



Inferenzstatistik

BSc Psychologie SoSe 2021

Prof. Dr. Dirk Ostwald

(0) Mathematische Grundlagen

Als bekannt vorausgesetzte Symbole

$A := B$	A ist definiert als B	$A = B$	A ist gleich B
$A \Leftrightarrow B$	A ist äquivalent mit B	$A \Rightarrow B$	Aus A folgt B
$a < b$	a ist kleiner als b	$a \leq b$	a ist kleiner oder gleich b
$a > b$	a ist größer als b	$a \geq b$	a ist größer oder gleich b
$n!$	Fakultät von n	$ a $	Betrag von a

Definition. Eine Definition (oder auch ein *Axiom*) ist ein Grundsatz eines logischen Systems, der innerhalb dieses Systems weder begründet noch deduktiv abgeleitet wird. Definitionen können nur nach ihrer Nützlichkeit innerhalb eines logischen Systems bewertet werden. Eine Definition lernt man am besten erst einmal auswendig und hinterfragt sie erst dann, wenn man ihren Nutzen in der Anwendung verstanden hat oder von diesem nicht überzeugt ist. Etwas Entspannung beim Umgang mit Definitionen ist generell hilfreich.

Theorem. Ein Theorem (oder auch *Satz*) ist eine Aussage, die mittels eines Beweises als richtig erkannt, dass heißt, aus Definitionen und/oder bereits bekannten Sätzen hergeleitet werden kann. Theoreme sind die “empirischen Ergebnisse” der Mathematik. Theoreme sind in der angewandten Mathematik oft für Berechnungen hilfreich, es lohnt sich also, sie auswendig zu lernen.

Beweis. Ein Beweis ist eine logische Argumentationskette, die auf bekannte Definitionen und Theoreme zurückgreift, um die Richtigkeit eines Theorems zu belegen. Kurze Beweise tragen oft zum Verständnis eines Theorems bei, lange Beweise eher nicht. **Beweise sind die Antwort auf die Frage warum eine mathematische Aussage gilt (“Warum ist das so?”).** Beweise lernt man nicht auswendig. Wenn sie lang sind, ist es sinnvoller sie beim ersten Lesen zu übergehen, um sich nicht in Details zu verlieren.

Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- Funktionen
- Differentialrechnung
- Analytische Optimierung
- Integralrechnung
- Selbstkontrollfragen

Literaturhinweise

Die Lektüre von Bärwolff (2017) wird empfohlen. Speziell werden zum Thema Mengenlehre die Abschnitte 1.2 und 1.5, zum Thema Funktionen die Abschnitte 1.3, 2.1 und 2.3, zum Thema Differentialrechnung Abschnitt 2.6, zum Thema Analytische Optimierung Abschnitt 2.10, und zum Thema Integralrechnung Abschnitt 2.13 (ohne 2.13.3) empfohlen.

Mathematische Grundlagen

- **Mengen**
- Summen, Produkte, Potenzen
- Funktionen
- Differentialrechnung
- Analytische Optimierung
- Integralrechnung
- Selbstkontrollfragen

Definition (Mengen und Mengendefinition)

Nach Cantor (1895) ist eine *Menge* definiert als “eine Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unsere Anschauung oder unseres Denken (welche die Elemente der Menge genannt werden) zu einem Ganzen”. Wir schreiben

$$m \in M \text{ bzw. } m \notin M \quad (1)$$

um auszudrücken, dass m ein Element bzw. kein Element von M ist. Zur Definition von Mengen gibt es mindestens folgende Möglichkeiten:

- (1) Auflisten der Elemente in geschweiften Klammern, z.B. $M := \{1, 2, 3\}$.
- (2) Angabe der Eigenschaften der Elemente, z.B. $M := \{x \in \mathbb{N} | x < 4\}$.
- (3) Gleichsetzen mit einer anderen eindeutig definieren Menge, z.B. $M := \mathbb{N}_3$.

Bemerkungen

- (2) wird als “ $x \in \mathbb{N}$, für die gilt, dass $x < 4$ ist” gelesen.
- Die Bedeutung von $\mathbb{N}$ und $\mathbb{N}_3$ wird im Folgenden erläutert.
- Mengen sind *ungeordnet*, d.h. $\{1, 2, 3\} = \{1, 3, 2\} = \{2, 3, 1\}$ etc.

Definition (Teilmengen und Mengengleichheit)

- Eine Menge A heißt *Teilmenge* einer Menge B , wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$. Ist A eine Teilmenge von B , so schreibt man

$$A \subseteq B \quad (2)$$

und nennt A *Untermenge* von B und B *Obermenge* von A .

- Eine Menge A heißt *echte Teilmenge* einer Menge B , wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$, es aber zumindest ein Element $b \in B$ gibt, für das gilt $b \notin A$. Ist A eine echte Teilmenge von B , so schreibt man

$$A \subset B. \quad (3)$$

- Zwei Mengen A und B heißen *gleich*, wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$, und wenn für jedes Element $b \in B$ gilt, dass auch $b \in A$. Sind die Mengen A und B gleich, so schreibt man

$$A = B. \quad (4)$$

Bemerkung

- Seien $A := \{1\}$, $B := \{1, 2\}$, $C := \{1, 2\}$. Dann gilt $A \subset B$, $B \subseteq C$, $C \subseteq B$ und $B = C$.

Definition (Kardinalität, leere Menge)

Die Anzahl der Elemente einer Menge M heißt *Kardinalität* und wird mit $|M|$ bezeichnet. Eine Menge mit Kardinalität 0 heißt *leere Menge* und wird mit $\emptyset$ bezeichnet.

Bemerkung

- Für $M := \{1, 2, 3\}$ gilt $|M| = 3$.

Definition (Potenzmenge)

Die Menge aller Teilmengen einer Menge M heißt *Potenzmenge von M* und wird mit $\mathcal{P}(M)$ bezeichnet. Die leere Untermenge von M und M selbst sind immer Elemente von $\mathcal{P}(M)$.

Bemerkungen

- Für $M := \{1, 2, 3\}$ gilt $\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.
- Allgemein gilt $|M| = n \Rightarrow |\mathcal{P}(M)| = 2^n$.

Definition (Mengenoperationen)

M und N seien zwei Mengen.

- Die *Vereinigung von M und N* ist definiert als die Menge

$$M \cup N := \{x | x \in M \text{ oder } x \in N\}, \quad (5)$$

wobei *oder* im inklusiven Sinne als *und/oder* zu verstehen ist.

- Der *Durchschnitt von M und N* ist definiert als die Menge

$$M \cap N := \{x | x \in M \text{ und } x \in N\}. \quad (6)$$

- Die *Differenz von M und N* ist definiert als die Menge

$$M \setminus N := \{x | x \in M \text{ und } x \notin N\}. \quad (7)$$

- Die *symmetrische Differenz von M und N* ist definiert als die Menge

$$M \Delta N := \{x | x \in M \text{ oder } x \in N, \text{ aber } x \notin M \cap N\}, \quad (8)$$

wobei *oder* hier also im exklusiven Sinne zu verstehen ist.

Bemerkung

- Für $M := \{1, 2, 3\}$ und $N := \{2, 3, 4, 5\}$ gilt

$$M \cup N = \{1, 2, 3, 4, 5\}, M \cap N = \{2, 3\}, M \setminus N = \{1\}, N \setminus M = \{4, 5\}, M \Delta N = \{1, 4, 5\}.$$

Definition (Zahlenmengen)

Es bezeichnen

- $\mathbb{N} := \{1, 2, 3, \dots\}$ die *natürlichen Zahlen*,
- $\mathbb{N}_n := \{1, 2, 3, \dots, n\}$ die *natürlichen Zahlen der Ordnung n* ,
- $\mathbb{N}^0 := \mathbb{N} \cup \{0\}$ die *natürlichen Zahlen* und Null,
- $\mathbb{Z} := \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ die *ganzen Zahlen*,
- $\mathbb{Q} := \{\frac{p}{q} | p, q \in \mathbb{Z}, q \neq 0\}$ die *rationalen Zahlen*,
- $\mathbb{R}$ die *reellen Zahlen*, und
- $\mathbb{C} := \{a + ib | a, b \in \mathbb{R}, i := \sqrt{-1}\}$ die *komplexen Zahlen*.

Bemerkungen

- $\mathbb{R}$ umfasst die rationalen Zahlen und die irrationalen Zahlen $\mathbb{R} \setminus \mathbb{Q}$ wie z.B. e , π und $\sqrt{2}$.
- Es gilt $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Definition (Intervalle)

Zusammenhängende Teilmengen der reellen Zahlen heißen *Intervalle*. Für $a, b \in \mathbb{R}$ unterscheidet man

- das *abgeschlossene Intervall*

$$[a, b] := \{x \in \mathbb{R} | a \leq x \leq b\}, \quad (9)$$

- das *offene Intervall*

$$]a, b[:= \{x \in \mathbb{R} | a < x < b\}, \quad (10)$$

- die *halboffenen Intervalle*

$$]a, b] := \{x \in \mathbb{R} | a < x \leq b\} \text{ und } [a, b[:= \{x \in \mathbb{R} | a \leq x < b\}. \quad (11)$$

Bemerkung

- Positiv Unendlich (∞) und negativ Unendlich ($-\infty$) sind keine Element von $\mathbb{R}$.
- Es gilt also immer $] - \infty, b]$ oder $] - \infty, b[$ bzw. $]a, \infty[$ oder $[a, \infty[$, sowie $\mathbb{R} =] - \infty, \infty[$.

