

# Mathematik für Maschinenbauer, Bauingenieure und Umwelttechniker III

Vorlesungsskriptum WS 2001/02 - WS 2002/03  
überarbeitet Februar 2008

R. Verfürth

Fakultät für Mathematik, Ruhr-Universität Bochum



## Inhaltsverzeichnis

Kapitel XI. Stochastik I	
Diskrete Modelle	7
XI.1. Modelle für Zufallsexperimente	7
XI.1.1. Endliche Wahrscheinlichkeitsräume	7
XI.1.2. Urnenmodelle	9
XI.1.3. Anwendungsbeispiele	10
XI.1.4. Die hypergeometrische Verteilung	12
XI.1.5. Multinomialkoeffizienten	13
XI.1.6. Identitäten für Binomialkoeffizienten	14
XI.2. Bedingte Wahrscheinlichkeit und Unabhängigkeit	14
XI.2.1. Definition bedingter Wahrscheinlichkeiten	14
XI.2.2. Eigenschaften	16
XI.2.3. Unabhängigkeit	18
XI.2.4. Produktexperimente	19
XI.2.5. Binomialverteilung	20
XI.2.6. Multinomialverteilung	21
XI.2.7. Geometrische Verteilung	21
XI.2.8. Negative Binomialverteilung	21
XI.3. Zufallsvariable, Erwartungswert, Varianz	21
XI.3.1. Zufallsvariable	21
XI.3.2. Gemeinsame Verteilung mehrerer Zufallsvariabler	22
XI.3.3. Unabhängigkeit	24
XI.3.4. Erwartungswert	24
XI.3.5. Varianz und Kovarianz	26
XI.3.6. Das schwache Gesetz der großen Zahlen	30
XI.4. Grundbegriffe der Schätztheorie	32
XI.4.1. Motivation	32
XI.4.2. Der allgemeine Rahmen von Schätzproblemen	33
XI.4.3. Maximum-Likelihood Schätzer	33
XI.4.4. Erwartungstreue	34
XI.4.5. Der mittlere quadratische Fehler	36
XI.5. Approximationen der Binomialverteilung	37
XI.5.1. Approximation von $n!$ und $b_{n,p}(k)$	38
XI.5.2. Der Satz von Moivre-Laplace	41
XI.5.3. Die Poisson-Approximation	43
XI.6. Tests	45
XI.6.1. Motivation	45

XI.6.2.	Grundbegriffe der Testtheorie	47
XI.6.3.	Zurück zur Motivation	48
Kapitel XII. Stochastik II		
	Allgemeine Modelle	51
XII.1.	Wahrscheinlichkeitsmaße mit Dichten	51
XII.1.1.	Ergebnismengen	51
XII.1.2.	$\sigma$ -Algebren	51
XII.1.3.	Wahrscheinlichkeitsmaße	52
XII.1.4.	Wahrscheinlichkeitsmaße mit Dichten	53
XII.1.5.	Gleichverteilung auf einem Intervall	53
XII.1.6.	Exponentialverteilung	54
XII.1.7.	Normalverteilung	54
XII.1.8.	Produkttdichten	54
XII.2.	Zufallsvariable und ihre Momente	55
XII.2.1.	Messbare Funktionen	55
XII.2.2.	Zufallsvariable	55
XII.2.3.	Unabhängigkeit	56
XII.2.4.	Erwartungswert	58
XII.2.5.	Varianz	60
XII.3.	Schätzverfahren	62
XII.3.1.	Maximum-Likelihood Schätzung	62
XII.3.2.	Die Methode der kleinsten Quadrate	64
XII.3.3.	Median	66
XII.4.	Tests	67
XII.4.1.	Vorbemerkungen	67
XII.4.2.	Der $t$ -Test	68
XII.4.3.	Der $\chi^2$ -Test	71
Kapitel XIII. Fourier-Analysis		
XIII.1.	Trigonometrische Polynome und Reihen	75
XIII.1.1.	Periodische Funktionen	75
XIII.1.2.	Trigonometrische Polynome	76
XIII.1.3.	Trigonometrische Reihen	77
XIII.2.	Fourier-Reihen	78
XIII.2.1.	Die Fourier-Reihe einer Funktion	78
XIII.2.2.	Rechenregeln	79
XIII.2.3.	Die Bessel-Ungleichung	80
XIII.2.4.	Konvergenz der Fourier-Reihe	82
XIII.2.5.	Anwendung auf gewöhnliche Differentialgleichungen	84
XIII.3.	Die Fourier-Transformation	85
XIII.3.1.	Definition	85
XIII.3.2.	Rechenregeln	86
XIII.3.3.	Existenz- und Eindeutigkeitssätze	88
Kapitel XIV. Partielle Differentialgleichungen		
		91

XIV.1.	Einführung	91
XIV.1.1.	Beispiele	91
XIV.1.2.	Bezeichnungen	95
XIV.1.3.	Lineare partielle Differentialgleichungen zweiter Ordnung	96
XIV.1.4.	Anfangs- und Randbedingungen	97
XIV.2.	Die Wärmeleitungsgleichung	98
XIV.2.1.	Vorbemerkungen	98
XIV.2.2.	Fortsetzung der Randwerte	100
XIV.2.3.	Lösung der homogenen Gleichung mit homogenen Randbedingungen	100
XIV.2.4.	Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen	103
XIV.2.5.	Integraldarstellung	105
XIV.2.6.	Höhere Raumdimensionen	105
XIV.3.	Die Wellengleichung	106
XIV.3.1.	Vorbemerkung	106
XIV.3.2.	Lösung der homogenen Gleichung mit homogenen Randbedingungen	107
XIV.3.3.	Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen	109
XIV.3.4.	Integraldarstellung	110
XIV.3.5.	Höhere Raumdimensionen	111
XIV.3.6.	Dämpfung	112
XIV.4.	Die Poissongleichung	112
XIV.4.1.	Das Eigenwertproblem im Rechteck	112
XIV.4.2.	Die Poissongleichung im Rechteck	114
XIV.4.3.	Das Eigenwertproblem im Kreis	116
XIV.4.4.	Die Poissongleichung im Kreis	118
	Zusammenfassung	123
	Index	127



## KAPITEL XI

### Stochastik I Diskrete Modelle

#### XI.1. Modelle für Zufallsexperimente

**XI.1.1. Endliche Wahrscheinlichkeitsräume.** Wir betrachten Zufallsexperimente mit endlich vielen möglichen Ausgängen. Diese werden beschrieben durch eine endliche, nicht leere Menge  $\Omega$ , deren Elemente  $\omega$  die Versuchsausgänge bezeichnen. Sie heißen ERGEBNISSE oder ELEMENTAREREIGNISSE.  $\Omega$  heißt ERGEBNISMENGE.

Die Teilmengen von  $\Omega$  sind die EREIGNISSE, die in dem Modell in Betracht gezogen werden. Genauer:

Wir identifizieren  $A \subset \Omega$  mit dem Ereignis, dass ein  $\omega \in A$  der beobachtete Versuchsausgang ist.

Dementsprechend bezeichnen  $A \cap B$  bzw.  $A \cup B$  die Ereignisse, dass  $A$  und  $B$  bzw.  $A$  oder  $B$  eintreten. Die leere Menge  $\emptyset$  heißt auch das UNMÖGLICHE EREIGNIS;  $\Omega$  ist das SICHERE EREIGNIS.

Jedem Ereignis ordnen wir eine Wahrscheinlichkeit zu. Die Menge aller möglichen Ereignisse ist die POTENZMENGE  $\mathcal{P}(\Omega)$ , d.h. die Menge aller Teilmengen von  $\Omega$ . Eine Abbildung  $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$  heißt WAHRSCHEINLICHKEITSVERTEILUNG oder WAHRSCHEINLICHKEITSMASS, wenn sie folgende Eigenschaften hat:

$P(\Omega) = 1$	(NORMIERUNG)
$P(A) \geq 0$	für alle $A$ (POSITIVITÄT)
$P(A \cup B) = P(A) + P(B)$	für alle disjunkten $A, B$ (ADDITIVITÄT)

$P(A)$  heißt die WAHRSCHEINLICHKEIT des Ereignisses  $A$ . Das Paar  $(\Omega, P)$  heißt der dem Zufallsexperiment zugeordnete WAHRSCHEINLICHKEITSRAUM.

**BEISPIEL XI.1.1.** Wir wollen die Wahrscheinlichkeit bestimmen, dass die Summe der bei zwei Würfeln eines Würfels erhaltenen Augenzahlen mindestens 10 ist. Wir können die Ergebnisse des Zufallsexperimentes „Zweimaliges Werfen eines Würfels“ durch die Paare  $(i, k)$

der beobachteten Augenzahlen beschreiben. Daher ist

$$\Omega = \{(i, k) : 1 \leq i, k \leq 6\}.$$

$\Omega$  hat 36 Elemente. Aus Symmetriegründen ist es naheliegend, sie alle als gleich wahrscheinlich zu betrachten. Jedes  $(i, k)$  hat also die Wahrscheinlichkeit  $\frac{1}{36}$ . Die Menge der Ergebnisse, für die die Summe  $i + k$  mindestens 10 ist, ist

$$A = \{(6, 6), (6, 5), (5, 6), (6, 4), (4, 6), (5, 5)\}.$$

Da  $A$  genau 6 Elemente hat, ist die gesuchte Wahrscheinlichkeit  $\frac{6}{36} = \frac{1}{6}$ .

Aus den obigen Eigenschaften eines Wahrscheinlichkeitsmaßes folgen leicht weitere Eigenschaften. Für  $A, B, A_i \in \mathcal{P}(\Omega)$  gilt

$$\begin{aligned} P(\Omega \setminus A) &= 1 - P(A) \\ P(\emptyset) &= 0 \\ A \subset B &\Rightarrow P(A) \leq P(B) \\ P(A \setminus B) &= P(A) - P(A \cap B) \end{aligned}$$

und

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) \quad \text{falls } A_1, \dots, A_n \text{ paarweise disjunkt} \\ P\left(\bigcup_{i=1}^n A_i\right) &\leq \sum_{i=1}^n P(A_i) \quad \text{für beliebige } A_1, \dots, A_n \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Aus diesen Eigenschaften folgt insbesondere

$$P(A) = \sum_{\omega \in A} P(\{\omega\}).$$

Die Wahrscheinlichkeit von  $A$  ist also die Summe der Wahrscheinlichkeiten der Ergebnisse, bei denen  $A$  eintritt.  $P$  ist also eindeutig bestimmt durch die Werte aller  $P(\{\omega\})$  mit  $\omega \in \Omega$ . Man schreibt auch  $P(\omega)$  statt  $P(\{\omega\})$ . Die Abbildung  $\omega \mapsto P(\omega)$  heißt WAHRSCHEINLICHKEITSFUNKTION. Die Wahrscheinlichkeitsverteilung kann also auch durch Angabe der zugehörigen Wahrscheinlichkeitsfunktion beschrieben werden. Offensichtlich gilt für jede Wahrscheinlichkeitsfunktion:

$$\begin{aligned} P(\omega) &\geq 0 \quad \text{für alle } \omega \in \Omega \\ \sum_{\omega \in \Omega} P(\omega) &= 1. \end{aligned}$$

Ein besonders wichtiger Spezialfall einer Wahrscheinlichkeitsfunktion ist die GLEICHVERTEILUNG auf  $\Omega$ , bei der alle Ergebnisse  $\omega \in \Omega$  gleich wahrscheinlich sind. In diesem Fall wird  $(\Omega, P)$  als LAPLACESCHER WAHRSCHEINLICHKEITSRAUM bezeichnet; das zugehörige Zufallsexperiment heißt LAPLACE-EXPERIMENT. Offensichtlich gilt dann:

$$\begin{aligned} P(\omega) &= \frac{1}{\text{card}(\Omega)} && \text{für alle } \omega \in \Omega, \\ P(A) &= \frac{\text{card}(A)}{\text{card}(\Omega)} && \text{für alle } A \subset \Omega. \end{aligned}$$

**XI.1.2. Urnenmodelle.** Viele Laplace-Experimente lassen sich durch Urnenmodelle beschreiben. Dabei stellen wir uns vor, dass wir aus einer Urne mit  $N$  Kugeln, die von 1 bis  $N$  nummeriert sind, sukzessive  $n$  Kugeln zufällig ziehen.

Offensichtlich sind zwei Arten des Ziehens zu unterscheiden:

- (1) MIT RÜCKLEGEN: Nach jedem Zug wird die gerade gezogene Kugel wieder in die Urne zurückgelegt; jede Kugel kann mehrmals gezogen werden.
- (2) OHNE RÜCKLEGEN: eine gezogene Kugel wird nicht zurückgelegt; jede Kugel kann nur einmal gezogen werden.

Ein Beispiel für (1) ist das mehrmalige Werfen eines Würfels; ein Beispiel für (2) ist die Ziehung der Lottozahlen.

Man kann das Ergebnis der Folge der Ziehungen dadurch beschreiben, dass man das  $n$ -Tupel  $(\omega_1, \dots, \omega_n)$  angibt, in dem  $\omega_i$  die Nummer der im  $i$ -ten Zug gezogenen Kugel ist. Offensichtlich kann man hier wiederum zwei Arten von Ziehen unterscheiden:

- (i) MIT REIHENFOLGE: Es kommt auf die Reihenfolge des Erscheinens an;  $(1, 2)$  und  $(2, 1)$  bezeichnen verschiedene Ergebnisse.
- (ii) OHNE REIHENFOLGE: Es kommt nicht auf die Reihenfolge des Erscheinens an;  $(1, 2)$  und  $(2, 1)$  bezeichnen gleiche Ergebnisse.

Ein Beispiel für (i) ist das Bestimmen einer Losnummer durch sukzessives Ziehen ihrer Ziffern; ein Beispiel für (ii) ist wieder die Ziehung der Lottozahlen.

Durch diese Unterscheidungen ergeben sich insgesamt vier verschiedene Ergebnismengen. Zu ihrer Beschreibung setzen wir

$$\mathbb{A} = \{1, 2, \dots, N\}.$$

(I) STICHPROBEN IN REIHENFOLGE MIT RÜCKLEGEN (1i):

$$\begin{aligned} \Omega_I &= \mathbb{A}^n, \\ \text{card}(\Omega_I) &= N^n. \end{aligned}$$

(II) STICHPROBEN IN REIHENFOLGE OHNE RÜCKLEGEN (2i):

$$\begin{aligned}\Omega_{II} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_i \neq \omega_j \text{ für } i \neq j\} \\ \text{card}(\Omega_{II}) &= \frac{N!}{(N-n)!} \\ &= N \cdot (N-1) \cdot \dots \cdot (N-n+1).\end{aligned}$$

(III) STICHPROBEN OHNE REIHENFOLGE OHNE RÜCKLEGEN (2ii):

$$\begin{aligned}\Omega_{III} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_1 < \omega_2 < \dots < \omega_n\} \\ \text{card}(\Omega_{III}) &= \binom{N}{n} \\ &= \frac{N!}{(N-n)!n!}.\end{aligned}$$

(IV) STICHPROBEN OHNE REIHENFOLGE MIT RÜCKLEGEN (1ii):

$$\begin{aligned}\Omega_{IV} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_1 \leq \omega_2 \leq \dots \leq \omega_n\} \\ \text{card}(\Omega_{IV}) &= \binom{N+n-1}{n}.\end{aligned}$$

Man kann diese Urnenmodelle auch alternativ wie folgt interpretieren:

Gefragt ist nach der Anzahl der Möglichkeiten,  $n$  Murmeln auf  $N$  Plätze zu verteilen. Dabei entspricht die Nummer der Ziehung der Nummer der Murmel und die Nummer der Kugel der Nummer des Platzes.

### XI.1.3. Anwendungsbeispiele.

BEISPIEL XI.1.2. Es werden vier völlig gleich aussehende Würfel gleichzeitig geworfen. Welches ist die Wahrscheinlichkeit  $p$ , dass die vier erscheinenden Augenzahlen verschieden sind?

Da ein gleichzeitiger Wurf von vier Würfeln dem viermaligen Werfen eines Würfels entspricht, handelt es sich um ein Experiment mit Rücklegen. Da die Würfel nicht unterscheidbar sind, könnte man versucht sein, ein Modell ohne Reihenfolge zu betrachten. Dies ist aber *falsch*: Dem Resultat, dass die Augenzahlen 1, 2, 3 und 4 auftreten, entspricht in diesem Modell nur ein Ereignis, da es nicht auf die Reihenfolge ankommt. Tatsächlich entsprechen diesem Resultat jedoch die  $4! = 24$  Permutationen der Menge  $\{1, 2, 3, 4\}$ . Das *richtige* Modell ist also dasjenige mit Berücksichtigung der Reihenfolge. Daher sind alle möglichen Fälle  $\Omega_I$  mit  $N = 6$  und  $n = 4$ . Die günstigen Fälle (alle Augenzahlen verschieden) sind  $\Omega_{II}$  mit  $N = 6$  und  $n = 4$ . Daher ist die gesuchte Wahrscheinlichkeit

$$\begin{aligned}p &= \frac{6!}{6^4} \\ &= \frac{3 \cdot 4 \cdot 5 \cdot 6}{6^4}\end{aligned}$$

$$= \frac{5}{18}.$$

BEISPIEL XI.1.3. Wie groß ist die Wahrscheinlichkeit  $p$ , dass mindestens 2 von 25 Schülern einer Klasse am gleichen Tag Geburtstag haben?

Als Ereignisraum können wir  $\Omega_I$  mit  $n = 25$  und  $N = 365$  wählen. Das Ergebnis  $(\omega_1, \dots, \omega_{25})$  bedeutet, dass Schüler Nummer 1 am  $\omega_1$ -ten Tag Geburtstag hat, Schüler Nummer 2 am  $\omega_2$ -ten Tag usw. Das interessierende Ereignis ist das Komplement des Ereignisses  $\Omega_{II}$ , das alle Schüler an verschiedenen Tagen Geburtstag haben. Also ist

$$p = 1 - P(\Omega_{II}).$$

Für  $P(\Omega_{II})$  erhalten wir

$$\begin{aligned} P(\Omega_{II}) &= \frac{N!}{(N-n)! N^n} \\ &= \frac{N}{N} \cdot \frac{N-1}{N} \cdot \frac{N-2}{N} \cdots \frac{N-n+1}{N} \\ &= 1 \cdot \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right). \end{aligned}$$

Wenn wir jeden Ausdruck  $\left(1 - \frac{k}{N}\right)$  durch  $\exp\left(-\frac{k}{N}\right)$  approximieren, erhalten wir

$$\begin{aligned} P(\Omega_{II}) &\approx \exp\left(-\sum_{k=1}^{n-1} \frac{k}{N}\right) \\ &= \exp\left(-\frac{n(n-1)}{2N}\right) \\ &= \exp\left(-\frac{600}{730}\right) \\ &\approx 0.44. \end{aligned}$$

Damit ergibt sich  $p \approx 0.56$ . Der exakte Wert ohne diese Approximation ist  $p = 0.568$ .

BEISPIEL XI.1.4. Beim Zahlenlotto „6 aus 49“ werden  $n = 6$  Kugeln aus  $N = 49$  Kugeln ohne Rücklegen gezogen. Dabei kommt es nicht auf die Reihenfolge an. Also ist die Ereignismenge  $\Omega_{III}$  mit

$$\begin{aligned} \text{card}(\Omega_{III}) &= \binom{49}{6} \\ &= 13983816. \end{aligned}$$

Daher ist die Wahrscheinlichkeit  $p_6$  für „6 Richtige“

$$\begin{aligned} p_6 &= \frac{1}{\text{card}(\Omega_{III})} \\ &\approx 7.1511 \cdot 10^{-8}. \end{aligned}$$

Wie groß ist die Wahrscheinlichkeit  $p_4$  für „genau 4 Richtige“?

Seien  $\tilde{\omega}_1, \dots, \tilde{\omega}_6$  die gezogenen Zahlen. Dann können wir die Menge aller gesuchten Ergebnisse wie folgt erzeugen: Wir ziehen zunächst 4 Kugeln aus  $\{\tilde{\omega}_1, \dots, \tilde{\omega}_6\}$  und dann zwei Kugeln aus den 43 Kugeln  $\{1, \dots, 49\} \setminus \{\tilde{\omega}_1, \dots, \tilde{\omega}_6\}$ . Für das erste Ziehen haben wir

$$\binom{6}{4} = 15$$

Möglichkeiten, für das zweite Ziehen

$$\binom{43}{2} = 21 \cdot 43 = 903$$

Möglichkeiten. Daher ist

$$\begin{aligned} p_4 &= \frac{15 \cdot 903}{13983816} \\ &= \frac{3 \cdot 5 \cdot 43}{7 \cdot 44 \cdot 46 \cdot 47} \\ &\approx 9.68619 \cdot 10^{-4}. \end{aligned}$$

Analog kann man die Wahrscheinlichkeit  $p_5$  für „genau 5 Richtige“ berechnen. Die Wahrscheinlichkeit für „mindestens 4 Richtige“ ist dann

$$p_4 + p_5 + p_6 \approx 9.8714 \cdot 10^{-4}.$$

BEISPIEL XI.1.5. Auf wie viele Arten können sich zwei nicht unterscheidbare Spatzen auf vier Telegraphenleitungen verteilen?

Dies ist  $\Omega_{IV}$  mit  $N = 4$  und  $n = 2$ . Dementsprechend gibt es

$$\binom{5}{2} = 10$$

Möglichkeiten.

**XI.1.4. Die hypergeometrische Verteilung.** In Beispiel XI.1.4 haben wir einen Spezialfall dieser wichtigen Verteilung kennen gelernt. Für den allgemeinen Fall betrachten wir eine Urne, die  $S$  schwarze und  $W$  weiße Kugeln enthält, insgesamt  $N = S + W$  Kugeln. Es werden  $n \leq S + W$  Kugeln ohne Zurücklegen gezogen. Die Wahrscheinlichkeit, dass dabei  $s$  schwarze und  $w = n - s$  weiße Kugeln gezogen werden, beträgt

$$h(s; n, N, S) = \frac{\binom{S}{s} \binom{W}{w}}{\binom{S+W}{n}} \quad (0 \leq s \leq n).$$

Diese Verteilung heißt HYPERGEOMETRISCHE VERTEILUNG.

BEISPIEL XI.1.6. Wie groß ist die Wahrscheinlichkeit, dass beim Skatspiel ein Spieler genau 3 Asse erhält?

Dies entspricht obiger Situation mit  $S = 4$  (Zahl aller Asse),  $W =$

28 (Zähler aller anderen Karten),  $s = 3$  (Zahl der Asse des Spielers) und  $n = 10$  (Zahl der Karten des Spielers). Daher ist die gesuchte Wahrscheinlichkeit

$$\begin{aligned} p &= \frac{\binom{4}{3} \binom{28}{7}}{\binom{32}{10}} \\ &= \frac{4 \cdot 22 \cdot 8 \cdot 9 \cdot 10}{32 \cdot 31 \cdot 30 \cdot 29} \\ &= \frac{66}{899} \\ &\approx 7.34\%. \end{aligned}$$

Die Wahrscheinlichkeit, dass mindestens ein Spieler 3 Asse erhält, ist dreimal so groß, da keine zwei Spieler gleichzeitig drei Asse erhalten können.

**XI.1.5. Multinomialkoeffizienten.** Die Binomialkoeffizienten  $\binom{n}{k}$  beschreiben, auf wie viele Arten man eine Menge von  $n$  nummerierten Kugeln so in zwei Gruppen einteilen kann, dass die erste Gruppe  $k$  Kugeln enthält. Wie viele Möglichkeiten gibt es, die Zahlen  $1, \dots, n$  so in  $r$  Gruppen einzuteilen, dass die erste Gruppe  $k_1$  Elemente hat, die zweite  $k_2$  Elemente usw.? Dabei muss man natürlich  $k_1 + \dots + k_r = n$  voraussetzen. Man kann zuerst auf  $\binom{n}{k_1}$  Arten die erste Gruppe auswählen, dann auf  $\binom{n-k_1}{k_2}$  Arten die zweite Gruppe usw. Insgesamt erhält man

$$\begin{aligned} &\binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \binom{n-k_1-k_2}{k_3} \cdot \dots \cdot \binom{n-k_1-\dots-k_{r-1}}{k_r} \\ &= \frac{n!}{k_1!k_2!\dots k_r!} \end{aligned}$$

Möglichkeiten. Diesen Ausdruck bezeichnet man als MULTINOMIALKOEFFIZIENTEN und kürzt ihn mit  $\binom{n}{k_1, k_2, \dots, k_r}$  ab:

$$\boxed{\binom{n}{k_1, k_2, \dots, k_r} = \frac{n!}{k_1!k_2!\dots k_r!} \quad \text{mit } k_1 + \dots + k_r = n.}$$

BEISPIEL XI.1.7. 26 Schulkinder haben einen Fußball, vier Tennisschläger, ein Fußballfeld und einen Tennisplatz. Die Zahl der Einteilungen in zwei Fußballmannschaften mit je 11 Spielern und in zwei Tennisteamen mit je 2 Spielern beträgt

$$\begin{aligned} \binom{26}{11, 11, 2, 2} &= \frac{26!}{11!11!2!2!} \\ &= 2 \cdot 5 \cdot 12 \cdot 13 \cdot 14 \cdot 15 \cdot 17 \cdot 19 \cdot 23 \cdot 26 \\ &\approx 6.327 \cdot 10^{10}. \end{aligned}$$

**XI.1.6. Identitäten für Binomialkoeffizienten.** Aus der Binomischen Formel (vgl. Abschnitt I.2.8 (S. 19, Teil I))

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

ergibt sich durch Einsetzen von  $x = 1, y = 1$  bzw.  $x = -1, y = 1$  bzw. durch Ableiten nach  $x$  und anschließendes Einsetzen von  $x = 1, y = 1$ :

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} &= 2^n, \\ \sum_{k=0}^n (-1)^k \binom{n}{k} &= 0, \\ \sum_{k=1}^n k \binom{n}{k} &= n2^{n-1}. \end{aligned}$$

In Abschnitt I.2.8 (S. 19, Teil I) haben wir bereits die folgenden Identitäten kennen gelernt:

$$\begin{aligned} \binom{n}{k} &= \binom{n}{n-k}, \\ \binom{n}{k} &= \binom{n-1}{k} + \binom{n-1}{k-1}. \end{aligned}$$

Da die Summe der Wahrscheinlichkeiten  $h(s; n, N, S)$  aus Abschnitt XI.1.4 (S. 12) gleich 1 ist, ergibt sich schließlich

$$\sum_{s=0}^n \binom{S}{s} \binom{W}{n-s} = \binom{S+W}{n} \quad \text{mit } n \leq S+W.$$

## XI.2. Bedingte Wahrscheinlichkeit und Unabhängigkeit

**XI.2.1. Definition bedingter Wahrscheinlichkeiten.** Häufig steht, bevor das Ergebnis eines Zufallsexperimentes bekannt ist, schon die Information zur Verfügung, dass das Ergebnis zu einer gewissen Teilmenge des Ereignisraumes gehört. Z.B. sieht ein Spieler beim Skat seine eigenen Karten. Interessiert sich Spieler 1 für das Ereignis  $A$ , dass Spieler 2 zwei Asse hat, so wird er zunächst seine eigenen Asse zählen. Hat er selbst drei oder vier Asse, so ist für ihn die Wahrscheinlichkeit von  $A$  natürlich 0; hat er maximal zwei Asse, so ist die Wahrscheinlichkeit von  $A$  für ihn positiv.

Für Laplace-Experimente ist der Ansatz für die Definition bedingter Wahrscheinlichkeiten sehr nahe liegend. Waren ursprünglich alle Ergebnisse  $\omega \in \Omega$  gleich wahrscheinlich und erhält man nun die Information, dass  $\omega \in B$  liegt, so ordnen wir den Ergebnissen in  $\Omega \setminus B$  die bedingte Wahrscheinlichkeit 0 zu und betrachten die Ergebnisse aus  $B$  als gleich wahrscheinlich unter der bedingten Wahrscheinlichkeit. Dies bedeutet, dass für  $A \subset \Omega$  die bedingte Wahrscheinlichkeit von  $A$  bei gegebenem  $B$  den Wert

$$P(A|B) = \frac{\text{card}(A \cap B)}{\text{card}(B)}$$

erhält. Aus

$$P(A \cap B) = \frac{\text{card}(A \cap B)}{\text{card}(\Omega)}$$

und

$$P(B) = \frac{\text{card}(B)}{\text{card}(\Omega)}$$

ergibt sich daher in diesem Fall

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Wir definieren nun auch für beliebige Wahrscheinlichkeitsräume  $(\Omega, P)$  und für beliebige Ereignisse  $B$  mit  $P(B) > 0$  die **BEDINGTE WAHRSCHEINLICHKEIT**  $P(A|B)$  von  $A$  bei gegebenem  $B$  durch diese Formel.

**BEISPIEL XI.2.1.** Aus einer Urne, die zwei schwarze und drei weiße Kugeln enthält, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Wir interessieren uns für die bedingte Wahrscheinlichkeit des Ereignisses  $A$ , dass die zweite gezogene Kugel schwarz ist, bei gegebenem Ereignis  $B$ , dass die erste gezogene Kugel weiß ist. Da nach dem ersten Ziehen drei von vier Kugeln schwarz sind, sollte diese Wahrscheinlichkeit  $\frac{3}{4}$  betragen. Um die Richtigkeit unserer Definition nachzuprüfen, nummerieren wir die weißen Kugeln mit 1 und 2 und die schwarzen Kugeln mit 3, 4 und 5. Dann haben die interessierenden Ereignisse die Form

$$A \cap B = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5)\},$$

$$B = \{(1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 3), (2, 4), (2, 5)\}.$$

Es ergibt sich

$$P(A|B) = \frac{\text{card}(A \cap B)}{\text{card}(B)} = \frac{6}{8} = \frac{3}{4}.$$

**XI.2.2. Eigenschaften.** In der Praxis wird häufig nicht  $P(A|B)$  aus  $P(A \cap B)$  und  $P(B)$  berechnet, sondern umgekehrt  $P(A \cap B)$  aus  $P(A|B)$  und  $P(B)$ . Aus der Definition von  $P(A|B)$  ergibt sich nämlich die **PRODUKTFORMEL FÜR WAHRSCHEINLICHKEITEN**

$$P(A \cap B) = P(B)P(A|B).$$

**BEISPIEL XI.2.2.** In Beispiel XI.2.1 können wir zur Berechnung von  $P(A \cap B)$  wie folgt argumentieren: Da zu Beginn 2 von 5 Kugeln weiß sind, ist  $P(B) = \frac{2}{5}$ . Da nach Eintreten von  $B$  3 von 4 Kugeln schwarz sind, ist  $P(A|B) = \frac{3}{4}$ . Also muss

$$P(A \cap B) = \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{10}$$

sein.

Die obige Produktformel lässt sich wie folgt verallgemeinern: Sind  $A_1, \dots, A_k$  Ereignisse mit  $P(A_1 \cap \dots \cap A_k) > 0$ , so ist

$$P(A_1 \cap \dots \cap A_k) = P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_k|A_1 \cap \dots \cap A_{k-1}).$$

**BEISPIEL XI.2.3.** Wir wollen die Wahrscheinlichkeit bestimmen, dass beim Skat jeder der drei Spieler genau ein Ass hat. Bezeichne dazu mit  $A_i$  das Ereignis, dass Spieler  $i$  genau ein Ass erhält. Aus Symmetriegründen können wir annehmen, dass Spieler 1 die ersten 10 ausgeteilten Karten erhält, Spieler 2 die nächsten 10, dann Spieler 3 zehn, und die letzten 2 Karten in den Stock kommen. Dann ist die gesuchte Wahrscheinlichkeit

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2).$$

Es ist

$$P(A_1) = \frac{\binom{4}{1} \binom{28}{9}}{\binom{32}{10}}.$$

Da nach Austeilen der Karten an Spieler 1 noch 22 Karten incl. 3 Assen zu verteilen sind, ist

$$P(A_2|A_1) = \frac{\binom{3}{1} \binom{19}{9}}{\binom{22}{10}}.$$

Analog ist

$$P(A_3|A_1 \cap A_2) = \frac{\binom{2}{1} \binom{10}{9}}{\binom{12}{10}}.$$

Damit ergibt sich

$$P(A_1 \cap A_2 \cap A_3) = \frac{50}{899} \\ \approx 5.56\%.$$

Folgende Eigenschaften bedingter Wahrscheinlichkeiten sind sehr nützlich:

- (1) Sei  $P(B) > 0$ . Durch  $P_B(A) = P(A|B)$  wird ein Wahrscheinlichkeitsmaß auf  $\Omega$  definiert. Ist  $A \subset \Omega \setminus B$  oder  $P(A) = 0$ , so ist  $P(A|B) = 0$ .
- (2) (FORMEL VON DER TOTALEN WAHRSCHEINLICHKEIT) Sind  $B_1, B_2, \dots$  paarweise disjunkte Teilmengen von  $\Omega$ , deren Vereinigung  $\Omega$  ist, so gilt für jedes Ereignis  $A$

$$P(A) = \sum_k P(B_k)P(A|B_k).$$

- (3) (FORMEL VON BAYES) Ist  $P(A) > 0$  und gelten die Voraussetzungen von (2), so gilt für jedes  $i$

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_k P(B_k)P(A|B_k)}.$$

- (4) Ist  $C$  die Vereinigung der paarweise disjunkten Ereignisse  $C_1, C_2, \dots$  mit  $P(C_i) > 0$  für alle  $i$  und sind die  $P(A|C_i)$  alle gleich, so ist  $P(A|C) = P(A|C_1)$ .

BEISPIEL XI.2.4. Wir greifen Beispiel XI.2.1 (S. 15) auf. Wie groß ist die Wahrscheinlichkeit für das Ereignis  $C$ , dass beide ohne Zurücklegen gezogenen Kugeln die gleiche Farbe haben?

Sei  $B_1$  das Ereignis, dass die erste gezogene Kugel weiß ist, und  $B_2$  das Ereignis, dass sie schwarz ist. Wenn die erste gezogene Kugel weiß ist, ist beim zweiten Zug nur eine von vier Kugeln weiß. Damit ist  $P(C|B_1) = \frac{1}{4}$ . Analog ergibt sich  $P(C|B_2) = \frac{1}{2}$ . Wegen  $P(B_1) = \frac{2}{5}$  und  $P(B_2) = \frac{3}{5}$  ergibt sich mit der Formel von der totalen Wahrscheinlichkeit

$$P(C) = P(B_1)P(C|B_1) + P(B_2)P(C|B_2) \\ = \frac{2}{5} \cdot \frac{1}{4} + \frac{3}{5} \cdot \frac{1}{2} \\ = \frac{4}{10} \\ = \frac{2}{5}.$$

BEISPIEL XI.2.5. Eine Krankheit komme bei ca. 0.5% der Bevölkerung vor. Ein Test zur Erkennung der Krankheit führt bei 99% der Kranken zu einer Reaktion, aber auch bei 2% der Gesunden. Wie groß ist die Wahrscheinlichkeit, dass eine Person, bei der die Reaktion auftritt, tatsächlich krank ist?

