $x$  äquivalent wie folgt kennzeichnen:

$$y = \begin{cases} 1, & \text{if } f_S(x^T \cdot w) \geq 1/2, \\ 0, & \text{else.} \end{cases}$$

Je nach Aktivierungsfunktion untersucht man dann, wie ein Neuron die Gewichte erfolgreich lernen bzw. trainieren kann. Um die Verwendung der Sigmoid-Aktivierungsfunktion  $f_S$  zu rechtfertigen, berechnen wir zunächst ihre Inverse. Für  $p \in [0, 1]$  und  $z \in \mathbb{R}$  gilt:

$$p = \frac{1}{1 + e^{-z}} \quad \text{genau dann, wenn} \quad z = \ln \frac{p}{1 - p}.$$

Der letztere Ausdruck wird als *Logit* bezeichnet:

$$\text{logit}(p) = \ln \frac{p}{1 - p}.$$

Er nimmt den Logarithmus der Chancen, die die Wahrscheinlichkeit  $p$  für Erfolg und die Wahrscheinlichkeit  $1 - p$  für Misserfolg verknüpfen. Äquivalent können wir die Aktivierung eines Neurons

$$y = f_S(x^T \cdot w)$$

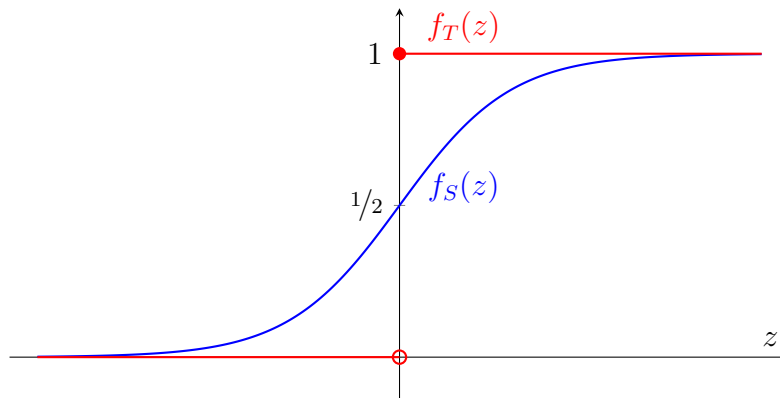


Abbildung II.3: Aktivierungsfunktionen

wie folgt mithilfe des Logits umformulieren:

$$x^T \cdot w = \text{logit}(y).$$

Die logistische Regression repräsentiert somit linear die *Log-Odds*. Es ist interessant zu bemerken, dass hier der Logarithmus wesentlich ist. Seine Verwendung beruht auf dem Weber-Fechner-Gesetz aus der Psychophysik. Das *Weber-Fechner-Gesetz* bezieht sich auf die menschliche Wahrnehmung, genauer gesagt auf die Beziehung zwischen der tatsächlichen Veränderung eines physikalischen Reizes und der wahrgenommenen Veränderung. Dies umfasst Reize aller Sinne, wie Sehen, Hören, Geschmack, Berührung und Geruch. Es besagt, dass die subjektive Empfindung proportional zum Logarithmus der Reizintensität ist. In unserem Fall wird die Reizintensität durch die Chancen  $\frac{y}{1-y}$  ausgedrückt, und die Wahrnehmung  $x^T \cdot w$  wird als linear bezüglich der unbekanntes Gewichte angenommen.

## 4 Stimulus und Reaktion

Bei der Anpassung auf dem Markt bestehender Produkte ist es zwingend erforderlich, die Kundenreaktionen darauf voraussehen zu können. Die Produkteigenschaften (wie z. B. Zucker-, Fettgehalt oder Farbintensität) werden dabei als physische Stimuli aufgefasst, auf welche Kunden in ihrem Kaufverhalten positiv bzw. negativ reagieren. Solche psychischen Reaktionen müssen quantitativ erfasst werden, um basierend darauf eine vorteilhafte Produktgestaltung zu ermöglichen. In der Verhaltenspsychologie ist daher die folgende Fragestellung zentral:

Wie hängen psychische Reaktionen von physischen Stimuli ab?

Um den gleichen Wahrnehmungseffekt zu erzielen, muss man den Reiz, bezogen auf das aktuelle Niveau, ändern. Wenn man also ein Produkt mit einem sehr hohen Zuckergehalt konsumiert, wird die für eine unterscheidbare Wahrnehmung erforderliche Änderung des Zuckergehalts umso höher ausfallen. Wir bezeichnen diese konstante Reaktionsänderung mit  $\Delta R$  und die dafür benötigte Stimulusänderung mit  $\Delta S$ . Das von Ernst Weber im Jahr 1834 aufgestellte und nach ihm benannte *Weber-Gesetz* aus der Psychophysik besagt, dass  $\Delta S$  proportional zum aktuellen Stimulus  $S$  ist:

$$\Delta S = k \cdot S,$$

wobei  $k > 0$  ein Proportionalitätsfaktor ist. Aus dieser Formel folgt insbesondere: Ist der Stimulus  $S$  verhältnismäßig stark, muss die konstante Reaktionsänderung  $\Delta R$  hervorrufende Stimulusänderung  $\Delta S$  entsprechend ausgeprägt sein. Für die Geschmackswahrnehmung liegt experimentell diese Stimulusänderung bei etwa 15 % des aktuellen Intensitätsniveaus  $S$ , d. h.  $k = 0,15$ . Wir vergleichen exemplarisch die Wahrnehmungen der Getränke mit einem unterschiedlichen Zuckergehalt in Tabelle II.1. Um denselben geschmacklichen Unterschied Kunden