Definition (Kartesische Produkte)

M und N seien zwei Mengen. Dann ist das *Kartesische Produkt der Mengen M und N* die Menge aller geordneten Tupel (m, n) mit $m \in M$ und $n \in N$, formal

$$M \times N := \{(m, n) | m \in M, n \in N\}. \quad (12)$$

Das Kartesische Produkt einer Menge M mit sich selbst wird bezeichnet mit

$$M^2 := M \times M. \quad (13)$$

Seien weiterhin $M_1, M_2, \dots, M_n$ Mengen. Dann ist das *Kartesische Produkt der Mengen $M_1, \dots, M_n$* die Menge aller geordneten n -Tupel $(m_1, \dots, m_n)$ mit $m_i \in M_i$ für $i = 1, \dots, n$, formal

$$\prod_{i=1}^n M_i := M_1 \times \dots \times M_n := \{(m_1, \dots, m_n) | m_i \in M_i \text{ für } i = 1, \dots, n\}. \quad (14)$$

Das n -fache Kartesische Produkt einer Menge M mit sich selbst wird bezeichnet mit

$$M^n := \prod_{i=1}^n M := \{(m_1, \dots, m_n) | m_i \in M\}. \quad (15)$$

Definition (Die Menge $\mathbb{R}^n$)

Das n -fache Kartesische Produkt der reellen Zahlen mit sich selbst wird bezeichnet mit

$$\mathbb{R}^n := \prod_{i=1}^n \mathbb{R} := \{x := (x_1, \dots, x_n) | x_i \in \mathbb{R}\} \quad (16)$$

und " $\mathbb{R}$ hoch n " gesprochen. Wir schreiben die Elemente von $\mathbb{R}^n$ typischerweise als Spalten

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (17)$$

und nennen sie *n-dimensionale Vektoren*. Zu Abgrenzung nennen wir die Elemente von $\mathbb{R}^1 = \mathbb{R}$ auch *Skalare*.

Bemerkung

- Ein Beispiel für $x \in \mathbb{R}^4$ ist $x = \begin{pmatrix} 0.16 \\ 1.76 \\ 0.23 \\ 7.10 \end{pmatrix}$.

Mathematische Grundlagen

- Mengen
- **Summen, Produkte, Potenzen**
- Funktionen
- Differentialrechnung
- Analytische Optimierung
- Integralrechnung
- Selbstkontrollfragen

Definition (Summenzeichen)

Summen werden oft mithilfe des *Summenzeichens*

$$\sum_{i=1}^n x_i = x_1 + x_2 + \cdots + x_n \quad (18)$$

dargestellt. Dabei stehen

- Σ für das griechische Sigma, mnemonisch für Summe,
- das Subskript $i = 1$ für den Laufindex der Summanden und den Startindex,
- das Superskript n für den Endindex,
- $x_1, x_2, \dots, x_n$ für die Summanden.

Bemerkung

- Nur mit Subskript und Superskripten macht das Summenzeichen Sinn.
- Die Bezeichnung des Laufindex ist irrelevant, es gilt $\sum_{i=1}^n x_i = \sum_{j=1}^n x_j$.
- In seltenen Fällen wird der Laufindex auch als Element einer Indexmenge angegeben. Ist z.B. die Indexmenge $I := \{1, 5, 7\}$ definiert, so ist $\sum_{i \in I} x_i := x_1 + x_5 + x_7$.

Theorem (Rechenregeln für Summen)

Summen gleicher Summanden

$$\sum_{i=1}^n x = nx \quad (19)$$

Assoziativität bei Summen gleicher Länge

$$\sum_{i=1}^n x_i + \sum_{i=1}^n y_i = \sum_{i=1}^n (x_i + y_i) \quad (20)$$

Distributivität bei Multiplikation mit einer Konstante

$$\sum_{i=1}^n ax_i = a \sum_{i=1}^n x_i \quad (21)$$

Aufspalten von Summen mit $1 < m < n$

$$\sum_{i=1}^n x_i = \sum_{i=1}^m x_i + \sum_{i=m+1}^n x_i \quad (22)$$

Umindizierung

$$\sum_{i=0}^n x_i = \sum_{j=m}^{n+m} x_{j-m} \quad (23)$$

Definition (Produktzeichen)

Produkte werden oft mithilfe des *Produktzeichens*

$$\prod_{i=1}^n x_i = x_1 \cdot x_2 \cdot \dots \cdot x_n \quad (24)$$

dargestellt. Dabei stehen

- $\prod$ für das griechische Π , mnemonisch für *Produkt*,
- das Subskript $i = 1$ für den Laufindex der Produkterme und den Startindex,
- das Superskript n für den Endindex,
- $x_1, x_2, \dots, x_n$ für die Produktterme.

Bemerkungen

- Nur mit Subskript und Superskripten macht das Produktzeichen Sinn.
- Die Bezeichnung des Laufindex ist irrelevant, es gilt $\prod_{i=1}^n x_i = \prod_{j=1}^n x_j$.
- In seltenen Fällen wird der Laufindex auch als Element einer Indexmenge angegeben.
Ist z.B. die Indexmenge $J := \mathbb{N}_2^0$ definiert, so ist $\sum_{j \in J} x_j := x_0 + x_1 + x_2$.

Definition (Potenz)

Für $a \in \mathbb{R}$ und $n \in \mathbb{N}^0$ ist die n -te *Potenz von a* definiert durch

$$a^0 := 1 \text{ und } a^{n+1} := a^n \cdot a. \quad (25)$$

Weiterhin ist für $a \in \mathbb{R} \setminus 0$ und $n \in \mathbb{N}^0$ die *negative n -te Potenz von a* definiert durch

$$a^{-n} := (a^n)^{-1} := \frac{1}{a^n}. \quad (26)$$

a wird dabei *Basis* und n wird *Exponent* genannt.

Theorem (Rechenregel für Potenzen)

Für $a, b \in \mathbb{R}$ und $n, m \in \mathbb{Z}$ mit $a \neq 0$ bei negativen Exponenten gelten

$$a^n a^m = a^{n+m} \quad (27)$$

$$(a^n)^m = a^{nm} \quad (28)$$

$$(ab)^n = a^n b^n \quad (29)$$

Definition (n -te Wurzel)

Für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ ist die n -te Wurzel von a definiert als die Zahl r , so dass

$$r^n = a. \quad (30)$$

Theorem (Potenzschreibweise der n -ten Wurzel)

Es sei $a \in \mathbb{R}$, $n \in \mathbb{N}$, und r die n -te Wurzel von a . Dann gilt

$$r = a^{\frac{1}{n}} \quad (31)$$

Beweis

Es gilt

$$\left(a^{\frac{1}{n}}\right)^n = a^{\frac{1}{n}} \cdot a^{\frac{1}{n}} \cdot \dots \cdot a^{\frac{1}{n}} = a^{\sum_{i=1}^n \frac{1}{n}} = a^1 = a. \quad (32)$$

Also gilt mit der Definition der n -ten Wurzel, dass $r = a^{\frac{1}{n}}$.

□

Bemerkung

- Das Rechnen mit Quadratwurzeln wird durch $\sqrt{x} = x^{\frac{1}{2}}$ sehr erleichtert.

Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- **Funktionen**
- Differentialrechnung
- Analytische Optimierung
- Integralrechnung
- Selbstkontrollfragen

Definition (Funktion)

Eine *Funktion* (oder *Abbildung*) f ist eine Zuordnungsvorschrift, die jedem Element einer Menge D genau ein Element einer Zielmenge Z zuordnet. D wird dabei *Definitions-
menge* von f und Z wird *Zielmenge* von f genannt. Wir schreiben

$$f : D \rightarrow Z, x \mapsto f(x), \quad (33)$$

wobei $f : D \rightarrow Z$ gelesen wird als “die Funktion f bildet alle Elemente der Menge D eindeutig auf Elemente in Z ab” und $x \mapsto f(x)$ gelesen wird als “ x , welches ein Element von D ist, wird durch die Funktion f auf $f(x)$ abgebildet, wobei $f(x)$ ein Element von Z ist”. Der Pfeil $\rightarrow$ steht für die Abbildung zwischen den Mengen D und Z , der Pfeil $\mapsto$ steht für die Abbildung zwischen einem Element von D und einem Element von Z .

Bemerkungen

- Es ist zentral, zwischen der *Funktion* f als Zuordnungsvorschrift und einem *Wert* der Funktion $f(x)$ als Element von Z zu unterscheiden.
- x ist der Funktionsinput (auch *Argument* der Funktion genannt), $f(x)$ der Funktionsoutput.
- Üblicherweise folgt $f(x)$ die Definition der *funktionalen Form* von f , z.B.

$$f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto f(x) := x^2. \quad (34)$$

Definition (Bildmenge, Urbildmenge)

Es sei $f : D \rightarrow Z, x \mapsto f(x)$ eine Funktion und es seien $D' \subseteq D$ und $Z' \subseteq Z$.