Zur Lösung denken wir uns die Bevölkerung von 1 bis  $N$  nummeriert.  $B_1$  sei die Menge der Kranken und  $B_2$  diejenige der Gesunden. Also ist  $\text{card}(B_1) \approx 0.005N$  und  $\text{card}(B_2) \approx 0.995N$ .  $A$  bezeichne die Menge der Personen, bei denen der Test zur Reaktion führt. Dann ist  $\text{card}(A \cap B_1) \approx 0.99 \text{card}(B_1)$  und  $\text{card}(A \cap B_2) \approx 0.02 \text{card}(B_2)$ . Gesucht ist  $P(B_1|A)$ .

Wir wenden die Formel von Bayes an und ordnen jeder Person bei zufälliger Auswahl die Wahrscheinlichkeit  $\frac{1}{N}$  zu. Dann ist

$$\begin{aligned} P(B_1) &= 0.005, \\ P(B_2) &= 0.995, \\ P(A \cap B_1) &= 0.99 \cdot 0.005, \\ P(A \cap B_2) &= 0.02 \cdot 0.995. \end{aligned}$$

Damit ergibt sich mit der Formel von Bayes

$$\begin{aligned} P(B_1|A) &= \frac{P(A \cap B_1)}{P(A)} \\ &= \frac{0.99 \cdot 0.005}{0.99 \cdot 0.005 + 0.02 \cdot 0.995} \\ &= \frac{495}{2485} \\ &\approx 0.2. \end{aligned}$$

Von allen Personen, bei denen der Test eine Reaktion zeigt, sind also nur 20% tatsächlich erkrankt.

**XI.2.3. Unabhängigkeit.** Zwei Ereignisse  $A$  und  $B$  heißen UNABHÄNGIG, wenn  $P(A \cap B) = P(A)P(B)$  ist. Ist  $P(B) > 0$ , ist dies äquivalent zu  $P(A) = P(A|B)$ .

Unabhängigkeit von  $A$  und  $B$  drückt aus, dass  $A$  und  $B$  wahrheitstheoretisch in dem Sinn keinerlei Einfluss aufeinander haben, dass die Information „ $B$  tritt ein“ keinen Einfluss auf die Wahrscheinlichkeit von  $A$  hat. Dies muss man deutlich von der realen Beeinflussung unterscheiden.

BEISPIEL XI.2.6. Wir werfen einen Würfel zweimal. Sei  $A$  das Ereignis, dass die Summe der beiden Augenzahlen gerade ist, und  $B$  das Ereignis, dass die zweite Augenzahl gerade ist. Offensichtlich ist

$$P(A) = P(B) = \frac{1}{2}$$

und

$$P(A \cap B) = \frac{1}{4} = P(A)P(B).$$

Also sind die beiden Ereignisse unabhängig, obwohl  $B$  mitbestimmt, ob  $A$  eintritt.

Wir müssen den Begriff der Unabhängigkeit auf Familien von Ereignissen übertragen. Dazu sagen wir, dass für die endliche Familie  $\{A_i : i \in J\}$  von Ereignissen die **PRODUKTFORMEL** gilt, wenn

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

ist. Eine beliebige Familie  $\{A_i : i \in J\}$  von Ereignissen heißt **UNABHÄNGIG**, wenn für jede endliche Teilfamilie die Produktformel gilt.

Damit ergeben sich folgende Eigenschaften:

- (1) Jede Teilfamilie einer unabhängigen Familie von Ereignissen ist unabhängig. Eine Familie ist genau dann unabhängig, wenn jede endliche Teilfamilie unabhängig ist.
- (2) Ist  $\{A_i : i \in J\}$  eine Familie unabhängiger Ereignisse,  $k$  ein nicht zu  $J$  gehörender Index und  $P(A_k) = 0$  oder  $P(A_k) = 1$ , so ist auch  $\{A_i : i \in J \cup \{k\}\}$  unabhängig.
- (3) Ist  $\{A_i : i \in J\}$  unabhängig und  $B_i$  für jedes  $i$  eines der Ereignisse  $A_i$ ,  $\Omega \setminus A_i$ ,  $\emptyset$  oder  $\Omega$ , so ist  $\{B_i : i \in J\}$  unabhängig.
- (4) Ist  $J = \{1, \dots, n\}$  endlich, so ist  $\{A_i : i \in J\}$  genau dann unabhängig, wenn für jede Wahl von  $B_i \in \{A_i, \Omega \setminus A_i\}$  die Produktformel für  $B_1, \dots, B_n$  gilt.

**XI.2.4. Produktexperimente.** Wir nehmen an, dass wir schon Modelle  $(\Omega_1, P_1), \dots, (\Omega_n, P_n)$  für gewisse Zufallsexperimente kennen, und wollen nun ein Modell für das Experiment konstruieren, das in der unabhängigen Hintereinanderausführung aller dieser Telexperimente besteht. Dazu liegt es nahe, als Ereignisraum das kartesische Produkt

$$\Omega_1 \times \dots \times \Omega_n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \Omega_i, 1 \leq i \leq n\}$$

zu wählen und als Wahrscheinlichkeitsfunktion

$$P(\omega) = P_1(\omega_1) \cdot \dots \cdot P_n(\omega_n).$$

Man bezeichnet  $(\Omega, P)$  als **PRODUKT DER WAHRSCHEINLICHKEITS-RÄUME**  $(\Omega_i, P_i)$  und schreibt

$$\Omega = \Omega_1 \times \dots \times \Omega_n,$$

$$P = P_1 \times \dots \times P_n.$$

Ist  $A_i$  ein Ereignis im  $i$ -ten Experiment, so bezeichnet das kartesische Produkt  $A_1 \times \dots \times A_n$  das Ereignis in  $\Omega$ , dass für alle  $i$  im  $i$ -ten Teilexperiment  $A_i$  eintritt. Die Wahrscheinlichkeit für dieses Ereignis ist

$$P(A_1 \times \dots \times A_n) = P(A_1) \cdot \dots \cdot P(A_n).$$

**XI.2.5. Binomialverteilung.** Wir betrachten ein Experiment, das in der unabhängigen  $n$ -fachen Wiederholung eines Einzelexperimentes mit nur zwei verschiedenen möglichen Ausgängen besteht. Wir bezeichnen diese beiden Ausgänge mit 0 und 1. Dann ist  $\Omega_i = \{0, 1\}$  und  $\Omega = \{0, 1\}^n$ . Da die Teilexperimente Wiederholungen des gleichen Experimentes sind, sollen in allen Teilexperimenten die gleichen Wahrscheinlichkeiten auftreten.  $p = P_i(1)$  soll also nicht von  $i$  abhängen.  $p$  heißt oft **ERFOLGSWAHRSCHEINLICHKEIT**. Natürlich ist dann  $P_i(0) = 1 - p$ . Dann ist im Produktmodell  $P(\omega) = p^k(1 - p)^{n-k}$ , wenn  $k$  die Anzahl der Einsen in  $\omega$  ist. Ein Experiment dieser Form nennt man **BERNOULLI-EXPERIMENT** und  $P$  heißt **BERNOULLI-VERTEILUNG**.

Das Ereignis, dass insgesamt  $k$  Einsen auftreten wird durch

$$E_k = \{\omega \in \Omega : \sum_{i=1}^n \omega_i = k\}$$

beschrieben. Die Zahl der Elemente von  $E_k$  ist gleich der Zahl der Möglichkeiten, die  $k$  Zeitpunkte in  $\{1, \dots, n\}$  festzulegen, an denen die Einsen auftreten, also  $\binom{n}{k}$ . Es folgt

$$P(E_k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Die Terme

$$b_{n,p}(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{mit } 0 \leq k \leq n$$

bestimmen die Wahrscheinlichkeitsverteilung. Sie sind nicht negativ und haben die Summe 1. Man nennt sie **BINOMIALVERTEILUNG** mit Parametern  $n$  und  $p$  oder kurz  $b_{n,p}$ -Verteilung.

**BEISPIEL XI.2.7.** Wie groß ist die Wahrscheinlichkeit, bei 10 Würfeln eines Würfels 3 Sechsen zu erhalten?

Wir können die geworfenen Sechsen als Erfolge auffassen und ihnen den Wert 1 zuordnen; die anderen Ergebnisse sind Misserfolge und erhalten den Wert 0. Dann ist  $p = \frac{1}{6}$ . Die gesuchte Wahrscheinlichkeit ist

$$\binom{10}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^7 \approx 15.5\%.$$

BEISPIEL XI.2.8. Die Wahrscheinlichkeit für die Geburt eines Jungen sei  $p = 0.51$ . Aufeinanderfolgende Geburten seien unabhängig. Dann ist die Wahrscheinlichkeit, dass in einer Familie mit 4 Kindern 2 Jungen und 2 Mädchen vorkommen, gleich

$$\binom{4}{2} 0.51^2 0.49^2 \approx 37.47\%.$$

**XI.2.6. Multinomialverteilung.** Hier hat man wieder  $n$  unabhängige, identische Teilversuche, aber jeder Teilversuch hat nun  $r$  mögliche Ausgänge. Die Teilerperimente sind beschreibbar durch  $P_i(j) = p_j$ ,  $j = 1, \dots, r$ , wobei die Wahrscheinlichkeiten  $p_1, \dots, p_r$  beliebig vorgegeben sind mit  $p_j \geq 0$ ,  $j = 1, \dots, r$ , und  $p_1 + \dots + p_r = 1$ . Die Wahrscheinlichkeit, in den  $n$  Teilversuchen  $k_1$ -mal das Ergebnis 1,  $k_2$ -mal das Ergebnis 2 usw. zu erreichen, ist

$$\binom{n}{k_1, \dots, k_r} p_1^{k_1} \cdot \dots \cdot p_r^{k_r}.$$

**XI.2.7. Geometrische Verteilung.** Die Wahrscheinlichkeit, dass in einem Bernoulli-Experiment mit Erfolgswahrscheinlichkeit  $p$  der erste Erfolg genau im  $k$ -ten Versuch eintritt, ist  $p(1-p)^{k-1}$ ,  $k \geq 1$ . Man  $k$  als Ergebnis eines Experimentes auffassen, das darin besteht, zu beobachten, wann in einer Folge von Bernoulli-Experimenten der erste Erfolg eintritt. Die zugehörige Ereignismenge ist dann  $\Omega = \mathbb{N} \setminus \{0\}$ . Die zugehörige Wahrscheinlichkeitsfunktion ist  $P(k) = p(1-p)^{k-1}$ . Die hierdurch definierte Verteilung heißt GEOMETRISCHE VERTEILUNG. Man beachte, dass im Gegensatz zu allen bisherigen Beispielen der Ereignisraum jetzt nicht mehr endlich ist.

**XI.2.8. Negative Binomialverteilung.** Dies ist eine Verallgemeinerung der geometrischen Verteilung. Sei  $f(k; r, p)$  die Wahrscheinlichkeit, dass bei  $n \geq r + k$  Bernoulli-Experimenten mit Erfolgswahrscheinlichkeit  $p$  genau  $k$  Misserfolge dem  $r$ -ten Erfolg vorangehen. Man kann nachrechnen, dass

$$f(k; r, p) = \binom{k+r-1}{k} p^r (1-p)^k$$

ist. Die hierdurch bei festem  $r$  auf  $\Omega = \mathbb{N}$  definierte Verteilung nennt man NEGATIVE BINOMIALVERTEILUNG oder PASCAL-VERTEILUNG.

### XI.3. Zufallsvariable, Erwartungswert, Varianz

**XI.3.1. Zufallsvariable.** Sei  $(\Omega, P)$  ein endlicher Wahrscheinlichkeitsraum und  $\mathcal{X}$  eine beliebige Menge. Eine Abbildung  $X : \Omega \rightarrow \mathcal{X}$  heißt ( $\mathcal{X}$ -wertige) ZUFALLSVARIABLE.

BEMERKUNG XI.3.1. Da  $\Omega$  endlich ist, ist die Bildmenge  $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$  einer Zufallsvariablen  $X$  auf  $\Omega$  stets endlich.

BEISPIEL XI.3.2. Vor einer Wahl werden zufällig ausgewählte Bürger gefragt, welche Partei sie wählen wollen. Hier ist  $\Omega$  die Menge der Bürger,  $P$  die Gleichverteilung auf  $\Omega$ ,  $\mathcal{X}$  die Menge der wählbaren Parteien vereinigt mit der „Partei“ der Nichtwähler und  $X(\omega)$  die Partei, die Bürger  $\omega$  angeblich wählen will.

BEISPIEL XI.3.3. Wir interessieren uns für die Summe der Augenzahlen beim Würfeln mit zwei Würfeln. Hier ist  $\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$ ,  $P$  die Gleichverteilung auf  $\Omega$ ,  $\mathcal{X} = \{2, \dots, 12\}$  und  $X((i, j)) = i + j$ .

An Zufallsvariablen interessiert uns vor allem ihre Verteilungsfunktion. Diese gibt an, wie wahrscheinlich die einzelnen Ergebnisse von  $X$  sind. Genauer versteht man unter der VERTEILUNGSFUNKTION von  $X : \Omega \rightarrow \mathcal{X}$  das Wahrscheinlichkeitsmaß  $P_X$  auf  $\mathcal{X}_X = X(\Omega)$ , das definiert ist durch

$$P_X(x) = P(\{\omega \in \Omega : X(\omega) = x\}) \quad \text{mit } x \in \mathcal{X}_X.$$

$P_X(x)$  ist also die Wahrscheinlichkeit, dass  $X$  den Wert  $x$  annimmt.

Die Verteilungsfunktion einer Zufallsvariablen  $X$  kann durch ein STABDIAGRAMM veranschaulicht werden. Dazu erstellt man zunächst eine Liste der möglichen Werte  $x_k \in \mathcal{X}$  von  $X$ , berechnet für jedes  $x_k$  die Wahrscheinlichkeit  $P_X(x_k)$  und zeichnet dann senkrecht über den Punkten  $x_k$  der  $x$ -Achse Striche der Länge  $P_X(x_k)$ .

BEISPIEL XI.3.4. Die Verteilungsfunktion der Zufallsvariablen aus Beispiel XI.3.3 ist durch Tabelle XI.3.1 bestimmt. Das zugehörige Stabdiagramm ist in Abbildung XI.3.1 wiedergegeben.

TABELLE XI.3.1. Verteilungsfunktion der Zufallsvariablen aus Beispiel XI.3.3

$x_k$	2	3	4	5	6	7	8	9	10	11	12
$P_X(x_k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

### XI.3.2. Gemeinsame Verteilung mehrerer Zufallsvariabler.

Sind auf einem Wahrscheinlichkeitsraum  $(\Omega, P)$  mehrere Zufallsvariable  $X_1, \dots, X_n$  mit evtl. verschiedenen Wertebereichen  $\mathcal{X}_1, \dots, \mathcal{X}_n$  definiert, kann man sie zu einer einzigen Zufallsvariablen  $X$  mit Wertebereich

$$\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$$

zusammenfassen, indem man

$$X(\omega) = (X_1(\omega), \dots, X_n(\omega))$$

setzt. Die Verteilungsfunktion von  $X$  nennt man die GEMEINSAME VERTEILUNGSFUNKTION von  $X_1, \dots, X_n$ .

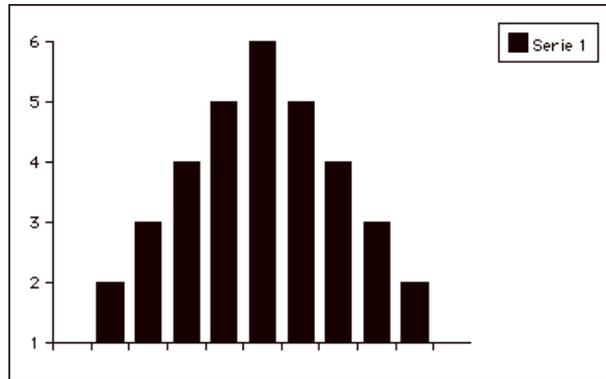


ABBILDUNG XI.3.1. Stabdiagramm der Zufallsvariablen aus Beispiel XI.3.3

BEISPIEL XI.3.5. Wir betrachten die Bernoulli-Verteilung  $P$  zu  $p \in (0, 1)$  auf  $\Omega = \{0, 1\}^n$  (vgl. Abschnitt XI.2.5 (S. 20)). Für  $\omega \in \Omega$  sei

$$S(\omega) = \omega_1 + \dots + \omega_n$$

die Zahl der Erfolge. Für  $\omega$  mit  $S(\omega) \geq 1$  setzen wir

$$N(\omega) = \min\{j \geq 1 : \omega_j = 1\}.$$

Dies ist die Wartezeit bis zum ersten Erfolg. Ist  $S(\omega) = 0$ , setzen wir  $N(\omega) = n + 1$ .  $S$  und  $N$  sind zwei Zufallsvariable auf  $(\Omega, P)$ . Ihre gemeinsame Verteilung wird beschrieben durch die Angabe der Wahrscheinlichkeiten

$$p(k, h) = P(S = k, N = h)$$

mit  $0 \leq k \leq n$  und  $1 \leq h \leq n + 1$ . Wir wollen diese Verteilung bestimmen.

Offenbar ist

$$p(0, n + 1) = (1 - p)^n$$

und

$$p(0, h) = 0$$

für  $1 \leq h \leq n$ . Ist  $S(\omega) = k \geq 1$  und  $N(\omega) = h$ , so muss gelten  $\omega_i = 0$  für  $i < h$  und  $\omega_h = 1$ , und es müssen genau  $k - 1$  Einsen unter  $\omega_{h+1}, \dots, \omega_n$  vorkommen. Es gibt  $\binom{n-h}{k-1}$  solche Elemente, und jedes hat die Wahrscheinlichkeit  $p^k(1-p)^{n-k}$ . Also ist

$$p(k, h) = \binom{n-h}{k-1} p^k (1-p)^{n-k}$$

für  $k \geq 1$ . Ist  $k - 1 > n - h$ , so gibt es kein solches  $\omega$ , und es ist  $\binom{n-h}{k-1} = 0$ . Die obige Formel bleibt also richtig.

**XI.3.3. Unabhängigkeit.** Sei  $(\Omega, P)$  ein endlicher Wahrscheinlichkeitsraum. Eine Familie  $X_i, i \in I$ , von Zufallsvariablen  $X_i : \Omega \rightarrow \mathcal{X}_i$  heißt UNABHÄNGIG, wenn für jede Wahl von  $A_i \subset \mathcal{X}_i, i \in I$ , die Ereignisse  $\{X_i \in A_i\}, i \in I$ , unabhängig sind.

Sei nun  $X_1, \dots, X_n$  eine unabhängige Familie von Zufallsvariablen. Dann gilt gemäß Abschnitt XI.2.3 (S. 18) speziell für alle  $(x_1, \dots, x_n) \in \mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i).$$

Die gemeinsame Verteilung von  $X_1, \dots, X_n$  ist also das Produkt der Verteilungen der  $X_i$ .

Die Unabhängigkeit von Zufallsvariablen bleibt unter der Komposition von Abbildungen erhalten: Sind  $X_i, i \in I$ , unabhängige Zufallsvariable und  $F_i : \mathcal{X}_i \rightarrow \mathcal{Y}_i$  beliebige Abbildungen mit beliebigen Wertebereichen, so sind die Zufallsvariablen  $F_i \circ X_i, i \in I$ , auch unabhängig.

**XI.3.4. Erwartungswert.** Sei  $(\Omega, P)$  ein endlicher Wahrscheinlichkeitsraum und  $X$  eine reellwertige Zufallsvariable auf  $\Omega$ . Dann ist der ERWARTUNGSWERT von  $X$  definiert durch

$$EX = E(X) = \sum_{\omega \in \Omega} X(\omega)P(\omega).$$

BEISPIEL XI.3.6. Für die Zufallsvariable „Summe der Augenzahlen beim Werfen von zwei Würfeln“ aus Beispiel XI.3.4 erhalten wir den Erwartungswert

$$\begin{aligned} EX &= \frac{1}{36} \{1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + 4 \cdot 5 + 5 \cdot 6 + 6 \cdot 7 \\ &\quad + 5 \cdot 8 + 4 \cdot 9 + 3 \cdot 10 + 2 \cdot 11 + 1 \cdot 12\} \\ &= \frac{252}{36} \\ &= 7. \end{aligned}$$

Sind  $X, Y$  reellwertige Zufallsvariable auf  $\Omega$  und ist  $\lambda \in \mathbb{R}$ , so gilt

$$\begin{aligned} E(\lambda X) &= \lambda E(X) \\ E(X + Y) &= E(X) + E(Y) \\ X, Y \text{ unabhängig} &\implies E(XY) = E(X)E(Y). \end{aligned}$$

Die ersten beiden Eigenschaften folgen direkt aus der Definition des Erwartungswertes. Zur Nachweis der dritten Eigenschaft bezeichnen wir mit  $x_i, i \in I$ , und  $y_j, j \in J$ , die (endlich vielen!) möglichen Werte von  $X$  bzw.  $Y$ . Dann folgt mit dem vorigen Abschnitt

$$\begin{aligned}
 E(XY) &= \sum_{\omega \in \Omega} X(\omega)Y(\omega)P(\omega) \\
 &= \sum_{i \in I} \sum_{j \in J} x_i y_j \underbrace{P(X = x_i, Y = y_j)}_{P(X=x_i)P(Y=y_j)} \\
 &= \sum_{i \in I} \sum_{j \in J} x_i y_j P(X = x_i)P(Y = y_j) \\
 &= \left\{ \sum_{i \in I} x_i P(X = x_i) \right\} \left\{ \sum_{j \in J} y_j P(Y = y_j) \right\} \\
 &= E(X)E(Y).
 \end{aligned}$$

BEISPIEL XI.3.7 (ERWARTUNGSWERT DER BINOMIALVERTEILUNG). Für die Zufallsvariable  $S$  aus Beispiel XI.3.5 erhalten wir

$$\begin{aligned}
 E(S) &= \sum_{k=0}^n k \underbrace{P(S = k)}_{= \binom{n}{k} p^k (1-p)^{n-k}} \\
 &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n k \underbrace{\binom{n}{k}}_{= n \binom{n-1}{k-1}} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n n \binom{n-1}{k-1} p p^{k-1} (1-p)^{(n-1)-(k-1)} \\
 &= np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell} \\
 &= np.
 \end{aligned}$$

Das gleiche Ergebnis erhalten wir einfacher mit folgendem „Trick“, der auch in anderen Situationen hilfreich ist: Definiere die Zufallsvariablen  $S_1, \dots, S_n$  durch die Vorschrift  $S_i(\omega) = 1$ , falls das  $i$ -te Telexperiment ein Erfolg ist,  $S_i(\omega) = 0$  sonst. Dann ist  $S = S_1 + \dots + S_n$ . Für jedes  $i$  ist aber offensichtlich  $E(S_i) = p$ . Daher ist

$$E(S) = \sum_{i=1}^n E(S_i) = np.$$

**BEISPIEL XI.3.8 (ERWARTUNGSWERT DER HYPERGEOMETRISCHEN VERTEILUNG).** In einer Urne befinden sich  $S$  schwarze und  $W$  weiße Kugeln, insgesamt  $N = S + W$  Kugeln. Wir ziehen daraus  $n \leq N$  Kugeln ohne Zurücklegen (vgl. Abschnitt XI.1.4 (S. 12)). Die Zufallsvariable  $X$  sei die Zahl der dabei gezogenen schwarzen Kugeln. Um  $E(X)$  zu berechnen, bezeichnen wir mit  $X_i$  die Zufallsvariable, die den Wert 1 liefert, wenn im  $i$ -ten Zug eine schwarze Kugel gezogen wird, und die sonst den Wert 0 gibt. Dann ist  $X = X_1 + \dots + X_n$ , und für jedes  $i$  gilt  $E(X_i) = \frac{S}{N}$ . Damit folgt

$$E(X) = n \frac{S}{N}.$$

**XI.3.5. Varianz und Kovarianz.** Seien  $X$  und  $Y$  zwei reellwertige Zufallsvariable auf einem endlichen Wahrscheinlichkeitsraum  $(\Omega, P)$ . Die VARIANZ von  $X$  ist definiert durch

$$\text{Var}(X) = E((X - E(X))^2).$$

Die Größe

$$\sigma_X = \sqrt{\text{Var}(X)}$$

heißt die STREUUNG oder STANDARDABWEICHUNG von  $X$ .

Die KOVARIANZ von  $X$  und  $Y$  ist definiert durch

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))).$$

Die Größe

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

heißt der KORRELATIONSKOEFFIZIENT von  $X$  und  $Y$ . Die Zufallsvariablen  $X$  und  $Y$  heißen UNKORRELIERT, wenn  $\text{Cov}(X, Y) = 0$  ist.

Die Varianz ist ein Maß für die Streuung einer Zufallsvariablen um den Erwartungswert  $EX$ , der ein Maß für den Mittelwert ist. Ist  $EX = 0$  und haben die Werte  $x_1, \dots, x_n$  von  $X$  alle die gleiche Wahrscheinlichkeit  $\frac{1}{n}$ , so ist

$$\sigma_X = \sqrt{\frac{1}{n} \sum x_i^2}.$$

Bis auf den Skalierungsfaktor  $\frac{1}{\sqrt{n}}$  ist also  $\sigma_X$  der Euklidische Abstand des Punktes  $(x_1, \dots, x_n)$  vom Ursprung.

Eine positive Kovarianz von  $X$  und  $Y$  bedeutet, dass  $X$  die Tendenz hat, dort groß zu sein, wo auch  $Y$  groß ist.

Sind  $X, Y, X_i$  reellwertige Zufallsvariable und  $a, b, c, d$  reelle Zahlen, so gelten folgende Eigenschaften:

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ \text{Var}(aX + b) &= a^2 \text{Var}(X) \\ \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ \text{Cov}(aX + b, cY + d) &= ac \text{Cov}(X, Y) \\ \text{Cov}(X, Y) &= \text{Cov}(Y, X) \\ |\text{Cov}(X, Y)| &\leq \sigma_X \sigma_Y \\ \text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ X, Y \text{ unabhängig} &\implies \text{Cov}(X, Y) = 0 \\ X_1, \dots, X_n \text{ unabhängig} &\implies \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) \end{aligned}$$

Zum Nachweis dieser Eigenschaften beginnen wir mit der dritten. Da der Erwartungswert einer konstanten Zufallsvariablen diese Konstante ist, erhalten wir

$$\begin{aligned} \text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY - XE(Y) - YE(X) + E(X)E(Y)) \\ &= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Die erste Eigenschaft ist der Spezialfall  $X = Y$  der dritten Eigenschaft. Aus der dritten Eigenschaft folgt  $\text{Cov}(X, Y) = 0$ , wenn eine der Zufallsvariablen  $X$  oder  $Y$  konstant ist. Aus der Definition der Kovarianz folgt zudem, dass die Abbildung  $X, Y \mapsto \text{Cov}(X, Y)$  bilinear ist. Damit folgt

$$\begin{aligned} \text{Cov}(aX + b, cY + d) &= \text{Cov}(aX, cY) + \text{Cov}(aX, d) \\ &\quad + \text{Cov}(b, cY) + \text{Cov}(b, d) \\ &= ac \text{Cov}(X, Y). \end{aligned}$$

Die zweite Eigenschaft ist der Spezialfall  $X = Y, a = c$  der soeben bewiesenen vierten Eigenschaft.

Die fünfte Eigenschaft folgt direkt aus der Definition.

Die sechste Eigenschaft ist eine Konsequenz der Cauchy-Schwarzschen

Ungleichung (vgl. Abschnitt II.6.6 (S. 103, Teil I)).

Zum Nachweis der siebten Eigenschaft setze  $\bar{X}_i = X_i - E(X_i)$ . Dann folgt

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= E\left(\left(\sum_{i=1}^n \bar{X}_i\right)^2\right) \\ &= E\left(\sum_{i=1}^n \bar{X}_i^2 + \sum_{i \neq j} \bar{X}_i \bar{X}_j\right) \\ &= \sum_{i=1}^n E(\bar{X}_i^2) + \sum_{i \neq j} E(\bar{X}_i \bar{X}_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j). \end{aligned}$$

Sind  $X$  und  $Y$  unabhängig, so auch  $X - E(X)$  und  $Y - E(Y)$ . Damit folgt die achte Eigenschaft:

$$\begin{aligned} \text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(X - E(X))E(Y - E(Y)) \quad (\text{wg. Unabhängigkeit}) \\ &= 0. \end{aligned}$$

Die letzte Eigenschaft schließlich folgt aus der siebten und achten Identität.

**BEISPIEL XI.3.9 (VARIANZ DER BINOMIALVERTEILUNG).** Wir betrachten die Zufallsvariable  $S$  aus Beispiel XI.3.5 (S. 23) und definieren die Zufallsvariablen  $S_i$  wie in Beispiel XI.3.7 (S. 25). Da sich  $S_i$  nur auf das  $i$ -te Telexperiment bezieht und die Telexperimente unabhängig sind, sind die  $S_i$  unabhängig. Daher ist

$$\begin{aligned} \text{Var}(S) &= \text{Var}(S_1 + \dots + S_n) \\ &= \text{Var}(S_1) + \dots + \text{Var}(S_n) \\ &= n \text{Var}(S_1), \end{aligned}$$

da jedes  $S_i$  die gleiche Varianz hat. Weiter ist

$$\begin{aligned} \text{Var}(S_1) &= E(S_1^2) - E(S_1)^2 \\ &= p - p^2 \\ &= p(1 - p). \end{aligned}$$

Insgesamt folgt

$$\text{Var}(S) = np(1 - p).$$

BEISPIEL XI.3.10 (VARIANZ DER HYPERGEOMETRISCHEN VERTEILUNG). Wir betrachten die Zufallsvariablen  $X$  und  $X_1, \dots, X_n$  aus Beispiel XI.3.8 (S. 26). Setze  $p = \frac{S}{N}$ . Alle  $X_i$  haben die gleiche Varianz:

$$\begin{aligned}\text{Var}(X_i) &= \text{Var}(X_1) \\ &= E(X_1^2) - E(X_1)^2 \\ &= p - p^2 \\ &= p(1 - p).\end{aligned}$$

Ebenso gilt aus Symmetriegründen für alle  $i \neq j$

$$\begin{aligned}\text{Cov}(X_i, X_j) &= \text{Cov}(X_1, X_2) \\ &= E(X_1 X_2) - E(X_1)E(X_2).\end{aligned}$$

Weiter ist

$$\begin{aligned}E(X_1 X_2) &= P(\{X_1 = 1\} \cap \{X_2 = 1\}) \\ &= P(\{X_1 = 1\})P(\{X_2 = 1\} | \{X_1 = 1\}) \\ &= \frac{S}{N} \frac{S-1}{N-1} \\ &= p \frac{S-1}{N-1}.\end{aligned}$$

Damit folgt

$$\begin{aligned}\text{Cov}(X_i, X_j) &= p \frac{S-1}{N-1} - p^2 \\ &= p \left( \frac{S-1}{N-1} - p \right)\end{aligned}$$

für  $i \neq j$  und

$$\begin{aligned}\text{Var}(X) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= np(1-p) + (n^2 - n)p \underbrace{\left( \frac{S-1}{N-1} - p \right)}_{= \frac{S-1}{N-1} - \frac{S}{N} = \frac{S-N}{N(N-1)}} \\ &= np(1-p) + n(n-1)p \underbrace{\frac{S-N}{N(N-1)}}_{= \frac{1}{N-1}(p-1)} \\ &= np(1-p) \left( 1 - \frac{n-1}{N-1} \right) \\ &= np(1-p) \frac{N-n}{N-1}.\end{aligned}$$

Tabelle XI.3.2 fasst die Ergebnisse der vorigen Beispiele zusammen.

TABELLE XI.3.2. Erwartungswert und Varianz einiger Verteilungen

Verteilung	Erwartungswert	Varianz
binomial	$np$	$np(1-p)$
hypergeometrisch $h(s; n, N, S), p = \frac{S}{N}$	$np$	$np(1-p)\frac{N-n}{N-1}$

**XI.3.6. Das schwache Gesetz der großen Zahlen.** Seien  $X_1, \dots, X_n$  unabhängige, reellwertige Zufallsvariable mit gleichem Erwartungswert und  $\text{Var}(X_i) \leq M$  für alle  $i$ . Dann gilt für alle  $\varepsilon > 0$ :

SCHWACHES GESETZ DER GROSSEN ZAHLEN:

$$P\left(\left|\frac{1}{n}(X_1 + \dots + X_n) - E(X_1)\right| \geq \varepsilon\right) \leq \frac{M}{\varepsilon^2 n} \xrightarrow{n \rightarrow \infty} 0.$$

Es besagt, dass bei immer häufigerer, unabhängiger Wiederholung eines Experimentes das Ergebnis im Mittel fast sicher gleich dem Erwartungswert ist.

BEISPIEL XI.3.11. Für die Zufallsvariablen  $S$  und  $S_1, \dots, S_n$  des Bernoulli-Experimentes aus den Beispielen XI.3.5 (S. 23) und XI.3.9 haben wir für alle  $i$

$$E(S_i) = p,$$

$$\text{Var}(S_i) = p(1-p) \leq \frac{1}{4}.$$

$h_n = \frac{1}{n}S$  ist die relative Häufigkeit, mit der ein Erfolg eintritt. Die Wahrscheinlichkeit, dass diese relative Häufigkeit um mehr als  $\varepsilon$  von dem Erwartungswert  $p$  abweicht, ist

$$P(|h_n - p| \geq \varepsilon) \leq \frac{1}{4\varepsilon^2 n}.$$

BEISPIEL XI.3.12. Wir betrachten ein für uns vorteilhaftes Spiel, bei dem wir auf Dauer fast sicher verlieren.

Die Spielregeln lauten:

- Der Anfangseinsatz beträgt 1 Euro.
- Es wird immer wieder eine Münze geworfen. Fällt Zahl geht die Hälfte des aktuellen Einsatzes verloren, fällt Kopf gewinnen wir  $\frac{2}{3}$  des aktuellen Einsatzes hinzu.

Da der Gewinn höher ist als der Verlust, ist dieses Spiel vorteilhaft für uns. Dennoch werden wir auf Dauer fast sicher verlieren.