Getränke	Zuckergehalt $S$	Änderung des Zuckergehalts $\Delta S = 0,15 \cdot S$
Coca-Cola	10,6 g in 100 ml	1,590 g in 100 ml
Pepsi	10,0 g in 100 ml	1,500 g in 100 ml
Fanta	9,4 g in 100 ml	1,410 g in 100 ml
Sprite	9,1 g in 100 ml	1,365 g in 100 ml

Tabelle II.1: Zuckergehalt in Getränken

wahrnehmen zu lassen, muss man in 100 ml Coca-Cola 1,590 g Zucker dazutun, wobei für Sprite lediglich 1,365 g ausreichen.

Wir bemühen uns nun um eine qualitative Beschreibung der Kopplung zwischen Stimulus und Reaktion. Gesucht ist eine Funktion  $R(S)$ , die zu jedem physischen *Stimulus*  $S$  die entsprechende *Reaktion*  $R(S)$  angibt. Von der *Wahrnehmungsfunktion*  $R(S)$  fordern wir, dass sie das Weber-Gesetz erfüllt. Für  $\Delta R$  existiert  $k > 0$ , sodass für alle  $S$  gilt:

$$\Delta R = R(S + \Delta S) - R(S) \quad \Rightarrow \quad \Delta S = k \cdot S.$$

Ändert sich der Stimulus um  $\Delta S$  vom Niveau  $S$  aus und ruft dabei die konstante Reaktionsänderung  $\Delta R$  gemäß der Wahrnehmungsfunktion  $R$  hervor, so ist  $\Delta S$  zu  $S$  proportional.

Es wird sich herausstellen, dass der Logarithmus diesen funktionalen Zusammenhang recht adäquat beschreibt. Für die Wahrnehmungsfunktion verfolgen wir den logarithmischen Ansatz von Gustav Fechner, den er 1860 vorschlug:

$$R(S) = \mu \cdot \ln \frac{S}{S_0},$$

wobei  $\mu > 0$  den Proportionalitätsfaktor und  $S_0 > 0$  den maximalen nicht wahrnehmbaren Stimulus angeben. Wir prüfen nach, ob die logarithmische Wahrnehmungsfunktion das Weber-Gesetz erfüllt, d. h. ob für  $\Delta R$  ein  $k > 0$  existiert, sodass für alle  $S$  gilt:

$$\Delta R = R(S + \Delta S) - R(S) \quad \Rightarrow \quad \Delta S = k \cdot S.$$

Wir fangen mit der Reaktionsänderung an:

$$\Delta R = R(S + \Delta S) - R(S) = \mu \cdot \ln \frac{S + \Delta S}{S_0} - \mu \cdot \ln \frac{S}{S_0} = \mu \cdot \ln \frac{S + \Delta S}{S}.$$

Daraus folgt sofort:

$$\frac{\Delta R}{\mu} = \ln \left( \frac{S + \Delta S}{S} \right) \quad \Leftrightarrow \quad e^{\frac{\Delta R}{\mu}} = \frac{S + \Delta S}{S} \quad \Leftrightarrow \quad \Delta S = \left( e^{\frac{\Delta R}{\mu}} - 1 \right) \cdot S.$$

Demnach gilt das Weber-Gesetz  $\Delta S = k \cdot S$  mit dem Proportionalitätsfaktor

$$k = e^{\frac{\Delta R}{\mu}} - 1 > 0.$$

Außerdem können wir  $\mu$  in der Abhängigkeit von  $k$  ausdrücken:

$$k = e^{\frac{\Delta R}{\mu}} - 1 \quad \Leftrightarrow \quad k + 1 = e^{\frac{\Delta R}{\mu}} \quad \Leftrightarrow \quad \ln(k + 1) = \frac{\Delta R}{\mu} \quad \Leftrightarrow \quad \mu = \frac{\Delta R}{\ln(k + 1)}.$$

Nach dem *Weber-Fechner-Gesetz* nimmt die Wahrnehmungsfunktion die folgende Gestalt an:

$$R(S) = \frac{\Delta R}{\ln(k + 1)} \cdot \ln \frac{S}{S_0},$$

wobei die Konstante  $k$  aus dem Weber-Gesetz experimentell bestimmt werden kann.