- Die Menge

$$f(D') := \{z \in Z \mid \text{Es gibt ein } x \in D' \text{ mit } z = f(x)\} \quad (35)$$

heißt die *Bildmenge* von D' und $f(D) \subseteq Z$ heißt der *Wertebereich* von f .

- Die Menge

$$f^{-1}(Z') := \{x \in D \mid f(x) \in Z'\} \quad (36)$$

heißt die *Urbildmenge* von Z' . $x \in D$ mit $z = f(x) \in Z$ heißt auch *Urbild* von z .

Bemerkung

- Wertebereich $f(D)$ und Zielmenge Z sind nicht notwendigerweise identisch.

Definition (Injektivität, Surjektivität, Bijektivität)

Es sei $f : D \rightarrow Z, x \mapsto f(x)$ eine Funktion.

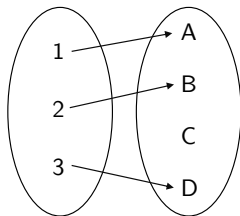
- Die Funktion f heißt *injektiv*, wenn es zu jedem Bild $z \in f(D)$ genau ein Urbild $x \in D$ gibt. Äquivalent gilt, dass f injektiv ist, wenn aus $x_1, x_2 \in D$ mit $x_1 \neq x_2$ folgt, dass $f(x_1) \neq f(x_2)$ ist.
- Die Funktion f heißt *surjektiv*, wenn $f(D) = Z$ gilt, wenn also jedes Element der Zielmenge Z ein Urbild in der Definitionsmenge D hat.
- Die Funktion f heißt *bijektiv*, wenn f injektiv und surjektiv ist.

Bemerkungen

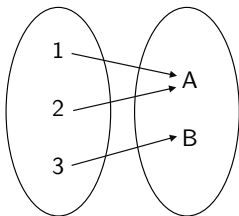
- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := x^2$ ist nicht injektiv, weil z.B. für $x_1 = 2 \neq -2 = x_2$ gilt, dass $f(x_1) = 2^2 = 4 = (-2)^2 = f(x_2)$. Weiterhin ist f auch nicht surjektiv, weil z.B. $-1 \in \mathbb{R}$ kein Urbild unter f hat.
- $f : [0, \infty[\rightarrow [0, \infty[, x \mapsto f(x) := x^2$ ist injektiv und surjektiv, also bijektiv.
- Bijektive Abbildungen heißen auch *eindeutige Funktionen* (engl. *one-to-one mappings*).

Beispiele

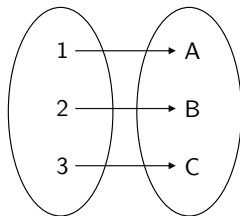
Nicht-surjektiv



Nicht-injektiv



Bijektiv



Definition (Verkettung von Funktionen)

Es seien $f : D \rightarrow R$ und $g : R \rightarrow S$ zwei Funktionen, wobei die Zielmenge von f mit der Definitionsmenge von g übereinstimmen. Dann ist durch

$$g \circ f : D \rightarrow S, x \mapsto (g \circ f)(x) := g(f(x)) \quad (37)$$

eine Funktion definiert, die die *Verkettung von f und g* genannt wird.

Bemerkungen

- $g \circ f$ bezeichnet die Funktion.
- $(g \circ f)(x)$ bezeichnet ein Element in S .
- Erst wird f auf x angewendet, dann wird g auf $f(x)$ angewendet.
- Für $f(x) := -x^2$ und $g(x) := \exp(x)$ ist $(g \circ f)(x) = \exp(-x^2)$.

Definition (Inverse Funktion)

Es sei $f : D \rightarrow R, x \mapsto f(x)$ eine bijektive Funktion. Dann heißt die Funktion f^{-1} mit

$$f^{-1} \circ f : D \rightarrow D, x \mapsto (f^{-1} \circ f)(x) := f^{-1}(f(x)) = x \quad (38)$$

inverse Funktion (oder Umkehrfunktion) von f .

Bemerkungen

- Weil f bijektiv ist, wird jedem $x \in D$ genau ein $z = f(x) \in Z$ zugeordnet.
- Jedem $z \in Z$ wird also auch genau ein $x \in D$ zugeordnet.
- Die inverse Funktion einer bijektiven Funktion ist also auch bijektiv.
- Die inverse Funktion von $f(x) := 2x =: y$ ist $f^{-1}(y) = y/2$.

Definition (Lineare Abbildung)

Eine Abbildung $f : D \rightarrow R, x \mapsto f(x)$ heißt *lineare Abbildung*, wenn für $x, y \in D$ und einen Skalar c gelten, dass

$$f(x + y) = f(x) + f(y) \text{ und } f(cx) = cf(x). \quad (39)$$

Eine Abbildung, für die obige Eigenschaften nicht gelten, heißt *nicht-lineare Abbildung*.

Bemerkungen

- Für $a \in \mathbb{R}$ ist die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := ax$ linear, weil

$$f(x + y) = a(x + y) = ax + ay = f(x) + f(y) \text{ und } f(cx) = acx = cax = cf(x). \quad (40)$$

- Für $a, b \in \mathbb{R}$ dagegen ist die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := ax + b$ nicht-linear, weil z.B. für $a := b := 1$ gilt, dass

$$f(x + y) = 1(x + y) + 1 = x + y + 1 \neq x + 1 + y + 1 = f(x) + f(y). \quad (41)$$

- Eine Abbildung der Form $f(x) := ax + b$ heißt *affin-lineare Abbildung*
- Abbildungen der Form $f(x) := ax + b$ werden auch als *lineare Funktionen* bezeichnet.

Definition (Funktionentypen)

In der statistischen Anwendung unterscheiden wir

- *univariate reellwertige Funktionen* der Form

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x), \quad (42)$$

- *multivariate reellwertige Funktionen* der Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto f(x) = f(x_1, \dots, x_n), \quad (43)$$

- *multivariate vektorwertige Funktionen* der Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}. \quad (44)$$

Bemerkung

- In der Physik werden multivariate reellwertige Funktionen auch *Skalarfelder* genannt.
- In der Physik werden multivariate vektorwertige Funktionen auch *Vektorfelder* genannt.
- In manchen Anwendungen kommen auch *matrixvariante matrixwertige Funktionen* zum Tragen.

Elementare univariate reellwertige Funktion der Statistik sind

- die Polynomfunktionen,
- die Exponentialfunktion,
- der Logarithmusfunktion,
- die Gammafunktion.

Wir skizzieren diese Funktionen im Folgenden.

Definition (Polynomfunktionen)

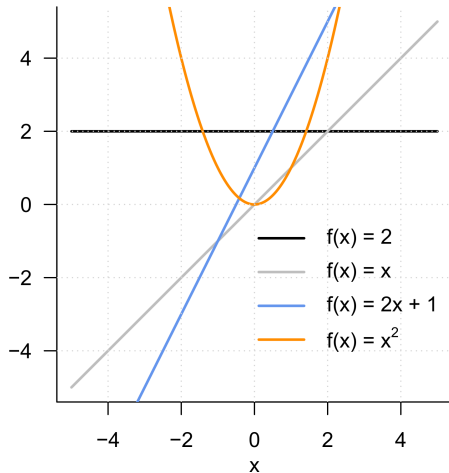
Eine Funktion der Form

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := \sum_{i=0}^k a_i x^i = a_0 + a_1 x^1 + a_2 x^2 + \cdots + a_k x^k \quad (45)$$

heißt *Polynomfunktion* k -ten Grades mit Koeffizienten $a_0, a_1, \dots, a_k \in \mathbb{R}$. Typische Polynomfunktionen sind in nachfolgender Tabelle aufgelistet

Name	Funktionale Form	Koeffizienten
Konstante Funktion	$f(x) = a$	$a_0 := a, a_i := 0, i > 0$
Identitätsfunktion	$f(x) = x$	$a_0 := 0, a_1 := 1, a_i := 0, i > 1$
Lineare Funktion	$f(x) = ax + b$	$a_0 := b, a_1 := a, a_i := 0, i > 1$
Quadratfunktion	$f(x) = x^2$	$a_0 := 0, a_1 := 0, a_2 := 1, a_i := 0, i > 2$

Graphen typischer Polynomfunktionen



Theorem (Exponentialfunktion und ihre Eigenschaften)

Die *Exponentialfunktion* ist definiert als

$$\exp : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \exp(x) := e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \quad (46)$$

Die Exponentialfunktion hat folgende Eigenschaften:

Wertebereich	$x \in]-\infty, 0[\Rightarrow \exp(x) \in]0, 1[$ $x \in]0, \infty[\Rightarrow \exp(x) \in]1, \infty[$
Monotonie	$x < y \Rightarrow \exp(x) < \exp(y)$
Spezielle Werte	$\exp(0) = 1$ und $\exp(1) = e$
Exponentialeigenschaften	$\exp(x + y) = \exp(x) \exp(y)$ $\exp(x - y) = \frac{\exp(x)}{\exp(y)}$ $\exp(x) \exp(-x) = 1$

Bemerkungen

- Für Beweise der Eigenschaften wird auf die einschlägige Literatur verwiesen.
- Die Exponentialeigenschaften sind beim Rechnen mit der Normalverteilung zentral.
- $e \approx 2.71$ heißt *Eulersche Zahl*.