Um dies einzusehen, bezeichnen wir mit  $X_{n-1}$  den Einsatz vor dem  $n$ -ten Wurf,  $n \geq 1$ . Dann ist  $X_0 = 1$  Euro und  $X_n = \frac{1}{2}X_{n-1}$ , wenn im  $n$ -ten Wurf Zahl kommt, und  $X_n = \frac{5}{3}X_{n-1}$ , wenn im  $n$ -ten Wurf Kopf fällt. Setze  $Y_n = \frac{1}{2}$ , falls im  $n$ -ten Wurf Zahl kommt, und  $Y_n = \frac{5}{3}$ ,

falls im  $n$ -ten Wurf Kopf fällt. Dann ist  $X_n = Y_1 \cdot \dots \cdot Y_n$ . Die  $Y_i$  sind unabhängig. Es ist

$$\begin{aligned} E(Y_i) &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{5}{3} \\ &= \frac{13}{12} \\ &> 1. \end{aligned}$$

Daher ist

$$\begin{aligned} E(X_n) &= \prod_{i=1}^n E(Y_i) \\ &= \left(\frac{13}{12}\right)^n \\ &\xrightarrow{n \rightarrow \infty} \infty. \end{aligned}$$

Setze

$$\mu = E(\log Y_i).$$

Dann ist

$$\begin{aligned} \mu &= \frac{1}{2} \log\left(\frac{1}{2}\right) + \frac{1}{2} \log\left(\frac{5}{3}\right) \\ &< \frac{1}{2} \log\left(\frac{1}{2}\right) + \frac{1}{2} \log(2) \\ &= 0. \end{aligned}$$

Mit  $\varepsilon = \frac{1}{2}|\mu| = -\frac{1}{2}\mu$  gilt nach dem schwachen Gesetz der großen Zahlen

$$P\left(\left|\frac{1}{n}(\log(Y_1) + \dots + \log(Y_n)) - \mu\right| \leq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1.$$

Wegen

$$\log(X_n) = \log(Y_1) + \dots + \log(Y_n)$$

und

$$\left\{\left|\frac{1}{n} \log(X_n) - \mu\right| \leq -\frac{\mu}{2}\right\} \subset \left\{\frac{1}{n} \log(X_n) \leq \frac{\mu}{2}\right\}$$

folgt

$$P\left(\frac{1}{n} \log(X_n) \leq \frac{\mu}{2}\right) \xrightarrow{n \rightarrow \infty} 1.$$

Für große  $n$  gilt also mit einer Wahrscheinlichkeit nahe bei 1

$$X_n \leq e^{n\frac{\mu}{2}}.$$

Wegen  $\mu < 0$  strebt dieser Ausdruck exponentiell schnell gegen Null. Dieses scheinbar paradoxe Verhalten liegt daran, dass man mit sehr kleiner Wahrscheinlichkeit sehr große Gewinne machen kann.

Wir können uns das Ergebnis heuristisch auch leicht wie folgt klar

machen: Nach  $2n$  Spielrunden sollten wir etwa  $n$ -mal gewonnen und  $n$ -mal verloren haben. Dann beträgt unser Kapital

$$X_{2n} \approx \left(\frac{1}{2}\right)^n \left(\frac{5}{3}\right)^n = \left(\frac{5}{6}\right)^n.$$

Wegen  $\frac{5}{6} < 1$  strebt dieser Ausdruck für  $n \rightarrow \infty$  aber gegen Null.

#### XI.4. Grundbegriffe der Schätztheorie

**XI.4.1. Motivation.** Ein Teich enthält eine unbekannt Zahl  $N$  von Fischen, die geschätzt werden soll. Um dies zu erreichen, fangen wir  $W$  Fische, markieren sie mit einem weißen Fleck und setzen sie wieder aus. Wir warten eine Weile und fangen nun  $n$  Fische. Sei  $x$  die Zahl der markierten Fische in diesem Fang. Eine von dieser Zahl  $x$  abhängige Schätzung  $\hat{N}(x)$  für  $N$  erhalten wir mit folgender Plausibilitätsbetrachtung: Wenn  $x$  nicht zu klein ist, sollte der Anteil  $\frac{x}{n}$  der markierten Fische im zweiten Fang etwa dem Anteil  $\frac{W}{N}$  der markierten Fische am Gesamtbestand entsprechen, d.h.  $\frac{x}{n} \approx \frac{W}{N}$ . Da wir  $x$ ,  $n$  und  $W$  kennen, können wir diese Näherungsgleichung nach  $N$  auflösen und erhalten die Schätzung

$$N \approx \hat{N}(x) = \left\lceil \frac{Wn}{x} \right\rceil.$$

Dabei bezeichnet  $\lceil a \rceil$  die größte ganze Zahl  $\leq a$ .

Wir können diese Plausibilitätsbetrachtung auf ein etwas sichereres mathematisches Fundament stellen. Dazu beschreiben wir den zweiten Fang durch das Modell des Ziehens ohne Rücklegen von  $n$  Kugeln aus einer Urne mit  $W$  weißen und  $N - W$  schwarzen Kugeln. Die Wahrscheinlichkeit, dass  $x$  weiße Kugeln gezogen werden, ist dann

$$P_N(x) = \frac{\binom{W}{x} \binom{N-W}{n-x}}{\binom{N}{n}} \quad \text{mit } 0 \leq x \leq n.$$

Man beachte, dass die Abhängigkeit von  $N$  hier eine ganz andere ist als bisher:  $N$  ist nicht ein Ereignis, sondern ein unbekannter, zu bestimmender Parameter.

Der MAXIMUM-LIKELIHOOD ANSATZ zur Schätzung von  $N$  besagt, dass man bei gegebenem  $W$  und  $x$  den Wert von  $N$  so bestimmt, dass  $P_N(x)$  maximal wird. Um  $N$  so zu bestimmen, betrachten wir

$$\begin{aligned} \frac{P_N(x)}{P_{N-1}(x)} &= \frac{\binom{W}{x} \binom{N-W}{n-x} \binom{N-1}{n}}{\binom{N}{n} \binom{W}{x} \binom{N-1-W}{n-x}} \\ &= \frac{(N-W)(N-n)}{N(N-W-n+x)}. \end{aligned}$$

Offensichtlich gilt

$$P_N(x) > P_{N-1}(x)$$

$$\begin{aligned} \iff (N - W)(N - n) &> N(N - W - n + x) \\ \iff Wn &> Nx. \end{aligned}$$

Analoge Äquivalenzen gelten auch für die Beziehungen  $<$  und  $=$ . Daher liefert der Maximum-Likelihood Ansatz die gleiche Schätzung wie unser Plausibilitätsargument.

**XI.4.2. Der allgemeine Rahmen von Schätzproblemen.** Zur Beschreibung eines Schätzproblems mit endlichem Stichprobenraum benötigen wir:

- eine nicht leere, endliche Menge  $\mathcal{X}$ , den STICHPROBENRAUM,
- eine Familie  $\{P_\vartheta : \vartheta \in \Theta\}$  von Wahrscheinlichkeitsmaßen auf  $\mathcal{X}$  und
- eine zu schätzende Funktion  $g(\vartheta)$ .

$\mathcal{X}$  ist die Menge der möglichen Beobachtungsergebnisse. Durch die unterschiedliche Notation zu den vorigen Abschnitten –  $\mathcal{X}$  statt  $\Omega$  – soll betont werden, dass jedes  $x \in \mathcal{X}$  tatsächlich beobachtbar sein muss. Bei wahrscheinlichkeitstheoretischen Modellen treten dagegen auch Stichproben- bzw. Ereignisräume  $\Omega$  auf, deren Elemente  $\omega$  nicht beobachtbar sind.

BEISPIEL XI.4.1. Im Rahmen des vorigen Abschnittes ist  $\mathcal{X} = \{1, 2, \dots, n\}$ , wenn  $n$  die Zahl der Fische im zweiten Fang ist. Weiter ist  $\vartheta = N$  die unbekannte Zahl der Fische im Teich.  $P_\vartheta = P_N$  ist die hypergeometrische Verteilung  $h(\cdot; n, N, W)$  und die Funktion  $g$  ist gegeben durch  $g(\vartheta) = \vartheta = N$ .

In obigem Beispiel ist die zu schätzende Funktion  $g(\vartheta) = \vartheta$ . Ist dagegen z.B. die Varianz einer Binomialverteilung zu schätzen, ist  $\vartheta = p$  und  $g(p) = np(1 - p)$ .

Bezeichne mit  $\mathcal{Y}$  den Wertebereich der zu schätzenden Funktion  $g$ . Jede Abbildung  $T : \mathcal{X} \rightarrow \mathcal{Y}$  heißt ein SCHÄTZER von  $g$ . Diese Sprechweise lässt natürlich auch völlig unsinnige Schätzer zu.

Häufig deutet man in der Notation an, dass geschätzt wird, und setzt ein „Dach“ über die zu schätzende Größe.  $\hat{N}$  wäre also ein Schätzer für  $N$ ,  $\hat{p}$  ein Schätzer für  $p$  und  $\hat{g}$  ein Schätzer für  $g$ .

**XI.4.3. Maximum-Likelihood Schätzer.** Sei  $x \in \mathcal{X}$  eine feste Beobachtung. Die Funktion

$$L_x : \Theta \rightarrow [0, 1]$$

mit

$$L_x(\vartheta) = P_\vartheta(x)$$

nennt man LIKELIHOOD-FUNKTION. Wenn  $L_x$  einen Maximalwert in  $\hat{\vartheta} = \hat{\vartheta}(x)$  annimmt, d.h.

$$L_x(\hat{\vartheta}) = \sup\{L_x(\vartheta) : \vartheta \in \Theta\},$$

nennt man  $\hat{\vartheta}(x)$  eine MAXIMUM-LIKELIHOOD SCHÄTZUNG von  $\vartheta$  und  $g(\hat{\vartheta})$  eine Maximum-Likelihood Schätzung von  $g(\vartheta)$ .

Häufig ist  $\Theta$  ein Intervall in  $\mathbb{R}$ , und eine Maximum-Likelihood Schätzung kann wie in Abschnitt IV.2.1 (S. 136, Teil I) durch Differentiation gefunden werden. Dabei ist es oft einfacher mit der Funktion  $\mathcal{L}_x = \ln L_x$  zu arbeiten. Wegen der Monotonie des Logarithmus haben  $\mathcal{L}_x$  und  $L_x$  ihre Maxima an der gleichen Stelle.

BEISPIEL XI.4.2. In  $n$  Bernoulli-Experimenten soll die Erfolgswahrscheinlichkeit  $p$  aus der Zahl  $x$  der Erfolge geschätzt werden. Es ist

$$\begin{aligned} L_x(p) &= \binom{n}{x} p^x (1-p)^{n-x} \\ \mathcal{L}_x(p) &= \ln L_x(p) \\ &= \ln \binom{n}{x} + x \ln p + (n-x) \ln(1-p) \\ \frac{d}{dp} \mathcal{L}_x(p) &= \frac{x}{p} - \frac{n-x}{1-p}. \end{aligned}$$

Also liegt das eindeutige Extremum bei  $\frac{x}{n}$ . Wie man leicht nachprüft, ist dies auch ein Maximum.  $\frac{x}{n}$  ist die Maximum-Likelihood Schätzung für die Erfolgswahrscheinlichkeit  $p$ .

**XI.4.4. Erwartungstreue.** Ist  $T : \mathcal{X} \rightarrow \mathbb{R}$  ein Schätzer, so bezeichnen wir den Erwartungswert von  $T$  bzgl.  $P_\vartheta$  mit  $E_\vartheta(T)$ :

$$E_\vartheta(T) = \sum_{x \in \mathcal{X}} T(x) P_\vartheta(x).$$

Die Beobachtung des Experimentes, das dem Schätzproblem zugrunde liegt, können wir als eine Zufallsvariable  $X$  auffassen, d.h.  $X(x) = x$ . Mit dieser Notation können wir den Begriff der Erwartungstreue definieren:

Ein Schätzer  $\hat{g}$  für  $g(\vartheta)$  heißt ERWARTUNGSTREU, wenn für alle  $\vartheta \in \Theta$  gilt

$$E_\vartheta(\hat{g}(X)) = g(\vartheta).$$

Speziell heißt  $\hat{\vartheta}$  ein ERWARTUNGSTREUER SCHÄTZER von  $\vartheta$ , wenn für alle  $\vartheta \in \Theta$  gilt

$$E_\vartheta(\hat{\vartheta}(X)) = \vartheta.$$

Die Differenz

$$b(\vartheta, \hat{g}) = E_\vartheta(\hat{g}(X)) - g(\vartheta)$$

heißt der BIAS (wörtlich: Vorurteil) des Schätzers  $\hat{g}$ . Ein Schätzer ist also genau dann erwartungstreu (engl.: UNBIASED), wenn sein Bias gleich Null ist.

BEISPIEL XI.4.3. Ist  $X$  binomial verteilt mit den Parametern  $n$  und  $p$ , so ist  $E(\frac{X}{n}) = p$  (vgl. Abschnitt XI.3.5 (S. 26)). Daher ist die Schätzung  $\hat{p}(X) = \frac{X}{n}$  aus Beispiel XI.4.2 erwartungstreu.

BEISPIEL XI.4.4. Häufig erfolgt die Messung einer unbekanntes Größe  $\mu$  durch Ausführen von  $n$  unabhängigen Zufallsexperimenten, die  $\mu$  als Erwartungswert haben. Mathematisch bedeutet dies, dass wir  $n$  unabhängige Zufallsvariable  $X_1, \dots, X_n$  und eine unbekanntes Familie  $P_\vartheta$  von Wahrscheinlichkeiten haben, so dass  $E_\vartheta(X_i) = \mu$  ist für alle  $i$  und  $\vartheta$ . Sei

$$\begin{aligned} g_1(\vartheta) &= E_\vartheta(X_i) \\ &= \mu \end{aligned}$$

und

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

der MITTELWERT. Dann ist für jedes  $\vartheta$

$$\begin{aligned} E_\vartheta(\bar{X}) &= \frac{1}{n} \sum_{i=1}^n E_\vartheta(X_i) \\ &= \mu. \end{aligned}$$

Also ist der Mittelwert ein erwartungstreuer Schätzer für  $g_1$ . Haben die  $X_i$  alle die gleiche unbekanntes Varianz  $\sigma^2$ , so bestimmt die Familie  $P_\vartheta$  auch diese Größe. Setze in diesem Fall

$$\begin{aligned} g_2(\vartheta) &= \sigma^2 \\ &= \text{Var}(X_i). \end{aligned}$$

Wir wählen

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

als Schätzer für  $\sigma^2$ . (Beachte, dass durch  $n-1$  geteilt wird!) Da die  $X_i$  unabhängig sind, erhalten wir für jedes  $\vartheta$

$$\begin{aligned} E_\vartheta((\bar{X} - \mu)^2) &= \text{Var}_\vartheta(\bar{X}) \\ &= \frac{1}{n^2} \text{Var}_\vartheta(X_1 + \dots + X_n) \\ &= \frac{1}{n^2} (\text{Var}_\vartheta(X_1) + \dots + \text{Var}_\vartheta(X_n)) \\ &= \frac{\sigma^2}{n} \end{aligned}$$

und

$$\begin{aligned}
 E_{\vartheta}((X_i - \bar{X})^2) &= E_{\vartheta}(((X_i - \mu) - (\bar{X} - \mu))^2) \\
 &= E_{\vartheta}((X_i - \mu)^2) - 2E_{\vartheta}((X_i - \mu)(\bar{X} - \mu)) \\
 &\quad + E_{\vartheta}((\bar{X} - \mu)^2) \\
 &= \sigma^2 - \frac{2}{n} \sum_{j=1}^n \underbrace{E_{\vartheta}((X_i - \mu)(X_j - \mu))}_{=0 \text{ für } j \neq i} + \frac{\sigma^2}{n} \\
 &= \sigma^2 - \frac{2}{n} E_{\vartheta}((X_i - \mu)^2) + \frac{\sigma^2}{n} \\
 &= \sigma^2 - \frac{\sigma^2}{n} \\
 &= \sigma^2 \left(1 - \frac{1}{n}\right).
 \end{aligned}$$

Hieraus folgt

$$\begin{aligned}
 E_{\vartheta}(s^2) &= \frac{1}{n-1} \sum_{i=1}^n E_{\vartheta}((X_i - \bar{X})^2) \\
 &= \frac{1}{n-1} n \sigma^2 \left(1 - \frac{1}{n}\right) \\
 &= \sigma^2.
 \end{aligned}$$

Also ist  $s^2$  ein erwartungstreuer Schätzer für die Unbekannte Varianz  $\sigma^2$ .

**XI.4.5. Der mittlere quadratische Fehler.** Die wohl wichtigste Forderung an einen Schätzer  $T$  von  $g(\vartheta)$  ist wohl diejenige, dass der Fehler  $T(X) - g(\vartheta)$  der Schätzwerte „klein“ sein sollte. Um diese Größe quantitativ zu messen, betrachtet man den MITTLEREN QUADRATISCHEN FEHLER

$$R(\vartheta, T) = E_{\vartheta}((T(X) - g(\vartheta))^2).$$

Man kann den mittleren quadratischen Fehler auch durch die Varianz und den Bias ausdrücken. Berücksichtigt man, dass  $E_{\vartheta}(T) - g(\vartheta)$  eine Konstante ist, erhält man nämlich

$$\begin{aligned}
 &E_{\vartheta} \left( (T(X) - g(\vartheta))^2 \right) \\
 &= E_{\vartheta} \left( \left( (T(X) - E_{\vartheta}(T)) - (g(\vartheta) - E_{\vartheta}(T)) \right)^2 \right)
 \end{aligned}$$

$$\begin{aligned}
 &= E_{\vartheta} \left( \underbrace{\left( T(X) - E_{\vartheta}(T) \right)^2}_{=\text{Var}_{\vartheta}(T)} \right) \\
 &\quad - 2 \underbrace{E_{\vartheta} \left( \left( T(X) - E_{\vartheta}(T) \right) \left( g(\vartheta) - E_{\vartheta}(T) \right) \right)}_{=(g(\vartheta) - E_{\vartheta}(T)) E_{\vartheta}(T(X) - E_{\vartheta}(T)) = 0}} \\
 &\quad + E_{\vartheta} \left( \underbrace{\left( g(\vartheta) - E_{\vartheta}(T) \right)^2}_{=(g(\vartheta) - E_{\vartheta}(T))^2} \right) \\
 &= \text{Var}_{\vartheta}(T) + b(\vartheta, T)^2.
 \end{aligned}$$

Also ist der mittlere quadratische Fehler

$$R(\vartheta, T) = \text{Var}_{\vartheta}(T) + b(\vartheta, T)^2.$$

BEISPIEL XI.4.5. Für den Schätzer  $T(X) = \frac{X}{n}$  aus Beispiel XI.4.3 (S. 35) für die binomial verteilte Zufallsvariable  $X$  gilt gemäß Beispiel XI.4.3 und Abschnitt XI.3.5 (S. 26)

$$\begin{aligned}
 b(\vartheta, T) &= 0, \\
 \text{Var}_{\vartheta}(T) &= \frac{1}{n^2} \text{Var}_{\vartheta}(X) \\
 &= \frac{1}{n} \vartheta(1 - \vartheta).
 \end{aligned}$$

Also ist der mittlere quadratische Fehler

$$R(\vartheta, T) = \frac{1}{n} \vartheta(1 - \vartheta).$$

Für alle  $\vartheta \in \Theta = (0, 1)$  gilt

$$R(\vartheta, T) \leq \frac{1}{4n}.$$

BEISPIEL XI.4.6. Wir greifen Beispiel XI.4.4 (S. 35) auf. Der Bias des Mittelwertes  $\bar{X}$  ist Null; seine Varianz ist gemäß Beispiel XI.4.4 gleich  $\frac{\sigma^2}{n}$ , wenn alle  $X_i$  die gleiche Varianz  $\sigma^2$  haben. Also ist der mittlere quadratische Fehler

$$R(\vartheta, \bar{X}) = \frac{\sigma^2}{n}.$$

### XI.5. Approximationen der Binomialverteilung

Für große Werte von  $n$  ist die exakte Berechnung der Wahrscheinlichkeit

$$b_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

äußerst mühselig und anfällig für Rundungsfehler. Dies gilt vermehrt für Summen dieser Ausdrücke. Wir wollen daher in diesem Abschnitt Approximationen für  $b_{n,p}(k)$  bestimmen, die leichter und mit geringeren Rundungsfehlern berechnet werden können.

Zur Vereinfachung der Sprache führen wir die Notation

$$a_n \sim b_n \iff \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$$

ein.  $a_n \sim b_n$  bedeutet, dass es eine Nullfolge  $(\theta_n)_{n \in \mathbb{N}}$  gibt mit

$$b_n(1 - \theta_n) \leq a_n \leq b_n(1 + \theta_n)$$

für alle bis auf endlich viele  $n$ .

**XI.5.1. Approximation von  $n!$  und  $b_{n,p}(k)$ .** Unser wichtigstes Hilfsmittel ist die STIRLINGSISCHE FORMEL

$$n! \sim \sqrt{2\pi n} n^n e^{-n}.$$

Für die eingangs erwähnte Folge  $(\theta_n)_{n \in \mathbb{N}}$  kann man in diesem Fall zeigen:

$$\frac{1}{12n+1} \leq \theta_n \leq \frac{1}{12n}$$

für alle  $n$ .

Tabelle XI.5.1 gibt einen Eindruck von der Genauigkeit der Stirlingschen Formel.

TABELLE XI.5.1. Genauigkeit der Stirlingschen Formel

$n$	$n!$	$\sqrt{2\pi n} n^n e^{-n}$	$\frac{ n! - \sqrt{2\pi n} n^n e^{-n} }{n!}$
2	2	1.919	4.05%
5	120	118.019	1.65%
10	3628800	3598690	0.83%

Mit der Stirlingschen Formel folgt

$$\begin{aligned} b_{n,p}(k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &\sim \frac{\sqrt{2\pi n} n^n e^{-n}}{\sqrt{2\pi k} k^k e^{-k} \sqrt{2\pi(n-k)} (n-k)^{(n-k)} e^{-(n-k)}} p^k (1-p)^{n-k} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{n(1-p)}{n-k}\right)^{n-k}. \end{aligned}$$

Wir betrachten nun eine Folge  $(k_n)_{n \in \mathbb{N}}$  mit  $k_n \sim np$ . Dann ist  $n - k_n \sim n(1 - p)$  und

$$\begin{aligned} \sqrt{\frac{n}{k_n(n - k_n)}} &\sim \frac{1}{\sqrt{np(1 - p)}} \\ &= \frac{1}{\sigma_n}, \end{aligned}$$

wobei

$$\sigma_n = \sqrt{np(1 - p)}$$

die Standardabweichung der  $b_{n,p}$ -Verteilung ist (vgl. Abschnitt XI.3.5 (S. 26)). Wir wollen nun das Verhalten von

$$\chi(n, k_n) = \left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1 - p)}{n - k_n}\right)^{n - k_n}$$

studieren. Dazu bilden wir den Logarithmus und setzen

$$t_n = \frac{k_n}{n}.$$

Dann gilt  $t_n \rightarrow p$  für  $n \rightarrow \infty$  und

$$\begin{aligned} -\ln \chi(n, k_n) &= -k_n \ln\left(\frac{np}{k_n}\right) - (n - k_n) \ln\left(\frac{n(1 - p)}{n - k_n}\right) \\ &= n \left\{ \frac{k_n}{n} \ln\left(\frac{t_n}{p}\right) + \left(1 - \frac{k_n}{n}\right) \ln\left(\frac{1 - \frac{k_n}{n}}{1 - p}\right) \right\} \\ &= n \left\{ t_n \ln\left(\frac{t_n}{p}\right) + (1 - t_n) \ln\left(\frac{1 - t_n}{1 - p}\right) \right\}. \end{aligned}$$

Definiere die Funktion  $g : (0, 1) \rightarrow \mathbb{R}$  durch

$$g(t) = t \ln\left(\frac{t}{p}\right) + (1 - t) \ln\left(\frac{1 - t}{1 - p}\right).$$

Dann ist

$$\begin{aligned} g(p) &= 0, \\ g'(p) &= 0, \\ g''(p) &= \frac{1}{p} + \frac{1}{1 - p} \\ &= \frac{1}{p(1 - p)}. \end{aligned}$$

Nach der Taylor Formel ist daher

$$g(t) = \frac{1}{2p(1 - p)}(t - p)^2 + r(t - p) \quad \text{mit } |r(t - p)| \leq c|t - p|^3$$

in einer Umgebung von  $p$ .

Wir nehmen nun an, dass nicht nur  $t_n \rightarrow p$  gilt, sondern dass sogar  $n(t_n - p)^3 \rightarrow 0$  gilt. Dann folgt  $n|r(t_n - p)| \rightarrow 0$  und daher

$$\left| -\ln \chi(n, k_n) - \frac{n(t_n - p)^2}{2p(1-p)} \right| \xrightarrow{n \rightarrow \infty} 0.$$

Setzen wir

$$x(n, k_n) = \frac{k_n - np}{\sigma_n},$$

so ist

$$\frac{n(t_n - p)^2}{2p(1-p)} = \frac{x(n, k_n)^2}{2}.$$

Damit erhalten wir

$$\chi(n, k_n) \sim \exp\left(-\frac{x(n, k_n)^2}{2}\right).$$

Insgesamt haben wir somit unter obigen Voraussetzungen die Beziehung

$$b_{n,p}(k_n) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_n} \exp\left(-\frac{x(n, k_n)^2}{2}\right)$$

gezeigt.

Die Funktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

heißt DICHTER DER STANDARD-NORMALVERTEILUNG. Damit lässt sich unsere Approximation der  $b_{n,p}$ -Verteilung schreiben als

$$b_{n,p}(k) \sim \frac{1}{\sigma_n} \varphi(x(n, k))$$

mit

$$\sigma_n = \sqrt{np(1-p)},$$

$$x(n, k) = \frac{k - pn}{\sigma_n}.$$

Mit Hilfe von Ergebnissen der Funktionentheorie folgt

$$\int_{-\infty}^{\infty} \varphi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1.$$

**XI.5.2. Der Satz von Moivre-Laplace.** Wir wollen Summen von Wahrscheinlichkeiten  $b_{n,p}(k)$  für große  $n$  näherungsweise berechnen. Dazu definieren wir

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Die Funktion  $\Phi$  heißt die VERTEILUNGSFUNKTION DER STANDARD-NORMALVERTEILUNG. Sie hat folgende Eigenschaften

$$\begin{aligned} \lim_{x \rightarrow -\infty} \Phi(x) &= 0 \\ \lim_{x \rightarrow +\infty} \Phi(x) &= 1 \\ \Phi(0) &= \frac{1}{2} \\ \Phi(-x) &= 1 - \Phi(x) \quad \text{für } x \geq 0 \\ \Phi(x) &\leq \Phi(y) \quad \text{für } x \leq y \end{aligned}$$

Sie ist ausführlich tabelliert (z.B. in U. Krengel: Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg 2000, S. 248).

Sei nun  $S_n$  eine  $b_{n,p}$ -verteilte Zufallsvariable. Setze

$$\begin{aligned} S_n^* &= \frac{S_n - np}{\sigma_n} \\ &= \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}}. \end{aligned}$$

Dann ist  $E(S_n^*) = 0$  und  $\text{Var}(S_n^*) = 1$ .  $S_n^*$  heißt daher die STANDARDISIERTE oder NORMIERTE Form von  $S_n$ . Nimmt  $S_n$  den Wert  $k$  an, so hat  $S_n^*$  den Wert  $x(n, k)$  mit  $x(n, k)$  aus dem vorigen Abschnitt.

Der SATZ VON MOIVRE-LAPLACE besagt, dass  $\Phi$  eine gute Approximation für die Verteilung von  $S_n^*$  ist:

$$\lim_{n \rightarrow \infty} P(a \leq S_n^* \leq b) = \Phi(b) - \Phi(a).$$

**BEISPIEL XI.5.1.** Wie groß ist näherungsweise die Wahrscheinlichkeit, bei 600 Würfeln mit einem Würfel mindestens 90 und höchstens 100 Sechsen zu erhalten?

Es ist

$$n = 600$$

und

$$p = \frac{1}{6},$$

also

$$np = 100$$

und

$$\begin{aligned}\sigma_n &= \sqrt{600 \cdot \frac{1}{6} \cdot \frac{5}{6}} \\ &= \sqrt{\frac{250}{3}} \\ &\approx 9.13\end{aligned}$$

Damit folgt

$$\begin{aligned}P(90 \leq S_n \leq 100) &= P\left(\frac{90 - 100}{\sigma_n} \leq S_n^* \leq \frac{100 - 100}{\sigma_n}\right) \\ &\approx \Phi(0) - \Phi\left(-\frac{10}{9.13}\right) \\ &= 0.5 - \Phi(-1.095) \\ &= 0.5 - (1 - \Phi(1.095)) \\ &\approx 0.36.\end{aligned}$$

BEISPIEL XI.5.2. Wir wollen den Prozentsatz der Wähler der Partei  $A$  schätzen. Werden  $n$  Wähler befragt und sind darunter  $S_n$  Wähler der Partei  $A$ , so sei  $\frac{S_n}{n}$  der Schätzer für die Wahrscheinlichkeit  $p$ , dass ein zufällig ausgewählter Wähler die Partei  $A$  wählt. Wie groß muss  $n$  sein, damit die Wahrscheinlichkeit eines Irrtums um 1% nicht größer ist als 0.05?

Es muss also gelten

$$P(-0.01 \leq \frac{S_n}{n} - p \leq 0.01) \approx 0.95.$$

Mit

$$\sigma_n = \sqrt{np(1-p)}$$

erhalten wir

$$\begin{aligned}0.95 &\approx P\left(-\frac{0.01n}{\sigma_n} \leq S_n^* \leq \frac{0.01n}{\sigma_n}\right) \\ &\approx \Phi\left(\frac{0.01n}{\sigma_n}\right) - \Phi\left(-\frac{0.01n}{\sigma_n}\right) \\ &= 2\Phi\left(\frac{0.01n}{\sigma_n}\right) - 1\end{aligned}$$

also

$$\Phi\left(\frac{0.01n}{\sigma_n}\right) \approx 0.975.$$

Da  $\Phi$  streng monoton ist, existiert die Umkehrfunktion  $\Phi^{-1}$ . Es muss also gelten

$$\frac{0.01n}{\sigma_n} \approx \Phi^{-1}(0.975).$$

Aus einer Tabelle von  $\Phi$  entnehmen wir

$$\Phi^{-1}(0.975) \approx 1.96.$$

Also muss gelten

$$\begin{aligned} 1.96 &\approx \frac{0.01}{\sigma_n} \\ &= \frac{0.01\sqrt{n}}{\sqrt{p(1-p)}} \\ \implies \sqrt{n} &\approx 196\sqrt{p(1-p)} \\ \implies n &\approx 196^2 p(1-p). \end{aligned}$$

Wir kennen aber  $p$  nicht! Aber wir wissen, dass für alle möglichen  $p$ , nämlich  $0 \leq p \leq 1$ , gilt

$$p(1-p) \leq \frac{1}{4}.$$

Damit erhalten wir die Schätzung

$$\begin{aligned} n &\approx \frac{196^2}{4} \\ &= 98^2 \\ &= 9604. \end{aligned}$$

Wir müssen also ca. 9600 Wähler befragen.

Hätten wir z.B. die Zusatzinformation  $p \leq 0.1$  zur Verfügung, kämen wir wegen

$$\max_{0 \leq p \leq 0.1} p(1-p) = 0.1 \cdot 0.9 = 0.09$$

mit ca. 3450 Befragungen aus.

**XI.5.3. Die Poisson-Approximation.** In diesem Abschnitt geben wir eine andere Approximation der  $b_{n,p}$ -Verteilung an, die für kleine Werte von  $p$  besser ist als die Normalverteilung aus Abschnitt XI.5.1.

Eine Zufallsvariable  $X$  heißt POISSON-VERTEILT mit Parameter  $\lambda \geq 0$ , wenn gilt

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{mit } k = 0, 1, \dots$$

Ist  $(p_n)_{n \in \mathbb{N}}$  eine Folge in  $[0, 1]$  mit  $np_n \rightarrow \lambda$ , so gilt für jedes  $k$  die POISSON-APPROXIMATION

$$b_{n,p_n}(k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ \xrightarrow{n \rightarrow \infty} e^{-\lambda} \frac{\lambda^k}{k!}.$$

Um dies einzusehen, setze  $\lambda_n = np_n$ . Dann gilt für festes  $k$

$$\begin{aligned} b_{n,p_n}(k) &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{1}{k!} \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{n^k} \lambda_n^k \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\ &= \frac{1}{k!} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n} \lambda_n^k \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k}. \end{aligned}$$

Für  $n \rightarrow \infty$  streben die Quotienten  $\frac{n}{n}$ ,  $\frac{n-1}{n}$ ,  $\dots$ ,  $\frac{n-k+1}{n}$  alle gegen 1; nach Voraussetzung gilt

$$\lambda_n^k \rightarrow \lambda^k$$

und

$$\left(1 - \frac{\lambda_n}{n}\right)^{-k} \rightarrow 1.$$

Wegen der Stetigkeit der Exponentialfunktion gilt

$$\left(1 - \frac{\lambda_n}{n}\right)^n \rightarrow e^{-\lambda}.$$

Damit folgt die Poisson-Approximation.

TABELLE XI.5.2. Poisson-Approximation für die Wahrscheinlichkeit, dass unter 91 Personen genau  $k$  am heutigen Tage Geburtstag haben

$k$	$P(X = k) \approx \frac{0.25^k}{k!} e^{-0.25}$
1	0.1947
2	0.0243
5	$6.337 \cdot 10^{-6}$

BEISPIEL XI.5.3. In einem Hörsaal befinden sich 91 Studierende. Die Wahrscheinlichkeit  $p$ , heute Geburtstag zu haben, ist  $p = \frac{1}{365}$ . Die Zahl derer, die heute Geburtstag haben, ist praktisch Poissonverteilt mit  $\lambda = \frac{91}{365} \approx 0.25$ . Tabelle XI.5.2 gibt die mit der Poisson-Approximation genäherte Wahrscheinlichkeit  $P(X = k)$  an, dass heute genau  $k$  Personen aus dem Hörsaal Geburtstag haben.

BEISPIEL XI.5.4. Bei der Produktion von Blitzlichtlampen ist mit der Wahrscheinlichkeit  $p = 0.015$  eine Lampe schon bei der Produktion defekt. Wie groß muss man  $n$  wählen, damit in einem Karton mit  $n$  unabhängig gewählten Lampen mit Wahrscheinlichkeit  $\geq 0.8$  mindestens 100 funktionstüchtige Exemplare sind? Die Zahl  $n$  ist minimal zu wählen unter der Bedingung

$$0.8 \leq P(\text{„höchstens } n - 100 \text{ fehlerhafte Lampen“}) \\ = \sum_{k=0}^{n-100} b_{n,p}(k).$$

Approximieren wir die rechte Seite mit der Poisson-Verteilung zu  $\lambda_n = np$ , erhalten wir die Bedingung

$$0.8 \leq r_n = e^{-\lambda_n} \sum_{k=0}^{n-100} \frac{\lambda_n^k}{k!}.$$

Mit Hilfe eines Taschenrechners prüfen wir nach, dass  $n = 102$  die kleinste Zahl ist, die diese Bedingung erfüllt. Wir müssen also 102 Lampen in den Karton legen.

Würden wir die Wahrscheinlichkeit auf 0.9 erhöhen, müssten wir 103 Lampen in den Karton legen.

## XI.6. Tests

**XI.6.1. Motivation.** Eine Person behauptet, dass sie bei Milchkaffee schmecken kann, ob zuerst der Kaffee oder die Milch in die Tasse gegeben wurde. Wie können wir die Stichhaltigkeit dieser Behauptung nachprüfen?