Getränke	Zuckergehalt $S$	Geschmack $R(S) \approx 7,155 \cdot \ln S$
Coca-Cola	10,6 g in 100 ml	16,89
Pepsi	10,0 g in 100 ml	16,47
Fanta	9,4 g in 100 ml	16,03
Sprite	9,1 g in 100 ml	15,80

Tabelle II.2: Geschmäcker der Getränke

#### 4. Stimulus und Reaktion

---

Die Geschmacksfunktion für das obige Getränkebeispiel mit  $\Delta R = 1$ ,  $k = 0,15$ ,  $S_0 = 1$  lautet:

$$R(S) = \frac{1}{\ln(0,15 + 1)} \cdot \ln \frac{S}{S_0} = \frac{1}{\ln 1,15} \cdot \ln \frac{S}{1} \approx 7,155 \cdot \ln S.$$

Wir berechnen nun die Geschmäcker verschiedener Getränke in Tabelle II.2. In 100 ml Sprite wird  $10,6 - 9,1 = 1,5$  g weniger Zucker als in Coca-Cola hinzugefügt. Dadurch spart man rund 14,15 % Zucker. Der Geschmack verschlechtert sich aber um  $16,89 - 15,80 = 1,09$  Wahrnehmungseinheiten, was lediglich 6,45 % ausmacht. Es lohnt sich also durchaus, bei Zucker zu sparen, ohne Geschmack sonderlich einzubüßen. Diese Schlussfolgerung wird dadurch begründet, da es sich bei den Getränken um verhältnismäßig hohe Zuckergehalte handelt.

## 5 Anomalien in den Zahlen

Den Anomalien in den Zahlen liegt eine Beobachtung zugrunde, dass in vielen realen numerischen Datensätzen die führende Ziffer aller Wahrscheinlichkeit nach klein ausfällt. 1881 wurde diese Gesetzmäßigkeit von dem Astronomen und Mathematiker Simon Newcomb entdeckt und publiziert. Er hatte bemerkt, dass in den benutzten Büchern mit Logarithmentafeln die Seiten mit Tabellen mit Eins als erster Ziffer deutlich schmutziger waren als die anderen Seiten, weil sie offenbar öfter benutzt worden waren. Die Abhandlung Newcombs blieb unbeachtet und war schon in Vergessenheit geraten, als der Physiker Frank Benford dieselbe Gesetzmäßigkeit wiederentdeckte und sie 1938 erneut publizierte. Seither war sie nach ihm benannt, in neuerer Zeit wird aber durch die Bezeichnung "Newcomb-Benford'sches Gesetz" der ursprüngliche Entdecker ebenfalls bedacht. Die Existenz einer solchen Gesetzmäßigkeit war selbst unter Statistikern nicht vielen bewusst, bis der Mathematiker Theodore Hill versuchte, die Benford-Verteilung zur Lösung praktischer Probleme nutzbar zu machen, und sie dadurch wesentlich bekannter machte. Das Gesetz lässt sich etwa in Datensätzen über Einwohnerzahlen von Städten, Geldbeträge in der Buchhaltung, Naturkonstanten etc. beobachten. Wir wenden uns also der folgenden Fragestellung aus der Statistik:

Wie wahrscheinlich ist das Auftreten der führenden Ziffer  
1, 2, 3, 4, 5, 6, 7, 8 oder 9 in zufälligen Datensätzen?

Das *Newcomb-Benford'sche Gesetz* besagt, dass für empirisch gegebene Zahlen die Ziffer  $d \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$  mit Wahrscheinlichkeit

$$\mathbb{P}(d) = \log_{10}(d + 1) - \log_{10}(d)$$

als erste in der Dezimaldarstellung der Zahlen vorkommen wird. Die Größe  $\mathbb{P}(d)$  ist also proportional zum Abstand zwischen  $d$  und  $d + 1$  auf der logarithmischen Skala zur Basis 10. Zahlenmäßig bekommen wir konkret folgende Wahrscheinlichkeiten, siehe Tabelle II.3. Abbildung II.4 veranschaulicht die entsprechenden Wahrscheinlichkeiten in Form eines Säulendiagramms.

Ziffer $d$	1	2	3	4	5	6	7	8	9
Wahrscheinlichkeit $\mathbb{P}(d)$	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Tabelle II.3: Newcomb-Benford'sches Gesetz

Das Newcomb-Benford'sche Gesetz gilt für reale Datensätze, die genügend umfangreich sind und Zahlen enthalten, welche sich voneinander in der Größenordnung von mindestens 10000 unterscheiden. Die Daten müssen dementsprechend einigermaßen weit verteilt sein. Es besagt, dass die Auftretenswahrscheinlichkeit der Ziffernsequenzen in den Zahlen nicht gleichverteilt ist, sondern logarithmischen Gesetzen folgt. Der Grund für die erstaunlich weite Gültigkeit des Newcomb-Benford'schen Gesetzes liegt an dem Umstand, dass viele reale Datensätze log-normalverteilt sind, also nicht die Häufigkeiten der Daten selbst, sondern die Größenordnungen dieser Daten einer Normalverteilung folgen.