Theorem (Logarithmusfunktion und ihre Eigenschaften)

Die *Logarithmusfunktion* ist definiert als inverse Funktion der Exponentialfunktion,

$$\ln :]0, \infty[\rightarrow \mathbb{R}, x \mapsto \ln(x) \text{ mit } \ln(\exp(x)) = x \text{ für alle } x \in \mathbb{R}. \quad (47)$$

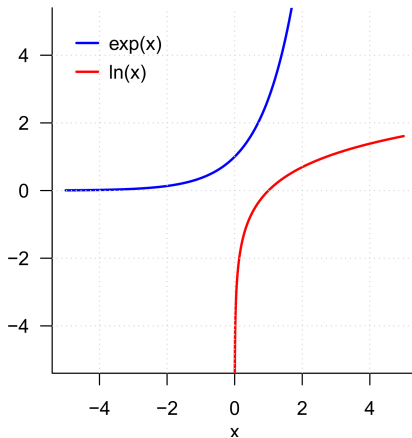
Die Logarithmusfunktion hat folgende Eigenschaften:

Wertebereich	$x \in]0, 1[\Rightarrow \ln(x) \in]-\infty, 0[$ $x \in]1, \infty[\Rightarrow \ln(x) \in]0, \infty[$
Monotonie	$x < y \Rightarrow \ln(x) < \ln(y)$
Spezielle Werte	$\ln(1) = 0$ und $\ln(e) = 1$.
Logarithmeneigenschaften	$\ln(xy) = \ln(x) + \ln(y)$ $\ln(x^c) = c \ln(x)$ $\ln\left(\frac{1}{x}\right) = -\ln(x)$

Bemerkungen

- Für Beweise der Eigenschaften wird auf die einschlägige Literatur verwiesen.
- Die Logarithmeneigenschaften sind beim Rechnen mit Log-Likelihood-Funktionen zentral.
- “Die Logarithmusfunktion wandelt Produkte in Summen und Potenzen in Produkte um.”

Graphen der Exponential- und Logarithmusfunktion



Theorem (Gammafunktion und ihre Eigenschaften)

Die *Gammafunktion* ist definiert durch

$$\Gamma : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \Gamma(x) := \int_0^{\infty} \xi^{x-1} \exp(-\xi) d\xi \quad (48)$$

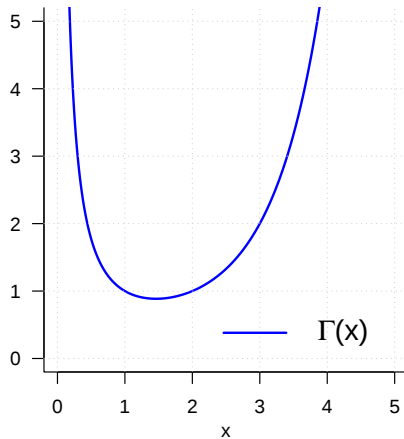
Die Gammafunktion hat folgende Eigenschaften:

Spezielle Werte	$\Gamma(1) = 1$ $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ $\Gamma(n) = (n-1)!$ für $n \in \mathbb{N}$.
Rekursionseigenschaft	Für $x > 0$ gilt $\Gamma(x+1) = x\Gamma(x)$

Bemerkungen

- Für Beweise der Eigenschaften wird auf die einschlägige Literatur verwiesen.
- Die Gammafunktion ist im Kontext von χ^2 -, t - und F -Verteilung zentral.

Graph der Gammafunktion auf $]0, 5[$



Zur Stetigkeit von Funktionen

- Auf eine tiefergehende Diskussion der Stetigkeit von Funktion wird verzichtet.
- Intuitiv bedeutet Stetigkeit, dass beliebig kleine Änderungen des Funktionswertes durch genügend kleine Änderungen des Funktionsargumentes erreicht werden können. Anderes ausgedrückt, heißt dies, dass die Funktion keine Sprünge hat, oder dass man “beim Zeichnen der Funktion den Bleistift nicht absetzen muss”.
- Eine Möglichkeit diese Intuition zu formalisieren, ist es zu fordern, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ die gegen ein x in der Definitionsmenge der Funktion konvergiert, die korrespondierende Folge von Funktionswerten $(f(x_n))_{n \in \mathbb{N}}$ in der Zielmenge der Funktion gegen den Funktionswert $f(x)$ konvergiert, und zwar für alle x der Definitionsmenge:

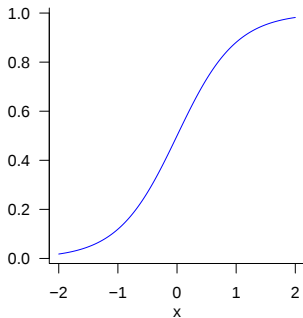
$f : D \rightarrow Z$ ist stetig $\Leftrightarrow$

Für alle $x \in D$ und $(x_n)_{n \rightarrow \infty} \subset D : \lim_{n \rightarrow \infty} x_n = x \Rightarrow \lim_{n \rightarrow \infty} (f(x_n)) = f(x)$.

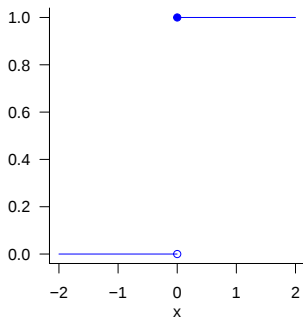
- Die Polynom-, Exponential-, Logarithmusfunktion sind stetige Funktionen.
- Nichtstetige Funktionen haben oft Fallunterscheidungen, z.B. $f(x) := \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$.

Zur Stetigkeit von Funktionen

Stetige Funktion



Nichtstetige Funktion



Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- Funktionen
- **Differentialrechnung**
- Analytische Optimierung
- Integralrechnung
- Selbstkontrollfragen

Definition (Ableitung)

Es sei $I \subseteq \mathbb{R}$ ein Intervall und

$$f : I \rightarrow \mathbb{R}, x \mapsto f(x) \quad (49)$$

eine univariate reellwertige Funktion. f heißt in $a \in I$ *differenzierbar*, wenn der Grenzwert

$$f'(a) := \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (50)$$

existiert. $f'(a)$ heißt dann die *Ableitung von f an der Stelle a* . Ist f differenzierbar für alle $x \in I$, so heißt f *differenzierbar* und die Funktion

$$f' : I \rightarrow \mathbb{R}, x \mapsto f'(x) \quad (51)$$

heißt *Ableitung von f*

Bemerkungen

- Für $h > 0$ heißt $\frac{f(a+h) - f(a)}{h}$ *Differenzquotient*.
- Der Differenzquotient misst die Änderung $f(a+h) - f(a)$ von f pro Strecke h .
- Für $h \rightarrow 0$ misst der Differenzquotient die Änderungsrate von f in a .
- $f'(a)$ ist eine Zahl, f' ist eine Funktion.
- Wir werden keine Grenzwertbildung zur Berechnung von Ableitungen benötigen.

Definition (Notation für Ableitungen univariater reellwertiger Funktionen)

Es sei f eine univariate reellwertige Funktion. Äquivalente Schreibweisen für die Ableitung von f und die Ableitung von f an einer Stelle x sind

- (1) Die *Lagrange-Notation* f' und $f'(x)$
- (2) Die *Newton-Notation* $\dot{f}$ und $\dot{f}(x)$.
- (3) Die *Leibniz-Notation* $\frac{df}{dx}$ und $\frac{df(x)}{dx}$.
- (4) Die *Euler-Notation* Df und $Df(x)$.

Bemerkungen

- Für univariate reellwertige Funktionen benutzen wir f' und $f'(x)$ als Bezeichner.
- In Berechnungen benutzen wir auch die “Operator-Schreibweise” $\frac{d}{dx}f(x)$.
- Wir verstehen $\frac{d}{dx}f(x)$ als den Auftrag, die Ableitung von f zu berechnen.

Definition (Höhere Ableitungen)

Es sei f eine univariate reellwertige Funktion und

$$f^{(1)} := f' \quad (52)$$

sei die Ableitung von f . Die k -te Ableitung von f ist rekursiv definiert durch

$$f^{(k)} := \left(f^{(k-1)}\right)' \text{ für } k \geq 0, \quad (53)$$

unter der Annahme, dass $f^{(k-1)}$ differenzierbar ist. Insbesondere ist die *zweite Ableitung* von f definiert durch die Ableitung von f' , also

$$f'' := (f')' \quad (54)$$

Bemerkung

- Wir schreiben auch $\frac{d^2}{dx^2} f(x)$ für den Auftrag, die zweite Ableitung von f zu bestimmen.
- Die nullte Ableitung $f^{(0)}$ von f ist f selbst.
- Üblicherweise schreibt man für $k < 4$ f' , f'' , f''' statt $f^{(1)}$, $f^{(2)}$, $f^{(3)}$.
- Im Allgemeinen benötigen wir nur f' und f'' .