Ein erster Ansatz geht auf Fisher (1935) zurück. Er schlägt folgende „Versuchsanordnung“ vor:

In vier Tassen wird zuerst Kaffee und dann Milch gegossen. Diese Tassen nennen wir Tassen vom Typ 1. In weitere vier Tassen wird zuerst Milch und dann Kaffee gegossen. Diese Tassen nennen wir vom Typ 2. Die Kaffee- und Milchmenge ist dabei natürlich bei allen Tassen die gleiche, und es wird gut umgerührt. Nun werden der Person die acht Tassen in beliebiger Reihenfolge (8! Möglichkeiten) vorgesetzt, und sie soll die vier Tassen vom Typ 1

identifizieren. Falls ihr das gelingt, wollen wir ihre Behauptung als wahr akzeptieren.

Die Wahrscheinlichkeit, dass die Person rein zufällig richtig rät, ist aufgrund der hypergeometrischen Verteilung aus Abschnitt XI.1.4 (S. 12) gleich

$$\binom{8}{4}^{-1} = \frac{1}{70}.$$

Wir geben der Person also nur mit dieser Wahrscheinlichkeit Recht, wenn ihre Behauptung nicht stimmt.

Schwieriger wird die Angelegenheit, wenn die Person behauptet, nicht unfehlbar zu sein, aber mit großer Sicherheit die Unterscheidung treffen zu können. Wenn wir ihr die Behauptung in diesem Fall schon dann glauben, wenn sie mindestens drei der vier Tassen vom Typ 1 richtig identifiziert, ist die Wahrscheinlichkeit für einen zufälligen Treffer schon

$$\begin{aligned} \frac{\binom{4}{4}\binom{4}{0} + \binom{4}{3}\binom{4}{1}}{\binom{8}{4}} &= \frac{1 + 4 \cdot 4}{70} \\ &= \frac{17}{70} \\ &\approx 0.24. \end{aligned}$$

Wie kann man in diesem Fall die Gewissheit, der Person auf die Schliche zu kommen, erhöhen, ohne sie dabei zu benachteiligen?

Neyman (1950) schlägt dazu die folgende verbesserte „Versuchsanordnung“ vor:

Der Person wird  $n$ -mal die Aufgabe gestellt, zwei Tassen, von denen eine vom Typ 1 und eine vom Typ 2 ist, korrekt zu klassifizieren. Die beiden Tassen werden ihr jeweils in einer zufälligen, durch Münzwurf bestimmten Reihenfolge vorgesetzt. Damit die Person unabhängig von früheren Entscheidungen urteilen kann, wird jeder Teilversuch an einem anderen Tag ausgeführt.

$X$  sei die Zahl der Tage, an denen die Person die Tassen richtig klassifiziert.

Ein mathematisches Modell für diese Versuchsanordnung ist die Annahme,  $X$  sei binomial verteilt mit Parametern  $n$  und  $p$ . Die Hypothese „die Person schwindelt und trifft ihre Entscheidung rein zufällig“ entspricht dem Fall  $p = \frac{1}{2}$ . Diese Hypothese nennt man die „Nullhypothese“. Die Alternative ist, dass die Person tatsächlich die behauptete Fähigkeit hat, und entspricht dem Fall  $p > \frac{1}{2}$ . Es wird nun eine Zahl  $t$  festgelegt, so dass  $P(X \geq t) \leq \alpha$  ist. Dabei ist  $\alpha$  eine vorgegebene, kleine, positive Schranke, z.B.  $\alpha = 0.05$ . Dies ist die „Sicherheitsmarge“,

mit der wir ausschließen wollen, dass wir der Person ungerechtfertigterweise auf den Leim gehen. Falls die Person nun mindestens  $t$ -mal die Tassen richtig klassifiziert, wollen wir ihr glauben.

Der Unterschied des Neymanschen Ansatzes zu dem von Fisher liegt darin, dass er präzise Aussagen über die Wahrscheinlichkeit erlaubt, die Nullhypothese zu akzeptieren, wenn die Alternative zutrifft.

**XI.6.2. Grundbegriffe der Testtheorie.** Von einem Testproblem spricht man, wenn eine zufällige Größe  $X$  mit einer unbekanntem Verteilung  $P_\vartheta$  beobachtet wird, und man aufgrund des beobachteten Wertes  $x$  von  $X$  entscheiden soll, ob  $P_\vartheta$  einer bestimmten Menge von Verteilungen angehört oder nicht.

Zur genaueren mathematischen Beschreibung sei  $\mathcal{X}$  die Menge der möglichen Werte der Zufallsvariablen  $X$  und  $\{P_\vartheta : \vartheta \in \Theta\}$  die Menge der in Betracht gezogenen Verteilungen von  $X$ . Unter diesen möglichen Verteilungen ist eine echte Teilmenge  $\{P_\vartheta : \vartheta \in H\}$  durch zusätzliche Bedingungen ausgezeichnet. Ein TEST ist eine Entscheidungsregel, die für jeden möglichen Wert von  $X$  festlegt, ob man sich für die HYPOTHESE „ $\vartheta \in H$ “ oder für die ALTERNATIVE „ $\vartheta \in \Theta \setminus H$ “ entscheiden soll. Man nennt auch kurz  $H$  die Hypothese und  $K = \Theta \setminus H$  die Alternative. Die Entscheidung für die Hypothese nennt man ANNAHME DER HYPOTHESE, und die Entscheidung für die Alternative nennt man VERWERFEN DER HYPOTHESE.

Ein Test ist also beschrieben durch Angabe der Menge  $R$  derjenigen Werte  $x$  von  $X$ , für die die Hypothese verworfen werden soll.  $R$  heißt VERWERFUNGSBEREICH oder KRITISCHER BEREICH des Tests.

Innerhalb des gewählten Modells sind zwei Arten von Fehlern möglich:

- Ist  $\vartheta \in H$  und wird die Hypothese verworfen, spricht man von einem FEHLER ERSTER ART.
- Ist  $\vartheta \in K$  und wird die Hypothese angenommen, spricht man von einem FEHLER ZWEITER ART.

Praktisch beschreibt man den Verwerfungsbereich  $R$  durch eine Funktion  $T(x)$ , die so gewählt wird, dass große Werte gegen die Hypothese sprechen.  $T$  heißt TESTSTATISTIK. Man wählt dann einen KRITISCHEN WERT  $t$  und verwirft die Hypothese, wenn  $T(x) \geq t$  ist. Es ist also

$$R = \{x : T(x) \geq t\}.$$

Bisher ist das Testproblem so formuliert, dass  $H$  und  $K$  symmetrische Rollen spielen. Dies ist aber in der Praxis in der Regel nicht der Fall. Vielmehr wird man die Hypothese so wählen, dass sie der etablierten Hypothese oder der bisherigen Erfahrung entspricht. Bei unserem einführenden Beispiel bedeutet dies, dass man als Hypothese

„die Person schwindelt“ nehmen wird. Bei der Einführung eines neuen Medikamentes würde man z.B. die Hypothese „das Medikament ist nicht wirksamer als die bekannten Mittel“ wählen.

Man zieht nur Verwerfungsbereiche in Betracht, für die die Wahrscheinlichkeit eines Fehlers erster Art durch eine vorgegebene kleine Zahl  $\alpha > 0$  begrenzt ist. Quantitative Aussagen erhält man durch die GÜTE-FUNKTION

$$\beta(\vartheta) = P_{\vartheta}(X \in R).$$

Sie ordnet jedem  $\vartheta$  die Verwerfungswahrscheinlichkeit unter  $P_{\vartheta}$  zu. Der Test hat das Niveau  $\alpha$ , wenn für alle  $\vartheta \in H$  die Ungleichung  $\beta(\vartheta) \leq \alpha$  gilt. Die Wahrscheinlichkeit eines Fehlers erster Art ist dann also maximal  $\alpha$ . Gebräuchliche Werte für  $\alpha$  sind 0.05, 0.02 oder 0.01.

**XI.6.3. Zurück zur Motivation.** In unserem motivierenden Beispiel des ersten Abschnittes ist das gesuchte Modell beschrieben durch

$$\begin{aligned} \mathcal{X} &= \{0, 1, \dots, n\}, \\ \Theta &= \left[\frac{1}{2}, 1\right], \\ \vartheta &= p, \\ P_{\vartheta}(X = x) &= b_{n,p}(x) \\ &= \binom{n}{x} p^x (1-p)^{n-x}. \end{aligned}$$

Die Hypothese ist

$$H = \left\{\frac{1}{2}\right\}$$

und die Alternative ist

$$K = \left(\frac{1}{2}, 1\right].$$

Der Verwerfungsbereich ist von der Form

$$R = \{x : x \geq t\}.$$

Sei

$$\beta(p|t, n) = P_p(X \geq t)$$

die Gütefunktion dieses Tests. Als Niveau wollen wir

$$\alpha = 0.05$$

wählen. Wir müssen dann für gegebenes  $n$  den Parameter  $t$  so wählen, dass

$$\beta\left(\frac{1}{2}|t, n\right) \leq 0.05$$

ist.

Für  $n = 5$  kommt nur  $t = 5$  in Frage. Denn für  $t = 4$  ist

$$\begin{aligned}\beta\left(\frac{1}{2}|4, 5\right) &= \left(\frac{1}{2}\right)^5 + \binom{5}{1}\left(\frac{1}{2}\right)^5 \\ &= \frac{6}{32} \\ &\approx 0.187 \\ &> 0.05.\end{aligned}$$

Die Person müsste also in jedem der 5 Versuche die Tassen richtig klassifizieren, damit wir ihr glauben.

Für  $n = 5$  ist

$$\beta(p|5, 5) = p^5.$$

Es ist

$$\beta(0.6|5, 5) \approx 0.08$$

und

$$\beta(0.9|5, 5) \approx 0.59.$$

Hätte also die Person eine Erfolgswahrscheinlichkeit von 0.6 pro Klassifizierung, würde sie doch nur mit Wahrscheinlichkeit 0.08 ihre Fähigkeit beweisen können. Selbst bei einer Erfolgswahrscheinlichkeit von 0.9, würde ihre Behauptung nur mit Wahrscheinlichkeit 0.59 akzeptiert. Die Gütefunktion zeigt also, ob der Test überhaupt in der Lage ist, eine Abweichung von der Hypothese anzuzeigen.

Ist z.B.  $p = 0.6$ , ist erst bei  $n = 42$  zu klassifizierenden Tassenpaaren die Wahrscheinlichkeit wenigstens  $\frac{1}{3}$ , dass die Behauptung der Person akzeptiert wird. Der kleinste Wert von  $t$  mit  $\beta\left(\frac{1}{2}|t, 42\right) \leq 0.05$  ist  $t = 27$ . Die Person müsste also mindestens 27 von 42 Tassenpaaren richtig klassifizieren.



## KAPITEL XII

### Stochastik II Allgemeine Modelle

#### XII.1. Wahrscheinlichkeitsmaße mit Dichten

**XII.1.1. Ergebnismengen.** Im vorigen Kapitel haben wir stets endliche Ergebnismengen  $\Omega$  betrachtet. Die Bedingung  $\text{card}(\Omega) < \infty$  lassen wir jetzt fallen und betrachten allgemeine Ergebnismengen  $\Omega$ . Die wichtigsten Beispiele für dieses Kapitel sind  $\Omega = \mathbb{R}$  und  $\Omega = \mathbb{R}^n$ .

**XII.1.2.  $\sigma$ -Algebren.** Bisher waren die Wahrscheinlichkeitsmaße stets auf der ganzen Potenzmenge  $\mathcal{P}(\Omega)$  definiert, d.h. jede Teilmenge von  $\Omega$  war ein mögliches Ereignis. Man kann zeigen, dass dies für unendliche Mengen wie  $\Omega = \mathbb{N}$  oder  $\Omega = \mathbb{R}$  zu Widersprüchen führt. Daher können wir nur noch Teilmengen  $\mathcal{A}$  von  $\mathcal{P}(\Omega)$  als mögliche Ereignismengen betrachten. Diese müssen aber gewisse Eigenschaften haben.

Der richtige Begriff ist hier derjenige der  $\sigma$ -Algebra. Eine Menge  $\mathcal{A} \subset \mathcal{P}(\Omega)$  von Teilmengen von  $\Omega$  heißt  $\sigma$ -ALGEBRA (über  $\Omega$ ), wenn folgende drei Eigenschaften erfüllt sind:

$$\begin{array}{l} \Omega \in \mathcal{A} \\ A \in \mathcal{A} \implies \Omega \setminus A \in \mathcal{A} \\ A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \end{array}$$

Sei nun  $\mathcal{B} \subset \mathcal{P}(\Omega)$ . Dann kann man zeigen, dass es genau eine  $\sigma$ -Algebra  $\mathcal{A}$  mit den folgenden beiden Eigenschaften gibt:

- $\mathcal{B} \subset \mathcal{A}$ .
- Für jede  $\sigma$ -Algebra  $\tilde{\mathcal{A}}$  mit  $\mathcal{B} \subset \tilde{\mathcal{A}}$  gilt  $\mathcal{A} \subset \tilde{\mathcal{A}}$ .

$\mathcal{A}$  ist gewissermaßen die kleinste  $\sigma$ -Algebra, die das Mengensystem  $\mathcal{B}$  enthält.  $\mathcal{A}$  heißt die von  $\mathcal{B}$  ERZEUGTE  $\sigma$ -ALGEBRA.

BEISPIEL XII.1.1. Sei  $\Omega = \mathbb{R}$ . Betrachte folgende Mengensysteme

$$\begin{aligned} \mathcal{B}_1 &= \{(-\infty, a] : a \in \mathbb{R}\}, \\ \mathcal{B}_2 &= \{(-\infty, a) : a \in \mathbb{R}\}, \\ \mathcal{B}_3 &= \{(a, b) : a, b \in \mathbb{R}, a < b\}, \\ \mathcal{B}_4 &= \{[a, b] : a, b \in \mathbb{R}, a \leq b\}. \end{aligned}$$

Alle vier Mengensysteme erzeugen die gleiche  $\sigma$ -Algebra  $\mathcal{B}$ . Diese heißt die  $\sigma$ -Algebra der BORELMENGEN von  $\mathbb{R}$ . Alle Intervalle, egal ob offen oder nicht, egal ob beschränkt oder nicht, sind in  $\mathcal{B}$  enthalten.

BEISPIEL XII.1.2. Sei  $\Omega = \mathbb{R}^n$ . Wir definieren die Relation  $\leq$  auf  $\mathbb{R}^n$  durch

$$x \leq y \iff x_i \leq y_i \text{ für alle } 1 \leq i \leq n.$$

Mit dieser Relation können wir Intervalle verallgemeinern:

$$[a, b] = \{x \in \mathbb{R}^n : a \leq x \leq b\}.$$

Für  $n = 2$  ist z.B.  $[a, b]$  das Rechteck, dessen linke untere Ecke die Koordinaten  $(a_1, a_2)$  hat, und dessen rechte obere Ecke die Koordinaten  $(b_1, b_2)$  hat. Das Mengensystem  $\{[a, b] : a, b \in \mathbb{R}^n\}$  aller dieser Intervalle erzeugt eine  $\sigma$ -Algebra  $\mathcal{B}$ . Diese heißt wieder  $\sigma$ -Algebra der BORELMENGEN von  $\mathbb{R}^n$ . Zu ihr gehören z.B. alle offenen und alle abgeschlossenen Mengen (vgl. Abschnitt VIII.2.1 (S. 35, Teil II)).

**XII.1.3. Wahrscheinlichkeitsmaße.** Sei nun  $\Omega$  eine beliebige Menge und  $\mathcal{A}$  irgendeine  $\sigma$ -Algebra über  $\Omega$ . Eine Abbildung  $P : \mathcal{A} \rightarrow [0, 1]$  heißt ein WAHRSCHEINLICHKEITSMASS auf  $\Omega$ , wenn folgende drei Bedingungen erfüllt sind:

$$P(\Omega) = 1$$

$$P(A) \geq 0 \quad \text{für alle } A \in \mathcal{A}$$

$$A_1, A_2, \dots \in \mathcal{A} \text{ paarweise disjunkt} \implies P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Ein WAHRSCHEINLICHKEITSRAUM ist ein Tripel  $(\Omega, \mathcal{A}, P)$  mit einer Menge  $\Omega$ , einer  $\sigma$ -Algebra  $\mathcal{A}$  und einem Wahrscheinlichkeitsmaß  $P$ . Die Elemente von  $\mathcal{A}$  heißen EREIGNISSE.

Im vorigen Kapitel konnten wir auf die Angabe der  $\sigma$ -Algebra  $\mathcal{A}$  verzichten, weil diese „per ordre de mufti“ stets die Potenzmenge  $\mathcal{P}(\Omega)$  war.

Eine wichtige Konsequenz aus den obigen Eigenschaften eines Wahrscheinlichkeitsmaßes ist:

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum.

- Ist  $B_1 \subset B_2 \subset \dots$  eine wachsende Folge von Ereignissen und  $B$  deren Vereinigung, so ist

$$P(B) = \lim_{i \rightarrow \infty} P(B_i).$$

- Ist  $C_1 \supset C_2 \supset \dots$  eine fallende Folge von Ereignissen und  $C$  deren Durchschnitt, so ist

$$P(C) = \lim_{i \rightarrow \infty} P(C_i).$$

**XII.1.4. Wahrscheinlichkeitsmaße mit Dichten.** Wir betrachten nun den Fall  $\Omega = \mathbb{R}$  und  $\mathcal{A} = \mathcal{B}$  die Menge der Borelmengen auf  $\mathbb{R}$ . Sei  $P$  ein Wahrscheinlichkeitsmaß auf  $\mathbb{R}$ . Dann existiert für jedes  $x \in \mathbb{R}$  die Zahl

$$F(x) = P((-\infty, x]).$$

Die Funktion  $F : \mathbb{R} \rightarrow [0, 1]$  ist monoton wachsend mit

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x) &= 0, \\ \lim_{x \rightarrow +\infty} F(x) &= 1. \end{aligned}$$

Außerdem existiert für jedes  $x \in \mathbb{R}$  der rechtsseitige Grenzwert  $F(x+0)$  (vgl. Abschnitt III.3.4 (S. 121, Teil I)) und stimmt mit  $F(x)$  überein. Die Funktion  $F$  heißt die zu  $P$  gehörende VERTEILUNGSFUNKTION.

Ist umgekehrt  $F : \mathbb{R} \rightarrow [0, 1]$  eine Funktion mit obigen Eigenschaften, so wird durch

$$P([a, b]) = F(b) - F(a)$$

ein Wahrscheinlichkeitsmaß auf  $\mathbb{R}$  definiert. Die gegebene Funktion  $F$  ist dann die zugehörige Verteilungsfunktion. In diesem Sinne besteht eine eindeutige Beziehung zwischen Wahrscheinlichkeitsmaßen auf  $\mathbb{R}$  und Verteilungsfunktionen.

Eine Funktion  $f : \mathbb{R} \rightarrow [0, \infty)$  mit

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

heißt DICHTE. Sie definiert durch

$$F(x) = \int_{-\infty}^x f(t) dt$$

eine Verteilungsfunktion und damit ein Wahrscheinlichkeitsmaß  $P$  auf  $\mathbb{R}$ . Die Funktion  $f$  heißt dann die zu  $P$  gehörende Dichte.

Nicht jedes Wahrscheinlichkeitsmaß auf  $\mathbb{R}$  besitzt eine Dichte. Aber Wahrscheinlichkeitsmaße mit Dichte sind für die Anwendungen besonders wichtig.

**XII.1.5. Gleichverteilung auf einem Intervall.** Seien  $a, b \in \mathbb{R}$  mit  $a < b$ . Die Funktion  $f : \mathbb{R} \rightarrow [0, \infty)$  mit

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{falls } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

ist eine Dichte. Sie heißt Dichte der GLEICHVERTEILUNG. Das zugehörige Wahrscheinlichkeitsmaß ordnet jedem Intervall  $I$  seinen relativen Anteil an  $[a, b]$  zu.

**XII.1.6. Exponentialverteilung.** Für jedes  $\lambda > 0$  wird durch

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

eine Dichte definiert. Sie heißt Dichte der EXPONENTIALVERTEILUNG.

**XII.1.7. Normalverteilung.** Die Funktion

$$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

heißt Dichte der NORMALVERTEILUNG mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ . Ist

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

die Funktion aus Abschnitt XI.5.1 (S. 38), so ist

$$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right).$$

Mit

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

und der Substitutionsregel (vgl. Abschnitt V.2.3 (S. 174, Teil I)) folgt daher

$$\int_a^b \varphi_{\mu, \sigma^2}(x) dx = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

**XII.1.8. Produktdichten.** Werden  $n$  Telexperimente durch Dichten  $f_1, \dots, f_n$  beschrieben, verwendet man

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

als Dichte für die Verteilung auf  $\mathbb{R}^n$ , die die unabhängige Hintereinanderausführung der Telexperimente beschreibt.

## XII.2. Zufallsvariable und ihre Momente

**XII.2.1. Messbare Funktionen.** Seien  $\Omega$  und  $\Omega'$  zwei beliebige Mengen und  $\mathcal{A}$  und  $\mathcal{A}'$  zwei beliebige  $\sigma$ -Algebren auf  $\Omega$  bzw.  $\Omega'$ . Eine Abbildung  $f : \Omega \rightarrow \Omega'$  heißt MESSBAR, wenn für jede Menge  $A' \in \mathcal{A}'$  das Urbild

$$f^{-1}(A') = \{x \in \Omega : f(x) \in A'\}$$

unter  $f$  in  $\mathcal{A}$  enthalten ist.

Man beachte, dass der Begriff der Messbarkeit von den  $\sigma$ -Algebren  $\mathcal{A}$  und  $\mathcal{A}'$  abhängt.

**BEISPIEL XII.2.1.** Sei  $\Omega$  eine endliche Menge und  $\mathcal{A} = \mathcal{P}(\Omega)$ . Dann ist für jede Menge  $\Omega'$  und jede  $\sigma$ -Algebra  $\mathcal{A}'$  auf  $\Omega'$  jede Abbildung  $f : \Omega \rightarrow \Omega'$  messbar. Aus diesem Grunde haben wir den Begriff der Messbarkeit in Kapitel XI nicht benötigt.

**BEISPIEL XII.2.2.** Sei  $\Omega = \mathbb{R}^m$ ,  $\Omega' = \mathbb{R}^n$  und  $\mathcal{A}$  und  $\mathcal{A}'$  die  $\sigma$ -Algebren der Borelmengen. Dann ist jede stetige Abbildung  $f : \Omega \rightarrow \Omega'$  messbar. Ist speziell  $m = 1$ , so ist jede stückweise stetige Funktion (vgl. Abschnitt V.1.1 (S. 165, Teil I)) messbar.

Sind  $f : \Omega \rightarrow \Omega'$  und  $g : \Omega' \rightarrow \Omega''$  messbar, dann ist auch die Komposition  $g \circ f : \Omega \rightarrow \Omega''$  messbar. Sind  $f_1, f_2 : \Omega \rightarrow \mathbb{R}$  messbar und  $\alpha_1, \alpha_2 \in \mathbb{R}$ , so sind auch  $\alpha_1 f_1 + \alpha_2 f_2$ ,  $f_1 \cdot f_2$ ,  $\min\{f_1, f_2\}$  und  $\max\{f_1, f_2\}$  messbar.

**XII.2.2. Zufallsvariable.** Seien  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $\Omega'$  eine beliebige Menge und  $\mathcal{A}'$  eine  $\sigma$ -Algebra auf  $\Omega'$ . Eine messbare Funktion  $X : \Omega \rightarrow \Omega'$  heißt ( $\Omega'$ -wertige) ZUFALLSVARIABLE. Jede Zufallsvariable  $X : \Omega \rightarrow \Omega'$  definiert durch

$$P_X(A') = P(X \in A')$$

für  $A' \in \mathcal{A}'$  ein Wahrscheinlichkeitsmaß auf  $\Omega'$ . Ist speziell  $\Omega' = \mathbb{R}$ , so lässt sich dieses Wahrscheinlichkeitsmaß durch seine Verteilungsfunktion (vgl. Abschnitt XII.1.4 (S. 53))

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x)$$

beschreiben.  $F_X$  heißt die VERTEILUNGSFUNKTION der reellwertigen Zufallsvariablen  $X$ . Besitzt  $F_X$  eine Dichte  $f$ , so heißt diese die DICHTE der Zufallsvariablen.

Seien nun  $X : \Omega \rightarrow \mathbb{R}$  eine reellwertige Zufallsvariable und  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  eine streng monoton wachsende, differenzierbare Funktion. Dann ist  $Y = \varphi \circ X$  auch eine reellwertige Zufallsvariable.  $X$  habe die Dichte  $f$ . Dann folgt aus der Substitutionsregel (vgl. Abschnitt V.2.3 (S. 174, Teil I)) für jedes  $y \in \mathbb{R}$

$$F_Y(y) = P(\varphi \circ X \leq y)$$

$$\begin{aligned}
&= P(X \leq \varphi^{-1}(y)) \quad \text{wegen der Monotonie} \\
&= \int_{-\infty}^{\varphi^{-1}(y)} f(x) dx \quad |x = \varphi^{-1}(t), dx = \frac{1}{\varphi'(\varphi^{-1}(t))} dt \\
&= \int_{-\infty}^y f(\varphi^{-1}(t)) \frac{1}{\varphi'(\varphi^{-1}(t))} dt.
\end{aligned}$$

Also hat  $\varphi \circ X$  die Dichte

$$g(y) = \frac{f(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))}.$$

Mit den gleichen Argumenten folgt, dass für eine streng monoton fallende, differenzierbare Funktion  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  die Zufallsvariable  $\psi \circ X$  die Dichte

$$g(y) = \frac{f(\psi^{-1}(y))}{|\psi'(\psi^{-1}(y))|}$$

hat. Speziell folgt:

- Hat  $X$  die Dichte  $f$ , so hat  $X + a$  die Dichte

$$g(y) = f(y - a).$$

- Ist  $c \neq 0$ , hat  $cX$  die Dichte

$$g(y) = \frac{1}{|c|} f\left(\frac{y}{c}\right)$$

**XII.2.3. Unabhängigkeit.** Seien  $\Omega_i, i \in I$ , beliebige Mengen und  $\mathcal{A}_i$   $\sigma$ -Algebren auf den  $\Omega_i$ , sowie  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X_i : \Omega \rightarrow \Omega_i$  Zufallsvariable. Dann heißen die  $X_i$  UNABHÄNGIG, wenn für alle  $A_i \in \mathcal{A}_i$  die Ereignisse  $\{X_i \in A_i\}$  unabhängig sind.

Diese Definition ist eine direkte Verallgemeinerung derjenigen aus Abschnitt XI.3.3 (S. 24). Für die praktische Rechnung ist sie aber wenig geeignet. Für den Spezialfall reellwertiger Zufallsvariablen mit Dichten hilft folgendes Ergebnis:

Die reellwertigen Zufallsvariablen  $X_1, \dots, X_n$  mit Dichten  $f_1, \dots, f_n$  sind genau dann unabhängig, wenn  $X = (X_1, \dots, X_n)$  die Produktdichte

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

besitzt.

Seien  $X_1$  und  $X_2$  zwei reellwertige Zufallsvariable mit Dichten  $f_1$  und  $f_2$ . Dann folgt aus Obigem, dass  $X_1 + X_2$  die Dichte

$$(f_1 * f_2)(u) = \int_{-\infty}^{\infty} f_1(u - v) f_2(v) dv$$

besitzt. Der Ausdruck  $f_1 * f_2$  heißt FALTUNG der Dichten  $f_1$  und  $f_2$ .

BEISPIEL XII.2.3. Betrachte zwei unabhängige, reellwertige Zufallsvariable  $X_1$  und  $X_2$ , die  $N(\mu_i, \sigma_i^2)$ -verteilt sind,  $i = 1, 2$ . Wir behaupten, dass ihre Summe  $X_1 + X_2$  dann  $N(\mu, \sigma^2)$ -verteilt ist mit

$$\mu = \mu_1 + \mu_2$$

und

$$\sigma^2 = \sigma_1^2 + \sigma_2^2.$$

Wir müssen also für alle  $u \in \mathbb{R}$  die Gleichung

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(u - \mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - v - \mu_1)^2}{2\sigma_1^2}\right) \exp\left(-\frac{(v - \mu_2)^2}{2\sigma_2^2}\right) dv \end{aligned}$$

beweisen. Diese ist äquivalent zu

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left[ \frac{(u - v - \mu_1)^2}{\sigma_1^2} + \frac{(v - \mu_2)^2}{\sigma_2^2} \right. \right. \\ & \quad \left. \left. - \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \right] \right) dv \end{aligned}$$

$$= 1.$$

Wir betrachten den Term in eckigen Klammern und erhalten mit der Abkürzung

$$\begin{aligned} x &= u - \mu_1 - \mu_2 \\ y &= v - \mu_2 \end{aligned}$$

die Beziehung

$$\begin{aligned} & \frac{(u - v - \mu_1)^2}{\sigma_1^2} + \frac{(v - \mu_2)^2}{\sigma_2^2} - \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{(x - y)^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} - \frac{x^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{x^2}{\sigma_1^2} - \frac{2xy}{\sigma_1^2} + \frac{y^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} - \frac{x^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x^2 - \frac{1}{\sigma_1^2} 2xy + \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 \sigma_2^2} y^2 \\ &= \left( \frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} x - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} y \right)^2. \end{aligned}$$

Daher liefert die Variablentransformation

$$t = \frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} (u - \mu_1 - \mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} (v - \mu_2)$$

$$dt = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} dv$$

für alle  $a < b$  und alle  $u \in \mathbb{R}$

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} \int_a^b \exp\left(-\frac{1}{2} \left[ \frac{(u-v-\mu_1)^2}{\sigma_1^2} + \frac{(v-\mu_2)^2}{\sigma_2^2} - \frac{(u-\mu_1-\mu_2)^2}{\sigma_1^2 + \sigma_2^2} \right]\right) dv \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}(u-\mu_1-\mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2}(b-\mu_2)}^{\frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}(u-\mu_1-\mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2}(a-\mu_2)} \exp\left(-\frac{1}{2}t^2\right) dt \\ &\xrightarrow[\substack{a \rightarrow -\infty \\ b \rightarrow \infty}]{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt \\ &= 1. \end{aligned}$$

Dies beweist die Behauptung.

**XII.2.4. Erwartungswert.** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X : \Omega \rightarrow \mathbb{R}$  eine reellwertige Zufallsvariable.

Falls  $X$  nur endlich viele Werte  $x_1, \dots, x_n$  annimmt, können wir  $X$  wie in Abschnitt XI.3.4 (S. 24) einen Erwartungswert durch

$$E(X) = \sum_{i=1}^n x_i P(X = x_i)$$

zuordnen.

Wir wollen diese Einschränkung gerne fallen lassen. Dazu nehmen wir zunächst an, dass  $X$  beschränkt ist, d.h. es gibt ein  $R > 0$  mit  $|X(\omega)| \leq R$  für alle  $\omega \in \Omega$ . Für  $k \in \mathbb{Z}$  und  $n \in \mathbb{N}^*$  definieren wir dann

$$\begin{aligned} A_{k,n} &= \left\{ \omega \in \Omega : \frac{k}{n} \leq X(\omega) < \frac{k+1}{n} \right\}, \\ \chi_{A_{k,n}}(\omega) &= \begin{cases} 1 & \text{falls } \omega \in A_{k,n}, \\ 0 & \text{sonst,} \end{cases} \\ X_n &= \sum_{k \in \mathbb{Z}} \frac{k}{n} \chi_{A_{k,n}}. \end{aligned}$$

Da  $X$  beschränkt ist, nimmt  $X_n$  nur endlich viele Werte an und es ist

$$E(X_n) = \sum_{k \in \mathbb{Z}} \frac{k}{n} P(A_{k,n}).$$

Man kann zeigen, dass die Folge  $(E(X_n))_{n \in \mathbb{N}}$  konvergiert. Den Grenzwert nennen wir Erwartungswert von  $X$  und bezeichnen ihn mit

$$E(X) = \lim_{n \rightarrow \infty} E(X_n).$$

Falls  $X$  die stückweise stetige Dichte  $f$  besitzt, erhalten wir für jedes  $n$

$$\begin{aligned} E(X_n) &= \sum_{k \in \mathbb{Z}} \frac{k}{n} P(A_{k,n}) \\ &= \sum_{k \in \mathbb{Z}} \frac{k}{n} \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x) dx. \end{aligned}$$

Man kann zeigen, dass dieser Ausdruck für  $n \rightarrow \infty$  gegen

$$\int_{-\infty}^{\infty} x f(x) dx$$

konvergiert. Daher erhalten wir in diesem Fall

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Wenn  $X$  nicht beschränkt ist, benutzen wir diese Identität als Definition des Erwartungswertes:

Die reellwertige Zufallsvariable  $X$  auf  $\Omega$  besitze die stückweise stetige Dichte  $f$ . Dann ist der ERWARTUNGSWERT von  $X$  definiert durch

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

sofern das uneigentliche Integral existiert.

BEISPIEL XII.2.4.  $X$  sei  $N(\mu, \sigma^2)$ -verteilt. Für  $a < 0 < b$  folgt

$$\begin{aligned} & \int_a^b x \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \left| t = \frac{x-\mu}{\sigma}, dt = \frac{1}{\sigma} dx \right. \\ &= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} (\sigma t + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \underbrace{t e^{-\frac{1}{2}t^2}}_{=-\frac{d}{dt} e^{-\frac{1}{2}t^2}} dt + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\ &= \frac{\sigma}{\sqrt{2\pi}} \left[ e^{-\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{b-\mu}{\sigma}\right)^2} \right] + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\ & \xrightarrow[a \rightarrow -\infty]{b \rightarrow \infty} \mu. \end{aligned}$$

Also hat  $X$  den Erwartungswert  $\mu$ .

BEISPIEL XII.2.5.  $X$  sei exponentiell verteilt mit Parameter  $\lambda$ . Für  $a > 0$  gilt

$$\begin{aligned} \int_0^a x \underbrace{\lambda e^{-\lambda x}}_{=-\frac{d}{dx} e^{-\lambda x}} dx &= -x e^{-\lambda x} \Big|_{x=0}^{x=a} + \int_0^a e^{-\lambda x} dx \\ &= -a e^{-\lambda a} - \frac{1}{\lambda} e^{-\lambda a} + \frac{1}{\lambda} \\ &\xrightarrow{a \rightarrow \infty} \frac{1}{\lambda}. \end{aligned}$$

Da  $X(x) = 0$  ist für  $x < 0$ , folgt

$$E(X) = \frac{1}{\lambda}.$$

Offensichtlich gilt für den Erwartungswert

$$E(aX + bY) = aE(X) + bE(Y),$$

wobei  $X, Y$  reellwertige Zufallsvariablen und  $a, b$  reelle Zahlen sind.