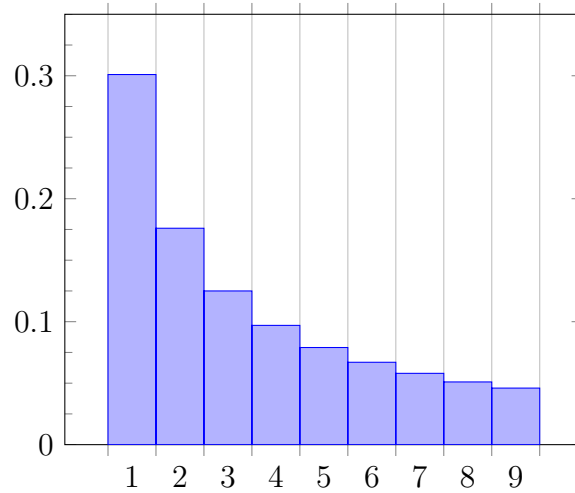


Abbildung II.4: Wahrscheinlichkeiten aus dem Newcomb-Benford'schen Gesetz

Das Benford'sche Gesetz gilt insbesondere für Zahlenmaterial, das natürlichen *Wachstumsprozessen* unterliegt. Das versuchen wir im Folgenden auszuarbeiten. Es sei dafür eine Folge von Potenzen  $\phi^n$ ,  $n = 1, 2, \dots$ , einer reellen Zahl  $\phi \in \mathbb{R}$  gegeben. Offenbar ist  $d$  die führende Ziffer von  $\phi^n$  genau dann wenn  $\phi^n \in [d \cdot 10^\ell, (d+1) \cdot 10^\ell)$  für mindestens ein  $\ell = 0, 1, \dots$  gilt. Da der Logarithmus monoton wachsend ist, ist das Letztere äquivalent zu  $n \log_{10} \phi \in [\ell + \log_{10} d, \ell + \log_{10}(d+1))$  für mindestens ein  $\ell = 0, 1, \dots$ . Mit anderen Worten liegt der gebrochene Anteil von  $n \log_{10} \phi$  im Intervall  $[\log_{10} d, \log_{10}(d+1))$ , oder mit Zeichen

$$\langle n \log_{10} \phi \rangle \in [\log_{10} d, \log_{10}(d+1)),$$

wobei  $\langle \cdot \rangle$  die Mantisse einer Zahl bezeichnet. Über die Verteilung solcher gebrochener Anteile  $\langle n\alpha \rangle$  für  $n = 1, 2, \dots$  mit einer irrationalen Zahl  $\alpha \in \mathbb{R}$  wurde seit über 100 Jahren ausgiebig geforscht. Das berühmte Resultat von Hermann Weyl von 1914 besagt, dass sie auf dem Intervall  $[0,1]$  gleichverteilt sind, d. h. für alle  $0 \leq a < b \leq 1$  gilt:

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : \langle n\alpha \rangle \in [a, b]\}}{N} = b - a,$$

wobei  $\#$  die Kardinalität einer endlichen Menge angibt. Das obige *Weyl-Kriterium* besagt, dass die Häufigkeit, mit welcher die Folgenglieder von gebrochenen Anteilen  $\langle n\alpha \rangle$  innerhalb des Intervalls  $[a, b]$  liegen, sich dessen Länge  $b - a$  nähert. Wendet man das Weyl-Kriterium auf die von uns studierte Folge mit  $\alpha = \log_{10} \phi$ ,  $a = \log_{10} d$  und  $b = \log_{10}(d+1)$ , so gilt:

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : \langle n \log_{10} \phi \rangle \in [\log_{10} d, \log_{10}(d+1)]\}}{N} = \log_{10}(d+1) - \log_{10} d.$$

Die linke Seite dieser Gleichheit entspricht im Grenzwert der Wahrscheinlichkeit, dass  $d$  die führende Ziffer von  $\phi^n$ ,  $n = 1, 2, \dots$  ist. Somit haben wir für solche Wachstumsprozesse das Newcomb-Benford'sche Gesetz nachgewiesen.



Lasst uns als Beispiel die Folge der Fibonacci-Zahlen auf ihre Gültigkeit für das Newcomb-Benford'sche Gesetz untersuchen. Bei der *Fibonacci-Folge* wird für die beiden ersten Zahlen der Wert 1 vorgegeben. Jede weitere Zahl ist die Summe ihrer beiden Vorgänger in der Folge. Die rekursive Formel dafür lautet also

$$f_n = f_{n-2} + f_{n-1}, \quad f_1 = f_2 = 1.$$

Die ersten 30 Fibonacci-Zahlen sind exemplarisch

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040.

Wir zählen in der Tabelle II.4, wie viele Male  $f_1, \dots, f_{30}$  die führende Ziffer  $1, \dots, 9$  haben und wie oft. Bei den Fibonacci-Zahlen ergeben also schon die Anfangsziffern der ersten 30 Zahlen eine Verteilung, die verblüffend gut das Newcomb-Benford'sche Gesetz wiederholt.

Erste Ziffer	1	2	3	4	5	6	7	8	9
Anzahl von 30	9	5	4	2	3	2	1	3	1
Häufigkeit	0.300	0.166	0.133	0.066	0.100	0.066	0.033	0.100	0.033

Tabelle II.4: Fibonacci-Zahlen und Newcomb-Benford'sches Gesetz

Wie lässt sich das denn erklären? Dies folgt daraus, dass die Fibonacci-Folge im Wesentlichen einen Wachstumsprozess darstellt. Das ist aus der *Formel von Moivre-Binet* ersichtlich:

$$f_n = \frac{\Phi^n - \Psi^n}{\sqrt{5}},$$

wobei  $\Phi = \frac{1+\sqrt{5}}{2}$  und  $\Psi = \frac{1-\sqrt{5}}{2}$  die beiden Lösungen der quadratischen Gleichung  $x^2 - x - 1 = 0$  sind. Zum Nachweis der Formel von Moivre-Binet ziehen wir die vollständige Induktion heran:

Induktionsanfang:  $f_1 = 1, \frac{\Phi^1 - \Psi^1}{\sqrt{5}} = 1.$

Induktionsvoraussetzung: Es gelte  $f_k = \frac{\Phi^k - \Psi^k}{\sqrt{5}}$  für alle  $k \leq n - 1.$