Theorem (Rechenregeln für Ableitungen)

Für $i = 1, \dots, n$ seien g_i reellwertige univariate differenzierbare Funktionen. Dann gelten folgende Rechenregeln für Ableitungen

(1) Summenregel

$$\text{Für } f(x) := \sum_{i=1}^n g_i(x) \text{ gilt } f'(x) = \sum_{i=1}^n g'_i(x). \quad (55)$$

(2) Produktregel

$$\text{Für } f(x) := g_1(x)g_2(x) \text{ gilt } f'(x) = g'_1(x)g_2(x) + g_1(x)g'_2(x). \quad (56)$$

(3) Quotientenregel

$$\text{Für } f(x) := \frac{g_1(x)}{g_2(x)} \text{ gilt } f'(x) = \frac{g'_1(x)g_2(x) - g_1(x)g'_2(x)}{g_2^2(x)} \quad (57)$$

(4) Kettenregel

$$\text{Für } f(x) := g_1(g_2(x)) \text{ gilt } f'(x) = g'_1(g_2(x))g'_2(x) \quad (58)$$

Bemerkung

- Für Herleitung und Beweise der Rechenregeln wird auf die einschlägige Literatur verwiesen.

Theorem (Ableitungen elementarer Funktionen)

Für elementare Funktionen der Statistik ergeben sich folgende Ableitungen

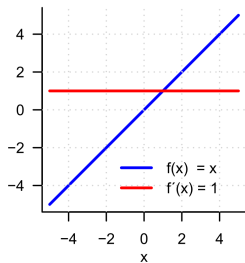
Name	Definition	Ableitung
Polynomfunktionen	$f(x) := \sum_{i=0}^n a_i x^i$	$f'(x) = \sum_{i=1}^n i a_i x^{i-1}$
Konstante Funktion	$f(x) := a$	$f'(x) = 0$
Identitätsfunktion	$f(x) := x$	$f'(x) = 1$
Lineare Funktion	$f(x) := ax + b$	$f'(x) = a$
Quadratfunktion	$f(x) := x^2$	$f'(x) = 2x$
Exponentialfunktion	$f(x) := \exp(x)$	$f'(x) = \exp(x)$
Logarithmusfunktion	$f(x) := \ln(x)$	$f'(x) = \frac{1}{x}$

Bemerkung

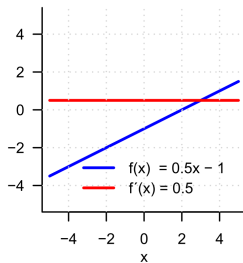
- Für Beweise wird auf die einschlägige Literatur verwiesen.

Beispiele

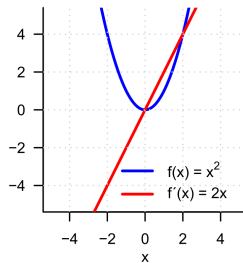
Identitätsfunktion



Lineare Funktion

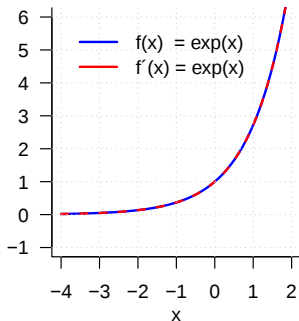


Quadratfunktion

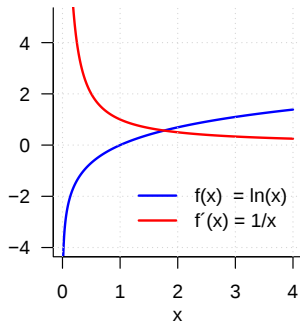


Beispiele

Exponentialfunktion



Logarithmusfunktion



Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- Funktionen
- Differentialrechnung
- **Analytische Optimierung**
- Integralrechnung
- Selbstkontrollfragen

Definition (Extremstellen und Extremwerte)

Es sei $U \subseteq \mathbb{R}$ und $f : U \rightarrow \mathbb{R}$ eine univariate reellwertige Funktion. f hat an der Stelle $x_0 \in U$

- ein *lokales Minimum*, wenn es ein Intervall $I :=]a, b[$ gibt mit $x_0 \in]a, b[$ und

$$f(x_0) \leq f(x) \text{ für alle } x \in I \cap U, \quad (59)$$

- ein *globales Minimum*, wenn gilt, dass

$$f(x_0) \leq f(x) \text{ für alle } x \in U, \quad (60)$$

- ein *lokales Maximum*, wenn es ein Intervall $I :=]a, b[$ gibt mit $x_0 \in]a, b[$ und

$$f(x_0) \geq f(x) \text{ für alle } x \in I \cap U, \quad (61)$$

- *lokales Maximum*, wenn gilt, dass

$$f(x_0) \geq f(x) \text{ für alle } x \in U. \quad (62)$$

Der Wert $x_0 \in U$ der Definitionsmenge von f heißt entsprechend *lokale* oder *globale Minimalstelle* oder *Maximalstelle*, der Funktionswert $f(x_0) \in \mathbb{R}$ heißt entsprechend *lokales* oder *globales Minimum* oder *Maximum*. Generell heißt der Wert $x_0 \in U$ *Extremstelle* und der Funktionswert $f(x_0) \in \mathbb{R}$ *Extremwert*.

Bemerkungen

- Extremstellen werden auch mit $\arg \min_{x \in I \cap U} f(x)$ oder $\arg \max_{x \in I \cap U} f(x)$ bezeichnet.
- Extremwerte werden auch mit $\min_{x \in I \cap U} f(x)$ oder $\max_{x \in I \cap U} f(x)$ bezeichnet.

Definition (Notwendige Bedingung für Extrema)

f sei eine univariate reellwertige Funktion. Dann gilt

$$x_0 \text{ ist Extremstelle von } f \Rightarrow f'(x_0) = 0. \quad (63)$$

Bemerkungen

- Wenn x_0 eine Extremstelle von f ist, dann ist die erste Ableitung von f in x_0 null.
- Sei zum Beispiel x_0 eine **lokale Maximalstelle** von f . Dann gilt
 - Links von x_0 steigt f an, rechts von x_0 fällt f ab.
 - In x_0 steigt f weder an, noch fällt f ab, also ist $f'(x_0) = 0$.

Definition (Hinreichende Bedingungen für lokale Extrema)

f sei eine zweimal differenzierbare univariate reellwertige Funktion.

- Wenn für $x_0 \in U \subseteq \mathbb{R}$

$$f'(x_0) = 0 \text{ und } f''(x_0) > 0 \quad (64)$$

gilt, dann hat f an der Stelle x_0 ein Minimum.

- Wenn für $x_0 \in U \subseteq \mathbb{R}$

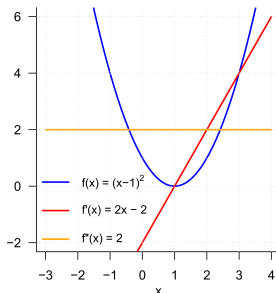
$$f'(x_0) = 0 \text{ und } f''(x_0) < 0 \quad (65)$$

gilt, dann hat f an der Stelle x_0 ein Maximum.

Bemerkung

- Eine Intuition vermittelt nachfolgende Abbildung.

Hinreichende Bedingungen für lokale Extrema



Hier ist offenbar $x_0 = 1$ eine lokale Minimalstelle von $f(x) = (x - 1)^2$. Man erkennt:

- Links von x_0 fällt f ab, rechts von x_0 steigt f an.
- In x_0 steigt f weder an, noch fällt f ab, also ist $f'(x_0) = 0$.
- Links und rechts von x_0 und in x_0 ist die Änderung f'' von f' positiv.
- Links von x_0 schwächt sich die Negativtaet von f' zu 0 ab.
- Rechts von x_0 verstärkt sich die Positivtaet von f' .

Definition (Standardverfahren der analytischen Optimierung)

f sei eine univariate reellwertige Funktion. Lokale Extremstellen von f können mit folgendem *Standardverfahren der analytischen Optimierung* identifiziert werden:

- (1) Berechnen der ersten und zweiten Ableitung von f .
- (2) Bestimmen von Nullstellen x^* von f' durch Auflösen von $f'(x^*) = 0$ nach x^* .
 $\Rightarrow$ Nullstellen von f' sind Kandidaten für Extremstellen von f .
- (3) Evaluation von $f''(x^*)$.
 $\Rightarrow$ Wenn $f''(x^*) > 0$, dann ist x^* lokale Minimumstelle von f .
 $\Rightarrow$ Wenn $f''(x^*) < 0$, dann ist x^* lokale Maximumstelle von f .
 $\Rightarrow$ Wenn $f''(x^*) = 0$, dann ist x^* keine Extremstelle von f .

Beispiel

- Wir betrachten die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := (x - 1)^2. \quad (66)$$

- Die erste Ableitung von f ergibt sich mit der Kettenregel zu

$$f'(x) = \frac{d}{dx} ((x - 1)^2) = 2(x - 1) \cdot \frac{d}{dx} (x - 1) = 2x - 2. \quad (67)$$

- Die zweite Ableitung von f ergibt sich zu

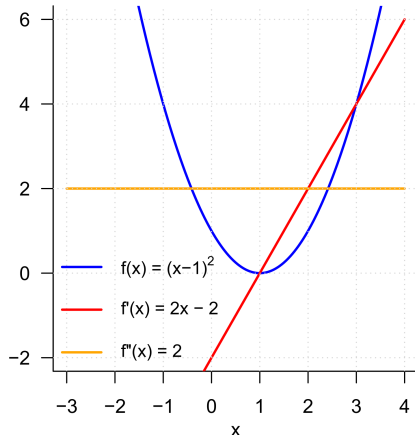
$$f''(x) = \frac{d}{dx} f'(x) = \frac{d}{dx} (2x - 2) = 2 > 0 \text{ für alle } x \in \mathbb{R}. \quad (68)$$

- Auflösen von $f'(x^*) = 0$ nach x^* ergibt

$$f'(x^*) = 0 \Leftrightarrow 2x^* - 2 = 0 \Leftrightarrow 2x^* = 2 \Leftrightarrow x^* = 1. \quad (69)$$

- $x^* = 1$ ist folglich eine Minimalstelle von f mit zugehörigen Minimalwert $f(1) = 0$.