**XII.2.5. Varianz.** Sei  $X$  eine reellwertige Zufallsvariable auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  mit einem endlichem Erwartungswert  $E(X)$ . Die VARIANZ von  $X$  ist definiert durch

$$\text{Var}(X) = E((X - E(X))^2),$$

sofern dieser Erwartungswert existiert. Die STANDARDABWEICHUNG ist dann definiert durch

$$\sigma = \sqrt{\text{Var}(X)}.$$

Wie in Abschnitt XI.3.5 (S. 26) gilt

$$\text{Var}(X) = E(X^2) - E(X)^2.$$

Hat daher  $X$  die stückweise stetige Dichte  $f$ , ergibt sich

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - \left[ \int_{-\infty}^{\infty} x f(x) dx \right]^2.$$

BEISPIEL XII.2.6. Sei  $X$   $N(\mu, \sigma^2)$ -verteilt. Für  $a < 0 < b$  erhalten wir wie in Beispiel XII.2.4 (S. 59)

$$\int_a^b x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \left| t = \frac{x-\mu}{\sigma}, dt = \frac{1}{\sigma} dx \right.$$

$$\begin{aligned}
 &= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} (\sigma t + \mu)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} t^2 e^{-\frac{1}{2}t^2} dt \\
 &\quad + \frac{2\sigma\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} t e^{-\frac{1}{2}t^2} dt \\
 &\quad + \frac{\mu^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} t(-e^{-\frac{1}{2}t^2}) \Big|_{t=\frac{a-\mu}{\sigma}}^{t=\frac{b-\mu}{\sigma}} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
 &\quad + \frac{2\sigma\mu}{\sqrt{2\pi}} (-e^{-\frac{1}{2}t^2}) \Big|_{t=\frac{a-\mu}{\sigma}}^{t=\frac{b-\mu}{\sigma}} \\
 &\quad + \frac{\mu^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
 &\xrightarrow[\substack{a \rightarrow -\infty \\ b \rightarrow \infty}]{} \sigma^2 + \mu^2.
 \end{aligned}$$

Da  $E(X) = \mu$  ist, folgt

$$\text{Var}(X) = \sigma^2.$$

BEISPIEL XII.2.7.  $X$  sei exponentiell verteilt mit Parameter  $\lambda$ . Für  $a > 0$  erhalten wir

$$\begin{aligned}
 \int_0^a x^2 \lambda e^{-\lambda x} dx &= x^2 (-e^{-\lambda x}) \Big|_{x=0}^{x=a} + 2 \int_0^a x e^{-\lambda x} dx \\
 &= -a^2 e^{-\lambda a} + \frac{2}{\lambda} x (-e^{-\lambda x}) \Big|_{x=0}^{x=a} + \frac{2}{\lambda} \int_0^a e^{-\lambda x} dx \\
 &= -a^2 e^{-\lambda a} - \frac{2}{\lambda} a e^{-\lambda a} - \frac{2}{\lambda^2} e^{-\lambda a} + \frac{2}{\lambda^2} \\
 &\xrightarrow{a \rightarrow \infty} \frac{2}{\lambda^2}.
 \end{aligned}$$

Wegen  $E(X) = \frac{1}{\lambda}$  ergibt sich

$$\text{Var}(X) = \frac{1}{\lambda^2}.$$

Tabelle XII.2.1 fasst die Ergebnisse der vorigen Beispiele zusammen.

Wie in Abschnitt XI.3.5 (S. 26) gilt die Rechenregel

$$\text{Var}(aX + b) = a^2 \text{Var}(X),$$

TABELLE XII.2.1. Erwartungswert und Varianz der Normal- und Exponentialverteilung

Verteilung	Erwartungswert	Varianz
$N(\mu, \sigma^2)$	$\mu$	$\sigma^2$
Exponential $\lambda$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

wobei  $X$  eine reellwertige Zufallsvariable und  $a, b$  reelle Zahlen sind. Ebenso gilt

$$X_1, \dots, X_n \text{ unabhängig}$$

$$\implies \text{Var}(X_1 + \dots + X_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n).$$

### XII.3. Schätzverfahren

**XII.3.1. Maximum-Likelihood Schätzung.** Es werde eine Zufallsvariable  $X$  mit Werten in  $\mathbb{R}^n$  beobachtet. Die Verteilung von  $X$  hänge von einem unbekanntem Parameter  $\vartheta \in \Theta$  ab. Wir nennen sie  $P_\vartheta$ .  $P_\vartheta$  habe für jedes  $\vartheta \in \Theta$  eine Dichte  $f(\cdot; \vartheta)$ . Dann ist für jedes  $x \in \mathbb{R}^n$   $P_\vartheta(x) = 0$ . Wir können also nicht, wie im diskreten Fall, aus der Betrachtung von  $P_\vartheta(x)$  Schätzer ableiten. Statt dessen definieren wir nun die Likelihood-Funktion durch die Dichte

$$L_x(\vartheta) = f(x; \vartheta).$$

Wie in Abschnitt XI.4.3 (S. 33) setzen wir

$$\mathcal{L}_x(\vartheta) = \ln L_x(\vartheta).$$

BEISPIEL XII.3.1. Die Zufallsvariablen  $X_1, \dots, X_n$  seien unabhängig und jeweils  $N(\mu, \sigma^2)$ -verteilt. Dann ist  $\vartheta = (\mu, \sigma)$ . Die Dichte von  $X_i$  ist

$$f_i(x; \vartheta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Daher hat  $X = (X_1, \dots, X_n)$  an der Stelle  $x = (x_1, \dots, x_n)$  die Produktdichte

$$\begin{aligned} f(x; \vartheta) &= \prod_{i=1}^n f_i(x; \vartheta) \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right). \end{aligned}$$

Der Maximum-Likelihood Schätzer  $\hat{\vartheta} = \hat{\vartheta}(x)$  ist wieder der Parameterwert, der  $L_x(\vartheta)$  bzw.  $\mathcal{L}_x(\vartheta)$  maximiert. Aus obiger Darstellung der Dichte erhalten wir

$$\mathcal{L}_x(\vartheta) = -n \ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

und somit

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}_x(\vartheta) &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \\ &= -\frac{n\mu}{\sigma^2} + \frac{1}{\sigma^2} \sum_{i=1}^n x_i, \\ \frac{\partial}{\partial \sigma} \mathcal{L}_x(\vartheta) &= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2. \end{aligned}$$

Wir müssen nun drei Fälle unterscheiden:

1.  $\mu$  IST UNBEKANNT, ABER  $\sigma^2 = \sigma_0^2$  IST BEKANNT: Dann ist  $\Theta = \{(\mu, \sigma_0) : \mu \in \mathbb{R}\}$ , und wir müssen eine Nullstelle  $\hat{\mu}$  von  $\frac{\partial}{\partial \mu} \mathcal{L}_x(\mu, \sigma_0)$  finden. Aus der Formel für  $\frac{\partial}{\partial \mu} \mathcal{L}_x$  folgt sofort

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Indem wir die zweite Ableitung bilden, sehen wir, dass  $\hat{\mu}$  wirklich ein Maximum ist.

2.  $\mu = \mu_0$  IST BEKANNT, ABER  $\sigma^2$  IST UNBEKANNT: Jetzt ist  $\Theta = \{(\mu_0, \sigma) : \sigma > 0\}$ , und wir müssen eine Nullstelle  $\hat{\sigma}$  von  $\frac{\partial}{\partial \sigma} \mathcal{L}_x(\mu_0, \sigma)$  finden. Aus der Formel für  $\frac{\partial}{\partial \sigma} \mathcal{L}_x$  ergibt sich

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2.$$

Wieder folgt durch Bilden der zweiten Ableitung, dass  $\hat{\sigma}^2$  tatsächlich ein Maximum ist.

3.  $\mu$  UND  $\sigma^2$  SIND BEIDE UNBEKANNT: Die Gleichungen

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}_x(\mu, \sigma) &= -\frac{1}{\sigma^2} \left[ n\mu - \sum_{i=1}^n x_i \right] \\ &= 0, \\ \frac{\partial}{\partial \sigma} \mathcal{L}_x(\mu, \sigma) &= \frac{1}{\sigma^3} \left[ -n\sigma^2 + \sum_{i=1}^n (x_i - \mu)^2 \right] \end{aligned}$$

$$= 0$$

für einen kritischen Punkt liefern die Lösungen

$$\begin{aligned}\hat{\mu} &= \bar{x}, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.\end{aligned}$$

Durch Bilden der Hesse-Matrix sehen wir, dass diese Werte wirklich das Maximum liefern.

Der Schätzer  $\hat{\mu}$  und der Schätzer  $\hat{\sigma}^2$  bei bekanntem Erwartungswert  $\mu_0$  sind erwartungstreu. Der Schätzer  $\hat{\sigma}^2$  im Fall 3 dagegen ist nicht erwartungstreu. Analog zu Abschnitt XI.4.4 (S. 34) liefert in diesem Fall

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

einen erwartungstreuen Schätzer für die Varianz.

**XII.3.2. Die Methode der kleinsten Quadrate.** Oft stellt sich das Problem, eine Gerade, eine Parabel oder eine andere „einfache“ Funktion einer gegebenen Menge von Messwerten anzupassen. Zum Beispiel kann eine Größe  $y$  in Abhängigkeit von einer Größe  $x$  gemessen worden sein und es liegen nun  $n$  Messergebnisse  $(x_1, y_1), \dots, (x_n, y_n)$  vor. Wenn die Messergebnisse relativ gut auf einer Geraden liegen, können wir einen linearen Zusammenhang der beobachteten Größen vermuten, der durch Messfehler  $z_i$  gestört ist. Dann wäre

$$y_i = \alpha + \beta x_i + z_i \quad , i = 1, \dots, n,$$

mit zu bestimmenden Parametern  $\alpha$  und  $\beta$ .

Allgemeiner nehmen wir an, dass  $\vartheta_1, \dots, \vartheta_p$  unbekannte Parameter sind und dass für bekannte Funktionen  $\varphi_i$  der gemessene Wert bei der  $i$ -ten Messung von der Form ist

$$y_i = \varphi_i(\vartheta_1, \dots, \vartheta_p) + z_i$$

mit einem Messfehler  $z_i$ . Im Beispiel der Geraden ist

$$\begin{aligned}p &= 2, \\ \vartheta_1 &= \alpha, \\ \vartheta_2 &= \beta, \\ \varphi_i(\vartheta_1, \vartheta_2) &= \vartheta_1 + \vartheta_2 x_i.\end{aligned}$$

Die METHODE DER KLEINSTEN QUADRATE besteht darin, die Parameter  $\vartheta_1, \dots, \vartheta_p$  so zu bestimmen, dass die Größe

$$Q = \sum_{i=1}^n [y_i - \varphi_i(\vartheta_1, \dots, \vartheta_p)]^2$$

minimal wird.

Dieser Ansatz kann ad hoc ohne Statistik formuliert werden und wird häufig auch so angewandt. Wir wollen nun zeigen, dass dieser Ansatz statistisch fundiert ist. Dazu nehmen wir an, dass die Messfehler  $z_i$  Realisierungen von unabhängigen  $N(0, \sigma^2)$ -verteilten Zufallsvariablen  $Z_1, \dots, Z_n$  sind. Dann sind die  $y_i$  Realisierungen von unabhängigen Zufallsvariablen  $Y_1, \dots, Y_n$ , die  $N(\varphi_i(\vartheta_1, \dots, \vartheta_p), \sigma^2)$ -verteilt sind. Daher hat  $Y = (Y_1, \dots, Y_n)$  die Dichte

$$\left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - \varphi_i(\vartheta_1, \dots, \vartheta_p)]^2 \right).$$

Diese Dichte ist genau dann maximal, wenn obige Größe  $Q$  minimal ist. Die Methode der kleinsten Quadrate ist also gerade der Maximum-Likelihood Schätzer für die Parameter  $\vartheta_1, \dots, \vartheta_p$ .

BEISPIEL XII.3.2 (REGRESSIONSGERADE). Wir wollen eine Gerade  $y = \alpha + \beta x$  mit der Methode der kleinsten Quadrate an Messdaten  $(x_1, y_1), \dots, (x_n, y_n)$  anpassen. Dann ist

$$Q = Q(\alpha, \beta) = \sum_{i=1}^n (y_i - \beta x_i - \alpha)^2.$$

Die Gleichungen für einen kritischen Punkt  $\hat{\alpha}, \hat{\beta}$  lauten

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} Q(\hat{\alpha}, \hat{\beta}) = -2 \sum_{i=1}^n (y_i - \hat{\beta} x_i - \hat{\alpha}) \\ 0 &= \frac{\partial}{\partial \beta} Q(\hat{\alpha}, \hat{\beta}) = -2 \sum_{i=1}^n (y_i - \hat{\beta} x_i - \hat{\alpha}) x_i. \end{aligned}$$

Zur Abkürzung setzen wir

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i, & \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i, \\ \overline{x^2} &= \frac{1}{n} \sum_{i=1}^n x_i^2, & \overline{xy} &= \frac{1}{n} \sum_{i=1}^n x_i y_i. \end{aligned}$$

Nach Multiplikation mit  $-\frac{1}{2n}$  nehmen dann obige Bestimmungsgleichungen die Form an

$$\begin{aligned} 0 &= \bar{y} - \widehat{\beta}\bar{x} - \widehat{\alpha} \\ 0 &= \overline{xy} - \widehat{\beta}\overline{xx} - \widehat{\alpha}\bar{x}. \end{aligned}$$

Dies liefert die Lösung

$$\begin{aligned} \widehat{\beta} &= \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{xx} - \bar{x}\bar{x}}, \\ \widehat{\alpha} &= \bar{y} - \widehat{\beta}\bar{x}. \end{aligned}$$

Die Gerade  $y = \widehat{\alpha} + \widehat{\beta}x$  heißt **REGRESSIONSGERADE**.  
Mit den Abkürzungen

$$\begin{aligned} s_{xy} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ s_{xx} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \end{aligned}$$

erhält man wegen

$$\begin{aligned} s_{xy} &= \frac{n}{n-1} \overline{xy} - \frac{n}{n-1} \bar{x}\bar{y}, \\ s_{xx} &= \frac{n}{n-1} \overline{xx} - \frac{n}{n-1} \bar{x}\bar{x} \end{aligned}$$

die alternative Darstellung

$$\widehat{\beta} = \frac{s_{xy}}{s_{xx}}.$$

**XII.3.3. Median.** Wenn in einem schweizer Bergdorf fünfzig Einheimische und fünf zugezogene Millionäre leben, ist es für die Einheimischen wenig befriedigend, wenn man ihnen erklärt, das durchschnittliche Einkommen in diesem Dorf sei hoch. Dieses Beispiel zeigt, dass der Erwartungswert manchmal kein guter Maßstab für das „Zentrum“ einer Verteilung ist. Einige wenige „Ausreißer“ können ihn stark verfälschen. Man betrachtet daher auch andere Maßzahlen. Die bekannteste davon ist der Median.

Ist  $Z$  eine reellwertige Zufallsvariable, so heißt *jede* Zahl  $\mu_m$  mit

$$P(Z \geq \mu_m) \geq \frac{1}{2}$$

und

$$P(Z \leq \mu_m) \geq \frac{1}{2}$$

ein MEDIAN von  $Z$ .

Man beachte, dass der Median  $\mu_m$  nicht notwendig eindeutig bestimmt ist. Eine Mehrdeutigkeit tritt genau dann auf, wenn es ein Intervall  $[a, b]$  gibt mit  $a < b$  und  $P(Z \geq b) = \frac{1}{2}$  und  $P(Z \leq a) = \frac{1}{2}$ . Dann ist jede Zahl in dem Intervall  $[a, b]$  ein Median. Gibt es dagegen eine Zahl  $c$  mit  $P(Z \geq c+t) = P(Z \leq c-t)$  für alle  $t > 0$ , so ist  $c$  der eindeutige Median von  $Z$ . In diesem Fall heißt  $Z$  symmetrisch, und es ist  $c = E(Z)$ .

Für  $n$  Messwerte  $x_1, \dots, x_n$  können wir den EMPIRISCHEN MEDIAN  $\hat{\mu}_m$  wie folgt ermitteln:

Bezeichne mit  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$  die der Größe nach umgeordneten  $x_i$ . Dann ist

$$\hat{\mu}_m = \begin{cases} x_{(\frac{n+1}{2})} & \text{falls } n \text{ ungerade,} \\ \frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}] & \text{falls } n \text{ gerade.} \end{cases}$$

BEISPIEL XII.3.3. Für  $x_1 = 3.1, x_2 = 4.5, x_3 = 4.5, x_4 = 2.8, x_5 = 2.6, x_6 = 9.8$  erhalten wir  $x_{(1)} = 2.6, x_{(2)} = 2.8, x_{(3)} = 3.1, x_{(4)} = 4.5, x_{(5)} = 4.5, x_{(6)} = 9.8$  und  $\hat{\mu}_m = \frac{1}{2}[x_{(3)} + x_{(4)}] = \frac{1}{2}[3.1 + 4.5] = 3.8$ . Der Mittelwert ist  $\bar{x} = 4.55$ .

## XII.4. Tests

**XII.4.1. Vorbemerkungen.** Wir beobachten eine, im allgemeinen vektorwertige, Zufallsvariable  $X$ , deren Verteilung einer Familie  $\{P_\vartheta : \vartheta \in \Theta\}$  angehört.  $\Theta$  ist die disjunkte Vereinigung der Mengen  $H$  und  $K$ , der Hypothese und der Alternative. Aufgrund des beobachteten Wertes  $x$  von  $X$  soll entschieden werden, ob der Parameter  $\vartheta$  der zu  $X$  gehörenden Verteilung in der Menge  $H$  liegt oder nicht.

Hierzu bilden wir den LIKELIHOOD-QUOTIENTEN

$$q(x) = \frac{\sup\{L_x(\vartheta) : \vartheta \in K\}}{\sup\{L_x(\vartheta) : \vartheta \in H\}}.$$

Offensichtlich ist  $0 \leq q(x)$  für alle  $x$ .  $q(x) \approx 0$  spricht für die Hypothese,  $q(x) \gg 0$  spricht für die Alternative.

Ein TEST ist nun eine messbare Abbildung  $\varphi$  des Wertebereiches  $\mathcal{X}$  von  $X$  in  $[0, 1]$ . Wird  $x$  beobachtet, besagt  $\varphi(x) = 1$ , dass die Hypothese verworfen werden soll, und  $\varphi(x) = 0$ , dass sie angenommen werden soll. Im Fall  $0 < \varphi(x) < 1$  soll ein zusätzliches Zufallsexperiment mit Wahrscheinlichkeit  $\varphi(x)$  zur Verwerfung führen.

Ein Test  $\varphi$  heißt LIKELIHOOD-QUOTIENTEN TEST, wenn es ein  $c > 0$  gibt mit den Eigenschaften

- $q(x) > c \implies \varphi(x) = 1$  und
- $q(x) < c \implies \varphi(x) = 0$ .

**XII.4.2. Der  $t$ -Test.** Wir beobachten Zufallsvariable  $X_1, \dots, X_n$ , die unabhängig und  $N(\mu, \sigma^2)$ -verteilt seien mit unbekanntem  $(\mu, \sigma^2)$ . Für ein gegebenes  $\mu_0$  sei zu testen, ob  $\mu = \mu_0$  ist oder nicht. Dann ist

$$\begin{aligned}\vartheta &= (\mu, \sigma) \\ \Theta &= \{(\mu, \sigma) : \mu \in \mathbb{R}, \sigma > 0\} \\ H &= \{(\mu_0, \sigma) : \sigma > 0\} \\ K &= \{(\mu, \sigma) : \mu \neq \mu_0, \sigma > 0\}.\end{aligned}$$

Die Dichte  $f(x; \vartheta)$  von  $X = (X_1, \dots, X_n)$  ist gemäß Abschnitt XII.3.1 (S. 62)

$$f(x; \vartheta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right).$$

Aus Stetigkeitsgründen ist

$$\sup\{f(x; \vartheta) : \vartheta \in K\} = \sup\{f(x; \vartheta) : \vartheta \in \Theta\}.$$

Damit ergibt sich aus Abschnitt XII.3.1 (S. 62), dass das Supremum an der Stelle

$$\begin{aligned}\hat{\mu} &= \bar{x} \\ &= \frac{1}{n} \sum_{i=1}^n x_i, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\end{aligned}$$

angenommen wird. Ebenso folgt, dass das Supremum  $\sup\{f(x; \vartheta) : \vartheta \in H\}$  an der Stelle

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2$$

angenommen wird. Daher ist der Likelihood-Quotient

$$\begin{aligned}q(x) &= \frac{f(x; \hat{\mu}, \hat{\sigma})}{f(x; \mu_0, \tilde{\sigma})} \\ &= \left(\frac{\tilde{\sigma}}{\hat{\sigma}}\right)^n \exp\left(-\frac{1}{2\hat{\sigma}^2} \underbrace{\sum_{i=1}^n (x_i - \hat{\mu})^2}_{=n\hat{\sigma}^2} + \frac{1}{2\tilde{\sigma}^2} \underbrace{\sum_{i=1}^n (x_i - \mu_0)^2}_{=n\tilde{\sigma}^2}\right) \\ &= \left(\frac{\tilde{\sigma}}{\hat{\sigma}}\right)^n.\end{aligned}$$

Ist  $\varphi$  irgendein Likelihood-Quotienten Test, gibt es daher ein  $c > 0$  mit

$$\begin{aligned}\varphi(x) &= 1 && \text{auf } \left\{ \left( \frac{\tilde{\sigma}}{\hat{\sigma}} \right)^n > c \right\} \\ \varphi(x) &= 0 && \text{auf } \left\{ \left( \frac{\tilde{\sigma}}{\hat{\sigma}} \right)^n < c \right\}.\end{aligned}$$

Mit

$$c' = c^{2/n}$$

ist dies äquivalent zu

$$\begin{aligned}\varphi(x) &= 1 && \text{auf } \left\{ \frac{\tilde{\sigma}^2}{\hat{\sigma}^2} > c' \right\} \\ \varphi(x) &= 0 && \text{auf } \left\{ \frac{\tilde{\sigma}^2}{\hat{\sigma}^2} < c' \right\}.\end{aligned}$$

Es ist

$$\begin{aligned}\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} &= \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= 1 + \frac{(\bar{x} - \mu_0)^2}{\hat{\sigma}^2}.\end{aligned}$$

Wir definieren daher

$$T(x) = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s(x)}$$

mit

$$s(x)^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Dann folgt

$$\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} = 1 + \frac{1}{n-1} |T(x)|^2.$$

Also gibt es für jeden Likelihood-Quotienten Test  $\varphi$  ein  $t > 0$  mit

$$\varphi(x) = \begin{cases} 1 & \text{falls } |T(x)| > t, \\ 0 & \text{falls } |T(x)| < t. \end{cases}$$

Unser Testproblem ist damit auf die Bestimmung der Verteilung von  $T(x)$  zurückgeführt. Diese Verteilung heißt  $t$ -VERTEILUNG mit  $n - 1$  Freiheitsgraden oder kurz  $t_{n-1}$ -VERTEILUNG. Die zugehörige Dichte ist gegeben durch

$$h_{n-1}(x) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}}.$$

Dabei ist  $\Gamma$  die Eulersche Gammafunktion aus Abschnitt V.4.3 (S. 192, Teil I). Für  $n \rightarrow \infty$  konvergiert  $h_{n-1}(x)$  gegen die Dichte  $\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  der Standard-Normalverteilung.

Damit lautet der  $t$ -TEST ZUM NIVEAU  $\alpha$  AUF DIE HYPOTHESE EINER  $N(\mu_0, \sigma^2)$ -VERTEILTEN STICHPROBE:

Berechne die Größen

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s(x) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$T(x) = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s(x)}.$$

Bestimme aus einer Tabelle der  $t_{n-1}$ -Verteilung das  $(1 - \frac{\alpha}{2})$ -QUANTIL  $t_{n-1, 1-\frac{\alpha}{2}}$  aus der Bedingung

$$\int_{-\infty}^{t_{n-1, 1-\frac{\alpha}{2}}} h_{n-1}(x) dx = 1 - \frac{\alpha}{2}.$$

Falls  $|T(x)| \leq t_{n-1, 1-\frac{\alpha}{2}}$  ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEISPIEL XII.4.1. An 15 Glasfaserplatten einer bestimmten Stärke wird die Wärmeleitfähigkeit gemessen. Es ergibt sich der Mittelwert  $\bar{x} = 17.1$  und der Wert  $s^2 = 0.36$ . Zu testen ist die Hypothese  $\mu = 17$  zu dem Niveau  $\alpha = 0.1$ . Aus einer Tabelle der  $t_{14}$ -Verteilung lesen wir

das 0.95-Quantil  $t_{14,0.95} = 1.76$  ab. Aus den gegebenen Daten ergibt sich

$$T(x) = \frac{\sqrt{14}(17.1 - 17)}{\sqrt{0.36}} \approx 0.624.$$

Also kann die Hypothese akzeptiert werden.

**XII.4.3. Der  $\chi^2$ -Test.** Wir betrachten folgendes Testproblem: Es werden  $n$  unabhängige, gleichartige Teilerperimente ausgeführt. Jedes hat  $r \geq 2$  mögliche Ausgänge, und der  $i$ -te Ausgang hat die Wahrscheinlichkeit  $p_i$ . Der Parameter  $\vartheta = (p_1, \dots, p_r)$  ist unbekannt. Für einen gegebenen Wahrscheinlichkeitsvektor  $\pi = (\pi_1, \dots, \pi_r)$  ist zu testen, ob  $\vartheta = \pi$  ist.

Die Wahrscheinlichkeit, die Häufigkeiten  $x_1, \dots, x_r$  zu beobachten, ist nach der Multinomialverteilung

$$P_{\vartheta}(x) = \binom{n}{x_1, \dots, x_r} p_1^{x_1} \cdot \dots \cdot p_r^{x_r}.$$

Der Vektor  $x = (x_1, \dots, x_r)$  muss dabei natürlich der Bedingung  $x_1 + \dots + x_r = n$  genügen. Die Likelihood-Funktion ist  $L_x(\vartheta) = P_{\vartheta}(x)$ . Bei der Ermittlung des Maximums muss die Nebenbedingung  $p_1 + \dots + p_r = 1$  berücksichtigt werden. Dies führt auf das Gleichungssystem

$$\frac{\partial}{\partial p_1} \ln L_x(\vartheta) = \frac{\partial}{\partial p_2} \ln L_x(\vartheta) = \dots = \frac{\partial}{\partial p_r} \ln L_x(\vartheta).$$

Als Lösung ergibt sich der Schätzer  $\hat{p}_i(x) = \frac{x_i}{n}$ . Damit folgt für den Likelihood-Quotienten

$$q(x) = \frac{P_{\hat{\vartheta}}(x)}{P_{\pi}(x)} \approx \left( \frac{\hat{p}_1}{\pi_1} \cdot \dots \cdot \frac{\hat{p}_r}{\pi_r} \right)^{-\frac{1}{2}} \exp \left( \frac{1}{2} \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i} \right).$$

Wir definieren daher

$$V^2(x) = \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i}.$$

Gilt die Hypothese  $\vartheta = \pi$ , so ist nach dem Gesetz der großen Zahlen  $\hat{p}_i(x) = \frac{x_i}{n}$  mit Wahrscheinlichkeit nahe 1 gleich  $\pi_i$ . Daher ist in diesem Fall

$$q(x) \approx \exp\left(\frac{1}{2}V^2(x)\right).$$

Man kann zeigen, dass  $V^2(x)$  durch die sog.  $\chi^2$ -VERTEILUNG mit  $r - 1$  Parametern, kurz  $\chi_{n-1}^2$ -VERTEILUNG approximiert wird. Diese hat die Dichte

$$g_{r-1}(x) = \begin{cases} \frac{1}{2^{\frac{r-1}{2}} \Gamma(\frac{r-1}{2})} x^{\frac{r-1}{2}-1} e^{-\frac{x}{2}} & \text{falls } x > 0, \\ 0 & \text{sonst.} \end{cases}$$

Damit lautet der  $\chi^2$ -TEST ZUM NIVEAU  $\alpha$ :

Berechne die Größe

$$V^2(x) = \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i}.$$

Bestimme aus einer Tabelle der  $\chi_{r-1}^2$ -Verteilung das  $(1-\alpha)$ -QUANTIL  $\chi_{r-1,1-\alpha}^2$  aus der Bedingung

$$\int_0^{\chi_{r-1,1-\alpha}^2} g_{r-1}(x) dx = 1 - \alpha.$$

Falls  $V^2(x) < \chi_{r-1,1-\alpha}^2$  ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEISPIEL XII.4.2. Man vermutet, dass die Blütenfarben rot, rosa und weiß einer Rosenart mit den Wahrscheinlichkeiten  $\frac{1}{4}$ ,  $\frac{1}{2}$  und  $\frac{1}{4}$  vererbt werden. Unter einer Auswahl von 320 Nachkommen beobachtet man 102 rote, 156 rosa und 62 weiße Blüten. Die Hypothese soll mit dem  $\chi^2$ -Test zum Niveau  $\alpha = 0.1$  getestet werden. Es ist

$$\pi = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$

und

$$(\hat{p}_1, \hat{p}_2, \hat{p}_3) = \left(\frac{102}{320}, \frac{156}{320}, \frac{62}{320}\right).$$

Wir erhalten den Wert

$$V^2(x) = \frac{(102 - 80)^2}{80} + \frac{(156 - 160)^2}{160} + \frac{(62 - 80)^2}{80} \\ \approx 10.2.$$

Aus einer Tabelle der  $\chi^2$ -Verteilung mit 2 Parametern erhalten wir das 0.9-Quantil

$$\chi_{2,0.9}^2 = 4.61.$$

Daher muss die Hypothese verworfen werden.

BEISPIEL XII.4.3. Es wird behauptet, dass auf einer Pferderennbahn die Startposition einen Einfluss auf die Gewinnwahrscheinlichkeit hat. In 144 Rennen hatten die Sieger die Startposition 1, 2, ..., 8 mit

den Häufigkeiten 29, 19, 18, 25, 17, 10, 15, 11. Wir wollen die Hypothese, dass alle Startpositionen die gleiche Gewinnwahrscheinlichkeit haben, zu dem Niveau  $\alpha = 0.05$  testen. Aus einer Tabelle der  $\chi^2$ -Verteilung mit 7 Freiheitsgraden lesen wir das 0.95-Quantil

$$\chi_{7,0.95}^2 = 20.28$$

ab. Da  $n\pi_i = 16$  ist für alle  $i$ , erhalten wir mit den beobachteten Werten

$$\begin{aligned} V^2(x) &= \frac{1}{16} \{(29 - 16)^2 + (19 - 16)^2 + (18 - 16)^2 + (25 - 16)^2 \\ &\quad + (17 - 16)^2 + (10 - 16)^2 + (15 - 16)^2 + (11 - 16)^2\} \\ &\approx 20.375. \end{aligned}$$

Also wird die Hypothese verworfen.

Die  $\chi^2$ -Verteilung wird auch für das Testen der Varianz einer Normalverteilung genutzt. Der  $\chi^2$ -TEST ZUM NIVEAU  $\alpha$  AUF DIE HYPOTHESE EINER  $N(\mu, \sigma_0^2)$ -VERTEILTEN STICHPROBE lautet:

Berechne die Größen

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s(x)^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$T(x) = \frac{(n-1)s(x)^2}{\sigma_0^2}.$$

Bestimme aus einer Tabelle der  $\chi_{n-1}^2$ -Verteilung die  $(1 - \frac{\alpha}{2})$ - und  $\frac{\alpha}{2}$ - QUANTILE  $\chi_{n-1, 1-\frac{\alpha}{2}}^2$  und  $\chi_{n-1, \frac{\alpha}{2}}^2$  aus den Bedingungen

$$\int_{-\infty}^{\chi_{n-1, 1-\frac{\alpha}{2}}^2} g_{n-1}(x) dx = 1 - \frac{\alpha}{2},$$

$$\int_{-\infty}^{\chi_{n-1, \frac{\alpha}{2}}^2} g_{n-1}(x) dx = \frac{\alpha}{2}.$$

Falls

$$\chi_{n-1, \frac{\alpha}{2}}^2 \leq T(x) \leq \chi_{n-1, 1-\frac{\alpha}{2}}^2$$

ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEMERKUNG XII.4.4. Bei dem oben dargestellten Test wird auf  $\sigma^2 = \sigma_0^2$  getestet. Falls auf  $\sigma^2 < \sigma_0^2$  getestet werden soll, lautet das Testkriterium

$$T(x) \leq \chi_{n-1, 1-\alpha}^2.$$

Falls auf  $\sigma^2 > \sigma_0^2$  getestet werden soll, lautet das Testkriterium

$$T(x) \geq \chi_{n-1, \alpha}^2.$$

BEISPIEL XII.4.5. Nach Angaben des Herstellers eines bestimmten PKW-Typs ist der Benzinverbrauch im Stadtverkehr annähernd normalverteilt mit Erwartungswert 9.5 l/100km und Streuung 2.5 l/100km. Zur Überprüfung dieser Angaben testet eine Verbraucherorganisation 25 PKWs und misst einen Durchschnittsverbrauch von 9.9 l/100km mit einer Streuung von 3.5 l/100km. Das Niveau des Tests soll  $\alpha = 0.05$  sein.

Für den Erwartungswert ergeben diese Daten

$$\begin{aligned} T(x) &= \frac{\sqrt{25}(9.9 - 9.5)}{3.5} \\ &\approx 0.571 \\ t_{24, 0.975} &= 2.064. \end{aligned}$$

Daher ist die Aussage über den erwarteten Durchschnittsverbrauch zu akzeptieren.

Für die Varianz liefern diese Daten

$$\begin{aligned} T(x) &= \frac{24 \cdot 3.5^2}{2.5^2} \\ &\approx 47.04 \\ \chi_{24, 0.025}^2 &= 12.40 \\ \chi_{24, 0.975}^2 &= 39.36. \end{aligned}$$

Wegen  $T(x) > \chi_{24, 0.975}^2$  ist die Aussage über die Streuung der Verbrauchswerte zu verwerfen.

## KAPITEL XIII

### Fourier-Analysis

#### XIII.1. Trigonometrische Polynome und Reihen

**XIII.1.1. Periodische Funktionen.** Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  oder  $f : \mathbb{R} \rightarrow \mathbb{C}$  heißt PERIODISCH MIT PERIODE  $T$  oder kurz  $T$ -PERIODISCH, wenn  $T > 0$  ist und  $f(t + T) = f(t)$  gilt für alle  $t \in \mathbb{R}$ . Statt  $2\pi$ -periodisch sagen wir häufig auch kurz PERIODISCH.

BEISPIEL XIII.1.1. Eine konstante Funktion ist  $T$ -periodisch für jedes  $T > 0$ . Die Funktionen  $\sin t$ ,  $\cos t$ ,  $e^{it}$ ,  $\alpha \cos(nt) + \beta \sin(nt)$  sind alle  $2\pi$ -periodisch.

Es gelten folgende Rechenregel für  $T$ -periodische Funktionen:

- Mit  $T$  ist auch  $nT$  für jedes  $n \in \mathbb{N}^*$  eine Periode.
- Sind  $f, g$   $T$ -periodisch und  $\alpha, \beta \in \mathbb{C}$ , so ist auch  $\alpha f + \beta g$   $T$ -periodisch.
- Ist  $f$   $T$ -periodisch, so gilt

$$\int_a^{a+T} f(t) dt = \int_0^T f(t) dt \quad \text{für jedes } a \in \mathbb{R}.$$

Durch die Substitution  $x = \frac{2\pi}{T}t = \omega t$  mit  $\omega = \frac{2\pi}{T}$  wird eine  $T$ -periodische Funktion  $f$  in eine  $2\pi$ -periodische Funktion transformiert.