Induktionsschluss: Wir zeigen  $f_n = \frac{\Phi^n - \Psi^n}{\sqrt{5}}$ . Es gilt in der Tat:

$$\begin{aligned}
 f_{n-2} + f_{n-1} &\stackrel{\text{Induktions-}}{=} \frac{\Phi^{n-2} - \Psi^{n-2}}{\sqrt{5}} + \frac{\Phi^{n-1} - \Psi^{n-1}}{\sqrt{5}} = \frac{\Phi^{n-2}(1 + \Phi) - \Psi^{n-2}(1 + \Psi)}{\sqrt{5}} \\
 &\stackrel{\text{quadratische}}{=} \frac{\Phi^{n-2} \cdot \Phi^2 - \Psi^{n-2} \cdot \Psi^2}{\sqrt{5}} = \frac{\Phi^n - \Psi^n}{\sqrt{5}} = f_n.
 \end{aligned}$$

Aus  $|\Psi| < 1$  folgern wir, dass  $\Psi^n \rightarrow 0$  für  $n \rightarrow \infty$ . Deswegen dominiert in der Fibonacci-Folge  $f_n$  der Term  $\Phi^n$  für große  $n$ . Da dafür das Newcomb-Benford'sche Gesetz erfüllt ist, gilt es auch für die Fibonacci-Folge.

## 5. Anomalien in den Zahlen

---

Wir schließen mit der Bemerkung, dass das Newcomb-Benford'sche Gesetz Anwendung bei der Aufdeckung von Betrug bei der Bilanzerstellung, der Fälschung in Abrechnungen und generell zum raschen Auffinden eklatanter Unregelmäßigkeiten im Rechnungswesen findet. Dieses Gesetz wird dazu verwendet, möglichen Betrug in sozioökonomischen Daten aufzudecken, die zur Unterstützung öffentlicher Planungsentscheidungen vorgelegt werden. Ausgehend von der plausiblen Annahme, dass Personen, die Zahlen fälschen, dazu neigen, ihre Ziffern ziemlich gleichmäßig zu verteilen, sollte ein einfacher Vergleich der Häufigkeitsverteilung der ersten Ziffer der Daten mit der erwarteten Verteilung nach dem Newcomb-Benford'schen Gesetz alle anomalen Ergebnisse aufzeigen. So kann es auch bei der Aufdeckung von Datenfälschung in der Wissenschaft helfen. In den Vereinigten Staaten werden Beweise, die auf dem Newcomb-Benford'schen Gesetz beruhen, in Strafverfahren auf Bundes-, Landes- und Kommunalebene zugelassen.

## 6 Ausbreitung von Innovationen

Historisch betrachtet gibt es Zeiten, zu denen technologische Erfindungen schlagartig aufkommen. Eine Erfindung ist aber bei Weitem noch keine Innovation. Sie umfasst lediglich neue Konzeptentwicklungen in der Vormarktphase. Von *Innovation* im ökonomischen Sinne kann nur dann die Rede sein, wenn ihre Nützlichkeit erkannt und ein darauf basierendes Produkt eingeführt wird. Dabei kommt es oft vor, dass der Nutzen einer Innovation erst nach einer bestimmten Verzögerungszeit entdeckt wird. Viele hergestellte Objekte gelten im Moment ihrer Schaffung noch für zu vage, um gleich angenommen zu werden. Das Hergestellte wird erst in einem Anwendungs- und Kommunikationsprozess sinnvoll, welcher unter dem Begriff „Ausbreitung von Innovationen“ zusammengefasst wird. Für das Marketing innovativer Produkte wird also die folgende Frage relevant:

Wie verläuft die Ausbreitung von Innovationen?

Zur Modellierung der *Ausbreitung* eines innovativen Produktes bezeichnet  $y(t)$  die kumulierte Anzahl dessen Nutzer zum Zeitpunkt  $t$ . Die Gesamtanzahl potenzieller Nutzer, die sich evtl. für das Produkt interessieren könnten, sei  $m$ . Wir nehmen an, dass die momentane Änderungsrate  $y'(t)$  sich proportional zur Anzahl solcher Nutzer  $m - y(t)$  verhält, die das Produkt zum Zeitpunkt  $t$  noch nicht gekauft haben, d. h.

$$y'(t) \sim m - y(t).$$

Mit anderen Worten: Je mehr Nutzer von dem Produkt schon überzeugt sind, desto langsamer wächst die Nutzeranzahl. Es wird im Allgemeinen schwieriger, neue Kunden zu akquirieren, wenn der Markt schon gesättigt ist. Es sei ferner  $g(t)$  der *Ausbreitungskoeffizient* für den obigen Proportionalitätszusammenhang, d. h.

$$y'(t) = g(t) \cdot (m - y(t)).$$

Der Ausbreitungskoeffizient  $g(t)$  erfasst denjenigen Anteil potenzieller Nutzer  $m - y(t)$ , welche sich pro Zeiteinheit für das Produkt entscheiden. Je nach Form des Ausbreitungskoeffizienten unterscheidet man verschiedene Kommunikationsarten für die Ausbreitung von Innovationen:

- Die *hierarchische Kommunikation* erreicht potenzielle Nutzer durch mediale Werbekampagnen, die zentral gesteuert werden. Unternehmen versuchen dadurch, über ihr innovatives Produkt aufzuklären und seine Vorteile aufzuzeigen. Dies entspricht dem konstanten Ausbreitungskoeffizienten

$$g(t) = p,$$

wobei  $p > 0$  die Rate der hierarchischen Kommunikation ist. Die Ausbreitungsgleichung der hierarchischen Kommunikation lautet:

$$y'(t) = p \cdot (m - y(t)).$$

- Die *sphärische Kommunikation* erreicht potenzielle Nutzer durch Kunden-Werben-Kunden-Kampagnen, die zwar zentral organisiert werden, aber dezentral verlaufen. Unternehmen

## 6. Ausbreitung von Innovationen

---

wollen dadurch bereits gewonnene Nutzer motivieren, sich mit potenziellen Nutzern über das innovative Produkt auszutauschen und positive Erfahrungen zu kommunizieren. Dies entspricht dem linearen Ausbreitungskoeffizienten

$$g(t) = q \cdot \frac{y(t)}{m},$$

wobei  $q > 0$  die Rate der sphärischen Kommunikation ist. Die Ausbreitungsgleichung der sphärischen Kommunikation lautet:

$$y'(t) = q \cdot \frac{y(t)}{m} \cdot (m - y(t)).$$

In beiden Ansätze bezeichnet  $a$  die Anzahl der zum aktuellen Zeitpunkt akquirierten Nutzer. Der Anfangswert für die gesuchte Ausbreitung von Innovationen  $y(t)$  ist also

$$y(0) = a.$$

Wir leiten die Formel für die Ausbreitung von Innovationen über die hierarchische Kommunikation her. Dafür lösen wir die Differentialgleichung mit dem konstanten Ausbreitungskoeffizienten:

$$y' = p \cdot (m - y), \quad y(0) = a.$$

Wir schreiben sie mit Hilfe der Differenzialen auf:

$$\frac{dy}{dt} = p \cdot (m - y).$$

Anschließend wird die Separation der Variablen durchgeführt:

$$\frac{dy}{m - y} = p \cdot dt.$$

Wir integrieren diese Gleichung auf:

$$\int \frac{dy}{m - y} = \int p \, dt.$$

Es folgt mit einer noch zu bestimmenden Integrationskonstante  $C$ :

$$-\ln(m - y) = p \cdot t + C.$$

Wir lösen die obige Gleichheit nach  $y$  auf:

$$y(t) = m - e^{-C - p \cdot t}.$$

Der Anfangswert wird eingesetzt:

$$a = y(0) = m - e^{-C - p \cdot 0} = m - e^{-C}.$$

Daraus kann die Integrationskonstante ermittelt werden:

$$e^{-C} = m - a.$$

Die Lösung des Anfangswertproblems ist die sogenannte *exponentielle Sättigungsfunktion*:

$$y_{exp}(t) = m - (m - a) \cdot e^{-p \cdot t}.$$

Als Nächstes leiten wir die Formel für die Ausbreitung von Innovationen über die sphärische Kommunikation her. Dafür lösen wir die Differentialgleichung mit dem variablen Ausbreitungskoeffizienten:

$$y' = q \cdot \frac{y}{m} \cdot (m - y), \quad y(0) = a.$$

Wir wenden die Separation der Variablen an:

$$\int \frac{dy}{y \cdot (m - y)} = \int \frac{q}{m} dt.$$

Gesondert berechnen wir das Integral auf der linken Seite:

$$\int \frac{dy}{y \cdot (m - y)} = \frac{1}{m} \cdot \int \left( \frac{1}{y} + \frac{1}{m - y} \right) dy = \frac{1}{m} \cdot (\ln y - \ln(m - y)) = \frac{1}{m} \cdot \ln \frac{y}{m - y}.$$

Insgesamt gilt:

$$\frac{1}{m} \cdot \ln \frac{y(t)}{m - y(t)} = \frac{q}{m} \cdot t + C.$$

Nun setzen wir den Anfangswert  $y(0) = a$  hier ein:

$$\frac{1}{m} \cdot \ln \frac{a}{m - a} = C.$$

Hiermit folgt:

$$\frac{1}{m} \cdot \ln \frac{y(t)}{m - y(t)} = \frac{q}{m} \cdot t + \frac{1}{m} \cdot \ln \frac{a}{m - a}.$$

Wir lösen nach  $y$  auf, um die sogenannte *logistische Sättigungsfunktion* zu bekommen:

$$y_{log}(t) = \frac{m}{1 + \frac{m-a}{a} \cdot e^{-q \cdot t}}.$$

Die logistische Funktion dient als Modell der in der Soziologie üblichen Beschreibung eines Adoptionsprozesses, wie im klassischen Werk von Rogers dargelegt. Everett Rogers klassifiziert Kunden in fünf Typen, die sich durch den Übernahmzeitpunkt einer Innovation unterscheiden: Innovatoren I, frühe Übernehmer II, frühe Mehrheit III, späte Mehrheit IV, Nachzügler V – siehe Abbildung II.5.