Beispiel



Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- Funktionen
- Differentialrechnung
- Analytische Optimierung
- **Integralrechnung**
- Selbstkontrollfragen

Übersicht

- Integrale kommen in der W-Theorie und Statistik an vielen Stellen vor:
 - WDF Definition und Berechnen von Wahrscheinlichkeiten aus WDF
 - Beziehung von WDF und KVF
 - Erwartungswerte inklusive Varianz und Kovarianz
 - u.v.a.m.
- Ziel hier ist es, zentrale Begriffe aus der Schulmathematik zu wiederholen.
- In der Schulmathematik wird üblicherweise das *Riemannintegral* betrachtet.
- Zentrale Begriffe sind *Stammfunktion*, *unbestimmtes Integral*, *bestimmtes Integral*, *Hauptsätze der Differential- und Integralrechnung*, *uneigentliches Integral*, und *mehrdimensionales Integral*.
- Der Fokus liegt auf der Klärung von Begriffen und Symbolen, nicht dem Rechnen.
- Auf die Einführung des modernen Integralbegriffs nach Lebesgue wird verzichtet.

Definition (Stammfunktion, Unbestimmtes Integral)

Für ein Intervall $I \subset \mathbb{R}$ sei $f : I \rightarrow \mathbb{R}$ eine univariate reellwertige Funktion. Dann heißt eine differenzierbare Funktion $F : I \rightarrow \mathbb{R}$ mit der Eigenschaft

$$F' = f \quad (70)$$

Stammfunktion von f . Ist F eine Stammfunktion von f , dann heißt

$$\int f(x) dx := F(x) + c \text{ mit } c \in \mathbb{R} \quad (71)$$

unbestimmtes Integral der Funktion f .

Bemerkungen

- Die Ableitung der Stammfunktion F von f ist f .
- Das unbestimmte Integral ist die Gesamtheit aller Stammfunktionen von f .
- Die Konstante $c \in \mathbb{R}$ heißt *Integrationskonstante*, es gilt $\frac{d}{dx} c = 0$.
- Der Ausdruck $\int f(x) dx$ ist als $F(x) + c$ definiert
- In $\int f(x) dx$ haben $\int$ und dx keine eigentliche Bedeutung, $f(x)$ heißt *Integrand*.

Theorem (Stammfunktionen elementarer Funktionen)

Für elementare Funktionen der Statistik ergeben sich folgende Stammfunktionen

Name	Definition	Stammfunktion
Polynomfunktionen	$f(x) := \sum_{i=0}^n a_i x^i$	$F(x) = \sum_{i=0}^n \frac{a_i}{i+1} x^{i+1}$
Konstante Funktion	$f(x) := a$	$F(x) = ax$
Identitätsfunktion	$f(x) := x$	$F(x) = \frac{1}{2}x^2$
Lineare Funktion	$f(x) := ax + b$	$F(x) = \frac{1}{2}ax^2 + bx$
Quadratfunktion	$f(x) := x^2$	$F(x) = \frac{1}{3}x^3$
Exponentialfunktion	$f(x) := \exp(x)$	$F(x) = \exp(x)$
Logarithmusfunktion	$f(x) := \ln(x)$	$F(x) = x \ln x - x$

Bemerkung

- Beweise ergeben sich direkt durch Ableiten der Stammfunktionen.

Theorem (Rechenregeln für Stammfunktionen)

f und g seien univariate reellwertige Funktion, die Stammfunktionen besitzen, und g sei invertierbar. Dann gelten folgende Rechenregeln für die Bestimmung von Integralen

(1) Summenregel

$$\int a f(x) + b g(x) dx = a \int f(x) dx + b \int g(x) dx \text{ für } a, b \in \mathbb{R} \quad (72)$$

(2) Partielle Integration

$$\int f'(x) g(x) dx = f(x) g(x) - \int f(x) g'(x) dx \quad (73)$$

(3) Substitutionsregeln

$$\int f(g(x)) g'(x) dx = \int f(t) dt \text{ mit } t = g(x) \quad (74)$$

Bemerkung

- Für die Herleitung der Summenregel wird auf die einschlägige Literatur verwiesen.

Integralrechnung

Beweis

Die Rechenregel der partiellen Integration ergibt sich durch Integration der Produktregel der Differentiation. Wir erinnern uns, dass gilt

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x). \quad (75)$$

Integration beider Seiten der Gleichung und Berücksichtigung der Summenregel für Stammfunktionen ergibt dann

$$\begin{aligned} \int (f(x)g(x))' dx &= \int f'(x)g(x) + f(x)g'(x) dx \\ \Leftrightarrow f(x)g(x) &= \int f'(x)g(x) dx + \int f(x)g'(x) dx \\ \Leftrightarrow \int f'(x)g(x) dx &= f(x)g(x) - \int f(x)g'(x) dx \end{aligned} \quad (76)$$

Die Substitutionsregel ergibt sich für $F' = f$ durch Anwendung der Kettenregel der Differentiation auf die verkettete Funktion $F(g)$. Speziell gilt zunächst

$$(F(g(x)))' = F'(g(x))g'(x) = f(g(x))g'(x) \quad (77)$$

Integration beider Seiten der Gleichung

$$(F(g(x)))' = f(g(x))g'(x) \quad (78)$$

ergibt dann

$$\begin{aligned} \int (F(g(x)))' dx &= \int f(g(x))g'(x) dx \\ \Leftrightarrow F(g(x)) + c &= \int f(g(x))g'(x) dx \\ \Leftrightarrow \int f(g(x))g'(x) dx &= \int f(t) dt \text{ mit } t := g(x). \end{aligned} \quad (79)$$

Dabei ist die rechte Seite der letzten obigen Gleichung zu verstehen als $F(g(x)) + c$, also als Stammfunktion von f evaluiert an der Stelle $t := g(x)$. Das dt ist nicht durch $dg(x)$ zu ersetzen, sondern rein notationstechnischer Natur. $\square$

Bestimmtes Integral von $f : [a, b] \rightarrow \mathbb{R}$

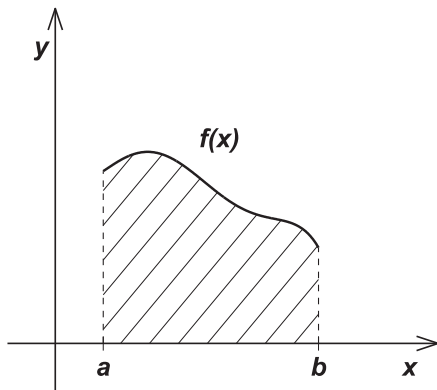


Abbildung 2.54 aus Bärwolff (2017)

Definition (Zerlegung eines Intervalls, Feinheit)

Es sei $[a, b] \subset \mathbb{R}$ ein Intervall und $x_0, x_1, x_2, \dots, x_n \in [a, b]$ eine Menge von Punkten mit

$$a =: x_0 < x_1 < x_2 \cdots < x_n := b \quad (80)$$

mit

$$\Delta x_i := x_i - x_{i-1} \text{ für } i = 1, \dots, n. \quad (81)$$

Dann heißt die Menge

$$Z := \{[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]\} \quad (82)$$

der durch $x_0, x_1, x_2, \dots, x_n$ definierten Teilintervalle von $[a, b]$ *Zerlegung von $[a, b]$* .
Weiterhin heißt

$$Z_{\max} := \max_{i \in n} \Delta x_i, \quad (83)$$

also die größte der Teilintervalllängen Δx_i , die *Feinheit* von Z .

Bemerkung

- Δx_i ist die Breite der der Streifen in untenstehender Abbildung.

Definition (Riemannsche Summe)

$f : [a, b] \rightarrow \mathbb{R}$ sei eine beschränkte Funktion auf $[a, b]$, d.h. $|f(x)| < c$ für $0 < c < \infty$ und alle $x \in [a, b]$, Z sei eine Zerlegung von $[a, b]$ mit Teilintervalllängen Δx_i für $i = 1, \dots, n$. Weiterhin sei für $i = 1, \dots, n$ ξ_i ein beliebiger Punkt im Teilintervall $[x_{i-1}, x_i]$ der Zerlegung Z . Dann heißt

$$R(Z) := \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (84)$$

Riemannsche Summe von f auf $[a, b]$ bezüglich der Zerlegung Z .

Bemerkungen

- Wählt man in jedem $[x_{i-1}, x_i]$ das Maximum von f , ergibt sich die *Riemannsche Obersumme*.
- Wählt man in jedem $[x_{i-1}, x_i]$ das Minimum von f , ergibt sich die *Riemannsche Untersumme*.
- Für $\Delta x_i \rightarrow 0, i = 1, \dots, n$ geht der Unterschied zwischen Ober- und Untersumme gegen 0.