Ist  $g$  eine gegebene Funktion auf dem Intervall  $[0, T]$  kann man sie auf drei Arten zu einer  $T$ -periodischen Funktion  $f$  auf  $\mathbb{R}$  fortsetzen (vgl. Abbildung XIII.1.1):

- DIREKTE FORTSETZUNG ( $T$ -PERIODISCH): Setze

$$f(t) = g(t - nT), \text{ falls } nT \leq t < (n + 1)T, n \in \mathbb{Z}.$$

- GERADE FORTSETZUNG ( $2T$ -PERIODISCH): Setze

$$\tilde{f}(t) = \begin{cases} g(t) & \text{für } 0 \leq t \leq T \\ g(-t) & \text{für } -T \leq t < 0 \end{cases}$$

und definiere  $f$  durch

$$f(t) = \tilde{f}(t - 2nT) \text{ falls } (2n - 1)T \leq t < (2n + 1)T, \\ n \in \mathbb{Z}.$$

- UNGERADE FORTSETZUNG ( $2T$ -PERIODISCH): Setze

$$\widehat{f}(t) = \begin{cases} g(t) & \text{für } 0 \leq t \leq T \\ -g(-t) & \text{für } -T \leq t < 0 \end{cases}$$

und definiere  $f$  durch

$$f(t) = \widehat{f}(t - 2nT) \text{ falls } (2n - 1)T \leq t < (2n + 1)T, \\ n \in \mathbb{Z}.$$

Ist  $g$  stetig, so ist im ersten Fall  $f$  genau dann stetig, wenn  $g(0) = g(T)$  ist. Im zweiten Fall ist  $f$  ohne weitere Zusatzbedingung an  $g$  stetig. Im dritten Fall schließlich ist  $f$  genau dann stetig, wenn  $g(0) = g(T) = 0$  ist.

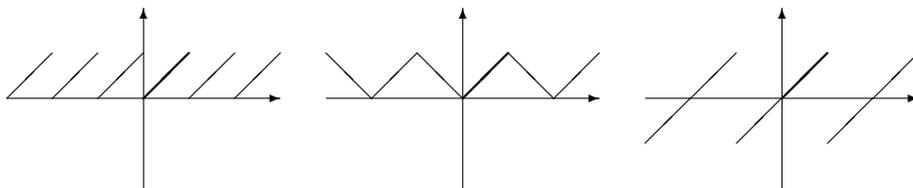


ABBILDUNG XIII.1.1. Periodische Fortsetzungen der Funktion  $g(x) = x$  auf  $[0, 1]$

**XIII.1.2. Trigonometrische Polynome.** Eine Funktion der Form

$$f(t) = \sum_{k=-N}^N c_k e^{ik\omega t}$$

mit  $\omega \neq 0$ ,  $c_{-N}, \dots, c_N \in \mathbb{C}$  und  $|c_{-N}| + |c_N| \neq 0$  nennt man ein **TRIGONOMETRISCHES POLYNOM** vom **GRAD**  $N$ . Es ist  $\frac{2\pi}{\omega}$ -periodisch. Wegen

$$e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t)$$

kann jedes trigonometrische Polynom vom Grad  $N$  in der Form

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^N [a_k \cos(k\omega t) + b_k \sin(k\omega t)]$$

dargestellt werden. Für die Koeffizienten gelten die Umrechnungsformeln

$$\begin{aligned} c_0 &= \frac{1}{2}a_0, & c_k &= \frac{1}{2}(a_k - ib_k), & c_{-k} &= \frac{1}{2}(a_k + ib_k) \\ a_0 &= 2c_0, & a_k &= c_k + c_{-k}, & b_k &= i(c_k - c_{-k}). \end{aligned}$$

BEISPIEL XIII.1.2. Die Funktion

$$\begin{aligned} f(t) &= \sum_{k=-N}^N e^{ikt} \\ &= 1 + 2 \cos x + \dots + 2 \cos(Nx) \end{aligned}$$

ist ein trigonometrisches Polynom vom Grad  $N$ . Aus der geometrischen Summenformel

$$\sum_{k=0}^{n-1} q^k = \begin{cases} \frac{q^n - 1}{q - 1} & \text{falls } q \neq 1 \\ n & \text{falls } q = 1 \end{cases}$$

folgt die Darstellung

$$\sum_{k=-N}^N e^{ikt} = \begin{cases} 2N + 1 & \text{falls } t = 2n\pi, n \in \mathbb{Z}, \\ \frac{\sin((N + \frac{1}{2})t)}{\sin(\frac{1}{2}t)} & \text{sonst.} \end{cases}$$

Sei  $f$  ein trigonometrisches Polynom vom Grad  $N$  und  $T = \frac{2\pi}{\omega}$  seine Periode. Dann gelten folgende Eigenschaften:

- $f$  hat in  $[0, T)$  höchstens  $2N$  Nullstellen.
- $f(t) = 0$  für alle  $t \in \mathbb{R} \iff c_{-N} = \dots = c_N = 0$ .
- $f$  ist reellwertig  $\iff c_k = \overline{c_{-k}}$  für  $k = 0, \dots, N$  (insb. ist  $c_0 \in \mathbb{R}$ ).
- Für  $-N \leq k \leq N$  und  $0 \leq n \leq N$  ist

$$\begin{aligned} c_k &= \frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt, \\ a_n &= \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt, \\ b_n &= \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt. \end{aligned}$$

Die vierte Eigenschaft folgt aus der Orthogonalitätsbeziehung (vgl. Beispiel V.2.6 (S. 175, Teil I))

$$\int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt = \begin{cases} T & \text{falls } k = \ell, \\ 0 & \text{falls } k \neq \ell. \end{cases}$$

**XIII.1.3. Trigonometrische Reihen.** Einen Ausdruck der Form

$$\sum_{k \in \mathbb{Z}} c_k e^{ik\omega t}$$

nennt man eine TRIGONOMETRISCHE REIHE. Die  $N$ -te PARTIALSUMME ist gegeben durch

$$S_N(t) = \sum_{k=-N}^N c_k e^{ik\omega t}.$$

Falls die Folge  $(S_N(t))_{N \in \mathbb{N}}$  für jedes  $t \in [0, T]$  konvergiert, wird durch ihren Grenzwert eine  $T$ -periodische Funktion  $f$  definiert. Man sagt dann,  $f$  werde durch die trigonometrische Reihe  $\sum_{k \in \mathbb{Z}} c_k e^{ik\omega t}$  dargestellt.

BEISPIEL XIII.1.3. Die Partialsummen der trigonometrischen Reihe  $\sum_{k \in \mathbb{Z}} e^{ikt}$  sind für kein  $t \in \mathbb{R}$  konvergent. Denn gemäß Beispiel XIII.1.2 erhalten wir für  $\frac{t}{2\pi} \in \mathbb{Z}$  die Folge  $(2N+1)_{N \in \mathbb{N}}$  und für  $\frac{t}{2\pi} \notin \mathbb{Z}$  die Folge  $(\frac{\sin((N+\frac{1}{2})t)}{\sin(\frac{1}{2}t)})_{N \in \mathbb{N}}$ . Beide Folgen sind aber nicht konvergent.

## XIII.2. Fourier-Reihen

**XIII.2.1. Die Fourier-Reihe einer Funktion.** Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  sei stückweise stetig und  $T$ -periodisch,  $T > 0$ . Die Zahlen

$$\hat{f}_k = \frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt, \quad \text{mit } \omega = \frac{2\pi}{T}, k \in \mathbb{Z}$$

heißen die (KOMPLEXEN) FOURIER-KOEFFIZIENTEN von  $f$ . Die mit ihnen gebildete trigonometrische Reihe

$$\sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t}$$

heißt die FOURIER-REIHE von  $f$ . Wir schreiben

$$f \sim \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t},$$

um zu betonen, dass die rechts stehende trigonometrische Reihe die Fourier-Reihe der links stehenden Funktion ist. Hierdurch *ist nichts über die Konvergenz der Reihe ausgesagt*.

Wie in Abschnitt XIII.1.2 (S. 76) ist

$$\sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t} = \frac{a_0}{2} + \sum_{n \in \mathbb{N} \setminus \{0\}} [a_n \cos(n\omega t) + b_n \sin(n\omega t)]$$

mit

$$\begin{aligned} a_n &= \hat{f}_n + \hat{f}_{-n} = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \\ b_n &= i(\hat{f}_n - \hat{f}_{-n}) = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt. \end{aligned}$$

Die Zahlen  $a_n, b_n$  heißen die REELLEN FOURIER-KOEFFIZIENTEN von  $f$ .

Die Fourier-Koeffizienten einer  $T$ -periodischen Schwingung  $f$  haben folgende Deutung:

- $\widehat{f}_0$ : Arithmetischer Mittelwert (Gleichspannungsanteil),
- $2\widehat{f}_1$ : Komplexe Amplitude der Grundschwingung,
- $2\widehat{f}_n$ : Komplexe Amplituden der  $n$ -ten Oberschwingung,  $n \geq 2$ ,
- $\frac{\sum_{n=2}^{\infty} |\widehat{f}_n|^2}{\sum_{n=1}^{\infty} |\widehat{f}_n|^2}$ : Oberschwingungsanteil (Klirrfaktor).

BEISPIEL XIII.2.1 (RECHTECKSCHWINGUNG). Definiere  $f : [0, 2\pi) \rightarrow \mathbb{R}$  durch

$$f(t) = \begin{cases} 1 & \text{für } 0 \leq t < \pi, \\ -1 & \text{für } \pi \leq t < 2\pi. \end{cases}$$

Dann folgt

$$\begin{aligned} a_n &= \frac{1}{\pi} \left\{ \int_0^{\pi} \cos(nt) dt - \int_{\pi}^{2\pi} \cos(nt) dt \right\} \\ &= 0 \\ b_n &= \frac{1}{\pi} \left\{ \int_0^{\pi} \sin(nt) dt - \int_{\pi}^{2\pi} \sin(nt) dt \right\} \\ &= \begin{cases} \frac{4}{n\pi} & \text{falls } n \text{ ungerade} \\ 0 & \text{falls } n \text{ gerade} \end{cases} \end{aligned}$$

Also lautet die Fourier-Reihe

$$\frac{4}{\pi} \sum_{n \in \mathbb{N}} \frac{\sin((2n+1)x)}{2n+1}.$$

**XIII.2.2. Rechenregeln.** Im Folgenden seien  $f, g : \mathbb{R} \rightarrow \mathbb{C}$  stückweise stetige,  $T$ -periodische Funktionen mit Fourier-Reihen

$$\begin{aligned} f &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_k e^{ik\omega t}, \\ g &\sim \sum_{k \in \mathbb{Z}} \widehat{g}_k e^{ik\omega t} \end{aligned}$$

und  $\omega = \frac{2\pi}{T}$ .

Wegen der Linearität des Integrals gilt für  $\alpha, \beta \in \mathbb{C}$

$$\begin{aligned} (\widehat{\alpha f + \beta g})_k &= \alpha \widehat{f}_k + \beta \widehat{g}_k \quad \text{für alle } k \in \mathbb{Z} \\ \alpha f + \beta g &\sim \sum_{k \in \mathbb{Z}} (\alpha \widehat{f}_k + \beta \widehat{g}_k) e^{ik\omega t}. \end{aligned}$$

Mit der Linearität der komplexen Konjugation und der Substitution  $t \mapsto -t$  folgt

$$\begin{aligned}\overline{f(t)} &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_{-k} e^{ik\omega t} \\ f(-t) &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_{-k} e^{ik\omega t}.\end{aligned}$$

Mit den Substitutionen  $t \mapsto ct$ ,  $c > 0$  und  $t \mapsto t + a$  ergibt sich

$$\begin{aligned}f(ct) &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_k e^{ik(\omega c)t} \\ f(t+a) &\sim \sum_{k \in \mathbb{Z}} (e^{ik\omega a} \widehat{f}_k) e^{ik\omega t}.\end{aligned}$$

Aus der Definition der Fourier-Koeffizienten folgt

$$e^{i\omega n t} f(t) \sim \sum_{k \in \mathbb{Z}} \widehat{f}_{k-n} e^{ik\omega t}.$$

Mittels partieller Integration ergibt sich

$$f'(t) \sim \sum_{k \in \mathbb{Z}} (ik\omega \widehat{f}_k) e^{ik\omega t}.$$

Das unbestimmte Integral einer  $T$ -periodischen Funktion  $f$  ist nur dann wieder  $T$ -periodisch, wenn der Mittelwert von  $f$  verschwindet, d.h.  $\widehat{f}_0 = 0$ . In diesem Fall gilt:

$$\int_0^t f(s) ds \sim -\frac{1}{T} \int_0^T s f(s) ds + \sum_{k \in \mathbb{Z} \setminus \{0\}} \left( \frac{1}{ik\omega} \widehat{f}_k \right) e^{ik\omega t}.$$

**XIII.2.3. Die Bessel-Ungleichung.** Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  sei stückweise stetig und  $T$ -periodisch. Für  $N \in \mathbb{N}^*$  bezeichnen wir mit  $S_N$  die  $N$ -te Partialsumme ihrer Fourier-Reihe

$$S_N(t) = \sum_{k=-N}^N \widehat{f}_k e^{ik\omega t}$$

$$= \frac{a_0}{2} + \sum_{k=1}^N [a_k \cos(k\omega t) + b_k \sin(k\omega t)].$$

Aus der Definition der Fourier-Koeffizienten folgt

$$\begin{aligned} \frac{1}{T} \int_0^T f(t) \overline{S_N(t)} dt &= \sum_{k=-N}^N \widehat{f}_k \underbrace{\frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt}_{=\widehat{f}_k} \\ &= \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Aus der Orthogonalitätsbeziehung

$$\frac{1}{T} \int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt = \begin{cases} 1 & \text{falls } k = \ell \\ 0 & \text{falls } k \neq \ell \end{cases}$$

aus Abschnitt XIII.1.2 (S. 76) ergibt sich

$$\begin{aligned} \frac{1}{T} \int_0^T S_N(t) \overline{S_N(t)} dt &= \sum_{k=-N}^N \sum_{\ell=-N}^N \widehat{f}_k \overline{\widehat{f}_\ell} \frac{1}{T} \int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt \\ &= \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Mit diesen beiden Ergebnissen erhalten wir

$$\begin{aligned} 0 &\leq \frac{1}{T} \int_0^T |f(t) - S_N(t)|^2 dt \\ &= \frac{1}{T} \int_0^T (f(t) - S_N(t)) \overline{(f(t) - S_N(t))} dt \\ &= \frac{1}{T} \int_0^T |f(t)|^2 dt - \frac{1}{T} \int_0^T f(t) \overline{S_N(t)} dt \\ &\quad - \underbrace{\frac{1}{T} \int_0^T \overline{f(t)} S_N(t) dt}_{=\overline{\int_0^T f(t) \overline{S_N(t)} dt}} + \frac{1}{T} \int_0^T S_N(t) \overline{S_N(t)} dt \\ &= \frac{1}{T} \int_0^T |f(t)|^2 dt - \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Also gilt für alle  $N \in \mathbb{N}^*$

$$\sum_{k=-N}^N |\widehat{f}_k|^2 \leq \frac{1}{T} \int_0^T |f(t)|^2 dt.$$

Dies beweist die sog. BESSEL-UNGLEICHUNG

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} [|a_n|^2 + |b_n|^2] = 2 \sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 \leq \frac{2}{T} \int_0^T |f(t)|^2 dt.$$

Da die Koeffizienten einer konvergenten Reihe eine Nullfolge bilden (vgl. Abschnitt VII.1.3 (S. 8, Teil II)), folgt hieraus das sog. RIEMANN-LEMMA:

$$\begin{aligned} \lim_{k \rightarrow \infty} |\widehat{f}_k| &= 0, & \lim_{k \rightarrow -\infty} |\widehat{f}_k| &= 0, \\ \lim_{n \rightarrow \infty} |a_n| &= 0, & \lim_{n \rightarrow \infty} |b_n| &= 0. \end{aligned}$$

Es gilt sogar folgende Verschärfung:

Die  $T$ -periodische Funktion  $f$  sei  $(m-1)$ -mal differenzierbar auf  $\mathbb{R}$  und  $f^{(m)}$  sei stückweise stetig auf  $[0, T]$ . Dann gibt es eine Zahl  $M > 0$  mit

$$|\widehat{f}_k| \leq \frac{M}{|k|^{m+1}} \text{ für alle } k \in \mathbb{Z} \setminus \{0\}.$$

**XIII.2.4. Konvergenz der Fourier-Reihe.** Sei wieder  $f : \mathbb{R} \rightarrow \mathbb{C}$  eine stückweise stetige,  $T$ -periodische Funktion. Es gilt folgende Verschärfung der Bessel-Ungleichung:

$$\sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 = \frac{1}{T} \int_0^T |f(t)|^2 dt.$$

Dies ist die sog. PARSEVAL-GLEICHUNG.

Aus ihr und dem Beweis der Bessel-Ungleichung folgt, dass die Fourier-Reihe im quadratischen Mittel gegen  $f$  konvergiert

$$\lim_{N \rightarrow \infty} \frac{1}{T} \int_0^T |f(t) - S_N(t)|^2 dt = 0.$$

Bezüglich der punktweisen Konvergenz ist die Situation komplizierter (vgl. Beispiele XIII.1.3 (S. 78) und XIII.2.1 (S. 79)). Es gilt:

Die Funktion  $f$  sei stückweise differenzierbar auf  $[0, T]$ . Dann konvergiert die Fourier-Reihe von  $f$  für jedes  $t \in \mathbb{R}$ . Mit  $\omega = \frac{2\pi}{T}$  ist

$$\sum_{k=-\infty}^{\infty} \widehat{f}_k e^{ik\omega t} = \frac{1}{2}[f(t+0) + f(t-0)].$$

BEISPIEL XIII.2.2. Betrachte die Funktion  $f(t) = t^2$  auf  $[-\pi, \pi]$  und setze sie  $2\pi$ -periodisch fort. Die so auf ganz  $\mathbb{R}$  definierte Funktion ist stetig und stückweise differenzierbar. Für die Fourier-Koeffizienten folgt

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \sin(nt) dt \\ &= 0 && \text{wg. Symmetrie} \\ a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \\ &= 2 \frac{\pi^2}{3} \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \cos(nt) dt \\ &= \frac{2}{\pi} \int_0^{\pi} t^2 \cos(nt) dt \\ &= \frac{2}{n\pi} t^2 \sin(nt) \Big|_{t=0}^{t=\pi} - \frac{4}{n\pi} \int_0^{\pi} t \sin(nt) dt \\ &= -\frac{4}{n\pi} \int_0^{\pi} t \sin(nt) dt \\ &= \frac{4}{n^2\pi} t \cos(nt) \Big|_{t=0}^{t=\pi} - \frac{4}{n^2\pi} \int_0^{\pi} \cos(nt) dt \\ &= \frac{4}{n^2} (-1)^{n-1}. \end{aligned}$$

Also gilt für alle  $t \in [-\pi, \pi]$

$$t^2 = \frac{\pi^2}{3} - 4 \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\cos(nt)}{n^2}.$$

Wegen

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} (t^2)^2 dt &= \frac{1}{\pi} \int_0^{\pi} t^4 dt \\ &= \frac{1}{5} \pi^4 \end{aligned}$$

folgt aus der Parseval-Gleichung

$$\begin{aligned}
 \sum_{n=1}^{\infty} \frac{1}{n^4} &= \frac{1}{16} \sum_{n=1}^{\infty} |a_n|^2 \\
 &= \frac{1}{16} \left\{ \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} |a_n|^2 - \frac{2}{9} \pi^4 \right\} \\
 &= \frac{2}{16} \sum_{k=-\infty}^{\infty} |\hat{f}_k|^2 - \frac{1}{72} \pi^4 \\
 &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{\pi} (t^2)^2 dt - \frac{1}{72} \pi^4 \\
 &= \frac{1}{40} \pi^4 - \frac{1}{72} \pi^4 \\
 &= \frac{\pi^4}{90}.
 \end{aligned}$$

**XIII.2.5. Anwendung auf gewöhnliche Differentialgleichungen.** Wir betrachten das Randwertproblem

$$\begin{aligned}
 a\ddot{x} + b\dot{x} + cx &= f(t) && \text{in } (0, T) \\
 x(0) &= x(T) \\
 \dot{x}(0) &= \dot{x}(T)
 \end{aligned}$$

mit  $a, b, c, \in \mathbb{R}$ ,  $a \neq 0$ , und einer stückweise stetigen Funktion  $f$ . Jede Lösung  $x$  dieses Randwertproblems kann  $T$ -periodisch festgesetzt werden und liefert eine zweimal differenzierbare  $T$ -periodische Funktion auf  $\mathbb{R}$  mit stückweise stetiger zweiter Ableitung. Daher gilt für alle  $t$

$$x(t) = \sum_{k=-\infty}^{\infty} \hat{x}_k e^{ik\omega t}$$

mit  $\omega = \frac{2\pi}{T}$ . Bilden wir die Fourier-Reihen der linken und der rechten Seite der gDgl, erhalten wir aufgrund der Rechenregeln für Fourier-Reihen

$$\begin{aligned}
 f &\sim \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t} \\
 a\ddot{x} + b\dot{x} + cx &\sim \sum_{k \in \mathbb{Z}} [a(ik\omega)^2 + b(ik\omega) + c] \hat{x}_k e^{ik\omega t}.
 \end{aligned}$$

Also muss für jede Lösung der gDgl gelten

$$\hat{x}_k = \frac{\hat{f}_k}{a(ik\omega)^2 + b(ik\omega) + c}.$$

Dies bestimmt die Fourier-Reihe der Lösung und damit die Lösung.

### XIII.3. Die Fourier-Transformation

**XIII.3.1. Definition.** Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  heißt **FOURIER-TRANSFORMIERBAR**, wenn für jedes  $\omega \in \mathbb{R}$  der Grenzwert

$$\int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt = \lim_{R \rightarrow \infty} \int_{-R}^R e^{-i\omega t} f(t) dt$$

existiert. In diesem Fall heißt die Funktion  $\mathcal{F}(f) : \mathbb{R} \rightarrow \mathbb{C}$  mit

$$\mathcal{F}(f)(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt$$

die **FOURIER-TRANSFORMIERTE** von  $f$ . Statt  $\mathcal{F}(f)$  schreibt man gelegentlich auch  $\hat{f}$ . Die **INVERSE FOURIER-TRANSFORMATION** einer Funktion  $F : \mathbb{R} \rightarrow \mathbb{C}$  ist definiert durch

$$\mathcal{F}^{-1}(F)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} F(\omega) d\omega.$$

Die Fourier-Transformation kann als Grenzübergang  $T \rightarrow \infty$  bei Bildung der Fourier-Reihen  $T$ -periodischer Funktionen gedeutet werden.

**BEISPIEL XIII.3.1.** Betrachte den Rechteckimpuls

$$f(t) = \begin{cases} 1 & \text{für } |t| \leq 1, \\ 0 & \text{für } |t| > 1. \end{cases}$$

Für jedes  $R > 1$  erhalten wir

$$\begin{aligned} \int_{-R}^R f(t) dt &= \int_{-1}^1 dt \\ &= 2 \end{aligned}$$

und für  $\omega \neq 0$

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_{-1}^1 e^{-i\omega t} dt \\ &= -\frac{1}{i\omega} e^{-i\omega t} \Big|_{t=-1}^{t=1} \\ &= -\frac{1}{i\omega} (e^{-i\omega} - e^{i\omega}) \\ &= 2 \frac{\sin(\omega)}{\omega}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \begin{cases} 2 & \text{für } \omega = 0, \\ 2 \frac{\sin(\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.2. Sei  $a > 0$  und

$$f(t) = \begin{cases} e^{-at} & \text{für } t \geq 0, \\ 0 & \text{für } t < 0. \end{cases}$$

Für  $R > 0$  und  $\omega \in \mathbb{R}$  erhalten wir

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_0^R e^{-(a+i\omega)t} dt \\ &= -\frac{1}{a+i\omega} e^{-(a+i\omega)t} \Big|_{t=0}^{t=R} \\ &= \frac{1}{a+i\omega} (1 - e^{-(a+i\omega)R}) \\ &\xrightarrow{R \rightarrow \infty} \frac{1}{a+i\omega}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \frac{1}{a+i\omega}, \quad \omega \in \mathbb{R}.$$

BEISPIEL XIII.3.3. Sei wieder  $a > 0$  und

$$f(t) = e^{-a|t|}.$$

Dann gilt für  $R > 0$

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_0^R e^{-(a+i\omega)t} dt + \int_{-R}^0 e^{(a-i\omega)t} dt \\ &= \frac{1}{a+i\omega} (1 - e^{-(a+i\omega)R}) + \frac{1}{a-i\omega} (1 - e^{-(a-i\omega)R}) \\ &\xrightarrow{R \rightarrow \infty} \frac{1}{a+i\omega} + \frac{1}{a-i\omega} \\ &= \frac{2a}{a^2 + \omega^2}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \frac{2a}{a^2 + \omega^2}, \quad \omega \in \mathbb{R}.$$

**XIII.3.2. Rechenregeln.** Im Folgenden sind  $f, g : \mathbb{R} \rightarrow \mathbb{C}$  stückweise stetige Funktionen mit

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty, \quad \int_{-\infty}^{\infty} |g(t)| dt < \infty.$$

Mit Hilfe der Rechenregeln für Integrale und partieller Integration kann man dann die folgenden Rechenregeln für die Fourier-Transformation beweisen.

$$\mathcal{F}(\alpha f + \beta g) = \alpha \mathcal{F}(f) + \beta \mathcal{F}(g) \quad (\text{Linearität}), \alpha, \beta \in \mathbb{C}$$

$$\mathcal{F}(\bar{f})(\omega) = \overline{\mathcal{F}(f)(-\omega)}$$

$$\mathcal{F}(f(ct))(\omega) = \frac{1}{|c|} \mathcal{F}(f)\left(\frac{\omega}{c}\right) \quad c \neq 0$$

$$\mathcal{F}(f(t-a))(\omega) = e^{-i\omega a} \mathcal{F}(f)(\omega) \quad a \in \mathbb{R}$$

$$\mathcal{F}(e^{i\Omega t} f(t))(\omega) = \mathcal{F}(f)(\omega - \Omega) \quad \Omega \in \mathbb{R}$$

$$\mathcal{F}(f')(\omega) = i\omega \mathcal{F}(f)(\omega)$$

$$\mathcal{F}(tf(t)) = i \frac{d}{d\omega} \mathcal{F}(f).$$

BEISPIEL XIII.3.4. Seien  $A \in \mathbb{R}$ ,  $T > 0$  und

$$g(t) = \begin{cases} A & \text{für } |t| \leq \frac{T}{2}, \\ 0 & \text{für } |t| > \frac{T}{2}. \end{cases}$$

Dann ist  $g(t) = Af(\frac{2}{T}t)$  mit  $f$  aus Beispiel XIII.3.1. Damit folgt aus den Rechenregeln und Beispiel XIII.3.1

$$\mathcal{F}(g)(\omega) = \begin{cases} AT & \text{für } \omega = 0, \\ 2A \frac{\sin(\frac{T}{2}\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.5. Sei  $a \in \mathbb{R}$ ,  $T > 0$  und

$$g(t) = \begin{cases} 1 & \text{für } |t-a| \leq T, \\ 0 & \text{für } |t-a| > T. \end{cases}$$

Dann ist  $g(t) = f(\frac{t-a}{T})$  mit  $f$  aus Beispiel XIII.3.1. Damit folgt aus den Rechenregeln und Beispiel XIII.3.1

$$\mathcal{F}(g)(\omega) = \begin{cases} 2T & \text{für } \omega = 0, \\ 2e^{-i\omega a} \frac{\sin(T\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.6. Für die Funktion

$$g(t) = \cos(\Omega t) e^{-|a|t}$$

mit  $\Omega \in \mathbb{R}$ ,  $a \in \mathbb{R} \setminus \{0\}$  gilt

$$g(t) = \frac{1}{2}(e^{i\Omega t} + e^{-i\Omega t})f(t)$$

mit  $f$  aus Beispiel XIII.3.3. Daher ist

$$\mathcal{F}(g)(\omega) = \frac{a}{a^2 + (\omega - \Omega)^2} + \frac{a}{a^2 + (\omega + \Omega)^2}.$$

BEISPIEL XIII.3.7. Für

$$f(t) = e^{-t^2}$$

gilt

$$\begin{aligned} tf(t) &= -\frac{1}{2} \frac{d}{dt}(e^{-t^2}) \\ &= -\frac{1}{2} f'(t). \end{aligned}$$

Damit folgt

$$\begin{aligned} -\frac{i}{2} \omega \mathcal{F}(f)(\omega) &= \mathcal{F}\left(-\frac{1}{2} f'\right)(\omega) \\ &= \mathcal{F}(tf(t))(\omega) \\ &= i \frac{d}{d\omega} \mathcal{F}(f)(\omega). \end{aligned}$$

Also ist  $\mathcal{F}(f)$  Lösung der gDgl (in  $\omega$ !)

$$\frac{dg}{d\omega} = -\frac{\omega}{2} g.$$

Mit der Methode der Trennung der Variablen (vgl. Abschnitt VI.2.1 (S. 217, Teil I)) folgt

$$\mathcal{F}(f)(\omega) = \mathcal{F}(f)(0) e^{-\frac{\omega^2}{4}}.$$

Weiter ist (vgl. Abschnitt XI.5.1 (S. 38))

$$\mathcal{F}(f)(0) = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}.$$

Daraus ergibt sich insgesamt

$$\mathcal{F}(e^{-t^2})(\omega) = \sqrt{\pi} e^{-\frac{\omega^2}{4}}.$$

**XIII.3.3. Existenz- und Eindeutigkeitsätze.** Wir nennen eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  absolut integrierbar, wenn sie stückweise stetig ist und das uneigentliche Integral  $\int_{-\infty}^{\infty} |f(t)| dt$  existiert.

Es gilt folgendes Kriterium für die Existenz der Fourier-Transformation:

Ist die Funktion  $f$  absolut integrierbar, so ist sie Fourier-transformierbar und es gilt die PARSEVAL-GLEICHUNG

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathcal{F}(f)(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |f(t)|^2 dt.$$

Es gilt folgendes Ergebnis für die inverse Fourier-Transformation:

Die Funktion  $f$  sei absolut integrierbar und stückweise differenzierbar. Dann gilt für alle  $t \in \mathbb{R}$

$$\frac{1}{2}[f(t+0) + f(t-0)] = \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R e^{i\omega t} \mathcal{F}(f)(\omega) d\omega.$$

Ist insbesondere

$$f(t) = \frac{1}{2}[f(t+0) + f(t-0)]$$

für alle  $t \in \mathbb{R}$ , so gilt

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f.$$



## KAPITEL XIV

# Partielle Differentialgleichungen

### XIV.1. Einführung

#### XIV.1.1. Beispiele.

BEISPIEL XIV.1.1 (MEMBRAN-, POISSONGLEICHUNG). Sei  $\Omega \subset \mathbb{R}^2$  eine offene, beschränkte Menge mit stückweise glattem Rand z.B. ein Kreissegment, das von einer dünnen Membran z.B. dem Trommelfell in seiner Ruhelage eingenommen wird. Auf die Membran wirke eine äußere Kraft  $f$ . Diese bewirkt eine Auslenkung  $u = u(x) = u(x_1, x_2)$  in vertikaler Richtung. Unter der Annahme, dass die Membran nicht dehnbar und die Auslenkung klein ist, folgt aus dem Prinzip von „Actio = Reactio“, dass die Auslenkung beschrieben wird durch

$$f = -\Delta u = -\frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} \quad \text{in } \Omega.$$

Dies ist eine PARTIELLE DIFFERENTIALGLEICHUNG (kurz PDGL), die sog. MEMBRAN- oder POISSONGLEICHUNG.

Falls die Membran am Rand  $\Gamma$  von  $\Omega$  eingespannt ist, gilt dort

$$u = 0 \quad \text{auf } \Gamma.$$

Dies ist eine RANDBEDINGUNG, die sog. (HOMOGENE) DIRICHLET-RANDBEDINGUNG.

Falls die Membran am Rand frei gelagert ist, gilt dort

$$\frac{\partial u}{\partial n} = 0 \quad \text{auf } \Gamma.$$

Dabei ist  $n$  der nach außen zeigende Einheitsnormalenvektor. Dies ist die sog. (HOMOGENE) NEUMANN-RANDBEDINGUNG.

BEISPIEL XIV.1.2 (PLATTEN-, BIHARMONISCHE GLEICHUNG). Wir ersetzen die Membran aus Beispiel XIV.1.1 durch eine dünne starre Platte und bezeichnen mit  $u$  die Auslenkung der Mittelebene. Dann folgt aus den gleichen physikalischen Prinzipien, dass  $u$  bestimmt wird durch

$$f = \Delta^2 u = \frac{\partial^4 u}{\partial x_1^4} + 2 \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} \quad \text{in } \Omega.$$

Dies ist eine pDgl vierter Ordnung, die sog. PLATTEN- oder BIHARMONISCHE GLEICHUNG.

Falls der Rand der Platte fest eingespannt ist, gilt zusätzlich die Randbedingung

$$u = \frac{\partial u}{\partial n} = 0 \quad \text{auf } \Gamma.$$

Ist der Rand dagegen frei gelagert, gilt die Randbedingung

$$u = \Delta u = 0 \quad \text{auf } \Gamma.$$

BEISPIEL XIV.1.3 (GASGLEICHUNG). Wir betrachten die rotationsfreie Strömung eines idealen, kompressiblen Gases. Aus der Rotationsfreiheit folgt für die Geschwindigkeit  $\mathbf{v}$  des Gases

$$\mathbf{v} = \nabla u$$

mit einem skalaren Potential  $u$ . Aus der Massenerhaltung folgt

$$\operatorname{div}(\rho \mathbf{v}) = 0,$$

wobei  $\rho = \rho(\mathbf{v})$  die Dichte des Gases ist. Da das Gas ideal ist, gilt die Zustandsgleichung

$$\rho(\mathbf{v}) = \left[ 1 - \frac{\gamma - 1}{2} |\mathbf{v}|^2 \right]^{\frac{1}{\gamma - 1}}$$

wobei  $\gamma > 1$  der spezifische Wärmekoeffizient ist. Insgesamt erfüllt damit das Potential  $u$  die pDgl

$$\operatorname{div} \left[ \left( 1 - \frac{\gamma - 1}{2} |\nabla u|^2 \right)^{\frac{1}{\gamma - 1}} \nabla u \right] = 0 \quad \text{in } \Omega.$$

Hinzu kommt die Randbedingung

$$u = u_0 \quad \text{auf } \Gamma$$

mit einer gegebenen Funktion  $u_0$ .