Ein Zahlenbeispiel hierzu sollte nicht fehlen. Es sei die Anzahl der potentiellen Nutzer  $m = 100$  und die Anzahl der Nutzer zum aktuellen Zeitpunkt  $a = 10$ . Die Rate der hierarchischen

## 6. Ausbreitung von Innovationen

---

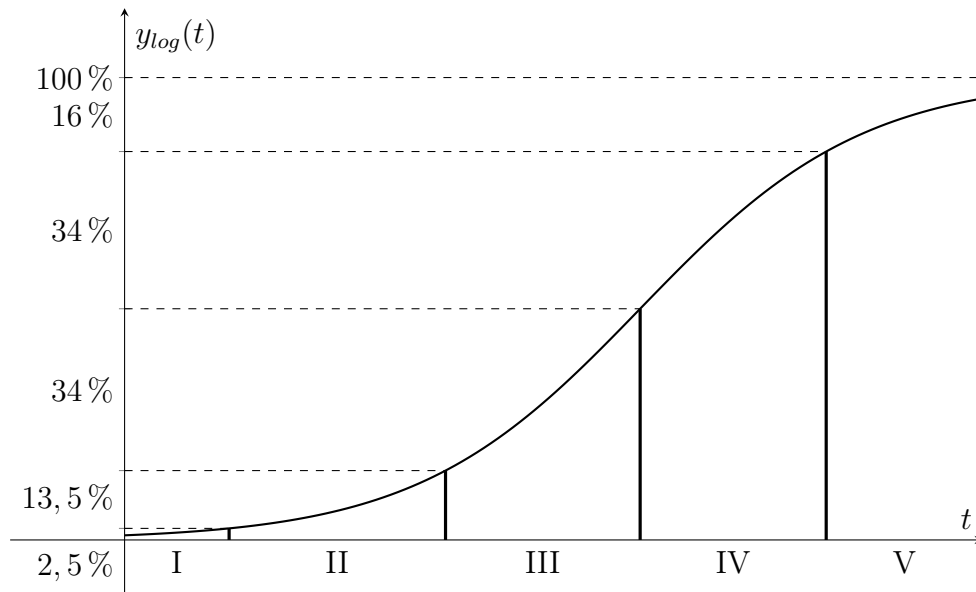


Abbildung II.5: Logistische Ausbreitung von Innovationen

Kommunikation sei  $p = 0,2$  und der sphärischen Kommunikation  $q = 0,9$ . Dies entspricht folgenden Ausbreitungstypen der Innovationen:

$$y_{exp}(t) = 100 - 90e^{-0,2 \cdot t}, \quad y_{log}(t) = \frac{100}{1 + 9e^{-0,9 \cdot t}}.$$

Wir vergleichen mittels welcher Kommunikation die Hälfte der potenziellen Nutzer schneller zu akquirieren ist. Aus beiden Ansätzen

$$50 = y_{exp}(t) = 100 - 90e^{-0,2 \cdot t}, \quad 50 = y_{log}(t) = \frac{100}{1 + 9e^{-0,9 \cdot t}}$$

folgt:

$$t_{exp} = -\frac{1}{0,2} \cdot \ln \frac{5}{9} \approx 2,94, \quad t_{log} = -\frac{1}{0,9} \cdot \ln \frac{1}{9} \approx 2,44.$$

Mittels sphärischer Kommunikation kann die Hälfte der potenziellen Nutzer um 0,5 Zeiteinheiten schneller als mittels hierarchischer Kommunikation überzeugt werden.

## 7 Risikoaversion eines Entscheiders

Die Begriffe Risikofreude, Risikoaversion und die dazwischen liegende Risikoneutralität treffen eine Aussage über die Risikobereitschaft eines Entscheiders. Diese Risikobereitschaft ist in vielen Fachgebieten von ausschlaggebender Bedeutung. So müssen Kreditinstitute bei der Anlageberatung vor der Abgabe einer Empfehlung die Risikobereitschaft des Anlegers erfragen, es sei denn, diese ist ihr aus einer langjährigen Geschäftsbeziehung oder dem bisherigen Anlageverhalten des Anlegers bereits bekannt. Auch bei allen übrigen Entscheidungsträgern wie Unternehmern, Managern, Führungskräften oder Privathaushalten spielt die Risikoeinstellung bei Entscheidungen die bedeutendste Rolle, denn die Entscheidung unter Ungewissheit, Entscheidung unter Unsicherheit und Entscheidung unter Risiko erfordern eine Risikoeinstellung des Entscheidungsträgers. Damit hängt also die folgende Frage aus der Entscheidungstheorie zusammen:

Wie quantifiziert man die Risikoaversion eines Entscheiders?

Ein Entscheider mit der Nutzenfunktion  $u : \mathbb{R} \rightarrow \mathbb{R}$  heißt *risikoavers*, falls der Nutzen vom Gewinn kleiner oder gleich dem Nutzen vom Verlust ist, d. h.

$$\forall x \in \mathbb{R}, \varepsilon > 0, \quad u(x + \varepsilon) - u(x) \leq u(x) - u(x - \varepsilon).$$

Wir wollen zeigen, dass dann die Nutzenfunktion konkav sein soll. In der Tat setze oben rechts  $y = x - \varepsilon$  und dividiere beide Seiten durch  $\varepsilon$ :

$$\forall x > y, \varepsilon > 0, \quad \frac{u(x + \varepsilon) - u(x)}{\varepsilon} \leq \frac{u(y + \varepsilon) - u(y)}{\varepsilon}.$$