Zerlegung und Riemann Ober- und Untersummen für $f : [a, b] \rightarrow \mathbb{R}$

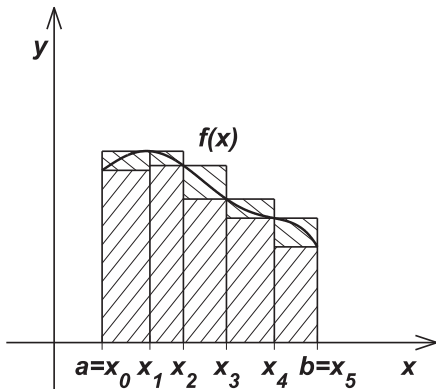


Abbildung 2.55 aus Bärwolff (2017)

Theorem (Riemannsches Integral)

$f : [a, b] \rightarrow \mathbb{R}$ sei eine beschränkte reellwertige Funktion auf $[a, b]$. Weiterhin sei für $Z_k, k = 1, 2, 3, \dots$ eine Folge von Zerlegungen von $[a, b]$ mit zugehörigen Feinheit $|Z_k|$. Wenn für jede Folge von Zerlegungen $Z_1, Z_2, \dots$ mit $|Z_k| \rightarrow 0$ für $k \rightarrow \infty$ und für beliebig gewählte Punkte $\xi_{ki}, i = 1, \dots, n$ im Teilintervall $[x_{k,i-1}, x_{k,i}]$ der Zerlegung Z_k gilt, dass die Folge der zugehörigen Riemannschen Summen $R(Z_1), R(Z_2), \dots$ gegen den gleichen Grenzwert strebt, dann heißt f auf $[a, b]$ *integrierbar*. Der entsprechende Grenzwert der Folge von Riemannschen Summen wird *bestimmtes Riemannsches Integral* genannt und mit

$$\int_a^b f(x) dx := \lim_{k \rightarrow \infty} R(Z_k) \text{ für } |Z_k| \rightarrow 0 \quad (85)$$

bezeichnet.

Bemerkungen

- Für $f > 0$ ist $\int_a^b f(x) dx$ der Flächeninhalt zwischen den $f(x)$ und der x -Achse
- Generell ist $\int_a^b f(x) dx$ der vorzeichenbehaftete Flächeninhalt den $f(x)$ und der x -Achse.
- Positive und negative Flächeninhalt gleichen einander aus.
- $\int_a^b f(x) dx$ ist als Mittelwert von f auf $[a, b]$ zu verstehen.

Theorem (Mittelwertsatz der Integralrechnung)

Für eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ existiert ein $\xi \in]a, b[$ mit

$$\int_a^b f(x) dx = f(\xi)(b - a) \quad (86)$$

- Statt eines Beweises verweisen wir auf untenstehende graphische Darstellung

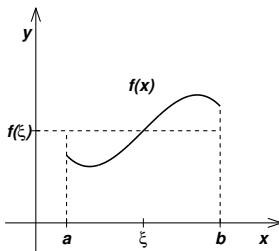


Abbildung 2.59 aus Bärwolff (2017)

Theorem (Rechenregeln für bestimmte Integrale)

Es seien f und g integrierbare Funktionen auf $[a, b]$. Dann gelten folgende Rechenregeln.

(1) Linearität. Für $c_1, c_2 \in \mathbb{R}$ gilt

$$\int_a^b (c_1 f(x) + c_2 g(x)) dx = c_1 \int_a^b f(x) dx + c_2 \int_a^b f(x) dx. \quad (87)$$

(2) Additivität. Für $a < c < b$ gilt

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (88)$$

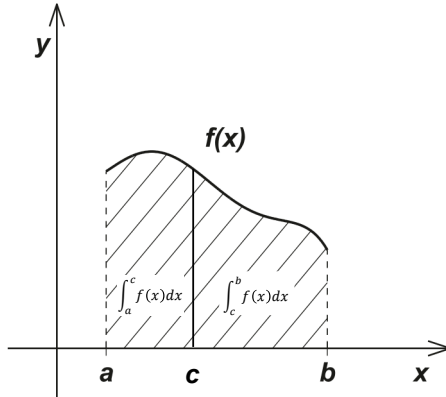
(2) Vorzeichenwechsel bei Umkehrung der Integralgrenzen

$$\int_a^b f(x) dx = - \int_b^a f(x) dx. \quad (89)$$

Bemerkung

- Für eine formale Herleitung verweisen wir auf die einschlägige Literatur.
- Eine graphische Darstellung der Additivität findet sich untenstehend.

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$



Nach Abbildung 2.55 aus Bärwolff (2017)

Theorem (Erster Hauptsatz der Differential- und Integralrechnung)

Ist $f : I \rightarrow \mathbb{R}$ eine auf dem Intervall $I \subset \mathbb{R}$ stetige Funktion, dann ist die Funktion

$$F : I \rightarrow \mathbb{R}, x \mapsto F(x) := \int_a^x f(t) dt \text{ mit } x, a \in I \quad (90)$$

eine Stammfunktion von f .

Beweis

Wir betrachten den Differenzquotienten

$$\frac{1}{h} (F(x+h) - F(x)) \quad (91)$$

Mit der Definition $F(x) := \int_a^x f(t) dt$ und der Additivität des bestimmten Integrals gilt dann

$$\frac{1}{h} (F(x+h) - F(x)) = \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt \quad (92)$$

Mit dem Mittelwertsatz der Integralrechnung gibt es also ein $\xi \in]x, x+h[$, so dass

$$\frac{1}{h} (F(x+h) - F(x)) = f(\xi) \quad (93)$$

Grenzwertbildung ergibt dann

$$\lim_{h \rightarrow 0} \frac{1}{h} (F(x+h) - F(x)) = \lim_{h \rightarrow 0} f(\xi) \text{ für } \xi \in]x, x+h[\Leftrightarrow F'(x) = f(x). \quad (94)$$

□

Theorem (Zweiter Hauptsatz der Differential- und Integralrechnung)

Ist F eine Stammfunktion einer stetigen Funktion $f : I \rightarrow \mathbb{R}$ auf einem Intervall I , so gilt für $a, b \in I$ mit $a \leq b$

$$\int_a^b f(x) dx = F(b) - F(a) =: F(x) \Big|_a^b \quad (95)$$

Beweis

Mit den Rechenregeln für bestimmte Integrale und dem ersten Hauptsatz der Differential- und Integralrechnung ergibt sich

$$F(b) - F(a) = \int_a^b f(t) dt - \int_a^a f(t) dt = \int_a^b f(x) dx \quad (96)$$

□

Definition (Uneigentliche Integrale)

$f : \mathbb{R} \rightarrow \mathbb{R}$ sei eine univariate reellwertige Funktion. Mit den Definitionen

$$\int_{-\infty}^b f(x) dx := \lim_{a \rightarrow -\infty} \int_a^b f(x) dx \text{ und } \int_a^{\infty} f(x) dx := \lim_{b \rightarrow \infty} \int_a^b f(x) dx \quad (97)$$

und der Additivität von Integralen

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^b f(x) dx + \int_b^{\infty} f(x) dx \quad (98)$$

wird der Begriff des bestimmten Integrals auf die unbeschränkten Integrationsintervalle $]-\infty, b]$, $[a, \infty[$ und $]-\infty, \infty[$ erweitert. Integrale mit unbeschränkten Integrationsintervallen heißen *uneigentliche Integrale*. Wenn die entsprechenden Grenzwerte existieren, sagt man, dass die uneigentlichen Integrale *konvergieren*.

Bemerkung

- Für die WDF f einer Zufallsvariable ist die Forderung $\int_{-\infty}^{\infty} f(x) dx = 1$ zentral.

Beispiel (Uneigentliches Integral)

Wir betrachten das uneigentliche Integral $\int_1^{\infty} \frac{1}{x^2} dx$. Nach den Festlegungen in der Definition uneigentlicher Integrale gilt

$$\int_1^{\infty} \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx. \quad (99)$$

Mit der Stammfunktion $F(x) = -x^{-1}$ von $f(x) = x^{-2}$ ergibt sich für das bestimmte Integral in obiger Gleichung

$$\int_1^b \frac{1}{x^2} dx = F(b) - F(1) = -\frac{1}{b} - \left(-\frac{1}{1}\right) = -\frac{1}{b} + 1. \quad (100)$$

Es ergibt sich also

$$\int_1^{\infty} \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \left(-\frac{1}{b} + 1\right) = -\lim_{b \rightarrow \infty} \frac{1}{b} + \lim_{b \rightarrow \infty} 1 = 0 + 1 = 1. \quad (101)$$

Definition (Mehrdimensionale Integrale)

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ sei eine multivariate reellwertige Funktion. Dann heißen Integrale der Form

$$\int_{[a_1, b_1] \times \dots \times [a_n, b_n]} f(x) dx = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_1 \dots dx_n \quad (102)$$

mehrdimensionale bestimmte Integrale auf Hyperrechtecken. Weiterhin heißen Integrale der Form

$$\int_{\mathbb{R}^n} f(x) dx = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \dots dx_n \quad (103)$$

mehrdimensionale uneigentliche Integrale.