BEISPIEL XIV.1.4 (WÄRMELEITUNGSGLEICHUNG). Sei  $\Omega \subset \mathbb{R}^3$  eine offene beschränkte Menge mit stückweise glattem Rand, z.B. ein Zylinder. Die Funktion  $u(x, t) : \Omega \times [0, \infty] \rightarrow \mathbb{R}$  beschreibe die Temperatur zur Zeit  $t$  im Punkt  $x$  des Körpers  $\Omega$ . Wenn dieser einer äußeren

Wärmequelle  $f$  ausgesetzt ist, wird der Verlauf der Temperatur durch die pDgl

$$f = \frac{\partial u}{\partial t} - \Delta u = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} - \frac{\partial^2 u}{\partial x_3^2} \quad \text{in } \Omega \times (0, \infty)$$

beschrieben. Dies ist die sog. WÄRMELEITUNGSGLEICHUNG. Diese Gleichung ist zu ergänzen durch eine Information über die anfängliche Temperaturverteilung, d.h. durch eine ANFANGSBEDINGUNG

$$u(x, 0) = u_0(x) \quad \text{in } \Omega$$

mit einer gegebenen Funktion  $u_0$ .

Wenn der Rand des Körpers künstlich auf einer bestimmten zeitlich nicht notwendig konstanten Temperatur gehalten wird, gilt zusätzlich die Randbedingung

$$u(x, t) = g_D(x, t) \quad \text{auf } \Gamma \times (0, \infty).$$

Ist der Rand des Körpers dagegen isoliert, d.h. findet dort kein Wärmefluss statt, gilt stattdessen die Randbedingung

$$\frac{\partial u}{\partial n}(x, t) = 0 \quad \text{auf } \Gamma \times (0, \infty).$$

Sei nun  $u$  eine hinreichend oft differenzierbare Lösung der Wärmeleitungsgleichung zu  $f = 0$  und der Randbedingung  $u = 0$  auf  $\Gamma \times (0, \infty)$ . Dann folgt mit dem Integralsatz von Gauß (vgl. Abschnitt IX.5.5 (S. 124, Teil II))

$$\begin{aligned} 0 &= \int_{\Omega} \left\{ \frac{\partial u}{\partial t} - \Delta u \right\} u dx \\ &= \int_{\Omega} \underbrace{\frac{\partial u}{\partial t} u}_{=\frac{1}{2} \frac{\partial}{\partial t}(u^2)} dx - \int_{\Omega} \underbrace{\Delta u u}_{=\operatorname{div}(u \nabla u) - \nabla u \cdot \nabla u} dx \\ &= \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} u^2 dx}_{=\frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} u^2 dx \right)} - \underbrace{\int_{\Omega} \operatorname{div}(u \nabla u) dx}_{=\int_{\Gamma} u \frac{\partial u}{\partial n} ds} + \int_{\Omega} |\nabla u|^2 dx \\ &= \frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} u^2 dx \right) - \int_{\Gamma} \underbrace{u}_{=0} \frac{\partial u}{\partial n} ds + \int_{\Omega} |\nabla u|^2 dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} u^2 dx \right) + \underbrace{\int_{\Omega} |\nabla u|^2 dx}_{\geq 0} \\
&\geq \frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} u^2 dx \right).
\end{aligned}$$

Also ist die „Energie“  $E : (0, \infty) \rightarrow \mathbb{R}$  mit

$$E(t) = \frac{1}{2} \int_{\Omega} u(x, t)^2 dx$$

eine monoton fallende Funktion.

**BEISPIEL XIV.1.5 (GRUNDWASSERSTRÖMUNG).** Die Funktion  $u(x, t)$  beschreibe die räumliche und zeitliche Verteilung einer Flüssigkeit, z.B. Grundwasser in einem porösen Medium  $\Omega \subset \mathbb{R}^3$ . Dann wird  $u$  bestimmt durch die pDgl

$$\frac{\partial u}{\partial t} - \operatorname{div}(D(x, u)\nabla u) + \mathbf{k}(x, u) \cdot \nabla u = f \quad \text{in } \Omega \times (0, \infty).$$

Dabei ist  $D(x, z) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}^{3 \times 3}$  die matrixwertige DIFFUSIVITÄT und  $\mathbf{k}(x, z) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}^3$  die vektorwertige KONDUKTIVITÄT des Mediums. Die Funktion  $f$  beschreibt die Zufuhr (Quellen) bzw. Entnahme (Brunnen) von Flüssigkeit. Die pDgl ist zu ergänzen durch Anfangs- und Randbedingungen ähnlich wie in Beispiel XIV.1.4.

**BEISPIEL XIV.1.6 (WELLENGLEICHUNG).** Wir betrachten wie in Beispiel XIV.1.1 eine dünne Membran, versetzen sie aber jetzt durch eine zeitlich veränderliche äußere Kraft in Schwingung. Falls die Auslenkung klein ist, tritt keine Dämpfung durch innere Reibung auf. Die Auslenkung  $u(x, t)$  am Ort  $x$  zur Zeit  $t$  wird dann beschrieben durch die pDgl

$$f = \frac{\partial^2 u}{\partial t^2} - \Delta u = \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{x_1^2} - \frac{\partial^2 u}{\partial x_2^2} \quad \text{in } \Omega \times (0, \infty).$$

Dies ist die sog. WELLENGLEICHUNG. Zusätzlich gelten auf  $\Gamma \times (0, \infty)$  Randbedingungen wie in Beispiel XIV.1.1, je nachdem ob die Membran eingespannt oder frei gelagert ist. Schließlich muss noch der Anfangszustand des Systems beschrieben werden. Dies geschieht durch die Anfangsbedingungen

$$\begin{aligned}
u(x, 0) &= u_0(x) \quad \text{in } \Omega \\
\frac{\partial u}{\partial t}(x, 0) &= u_1(x) \quad \text{in } \Omega
\end{aligned}$$

mit gegebenen Funktionen  $u_0$  und  $u_1$ .

Betrachte nun speziell den Fall  $f = 0$  mit homogener Dirichlet-Randbedingung, d.h.  $u(x, t) = 0$  auf  $\Gamma \times (0, \infty)$ . Sei  $u$  eine hinreichend oft differenzierbare Lösung des pDgl. Dann folgt mit dem Gaußschen Integralsatz (vgl. Abschnitt V.5.5 (S. 200, Teil I))

$$\begin{aligned} 0 &= \int_{\Omega} \left\{ \frac{\partial^2 u}{\partial t^2} - \Delta u \right\} \frac{\partial u}{\partial t} dx \\ &= \int_{\Omega} \underbrace{\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t}}_{=\frac{1}{2} \frac{d}{dt} \left( \left( \frac{\partial u}{\partial t} \right)^2 \right)} dx - \int_{\Omega} \underbrace{\Delta u \frac{\partial u}{\partial t}}_{=\operatorname{div}(\nabla u \frac{\partial u}{\partial t}) - \nabla u \cdot \nabla \left( \frac{\partial u}{\partial t} \right)} dx \\ &= \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} \left( \left( \frac{\partial u}{\partial t} \right)^2 \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} \left( \frac{\partial u}{\partial t} \right)^2 dx \right)} - \underbrace{\int_{\Omega} \operatorname{div}(\nabla u \frac{\partial u}{\partial t}) dx}_{=\int_{\Gamma} \frac{\partial u}{\partial n} \frac{\partial u}{\partial t} dx} + \underbrace{\int_{\Omega} \nabla u \cdot \nabla \left( \frac{\partial u}{\partial t} \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} |\nabla u|^2 \right) dx} \\ &= \frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} \left( \frac{\partial u}{\partial t} \right)^2 dx \right) - \underbrace{\int_{\Gamma} \frac{\partial u}{\partial n} \frac{\partial u}{\partial t} dx}_{=0} + \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} \left( |\nabla u|^2 \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left( \int_{\Omega} |\nabla u|^2 dx \right)} \\ &= \frac{1}{2} \frac{d}{dt} \left\{ \int_{\Omega} \left( \frac{\partial u}{\partial t} \right)^2 + |\nabla u|^2 dx \right\}. \end{aligned}$$

Also bleibt die „Energie“  $E : (0, \infty) \rightarrow \mathbb{R}$  mit

$$E(t) = \frac{1}{2} \int_{\Omega} \left( \frac{\partial u}{\partial t} \right)^2 + |\nabla u|^2 dx$$

erhalten.

**XIV.1.2. Bezeichnungen.** Sei  $G \subset \mathbb{R}^n$  eine offene Menge und  $u : G \rightarrow \mathbb{R}$  eine hinreichend oft differenzierbare Funktion. Für  $k \geq 1$  bezeichnen wir zur Abkürzung mit  $D^k u$  den „Vektor“ aller partieller Ableitungen der Ordnung  $k$  von  $u$ :

$$D^k u = \left\{ \frac{\partial^k u}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} : \alpha_1, \dots, \alpha_n \in \mathbb{N}, \alpha_1 + \dots + \alpha_n = k \right\}.$$

Ein DIFFERENTIALOPERATOR  $m$ -TER ORDNUNG,  $m \geq 1$ , hat die Form

$$u \mapsto \mathcal{D}(x, u(x), Du(x), \dots, D^m u(x)) \quad , x \in G.$$

Es heißt QUASILINEAR, wenn er darstellbar ist als

$$\begin{aligned} \mathcal{D}(x, u(x), \dots, D^m u(x)) &= A(x, u(x), \dots, D^m u(x)) \\ &\quad + B(x, u(x), \dots, D^{m-1} u(x)) \end{aligned}$$

mit

$$A(x, u(x), \dots, D^m u(x)) = \sum_{\substack{\alpha_1, \dots, \alpha_n \in \mathbb{N} \\ \alpha_1 + \dots + \alpha_n = m}} a_{\alpha}(x, u(x), \dots, D^{m-1} u(x))$$

$$\frac{\partial^m}{x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u(x).$$

$A$  heißt dann der HAUPTTEIL des Differentialoperators. Der Differentialoperator  $\mathcal{D}$  heißt LINEAR, wenn er von der Form ist

$$\mathcal{D}(x, u(x), \dots, D^m u(x)) = \sum_{0 \leq \alpha_1 + \dots + \alpha_n \leq m} a_\alpha(x) \frac{\partial^{\alpha_1 + \dots + \alpha_n}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u(x).$$

Die Funktionen  $a_\alpha$  heißen dann die Koeffizienten des Differentialoperators. Der Differentialoperator heißt LINEAR MIT KONSTANTEN Koeffizienten, wenn er linear ist und die Koeffizienten  $a_\alpha$  konstante Funktionen sind. Ist  $\mathcal{D}$  ein linearer Differentialoperator, so ist die Zuordnung  $u \mapsto \mathcal{D}(u)$  eine lineare Abbildung.

Eine PARTIELLE DIFFERENTIALGLEICHUNG  $m$ -TER ORDNUNG, kurz PDGL  $m$ -TER ORDNUNG, ist eine Gleichung der Form

$$\mathcal{D}(u) = f \quad \text{in } G$$

mit einem Differentialoperator  $\mathcal{D}$   $m$ -ter Ordnung und einer gegebenen Funktion  $f : \Omega \rightarrow \mathbb{R}$ . Eine pDgl heißt quasilinear bzw. linear bzw. linear mit konstanten Koeffizienten, wenn der Differentialoperator quasilinear bzw. linear bzw. linear mit konstanten Koeffizienten ist.

BEISPIEL XIV.1.7. Die pDglen der Beispiele XIV.1.1, XIV.1.3, XIV.1.4, XIV.1.5 und XIV.1.6 sind von zweiter Ordnung; diejenige von Beispiel XIV.1.2 ist von vierter Ordnung. Die pDglen der Beispiele XIV.1.1, XIV.1.2, XIV.1.4 und XIV.1.6 sind linear; diejenigen der Beispiele XIV.1.3 und XIV.1.5 sind quasilinear.

BEISPIEL XIV.1.8. Eine quasilineare pDgl erster Ordnung hat die allgemeine Form

$$a_1(x, u) \frac{\partial u}{\partial x_1} + \dots + a_n(x, u) \frac{\partial u}{\partial x_n} + a_0(x, u) = f \quad \text{in } G$$

mit gegebenen Funktionen  $a_0, a_1, \dots, a_n : G \times \mathbb{R} \rightarrow \mathbb{R}$  und  $f : G \rightarrow \mathbb{R}$ .

**XIV.1.3. Lineare partielle Differentialgleichungen zweiter Ordnung.** Eine lineare pDgl zweiter Ordnung hat die allgemeine Form

$$\sum_{i=1}^n \sum_{j=1}^n A_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{k=1}^n a_k(x) \frac{\partial u}{\partial x_k} + \alpha(x) u = f \quad \text{in } G.$$

Da es bei den zweiten Ableitungen einer zweimal stetig differenzierbaren Funktion nicht auf die Reihenfolge ankommt, kann man stets voraussetzen, dass  $A_{ij}(x) = A_{ji}(x)$  ist für alle  $x \in G$ . Die Matrix  $A(x) = (A_{ij}(x))_{1 \leq i, j \leq n}$  ist also für alle  $x$  symmetrisch und besitzt daher lauter reelle Eigenwerte (die von  $x$  abhängen!) (vgl. Abschnitt II.4.8 (S. 87, Teil I)). Je nach Vorzeichen dieser Eigenwerte werden pDglen zweiter Ordnung in drei Typen eingeteilt:

- ELLIPTISCH: die Eigenwerte sind für alle  $x \in G$  alle strikt positiv,
- PARABOLISCH: für alle  $x \in G$  sind  $n - 1$  Eigenwerte strikt positiv und ein Eigenwert gleich Null.
- HYPERBOLISCH: für alle  $x \in G$  sind  $n - 1$  Eigenwerte strikt positiv und ein Eigenwert strikt negativ.

BEMERKUNG XIV.1.9. Sind die Eigenwerte von  $A(x)$  für alle  $x \in G$  allesamt strikt negativ, erhält man durch Multiplikation der pDgl mit  $-1$  eine elliptische pDgl, die die gleiche Lösungsmenge hat wie die ursprüngliche pDgl. Analog geht man in den Fällen

- für alle  $x \in G$  sind  $n - 1$  Eigenwerte strikt negativ und ein Eigenwert gleich Null
- für alle  $x \in G$  sind  $n - 1$  Eigenwerte strikt negativ und ein Eigenwert strikt positiv

vor und erhält eine parabolische bzw. hyperbolische pDgl mit der gleichen Lösungsmenge wie die ursprüngliche pDgl.

Die Poisson-Gleichung aus Beispiel XIV.1.1 ist elliptisch. Die Wärmeleitungsgleichung aus Beispiel XIV.1.4 ist parabolisch. Die Wellengleichung aus Beispiel XIV.1.6 ist hyperbolisch. Physikalisch beschreiben diese drei Typen verschiedene Phänomene:

- ELLIPTISCH: Minimierung einer Energie,
- PARABOLISCH: Dissipation, d.h. zeitliche Abnahme einer Energie,
- HYPERBOLISCH: Energieerhaltung.

Durch eine geeignete Variablentransformation kann man bei parabolischen und hyperbolischen pDglen stets erreichen, dass der Eigenvektor zum nicht positiven Eigenwert durch  $(0, \dots, 0, 1)^T$  gegeben ist. Dementsprechend identifiziert man dann die letzte Komponente der transformierten Variablen mit der Zeit  $t$  und schreibt  $G$  in der Form  $\Omega \times (0, \infty)$  mit  $\Omega \subset \mathbb{R}^{n-1}$ .

**XIV.1.4. Anfangs- und Randbedingungen.** Wie in den Beispielen des ersten Abschnittes müssen bei pDglen zusätzliche Anfangs- und Randbedingungen gestellt werden. Für Gleichungen zweiter Ordnung gilt hierfür folgendes Schema:

- ELLIPTISCH: eine Randbedingung auf  $\Gamma$ .
- PARABOLISCH: eine Randbedingung auf  $\Gamma \times (0, \infty)$  und eine Anfangsbedingung zur Zeit  $t = 0$ .
- HYPERBOLISCH: eine Randbedingung auf  $\Gamma \times (0, \infty)$  und zwei Anfangsbedingungen zur Zeit  $t = 0$ .

Die wichtigsten Randbedingungen sind:

- DIRICHLET:  $u = g_D$  auf  $\Omega$  bzw. auf  $\Omega \times (0, \infty)$ ,

- NEUMANN:  $\sum_{i=1}^N \sum_{j=1}^N n_i A_{ij} \frac{\partial u}{\partial x_j} = g_N$  auf  $\Omega$  mit  $N = n$  bzw. auf  $\Omega \times (0, \infty)$  mit  $\Omega \subset \mathbb{R}^{n-1}$  und  $N = n - 1$ .

Dabei sind  $g_D$  und  $g_N$  wie  $f$  vorgegebene Funktionen. Dirichlet- und Neumann-Bedingungen können auch in der Form kombiniert werden, dass  $\Gamma$  in zwei disjunkte Stücke  $\Gamma_D$  und  $\Gamma_N$  zerfällt, auf denen Dirichlet- bzw. Neumann-Bedingungen gestellt werden. Für Gleichungen zweiter Ordnung dürfen die beiden Bedingungen aber nicht gleichzeitig auf demselben Randstück gefordert werden.

BEISPIEL XIV.1.10. Bei einer ringförmigen Membran, die am inneren Rand eingespannt und am äußeren Rand frei gelagert ist, ist  $\Omega$  von der Form

$$\Omega = \{x \in \mathbb{R}^2 : R_i^2 < x_1^2 + x_2^2 < R_a^2\}$$

und

$$\Gamma = \Gamma_i \cup \Gamma_a$$

mit

$$\Gamma_{i/a} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = R_{i/a}^2\}$$

Dann gilt auf  $\Gamma_i$  die Dirichlet-Bedingung  $u = 0$  und auf  $\Gamma_a$  die Neumann-Bedingung  $\frac{\partial u}{\partial n} = 0$ .

## XIV.2. Die Wärmeleitungsgleichung

**XIV.2.1. Vorbemerkungen.** Sei  $n \geq 1$  und  $\Omega \subset \mathbb{R}^n$  eine offene, beschränkte Menge. Falls  $n = 1$  ist, soll  $\Omega$  ein Intervall  $(a, b)$  sein. Der Rand  $\Gamma$  besteht dann aus den Punkten  $a$  und  $b$ . Falls  $n \geq 2$  ist, soll der Rand  $\Gamma$  von  $\Omega$  stückweise glatt sein.

Wir betrachten im Folgenden die Wärmeleitungsgleichung

$$\begin{array}{ll} \frac{\partial u}{\partial t} - \Delta u = f & \text{in } \Omega \times (0, \infty) \\ u = g & \text{auf } \Gamma \times (0, \infty) \\ u(x, 0) = u_0 & \text{in } \Omega \end{array}$$

mit bekannten, hinreichend oft differenzierbaren Funktionen  $f$ ,  $g$  und  $u_0$ . Für alle  $x \in \Gamma$  muss zudem die KOMPABILITÄTSBEDINGUNG

$$g(x, 0) = u_0(x)$$

erfüllt sein.

Da der Differentialoperator

$$u \mapsto \frac{\partial u}{\partial t} - \Delta u$$

linear ist, können wir die Lösung der Wärmeleitungsgleichung durch SUPERPOSITION aus Lösungen einfacherer Teilprobleme aufbauen. Genauer machen wir den Ansatz

$$u = v + w + G.$$

Dabei ist  $G : \Omega \times [0, \infty) \rightarrow \mathbb{R}$  eine Funktion mit

$$\begin{aligned} G(x, t) &= g(x, t) && \text{für alle } x \in \Gamma, t > 0 \\ G(x, 0) &= 0 && \text{für alle } x \in \Omega; \end{aligned}$$

$v$  soll eine Lösung der Wärmeleitungsgleichung

$$\begin{aligned} \frac{\partial v}{\partial t} - \Delta v &= 0 && \text{in } \Omega \times (0, \infty) \\ v &= 0 && \text{auf } \Gamma \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } \Omega \end{aligned}$$

sein; und  $w$  soll eine Lösung der Wärmeleitungsgleichung

$$\begin{aligned} \frac{\partial w}{\partial t} - \Delta w &= F && \text{in } \Omega \times (0, \infty) \\ w &= 0 && \text{auf } \Gamma \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } \Omega \end{aligned}$$

mit

$$F = f - \frac{\partial G}{\partial t} + \Delta G$$

sein. Wie man leicht nachrechnet löst  $u = v + w + G$  dann die ursprüngliche pDgl. Die Funktion  $G$  heißt eine FORTSETZUNG DER RANDWERTE. Die pDgl für  $v$  nennt man eine HOMOGENE WÄRMELEITUNGSGLEICHUNG MIT HOMOGENEN RANDBEDINGUNGEN, diejenige für  $w$  eine INHOMOGENE WÄRMELEITUNGSGLEICHUNG MIT HOMOGENEN RAND- UND ANFANGSBEDINGUNGEN.

In den folgenden drei Abschnitten erläutern wir detailliert, wie man im Fall  $n = 1$ , d.h. eine Raumdimension, die Funktionen  $G$ ,  $v$  und  $w$  bestimmt. Da sich die pDgl unter einer Translation  $x \mapsto x - a$  nicht ändert, können wir dabei stets annehmen, dass  $\Omega = (0, L)$  ist. Im letzten Abschnitt gehen wir dann auf die Modifikationen ein, die im Fall höherer Raumdimension  $n \geq 2$  erforderlich sind.

**XIV.2.2. Fortsetzung der Randwerte.** Diese ist im Fall  $n = 1$ ,  $\Gamma = \{0, L\}$  besonders einfach. Wir setzen

$$G(x, t) = \frac{L-x}{L}g(0, t) + \frac{x}{L}g(L, t) \quad 0 \leq x \leq L, t \geq 0.$$

Dann ist

$$F(x, t) = f(x, t) - \frac{L-x}{L} \frac{\partial g}{\partial t}(0, t) - \frac{x}{L} \frac{\partial g}{\partial t}(L, t) \quad 0 \leq x \leq L, t \geq 0.$$

In dem wichtigen Spezialfall, dass die Randbedingung  $g$  nicht von  $t$  abhängt, gilt daher  $F = f$ .

**XIV.2.3. Lösung der homogenen Gleichung mit homogenen Randbedingungen.** Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L). \end{aligned}$$

Zu ihrer Lösung machen wir den SEPARATIONSANSATZ

$$v(x, t) = X(x)T(t).$$

Im Folgenden bezeichnen wir mit  $'$  die Ableitung nach der Ortvariablen  $x$  und mit  $\dot{\phantom{x}}$  die Ableitung nach der Zeitvariablen  $t$ . Dann muss gelten

$$\begin{aligned} 0 &= \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} \\ &= \dot{T}(t)X(x) - T(t)X''(x) \quad \text{für alle } 0 < x < L, t > 0. \end{aligned}$$

Für alle  $x, t$  mit  $X(x)T(t) \neq 0$  können wir dann diese Gleichung durch  $X(x)T(t)$  dividieren und erhalten nach Umsortieren

$$\frac{X''(x)}{X(x)} = \frac{\dot{T}(t)}{T(t)}.$$

Da die linke Seite dieser Gleichung nur von  $x$ , die rechte Seite aber nur von  $t$  abhängt, müssen beide Ausdrücke konstant sein. Es muss also ein  $\lambda \in \mathbb{R}$  geben mit

$$\frac{X''(x)}{X(x)} = \lambda = \frac{\dot{T}(t)}{T(t)}$$

d.h.

$$\begin{aligned} X''(x) &= \lambda X(x), \\ \dot{T}(t) &= \lambda T(t). \end{aligned}$$

Wir müssen nun die drei Fälle  $\lambda > 0$ ,  $\lambda = 0$  und  $\lambda < 0$  unterscheiden.

FALL  $\lambda > 0$ : Gemäß Abschnitt VI.3.2 (S. 229, Teil I) lautet die allgemeine Lösung der gDgl für  $X$  in diesem Fall

$$X(x) = c_1 e^{-\sqrt{\lambda}x} + c_2 e^{\sqrt{\lambda}x}.$$

FALL  $\lambda = 0$ : In diesem Fall lautet gemäß Abschnitt VI.3.2 (S. 229, Teil I) die allgemeine Lösung für  $X$

$$X(x) = c_1 + c_2 x.$$

FALL  $\lambda < 0$ : Gemäß Abschnitt VI.3.2 (S. 229, Teil I) ergibt sich in diesem Fall für  $x$  die allgemeine Lösung

$$X(x) = c_1 \sin(\sqrt{|\lambda|x}) + c_2 \cos(\sqrt{|\lambda|x}).$$

Aus der Randbedingung

$$v(x, t) = 0 \quad \text{für } x \in \{0, L\}, t > 0$$

ergibt sich die Bedingung

$$X(x)T(t) = 0 \quad \text{für } x \in \{0, L\}, t > 0.$$

Diese Gleichung lässt zwei Lösungen zu:  $T(t) = 0$  für alle  $t$  oder

$$X(0) = X(L) = 0.$$

Die erste liefert offensichtlich die triviale Lösung  $v(x, t) = 0$  für alle  $x, t$ . Daher betrachten wir sinnvollerweise nur die zweite Bedingung für  $X$ . Diese führt in allen drei Fällen auf ein homogenes lineares Gleichungssystem für die Konstanten  $c_1$  und  $c_2$ , das eine vom Nullvektor verschiedene Lösung zulassen soll. Also muss die Determinante der zugehörigen Matrix verschwinden. In den drei Fällen erhalten wir folgende Determinanten:

FALL  $\lambda > 0$ :

$$\begin{aligned} \det \begin{pmatrix} e^{-\sqrt{\lambda}0} & e^{\sqrt{\lambda}0} \\ e^{-\sqrt{\lambda}L} & e^{\sqrt{\lambda}L} \end{pmatrix} &= \det \begin{pmatrix} 1 & 1 \\ e^{-\sqrt{\lambda}L} & e^{\sqrt{\lambda}L} \end{pmatrix} \\ &= e^{\sqrt{\lambda}L} - e^{-\sqrt{\lambda}L} \\ &= 2 \sinh(\sqrt{\lambda}L). \end{aligned}$$

FALL  $\lambda = 0$ :

$$\det \begin{pmatrix} 1 & 0 \\ 1 & L \end{pmatrix} = L$$

FALL  $\lambda < 0$ :

$$\begin{aligned} \det \begin{pmatrix} \sin(\sqrt{|\lambda|}0) & \cos(\sqrt{|\lambda|}0) \\ \sin(\sqrt{|\lambda|}L) & \cos(\sqrt{|\lambda|}L) \end{pmatrix} &= \det \begin{pmatrix} 0 & 1 \\ \sin(\sqrt{|\lambda|}L) & \cos(\sqrt{|\lambda|}L) \end{pmatrix} \\ &= -\sin(\sqrt{|\lambda|}L). \end{aligned}$$

Im ersten und zweiten Fall verschwindet die Determinante für keinen Wert von  $\lambda$ . (Beachte:  $\sinh(z)$  hat  $z = 0$  als einzige Nullstelle!) Im dritten Fall verschwindet die Determinante genau dann, wenn  $\sqrt{|\lambda|}L$  eine Nullstelle des Sinus ist, d.h.

$$\sqrt{|\lambda|}L = k\pi \quad , k \in \mathbb{Z} \setminus \{0\}$$

bzw. äquivalent

$$\lambda = - \left( \frac{k\pi}{L} \right)^2 \quad , k \in \mathbb{Z} \setminus \{0\}.$$

Für gegebenes  $\lambda$  hat die gDgl für  $T$  gemäß Abschnitt VI.2.1 (S. 217, Teil I) die allgemeine Lösung

$$T(t) = ce^{\lambda t}.$$

Insgesamt erhalten wir also mit unserem Ansatz die Funktionen

$$v_k(x, t) = e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L}x\right) \quad , k \in \mathbb{Z} \setminus \{0\}$$

als mögliche Lösungen der pDgl. Da die Funktionen  $v_k$  und  $v_{-k}$  sich für  $k \in \mathbb{N}^*$  nur um ein Vorzeichen unterscheiden lautet die allgemeine Lösung

$$v(x, t) = \sum_{k \in \mathbb{N}^*} c_k e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L}x\right).$$

Jede dieser Funktionen erfüllt konstruktionsgemäß die pDgl und die Randbedingungen. Wir müssen also nur noch die Anfangsbedingung erfüllen. Einsetzen von  $t = 0$  ergibt die Bedingung

$$u_0(x) = \sum_{k \in \mathbb{N}^*} c_k \sin\left(\frac{k\pi}{L}x\right).$$

Da nach Voraussetzung

$$u_0(0) = u_0(L) = 0$$

ist (vgl. Abschnitt XIV.2.1 (S. 98)) können wir  $u_0$  ungerade zu einer stetigen  $2L$ -periodischen Funktion auf  $\mathbb{R}$  fortsetzen (vgl. Abschnitt XIII.1.1 (S. 75)). Falls  $u_0$  zusätzlich differenzierbar ist, stimmt  $u_0$  mit seiner Fourier-Reihe überein, und die  $c_k$  sind die entsprechenden

Fourier-Koeffizienten von  $u_0$  (vgl. Abschnitt XIII.2.1 (S. 78) und XIII.2.4 (S. 82)). Insgesamt erhalten wir damit:

Die Funktion  $u_0$  sei differenzierbar und erfülle

$$u_0(0) = u_0(L) = 0$$

Dann ist die Lösung  $v$  der pDgl.

$$\begin{aligned} \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$v(x, t) = \sum_{k=1}^{\infty} c_k e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L} x\right)$$

mit

$$c_k = \frac{2}{L} \int_0^L u_0(x) \sin\left(\frac{k\pi}{L} x\right) dx.$$

**XIV.2.4. Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen.** Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } (0, L). \end{aligned}$$

Wegen des vorigen Abschnitts setzen wir die Funktion  $F$  für jedes  $t$  ungerade bzgl.  $x$   $2L$ -periodisch fort und entwickeln sie in eine Fourier-Reihe bzgl.  $x$ :

$$F(x, t) \sim \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin\left(\frac{k\pi}{L} x\right)$$

mit

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin\left(\frac{k\pi}{L} x\right) dx.$$

Für  $w$  machen wir den Ansatz

$$w(x, t) \sim \sum_{k \in \mathbb{N}^*} w_k(t) \sin\left(\frac{k\pi}{L}x\right).$$

Einsetzen in die pDgl liefert

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin\left(\frac{k\pi}{L}x\right) &\sim F(x, t) \\ &= \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} \\ &\sim \sum_{k \in \mathbb{N}^*} [\dot{w}_k(t) + \left(\frac{k\pi}{L}\right)^2 w_k(t)] \sin\left(\frac{k\pi}{L}x\right). \end{aligned}$$

Also ergibt sich für die Fourier-Koeffizienten  $w_k$  von  $w$  die gDgl

$$\dot{w}_k(t) = -\left(\frac{k\pi}{L}\right)^2 w_k(t) + \widehat{F}_k(t) \quad , k \in \mathbb{N}^*.$$

Wegen der Anfangsbedingung  $w(x, 0) = 0$  für  $x \in (0, L)$  muss gelten

$$w_k(0) = 0 \quad , k \in \mathbb{N}^*.$$

Damit ergibt sich mit der Methode der Variation der Konstanten (vgl. Abschnitt VI.2.2 (S. 221, Teil I))

$$w_k(t) = \int_0^t \exp\left(-\left(\frac{k\pi}{L}\right)^2(t-s)\right) \widehat{F}_k(s) ds \quad , k \in \mathbb{N}^*.$$

Falls  $F$  stetig ist, kann man zeigen, dass die Fourier-Reihe für  $w$  konvergiert. Damit folgt insgesamt:

Die Funktion  $F$  sei stetig. Dann ist die Lösung  $w$  der pDgl

$$\begin{aligned} \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } (0, L). \end{aligned}$$

gegeben durch

$$w(x, t) = \sum_{k=1}^{\infty} w_k(t) \sin\left(\frac{k\pi}{L}x\right)$$

mit

$$w_k(t) = \int_0^t \exp\left(-\left(\frac{k\pi}{L}\right)^2(t-s)\right) \widehat{F}_k(s) ds$$

und

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin\left(\frac{k\pi}{L}x\right) dx.$$

**XIV.2.5. Integraldarstellung.** Für  $x, \eta \in (0, L)$  und  $t \in (0, \infty)$  definieren wir die Funktion  $G(x, t, \eta)$  durch

$$G(x, t, \eta) = \frac{2}{L} \sum_{k=1}^{\infty} \exp\left(-\left(\frac{k\pi}{L}\right)^2 t\right) \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{k\pi}{L}\eta\right).$$

$G$  heißt WÄRMELEITUNGSKERN. Mit seiner Hilfe kann man die Ergebnisse der vorigen beiden Abschnitte wie folgt zusammenfassen:

Die Funktion  $u_0 : (0, L) \rightarrow \mathbb{R}$  sei differenzierbar und erfülle  
 $u_0(0) = u_0(L) = 0$ .

Die Funktion  $F : (0, L) \times (0, \infty) \rightarrow \mathbb{R}$  sei stetig. Dann ist die Lösung  $u$  der pDgl

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty), \\ u &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$\begin{aligned} u(x, t) &= \int_0^L G(x, t, \eta) u_0(\eta) d\eta \\ &+ \int_0^L \left\{ \int_0^t G(x, t-s, \eta) F(\eta, s) ds \right\} d\eta. \end{aligned}$$

**XIV.2.6. Höhere Raumdimensionen.** Im Fall  $n \geq 2$  gehen wir ähnlich wie in den Abschnitten XIV.2.3 (S. 100) und XIV.2.4 (S. 103) vor. Für die Funktion  $v$  machen wir wieder den Separationsansatz

$$v(x, t) = X(x)T(t).$$

Der einzige Unterschied ist, dass  $X$  nun eine Funktion in  $n$  Veränderlichen ist. Der Ansatz führt auf die Gleichungen

$$\begin{aligned} \dot{T}(t) &= \lambda T(t), \\ \Delta X &= \lambda X. \end{aligned}$$

Damit erhalten wir nun eine pDgl für  $X$ . Man nennt sie das EIGENWERTPROBLEM FÜR DEN LAPLACE-OPERATOR. Wir werden dieses

für spezielle Gebiete  $\Omega \subset \mathbb{R}^2$  in den Abschnitten XIV.4.1 (S. 112) und XIV.4.3 (S. 116) lösen. Allgemein kann man zeigen, dass jeder Eigenwert strikt negativ ist, dass insgesamt abzählbar viele Eigenwerte  $\lambda_k = -\omega_k^2, k \in \mathbb{N}$ , existieren und dass die zugehörigen Eigenfunktionen  $\varphi_k$  paarweise orthogonal sind, d.h.

$$\int_{\Omega} \varphi_k \varphi_\ell = 0$$

für  $k \neq \ell$ . Die „Fourier-Koeffizienten“

$$\int_{\Omega} u_0 \varphi_k dx \quad \text{und} \quad \int_{\Omega} F(x, t) \varphi_k(x) dx$$

übernehmen dann die Rolle der Fourier-Koeffizienten  $c_k$  und  $F_k(t)$  in den vorigen Abschnitten. Mit diesen Modifikationen übertragen sich dann die Ergebnisse der vorigen Abschnitte auf den Fall höherer Raumdimensionen.