Anschließend lassen wir  $\varepsilon$  gegen Null laufen:

$$u'(x) = \lim_{\varepsilon \rightarrow 0} \frac{u(x + \varepsilon) - u(x)}{\varepsilon} \leq \lim_{\varepsilon \rightarrow 0} \frac{u(y + \varepsilon) - u(y)}{\varepsilon} = u'(y).$$

Die erste Ableitung  $u'$  ist also monoton fallend, weswegen für die zweite Ableitung  $u'' \leq 0$  gilt. Wir wissen aber, dass eine zweimal stetig differenzierbare Funktion genau dann konkav ist, wenn ihre zweite Ableitung nichtpositiv ist. Geometrisch besagt die Konkavität einer Funktion, dass ihr Graph *nach oben* gekrümmt ist, siehe Abbildung II.6.

Um die Risikoaversion eines Entscheiders zu quantifizieren, bedient man sich in der Entscheidungstheorie des sogenannten *Arrow-Pratt-Maßes*:

$$r_A(x, u) = -\frac{u''(x)}{u'(x)}.$$

Das Arrow-Pratt-Maß gibt das Verhältnis zwischen der Krümmung der Nutzenfunktion und deren Wachstum an. Da die Nutzenfunktion monoton steigend ist, d. h.  $u'(x) > 0$ , und konkav, d. h.  $u''(x) \leq 0$ , gilt immer  $r_A(x, u) \geq 0$ . Je höher das Arrow-Pratt-Maß ist, desto risikoaverser der Entscheider.

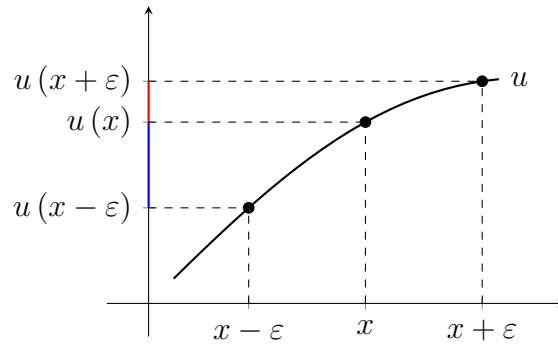


Abbildung II.6: Nutzenfunktion

Wir berechnen das Arrow-Pratt-Maß der Logarithmusfunktion:

$$r_A(x, \ln) = -\frac{\ln''(x)}{\ln'(x)} = -\left(-\frac{1}{x^2}\right) : \frac{1}{x} = \frac{1}{x}.$$

Aufgrund dieser Rechnung wird die logarithmische Nutzenfunktion auch *HARA-Funktion* (hyperbolic absolute risk aversion) genannt. Sie modelliert ein bekanntes Phänomen, dass die Risikoaversion mit zunehmendem Reichtum sinkt:

$$\forall x > y, \quad r_A(x, \ln) = \frac{1}{x} < \frac{1}{y} = r_A(y, \ln).$$

Mit anderen Worten nehmen Reichere viel lieber mehr Risiko auf sich als das Ärmere tun. Dass die Logarithmusfunktion diese Interpretation zulässt, macht sie zu einem oft verwendeten Modellierungswerkzeug in der Versicherungstheorie.

Die Formel des Arrow-Pratt-Maßes kann mit Hilfe des *Wahrscheinlichkeitsprämiums* begründet werden. Die Risikoaversion impliziert zunächst, dass der Nutzen  $x$  sicher zu bekommen höher liegt als der erwartete Nutzen von einer Lotterie zu 50% auf Gewinn  $x + \varepsilon$  bzw. Verlust  $x - \varepsilon$ :

$$u(x) \geq \frac{1}{2}u(x + \varepsilon) + \frac{1}{2}u(x - \varepsilon).$$

Das Wahrscheinlichkeitsprämium gibt an, um wie viel die Gewinnchancen in dieser Lotterie steigen müssen, damit der Entscheider den sicheren Gewinn als äquivalent dazu einzuschätzen beginnt. Die definierende Gleichung für das Wahrscheinlichkeitsprämium  $\pi(\varepsilon)$  ist somit

$$u(x) = \left(\frac{1}{2} + \pi(\varepsilon)\right) u(x + \varepsilon) + \left(\frac{1}{2} - \pi(\varepsilon)\right) u(x - \varepsilon).$$

Eine einfache, aber mühsame Rechnung zeigt, dass, wenn man diese Gleichung rechts wie links zweimal nach  $\varepsilon$  ableitet und  $\varepsilon = 0$  einsetzt, so gilt nach einer Umstellung der Terme:

$$4\pi'(0) = -\frac{u''(x)}{u'(x)}.$$



Im Wesentlichen ist also das Arrow-Pratt-Maß die Änderungsrate des Wahrscheinlichkeitsprämiums bzgl. des möglichen Gewinns bzw. Verlusts  $\varepsilon$ , wenn dieser infinitesimal wird. Wenn das Wahrscheinlichkeitsprämium mit einer hohen Sensitivität auf die Steigung des Gewinns bzw. Verlusts reagiert, d. h.  $\pi'(0)$  ist relativ groß, so zeugt auch das entsprechende Arrow-Pratt-Maß  $r_A(x, u)$  von einer erheblichen Risikoaversion.