Bemerkungen

- Man kann multivariate reellwertige Funktion nicht nur auf Hyperrechtecken, sondern im Prinzip auf beliebigen Hyperflächen integrieren. Dies kann sich jedoch oft schwierig gestalten.
- Der *Satz von Fubini* besagt, dass man mehrdimensionale Integrale in beliebiger Koordinatenfolge auswerten kann, also dass zum Beispiel

$$\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2. \quad (104)$$

- Für die WDF eines Zufallsvektors ist die Forderung $\int_{\mathbb{R}^n} f(x) dx = 1$ zentral.

Beispiel (Zweidimensionales Integral)

Wir betrachten das zweidimensionale bestimmte Integral der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto f(x_1, x_2) := x_1^2 + 4x_2 \quad (105)$$

auf dem Rechteck $[0, 1] \times [0, 1]$. In

$$\int_0^1 \int_0^1 x_1^2 + 4x_2 \, dx_1 \, dx_2 = \int_0^1 \left(\int_0^1 x_1^2 + 4x_2 \, dx_1 \right) dx_2 \quad (106)$$

betrachten wir zunächst das innere Integral. x_2 nimmt hier die Rolle einer Konstanten ein. Eine Stammfunktion von $g(x_1) := x_1^2 + 4x_2$ ist $G(x_1) = \frac{1}{3}x_1^3 + 4x_2x_1$, wie man sich durch Ableiten von G überzeugt. Es ergibt sich also für das innere Integral

$$\int_0^1 x_1^2 + 4x_2 \, dx_1 = G(1) - G(0) = \frac{1}{3} \cdot 1^3 + 4x_2 \cdot 1 - \frac{1}{3} \cdot 0^3 - 4x_2 \cdot 0 = \frac{1}{3} + 4x_2 \quad (107)$$

Betrachten des äußeren Integrals in Gleichung (106) ergibt dann mit der Stammfunktion $H(x_2) = \frac{1}{3}x_2 + 2x_2^2$ von $h(x_2) := \frac{1}{3} + 4x_2$, dass

$$\int_0^1 \int_0^1 x_1^2 + 4x_2 \, dx_1 \, dx_2 = \int_0^1 \frac{1}{3} + 4x_2 \, dx_2 = H(1) - H(0) = \frac{1}{3} \cdot 1 + 4 \cdot 1^2 - \frac{1}{3} \cdot 0 + 4 \cdot 0^2 = \frac{13}{3}. \quad (108)$$

Mathematische Grundlagen

- Mengen
- Summen, Produkte, Potenzen
- Funktionen
- Differentialrechnung
- Analytische Optimierung
- Integralrechnung
- **Selbstkontrollfragen**

Selbstkontrollfragen

1. Diskutieren Sie den Unterschiede zwischen Definitionen, Theoremen, und Beweisen.
2. Geben Sie die Definition einer Menge nach Cantor (1895) wieder.
3. Nennen Sie drei Möglichkeiten zur Definition einer Menge.
4. Für Mengen M, N erläutern Sie die Ausdrücke $m \in M, m \notin N, M \subseteq N, M \subset N$.
5. Definieren Sie den Begriff der Kardinalität einer Menge.
6. Definieren Sie den Begriff der Potenzmenge einer Menge.
7. Es sei $M := \{a, b\}$. Bestimmen Sie $\mathcal{P}(M)$.
8. Es seien $M := \{a, b\}, N := \{a, c, d\}$. Bestimmen Sie $M \cup N, M \cap N, M \setminus N, M \Delta N$.
9. Erläutern Sie die Symbole $\mathbb{N}, \mathbb{N}_n$, und $\mathbb{N}^0$.
10. Erläutern Sie den Unterschied zwischen $\mathbb{N}$ und $\mathbb{Z}$.
11. Erläutern Sie den Unterschied zwischen $\mathbb{R}$ und $\mathbb{Q}$.
12. Definieren Sie die Begriffe des abgeschlossenen, offenen, und halboffenen Intervalls.
13. Es seien M und N Mengen. Erläutern Sie die Notation $M \times N$.
14. Definieren sie die Menge $\mathbb{R}^n$.
15. Berechnen Sie die Summen $\sum_{i=1}^3 2$, $\sum_{i=1}^3 i^2$, und $\sum_{i=1}^3 \frac{2}{3}i$.
16. Schreiben Sie die Summe $1 + 3 + 5 + 7 + 9 + 11$ mithilfe des Summenzeichens.

Selbstkontrollfragen

17. Schreiben Sie die Summe $0 + 2 + 4 + 6 + 8 + 10$ mithilfe des Summenzeichens.
18. Definieren Sie das Produktzeichen.
19. Für $a \in \mathbb{R}$, definieren Sie die n te (negative) Potenz von a .
20. Berechnen Sie $2^2 \cdot 2^3$ und 2^5 . Geben Sie die zugehörige Potenzregel an.
21. Berechnen Sie 6^2 und $2^2 \cdot 3^2$. Geben Sie die zugehörige Potenzregel an.
22. Warum kann die n -te Wurzel von a als $a^{\frac{1}{n}}$ geschrieben werden?
23. Berechnen Sie $(\sqrt{2})^{\frac{2}{3}}$, $9^{\frac{1}{2}}$, und $4^{-\frac{1}{2}}$.
24. Erläutern Sie die Komponenten der Funktionsschreibweise $f : D \rightarrow Z, x \mapsto f(x)$.
25. Definieren Sie die Begriffe Bildmenge, Wertebereich, und Urbildmenge einer Funktion.
26. Definieren Sie die Begriffe Surjektivität, Injektivität, und Bijektivität einer Funktion.
27. Erläutern Sie, warum $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := x^2$ weder injektiv noch surjektiv ist.
28. Erläutern Sie, warum $f : [0, \infty[\rightarrow [0, \infty[, x \mapsto f(x) := x^2$ bijektiv ist.
29. Erläutern Sie die Komponenten der Schreibweise $g \circ f : D \rightarrow S, x \mapsto (g \circ f)(x)$.
30. Definieren Sie den Begriff der inversen Funktion.
31. Geben Sie die inverse Funktion von x^2 auf $[0, \infty[$ an.
32. Definieren Sie den Begriff der linearen Abbildung.

Selbstkontrollfragen

33. Definieren Sie die Begriffe der univariat-und multivariat-reellwertigen Funktion.
34. Definieren Sie Begriff der multivariaten vektorwertigen Funktion.
35. Skizzieren Sie die konstante Funktion für $a := 1$ und die Identitätsfunktion.
36. Für $a = 2$ und $b = 3$, skizzieren Sie die linear Funktion $f(x) = ax + b$.
37. Skizzieren Sie die Funktionen $f(x) := (x - 1)^2$ und $g(x) := (x + 3)^2$
38. Skizzieren Sie die Exponential- und Logarithmusfunktionen.
39. Geben Sie Exponentialeigenschaften der Exponentialfunktion an.
40. Geben Sie die Logarithmeneigenschaften der Logarithmusfunktion an.
41. Definieren Sie den Begriff der Ableitung $f'(a)$ einer Funktion f an einer Stelle a .
42. Definieren den Begriff der Ableitung f' einer Funktion f .
43. Erläutern Sie die Symbole $f'(x)$, $\dot{f}(x)$, $\frac{df(x)}{dx}$, und $\frac{d}{dx} f(x)$.
44. Definieren Sie den Begriff der zweiten Ableitung f'' einer Funktion f .
45. Geben Sie die Summenregel für Ableitungen wieder.
46. Geben Sie die Produktregel für Ableitungen wieder.
47. Geben Sie die Quotientenregel für Ableitungen wieder.
48. Geben Sie die Kettenregel für Ableitungen wieder.

50. Bestimmen Sie die Ableitung der Funktion $f(x) := 3x^2 + \exp(-x^2) - x \ln(x)$.
51. Bestimmen Sie die Ableitung der Funktion $f(x) := \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2$ für $\mu \in \mathbb{R}$.
52. Definieren Sie die Begriffe des globalen und lokalen Maximums/Minimums einer Funktion.
53. Geben Sie die notwendige Bedingung für ein Extremum einer Funktion wieder.
54. Geben Sie die hinreichende Bedingung für ein lokales Extremum einer Funktion wieder.
55. Geben Sie das Standardverfahren der analytischen Optimierung wieder.
56. Bestimmen Sie einen Extremwert von $f(x) := \exp(-\frac{1}{2}(x - \mu)^2)$ für $\mu \in \mathbb{R}$.
57. Definieren Sie den Begriff der Stammfunktion einer univariaten reellwertigen Funktion.
58. Definieren Sie den Begriff des unbestimmten Integrals einer univariaten reellwertigen Funktion.
59. Erläutern Sie den Begriff des Riemanschen Integrals.
60. Geben Sie den ersten Hauptsatz der Differential- und Integralrechnung wieder.
61. Geben Sie den zweiten Hauptsatz der Differential- und Integralrechnung wieder.
62. Erläutern Sie den Begriff des uneigentlichen Integrals.
63. Erläutern Sie den Begriff des mehrdimensionalen Integrals.
64. Berechnen Sie das bestimmte Integral $\int_0^1 2x \, dx$.

Literatur

- Bärwolff, G. (2017). *Höhere Mathematik für Naturwissenschaftler und Ingenieure*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen*, 46(4):481–512.