### XIV.3. Die Wellengleichung

**XIV.3.1. Vorbemerkung.** Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \Delta u &= f && \text{in } \Omega \times (0, \infty) \\ u &= g && \text{auf } \Gamma \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } \Omega \\ \frac{\partial u}{\partial t}(x, 0) &= u_1 && \text{in } \Omega. \end{aligned}$$

Dabei ist wie im vorigen Abschnitt  $\Omega$  das Intervall  $(0, L)$  mit Rand  $\Gamma = \{0, L\}$  oder eine beschränkte, offene Menge in  $\mathbb{R}^n, n \geq 2$ , mit stückweise glattem Rand. Für alle  $x \in \Gamma$  müssen zudem die KOMPABILITÄTSBEDINGUNGEN

$$\begin{aligned} g(x, 0) &= u_0(x) \\ \frac{\partial g}{\partial t}(x, 0) &= u_1(x) \end{aligned}$$

erfüllt sein.

Wir gehen wie im vorigen Paragraphen vor und bestimmen die Lösung  $u$  durch die SUPERPOSITION

$$u = v + w + G$$

mit einer Fortsetzung  $G$  der Randdaten, der Lösung  $v$  der homogenen Gleichung mit homogenen Randbedingungen, d.h.  $f = 0$ ,  $g = 0$ , und der Lösung  $w$  der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen, d.h.  $g = 0$ ,  $u_0 = 0$ ,  $u_1 = 0$ . Wieder betrachten wir detailliert den Fall einer Raumdimension und erläutern kurz die Modifikationen bei höheren Raumdimensionen. Die Fortsetzung  $G$  der Randwerte  $g$  wird wie in Abschnitt XIV.2.2 (S. 100) bestimmt. Die Funktion  $F$  ist jetzt gegeben durch

$$F = f - \frac{\partial^2 G}{\partial t^2} + \Delta G.$$

**XIV.3.2. Lösung der homogenen Gleichung mit homogenen Randbedingungen.** Wir betrachten die pDgl

$$\begin{aligned} \frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \\ \frac{\partial v}{\partial t}(x, 0) &= u_1 && \text{in } (0, L). \end{aligned}$$

Wie in Abschnitt XIV.2.3 (S. 100) machen wir den Separationsansatz

$$v(x, t) = X(x)T(t).$$

Dieser führt jetzt auf die gDglen

$$\begin{aligned} X''(x) &= \lambda X(x) \\ \ddot{T}(t) &= \lambda T(t) \end{aligned}$$

mit  $\lambda \in \mathbb{R}$ . Wie in Abschnitt XIV.2.3 (S. 100) folgt aus den Randbedingungen, dass  $X(0) = X(L) = 0$  sein muss und dass daher

$$\lambda = - \left( \frac{k\pi}{L} \right)^2, \quad k \in \mathbb{N}^*$$

und

$$X(x) = \sin\left(\frac{k\pi}{L}x\right)$$

gelten muss. Für die Funktion  $T$  ergibt sich gemäß Abschnitt VI.3.2 (S. 229, Teil I) jetzt die allgemeine Lösung

$$T(t) = c_k \cos\left(\frac{k\pi}{L}t\right) + s_k \sin\left(\frac{k\pi}{L}t\right).$$

Daher hat die pDgl die allgemeine Lösung

$$v(x, t) = \sum_{k \in \mathbb{N}^*} \left[ c_k \cos\left(\frac{k\pi}{L}t\right) + s_k \sin\left(\frac{k\pi}{L}t\right) \right] \sin\left(\frac{k\pi}{L}x\right).$$

Wir müssen noch die Anfangsbedingungen anpassen. Einsetzen von  $t = 0$  liefert die Bedingung

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} c_k \sin\left(\frac{k\pi}{L}x\right) &= v(x, 0) \\ &= u_0(x). \end{aligned}$$

Hieraus können die Koeffizienten  $c_k$  wie in Abschnitt XIV.2.3 (S. 100) bestimmt werden. Ableiten nach  $t$  und Einsetzen von  $t = 0$  liefert die Bedingung

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} s_k \frac{k\pi}{L} \sin\left(\frac{k\pi}{L}x\right) &= \frac{\partial v}{\partial t}(x, 0) \\ &= u_1(x). \end{aligned}$$

Zur Bestimmung der  $s_k$  setzen wir wieder  $u_1$  ungerade zu einer  $2L$ -periodischen Funktion fort (vgl. Abschnitt XIII.1.1 (S. 75)) und erhalten gemäß Abschnitt XIII.2.1 (S. 78)

$$u_1(x) \sim \sum_{k \in \mathbb{N}^*} \gamma_k \sin\left(\frac{k\pi}{L}x\right)$$

mit

$$\gamma_k = \frac{2}{L} \int_0^L u_1(x) \sin\left(\frac{k\pi}{L}x\right) dx.$$

Ein Vergleich der Fourier-Koeffizienten liefert dann

$$s_k = \frac{L}{k\pi} \gamma_k, \quad k \in \mathbb{N}^*.$$

Insgesamt erhalten wir

Die Funktionen  $u_0$  und  $u_1$  seien differenzierbar und mögen die Bedingungen

$$u_0(0) = u_0(L) = 0$$

$$u_1(0) = u_1(L) = 0$$

erfüllen. Dann ist die Lösung der pDgl

$$\begin{aligned} \frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

$$\frac{\partial v}{\partial t}(x, t) = u_1 \quad \text{in } (0, L)$$

gegeben durch

$$v(x, t) = \sum_{k=1}^{\infty} [c_k \cos(\frac{k\pi}{L}t) + s_k \sin(\frac{k\pi}{L}t)] \sin(\frac{k\pi}{L}x)$$

mit

$$c_k = \frac{2}{L} \int_0^L u_0(x) \sin(\frac{k\pi}{L}x) dx,$$

$$s_k = \frac{L}{k\pi} \frac{2}{L} \int_0^L u_1(x) \sin(\frac{k\pi}{L}x) dx.$$

**XIV.3.3. Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen.** Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} &= F \quad \text{in } (0, L) \times (0, \infty) \\ w &= 0 \quad \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 \quad \text{in } (0, L) \\ \frac{\partial w}{\partial t}(x, 0) &= 0 \quad \text{in } (0, L). \end{aligned}$$

Wie in Abschnitt XIV.2.4 (S. 103) setzen wir die Funktion  $F$  für jedes  $t$  ungerade bzgl.  $x$   $2L$ -periodisch fort und entwickeln sie in eine Fourierreihe bzgl.  $x$ :

$$F(x, t) \sim \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin(\frac{k\pi}{L}x)$$

mit

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin(\frac{k\pi}{L}x) dx.$$

Für  $w$  machen wir wieder den Ansatz

$$w(x, t) \sim \sum_{k \in \mathbb{N}^*} w_k(t) \sin(\frac{k\pi}{L}x).$$

Einsetzen in die pDgl ergibt dann

$$\sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin(\frac{k\pi}{L}x) \sim F$$

$$\begin{aligned}
&= \frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} \\
&\sim \sum_{k \in \mathbb{N}^*} [\ddot{w}_k(t) + (\frac{k\pi}{L})^2 w_k(t)] \sin(\frac{k\pi}{L} x).
\end{aligned}$$

Vergleich der Fourier-Koeffizienten liefert die gDgl

$$\ddot{w}_k(t) + (\frac{k\pi}{L})^2 w_k(t) = \widehat{F}_k(t)$$

für die Koeffizienten  $w_k$ . Zusammen mit den Anfangsbedingungen

$$w_k(0) = \dot{w}_k(0) = 0$$

ergibt sich gemäß Abschnitt VI.3.2 (S. 229, Teil I) und VI.3.3 (S. 232, Teil I) die Lösung

$$\begin{aligned}
w_k(t) &= -\cos(\frac{k\pi}{L} t) \int_0^t \frac{L}{k\pi} \sin(\frac{k\pi}{L} s) \widehat{F}_k(s) ds \\
&\quad + \sin(\frac{k\pi}{L} t) \int_0^t \frac{L}{k\pi} \cos(\frac{k\pi}{L} s) \widehat{F}_k(s) ds.
\end{aligned}$$

Insgesamt erhalten wir somit:

Die Lösung  $w$  der pDgl

$$\begin{aligned}
\frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\
w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\
w(x, 0) &= 0 && \text{in } (0, L) \\
\frac{\partial w}{\partial t}(x, 0) &= 0 && \text{in } (0, L)
\end{aligned}$$

ist gegeben durch

$$w(x, t) = \sum_{k \in \mathbb{N}^*}^{\infty} [-c_k(t) \cos(\frac{k\pi}{L} t) + s_k(t) \sin(\frac{k\pi}{L} t)] \sin(\frac{k\pi}{L} x)$$

mit

$$\begin{aligned}
c_k(t) &= \int_0^t \int_0^L \frac{L}{k\pi} \frac{2}{L} F(x, s) \sin(\frac{k\pi}{L} s) \sin(\frac{k\pi}{L} x) dx ds \\
s_k(t) &= \int_0^t \int_0^L \frac{L}{k\pi} \frac{2}{L} F(x, s) \cos(\frac{k\pi}{L} s) \sin(\frac{k\pi}{L} x) dx ds.
\end{aligned}$$

**XIV.3.4. Integraldarstellung.** Für  $x, \eta \in (0, L)$  und  $t \in (0, \infty)$  definieren wir die Funktion  $H(x, t, \eta)$  durch

$$H(x, t, \eta) = \frac{2}{L} \sum_{k=1}^{\infty} \frac{L}{k\pi} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{k\pi}{L}t\right) \sin\left(\frac{k\pi}{L}\eta\right).$$

$H$  heißt WELLENGLEICHUNGSKERN. Mit seiner Hilfe kann man die Ergebnisse der vorigen beiden Abschnitte wie folgt zusammenfassen:

Die Funktionen  $u_0$  und  $u_1$  seien differenzierbar und mögen die Bedingungen

$$u_0(0) = u_0(L) = 0$$

$$u_1(0) = u_1(L) = 0$$

erfüllen. Dann ist die Lösung  $u$  der pDgl

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ u &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } (0, L) \\ \frac{\partial u}{\partial t}(x, 0) &= u_1 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$\begin{aligned} u(x, t) &= \int_0^L \frac{\partial}{\partial t} H(x, t, \eta) u_0(\eta) d\eta \\ &+ \int_0^L H(x, t, \eta) u_1(\eta) d\eta \\ &+ \int_0^L \left\{ \int_0^t H(x, t-s, \eta) F(\eta, s) ds \right\} d\eta. \end{aligned}$$

**XIV.3.5. Höhere Raumdimensionen.** Das Vorgehen im Fall  $n \geq 2$  ist völlig analog zu Abschnitt XIV.2.6 (S. 105). Die Gleichung für  $T$  und  $x$  ist dann

$$\begin{aligned} \ddot{T}(t) &= \lambda T(t), \\ \Delta X &= \lambda X. \end{aligned}$$

Daher müssen die Fourier-Koeffizienten  $c_k$ ,  $s_k$  und  $\widehat{F}_k$  der Abschnitte XIV.3.2 und XIV.3.3 durch die entsprechenden Terme

$$\int_{\Omega} u_0 \varphi_k dx, \quad \int_{\Omega} u_1 \varphi_k dx, \quad \int_{\Omega} F(x, t) \varphi_k(x) dx$$

mit den Eigenfunktionen des Laplace-Operators ersetzt werden.

**XIV.3.6. Dämpfung.** Wir betrachten nun die pDgl

$$\frac{\partial^2 u}{\partial t^2} + \alpha \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f \quad \text{in } (0, L) \times (0, \infty)$$

mit den bisherigen Rand- und Anfangsbedingungen. Neu ist der Term  $\alpha \frac{\partial u}{\partial t}$ , der physikalisch eine Dämpfung beschreibt. Die Gleichungen für  $T$  in Abschnitt XIV.3.2 und  $w_k$  in Abschnitt XIV.3.3 haben jetzt die Form

$$\begin{aligned} \ddot{T}(t) + \alpha \dot{T}(t) &= -\left(\frac{k\pi}{L}\right)^2 T(t) \\ \ddot{w}_k(t) + \alpha \dot{w}_k(t) + \left(\frac{k\pi}{L}\right)^2 w_k(t) &= \widehat{F}_k(t). \end{aligned}$$

Diese können mit den Methoden der Abschnitte VI.3.2 (S. 229, Teil I) und VI.3.3 (S. 232, Teil I) gelöst werden. Sofern

$$\alpha^2 < \left(\frac{\pi}{L}\right)^2$$

ist, müssen die Funktion  $\sin(\frac{k\pi}{L}t)$  und  $\cos(\frac{k\pi}{L}t)$  nun durch die Funktionen

$$e^{-\frac{\alpha}{2}t} \cos\left(\frac{1}{2}\sqrt{\left(\frac{2k\pi}{L}\right)^2 - \alpha^2} t\right)$$

und

$$e^{-\frac{\alpha}{2}t} \sin\left(\frac{1}{2}\sqrt{\left(\frac{2k\pi}{L}\right)^2 - \alpha^2} t\right)$$

ersetzt werden.

#### XIV.4. Die Poissongleichung

**XIV.4.1. Das Eigenwertproblem im Rechteck.** Sei

$$\Omega = (0, L) \times (0, H) = \{(x, y) : 0 < x < L, 0 < y < H\}$$

das Rechteck mit Kantenlängen  $L$  und  $H$  und linker unterer Ecke im Ursprung. Wir betrachten das EIGENWERTPROBLEM:

$$\begin{array}{ll} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{auf } \partial\Omega. \end{array}$$

Gesucht sind dabei der EIGENWERT  $\lambda$  und die EIGENFUNKTION  $u : R \rightarrow \mathbb{R}$ . Dabei interessieren uns natürlich nur nicht triviale Lösungen, d.h. Eigenfunktionen  $u$  mit

$$\int_{\Omega} |u(x, y)|^2 dx dy > 0.$$

Vorab überlegen wir uns, dass alle Eigenwerte positiv sind. Ist nämlich

$$-\Delta u = \lambda u,$$

können wir die Gleichung mit  $u$  multiplizieren und über  $\Omega$  integrieren. Mit dem Gaußschen Integralsatz (vgl. Abschnitt IX.5.5 (S. 124, Teil II)) erhalten wir dann

$$\begin{aligned} \lambda \int_{\Omega} |u|^2 dx &= \int_{\Omega} \underbrace{-\Delta u u}_{=-\operatorname{div}(u \nabla u) + \nabla u \cdot \nabla u} dx \\ &= - \underbrace{\int_{\Omega} \operatorname{div}(u \nabla u) dx}_{=\int_{\partial \Omega} u \frac{\partial u}{\partial n} ds} + \int_{\Omega} |\nabla u|^2 dx \\ &= - \int_{\partial \Omega} \underbrace{u}_{=0} \frac{\partial u}{\partial n} ds + \int_{\Omega} |\nabla u|^2 dx \\ &= \int_{\Omega} |\nabla u|^2 dx. \end{aligned}$$

Offensichtlich ist

$$\int_{\Omega} |\nabla u|^2 dx \geq 0.$$

Wäre

$$\int_{\Omega} |\nabla u|^2 dx = 0,$$

folgte  $\nabla u(x) = 0$  für alle  $x \in \Omega$ . Daher wäre  $u$  konstant. Wegen der Randbedingung müsste diese Konstante gleich Null sein. Da wir aber nur an nicht trivialen Lösungen interessiert sind, folgt

$$\int_{\Omega} |\nabla u|^2 dx > 0.$$

Daher folgt aus obiger Gleichung  $\lambda > 0$ .

Man beachte, dass wir für diese Überlegung die spezielle Gestalt von  $\Omega$  nicht ausgenutzt haben. Sie gilt für jede offene Menge  $\Omega \subset \mathbb{R}^n$ ,  $n \geq 2$ , mit stückweise glattem Rand.

Wie in den vorigen Paragraphen machen wir einen Separationsansatz

$$u(x, y) = X(x)Y(y).$$

Im Folgenden bezeichnet  $'$  die Ableitung nach  $x$  und  $\dot{\phantom{x}}$  diejenige nach  $y$ . Wie in Abschnitt XIV.2.3 (S. 100) führt der Separationsansatz auf die gDglen

$$\begin{aligned} X''(x) &= \mu X(x), \\ \ddot{Y}(y) &= -(\lambda + \mu)Y(y) \end{aligned}$$

mit einer unbekanntes Konstanten  $\mu \in \mathbb{R}$ . Wegen der Randbedingung an  $u$  muss für die Lösungen  $X, Y$  dieser gDglen gelten

$$\begin{aligned} X(0) = X(L) &= 0, \\ Y(0) = Y(H) &= 0. \end{aligned}$$

Wie in Abschnitt XIV.2.3 (S. 100) folgt hieraus, dass

$$\mu = - \left( \frac{k\pi}{L} \right)^2, \quad k \in \mathbb{N}^*$$

und

$$X(x) = \sin\left(\frac{k\pi}{L}x\right)$$

sein muss. Ebenso folgt, dass

$$-(\lambda + \mu) = - \left( \frac{\ell\pi}{H} \right)^2, \quad \ell \in \mathbb{N}^*$$

und

$$Y(y) = \sin\left(\frac{\ell\pi}{H}y\right)$$

sein muss. Aus den Gleichungen für  $\lambda$  und  $\mu$  ergibt sich insbesondere

$$\lambda = \left( \frac{\ell\pi}{H} \right)^2 + \left( \frac{k\pi}{L} \right)^2, \quad k, \ell \in \mathbb{N}^*.$$

Insgesamt erhalten wir somit:

Die Lösungen  $\lambda$  und  $u$  des Eigenwertproblems

$$\begin{aligned} -\Delta u &= \lambda && \text{in } \Omega \\ u &= 0 && \text{auf } \partial\Omega \end{aligned}$$

mit  $\Omega = (0, L) \times (0, H)$  sind von der Form

$$\lambda = \left( \frac{\ell\pi}{H} \right)^2 + \left( \frac{k\pi}{L} \right)^2$$

$$u(x, y) = \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right)$$

mit  $k, \ell \in \mathbb{N}^*$ .

**XIV.4.2. Die Poissongleichung im Rechteck.** Sei wieder  $\Omega = (0, L) \times (0, H)$ . Wir betrachten die Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{array}{l} -\Delta u = f \quad \text{in } \Omega \\ u = g \quad \text{auf } \partial\Omega. \end{array}$$

Definiere die Fortsetzung  $G : \Omega \rightarrow \mathbb{R}$  der Randwerte  $g$  durch

$$\begin{aligned} G(x, y) &= g(x, 0)\left(1 - \frac{y}{H}\right) + g(x, H)\frac{y}{H} \\ &\quad + [g(0, y) - g(0, 0)\left(1 - \frac{y}{H}\right) - g(0, H)\frac{y}{H}]\left(1 - \frac{x}{L}\right) \\ &\quad + [g(L, y) - g(L, 0)\left(1 - \frac{y}{H}\right) + g(L, H)\frac{y}{H}]\frac{x}{L}. \end{aligned}$$

Dann löst

$$v = u - G$$

die Poissongleichung

$$\begin{array}{ll} -\Delta v = F & \text{in } \Omega \\ v = 0 & \text{auf } \partial\Omega \end{array}$$

mit homogenen Dirichlet-Randbedingungen und rechter Seite

$$F = f + \Delta G.$$

Für festes  $y \in (0, H)$  können wir die Funktion  $F$  bzgl.  $x$  ungerade zu einer  $2L$ -periodischen Funktion fortsetzen und die Fourier-Reihe bzgl.  $x$  dieser Fortsetzung wie in Abschnitt XIII.2.1 (S. 78) berechnen. Die Fourier-Koeffizienten sind dann Funktionen von  $y$ . Diese können ungerade zu  $2H$ -periodischen Funktionen fortgesetzt und in eine Fourier-Reihe bzgl.  $y$  entwickelt werden. Insgesamt erhalten wir eine doppelte Fourier-Reihe

$$F(x, y) \sim \sum_{k \in \mathbb{N}^*} \sum_{\ell \in \mathbb{N}^*} \widehat{F}_{k\ell} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right)$$

mit

$$\widehat{F}_{k\ell} = \frac{2}{L} \frac{2}{H} \int_0^L \int_0^H F(x, y) \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right) dx dy.$$

Für  $v$  machen wir nun den Ansatz

$$v(x, y) \sim \sum_{k \in \mathbb{N}^*} \sum_{\ell \in \mathbb{N}^*} v_{k\ell} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right).$$

Einsetzen in die pDgl und Vergleich der Fourier-Koeffizienten ergibt

$$\left[ \left(\frac{k\pi}{L}\right)^2 + \left(\frac{\ell\pi}{H}\right)^2 \right] v_{k\ell} = \widehat{F}_{k\ell} \quad k, \ell \in \mathbb{N}^*.$$

Unter geeigneten Stetigkeits- und Differenzierbarkeitsannahmen an  $f$  und  $g$  kann man zeigen, dass die Fourier-Reihe für  $v$  konvergiert. Man erhält dann insgesamt:

Die Lösung  $u$  der Poisson-Gleichung

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{auf } \partial\Omega \end{aligned}$$

im Rechteck  $\Omega = (0, L) \times (0, H)$  ist gegeben durch

$$u = v + G$$

mit

$$\begin{aligned} G(x, y) &= g(x, 0)\left(1 - \frac{y}{H}\right) + g(x, H)\frac{y}{H} \\ &+ [g(0, y) - g(0, 0)\left(1 - \frac{y}{H}\right) - g(0, H)\frac{y}{H}]\left(1 - \frac{x}{L}\right) \\ &+ [g(L, y) - g(L, 0)\left(1 - \frac{y}{H}\right) - g(L, H)\frac{y}{H}]\frac{x}{L} \end{aligned}$$

und

$$v(x, y) = \int_0^L \int_0^H G(x, y, \xi, \eta) F(\xi, \eta) d\xi d\eta$$

mit

$$\begin{aligned} G(x, y, \xi, \eta) &= \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{1}{\left(\frac{k\pi}{L}\right)^2 + \left(\frac{\ell\pi}{H}\right)^2} \sin\left(\frac{k\pi}{L}x\right) \cdot \\ &\cdot \sin\left(\frac{\ell\pi}{H}y\right) \sin\left(\frac{k\pi}{L}\xi\right) \sin\left(\frac{\ell\pi}{H}\eta\right) \end{aligned}$$

$$F = f + \Delta G.$$

#### XIV.4.3. Das Eigenwertproblem im Kreis. Sei nun

$$\Omega = B_R(0) = \{(x, y) : x^2 + y^2 < R^2\}$$

der Kreis um den Nullpunkt mit Radius  $R$ . Wir betrachten wieder das Eigenwertproblem aus Abschnitt XIV.4.1 (S. 112). Dazu führen wir ebene Polarkoordinaten ein

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad 0 \leq r \leq R, \quad 0 \leq \varphi < 2\pi.$$

Gemäß Beispiel VIII.2.19 (S. 45, Teil II) lautet die pDgl in den Polarkoordinaten dann

$$\begin{aligned} -\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} &= \lambda u \quad \text{in } (0, R) \times [0, 2\pi) \\ u &= 0 \quad \text{für } r = R. \end{aligned}$$

Wir machen wieder einen Separationsansatz

$$u(r, \varphi) = v(r)\Phi(\varphi)$$

und bezeichnen mit ' die Ableitung bzgl.  $r$  und mit  $\dot{\phantom{x}}$  diejenige bzgl.  $\varphi$ . Dann erhalten wir die beiden gDglen

$$\begin{aligned}\ddot{\Phi} &= \mu\Phi, \\ -r(rv')' - r^2\lambda v &= \mu v\end{aligned}$$

mit einer unbekanntenen Konstanten  $\mu$ . Aufgrund unseres Ansatzes muss die Funktion  $\Phi$   $2\pi$ -periodisch sein. Daher erhalten wir für  $\mu$  und  $\Phi$  die Lösungen

$$\begin{aligned}\mu &= 0, \\ \Phi(\varphi) &= 1\end{aligned}$$

und mit  $k \in \mathbb{N}^*$

$$\begin{aligned}\mu &= -k^2, \\ \Phi(\varphi) &= \cos(k\varphi)\end{aligned}$$

und

$$\Phi(\varphi) = \sin(k\varphi).$$

Damit lautet die Bestimmungsgleichung für  $v$  und  $\lambda$

$$r^2v'' + rv' + (r^2\lambda - k^2)v = 0$$

mit  $k \in \mathbb{N}$ . Da wir uns in Abschnitt XIV.4.1 (S. 112) schon überlegt haben, dass  $\lambda > 0$  sein muss, können wir den Ansatz

$$\begin{aligned}v(r) &= w(\rho), \\ \rho &= \sqrt{\lambda}r\end{aligned}$$

machen. Wir erhalten dann für  $w$  die gDgl

$$\rho^2 \frac{d^2w}{d\rho^2} + \rho \frac{dw}{d\rho} + (\rho^2 - k^2)w = 0.$$

Dies ist die Besselsche Differentialgleichung aus Beispiel IX.1.2 (S. 78, Teil II) mit der Lösung

$$\begin{aligned}w(\rho) &= J_k(\rho) \\ &= \frac{1}{\pi} \int_0^\pi \cos(\rho \sin t - kt) dt.\end{aligned}$$

Hieraus ergibt sich für  $v$

$$\begin{aligned}v(r) &= J_k(\sqrt{\lambda}r) \\ &= \frac{1}{\pi} \int_0^\pi \cos(\sqrt{\lambda}r \sin t - kt) dt.\end{aligned}$$

Der Eigenwert  $\lambda$  wird durch die Randbedingung

$$v(R) = 0$$

festgelegt.

Insgesamt erhalten wir:

Die Eigenfunktionen des Laplace-Operators auf dem Kreis  $B_R(0)$  sind von der Form

$$J_k(\sqrt{\lambda_k}r) \cos(k\varphi) \quad , k \in \mathbb{N},$$

$$J_\ell(\sqrt{\lambda_\ell}r) \sin(\ell\varphi) \quad , \ell \in \mathbb{N}^*$$

mit

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) dt.$$

Die Eigenwerte  $\lambda_k$ ,  $k \in \mathbb{N}$ , sind festgelegt durch die Bedingung

$$J_k(\sqrt{\lambda_k}R) = 0.$$

**XIV.4.4. Die Poissongleichung im Kreis.** Sei wieder  $\Omega = B_R(0)$  der Kreis um Null mit Radius  $R$ . Wir betrachten die Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{auf } \partial\Omega. \end{aligned}$$

In Polarkoordinaten lautet diese pDgl

$$-\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = f \quad \text{für } 0 \leq r < R, 0 \leq \varphi < 2\pi$$

$$u(R, \varphi) = g(\varphi) \quad \text{für } 0 \leq \varphi < 2\pi.$$

Wir setzen nun die Randdaten  $g$  fort und definieren dazu

$$G(r, \varphi) = \frac{r}{R} g(\varphi) \quad 0 \leq r \leq R, \quad 0 \leq \varphi < 2\pi.$$

Für

$$v = u - G$$

erhalten wir dann die pDgl

$$-\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial v}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 v}{\partial \varphi^2} = F \quad \text{für } 0 \leq r < R, 0 \leq \varphi < 2\pi$$

$$v(R, \varphi) = 0 \quad \text{für } 0 \leq \varphi < 2\pi$$

mit homogenen Dirichlet Randbedingungen und rechter Seite

$$F = f + \Delta G.$$

Wir entwickeln  $F$  in eine Fourier-Reihe bzgl.  $\varphi$

$$F(r, \varphi) \sim \sum_{k \in \mathbb{Z}} \hat{F}_k(r) e^{ik\varphi}$$

mit

$$\widehat{F}_k(r) = \frac{1}{2\pi} \int_0^{2\pi} F(r, \varphi) e^{-ik\varphi} d\varphi$$

und machen für  $v$  den Fourier-Ansatz

$$v(r, \varphi) \sim \sum_{k \in \mathbb{Z}} v_k(r) e^{ik\varphi}.$$

Einsetzen in die pDgl und Vergleich der Fourier-Koeffizienten liefert für die  $v_k$  die gDglen

$$-\frac{1}{r}(rv'_k)' + \frac{k^2}{r^2}v_k = \widehat{F}_k \quad , k \in \mathbb{Z}.$$

Dies sind lineare gDglen 2. Ordnung mit variablen Koeffizienten. Wir gehen zur Lösung ähnlich vor wie in Abschnitt VI.3 (S. 229, Teil I) und bestimmen zunächst die allgemeine Lösung der homogenen Gleichung. Sei dazu  $k \in \mathbb{Z} \setminus \{0\}$  beliebig, aber fest gewählt. Wir machen den Ansatz

$$w_k(r) = r^\alpha$$

für eine Lösung der homogenen Gleichung und erhalten die Bestimmungsgleichungen

$$\begin{aligned} 0 &= \frac{1}{r}(r\alpha r^{\alpha-1})' + \frac{k^2}{r^2}r^\alpha \\ &= -\alpha^2 r^{\alpha-2} + k^2 r^{\alpha-2} \\ &= (k^2 - \alpha^2)r^{\alpha-2}. \end{aligned}$$

Also lautet die allgemeine Lösung der homogenen Gleichung

$$w_k(r) = a_k r^k + b_k r^{-k}.$$

Für  $k = 0$  erhalten wir die allgemeine Lösung

$$w_0(r) = a_0 + b_0 \ln r.$$

Für uns sind aber nur die Lösungen von Interesse, die für  $r \rightarrow 0$  beschränkt bleiben. Daher machen wir zur Lösung der inhomogenen Gleichung den Variation-der-Konstanten-Ansatz

$$v_k(r) = a_k(r) r^{|k|} \quad , k \in \mathbb{Z}.$$

Einsetzen in die gDgl ergibt dann mit ein wenig Rechnung für alle  $k \in \mathbb{Z}$  die Bestimmungsgleichung

$$-\frac{1}{r^{|k|+1}} [r^{2|k|+1} a'_k]' = \widehat{F}_k$$

für die  $a_k$ . Wegen der Randbedingung  $a_k(R) = 0$  ergibt sich hieraus die Lösung

$$\begin{aligned} v_0(r) &= a_0(r) \\ &= \int_r^R \frac{1}{s} \int_0^s \tau \widehat{F}_0(\tau) d\tau ds \end{aligned}$$

$$\begin{aligned}
&= -\ln s \int_0^s \tau \widehat{F}_0(\tau) d\tau \Big|_{s=r}^{s=R} \\
&\quad + \int_r^R \ln s - s \widehat{F}_0(s) ds \\
&= -\ln R \int_0^R s \widehat{F}_0(s) ds \\
&\quad + \ln r \int_0^r s \widehat{F}_0(s) ds \\
&\quad + \int_r^R s \ln s \widehat{F}_0(s) ds
\end{aligned}$$

und für  $k > 0$

$$\begin{aligned}
v_k(r) &= r^{|k|} a_k(r) \\
&= -r^{|k|} \int_r^R \frac{1}{s^{2|k|+1}} \int_0^s \tau^{|k|+1} \widehat{F}_k(\tau) d\tau ds \\
&= -r^{|k|} \left\{ -\frac{1}{2|k|} s^{-2|k|} \int_0^s \tau^{|k|+1} \widehat{F}_k(\tau) d\tau \Big|_{s=r}^{s=R} \right. \\
&\quad \left. + \int_r^R \frac{1}{2|k|} s^{-2|k|} s^{|k|+1} \widehat{F}_k(s) ds \right\} \\
&= \frac{1}{2|k|} r^{|k|} R^{-2|k|} \int_0^R s^{|k|+1} \widehat{F}_k(s) ds \\
&\quad - \frac{1}{2|k|} r^{-|k|} \int_0^r s^{|k|+1} \widehat{F}_k(s) ds \\
&\quad - \frac{1}{2|k|} r^{|k|} \int_0^R s^{-|k|+1} \widehat{F}_k(s) ds.
\end{aligned}$$

Unter geeigneten Voraussetzungen an  $f$  und  $g$  kann man wieder zeigen, dass die Fourier-Reihe für  $v$  konvergiert. Insgesamt erhalten wir:

Die Lösung  $u$  der Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{aligned}
-\Delta u &= f && \text{in } \Omega \\
u &= g && \text{auf } \partial\Omega
\end{aligned}$$

im Kreis um Null mit Radius  $R$  ist gegeben durch

$$u = v + G$$

mit

$$G(r, \varphi) = \frac{r}{R} g(\varphi),$$

$$v(r, \varphi) = \sum_{k=-\infty}^{\infty} v_k(r) e^{ik\varphi}$$

und

$$\begin{aligned} v_0(r) &= -\ln R \int_0^R s \widehat{F}_0(s) ds \\ &\quad + \ln r \int_0^r s \widehat{F}_0(s) ds \\ &\quad + \int_r^R s \ln s \widehat{F}_0(s) ds \end{aligned}$$

$$\begin{aligned} v_k(r) &= \frac{1}{2|k|} \left\{ \left( \frac{r}{R^2} \right)^{|k|} \int_0^R s^{|k|+1} \widehat{F}_k(s) ds \right. \\ &\quad \left. - r^{-|k|} \int_0^r s^{|k|+1} \widehat{F}_k(s) ds \right. \\ &\quad \left. - r^{|k|} \int_r^R s^{-|k|+1} \widehat{F}_k(s) ds \right\}, \quad k \neq 0, \end{aligned}$$

$$\begin{aligned} \widehat{F}_k(r) &= \frac{1}{2\pi} \int_0^{2\pi} F(r, \varphi) e^{-ik\varphi} d\varphi, \quad k \in \mathbb{Z}, \\ F &= f + \Delta G. \end{aligned}$$



# Zusammenfassung

## XI Stochastik I: Diskrete Modelle

### 1. Modelle für Zufallsexperimente

Ergebnisse oder Elementarereignisse; Ergebnismenge; Ereignisse; Potenzmenge; Wahrscheinlichkeitsverteilung oder -maß; Wahrscheinlichkeit; Wahrscheinlichkeitsraum; Wahrscheinlichkeitsfunktion; Laplace-Experimente; Urnenmodelle: Stichproben in Reihenfolge mit Rücklegen, Stichproben in Reihenfolge ohne Rücklegen, Stichproben ohne Reihenfolge ohne Rücklegen, Stichproben ohne Reihenfolge mit Rücklegen; Anwendungsbeispiele; hypergeometrische Verteilung; Multinomialkoeffizienten; Identitäten für Binomialkoeffizienten

### 2. Bedingte Wahrscheinlichkeit und Unabhängigkeit

Bedingte Wahrscheinlichkeiten; Beispiele; Produktformel; Formel der totalen Wahrscheinlichkeit; Formel von Bayes; Anwendungen; Unabhängigkeit; Produktexperimente; Produkte von Wahrscheinlichkeitsräumen; Bernoulliexperiment und -verteilung; Binomialverteilung; Multinomialverteilung; geometrische Verteilung; negative Binomialverteilung oder Pascalverteilung

### 3. Zufallsvariable, Erwartungswert, Varianz

Zufallsvariable; Beispiele; Stabdiagramme; gemeinsame Verteilungsfunktion; Unabhängigkeit; Erwartungswert; Erwartungswert der Binomialverteilung; Erwartungswert der hypergeometrischen Verteilung; Varianz; Kovarianz; Streuung; Standardabweichung; Korrelationskoeffizient; unkorrelierte Zufallsvariable; Varianz der Binomialverteilung; Varianz der hypergeometrischen Verteilung; schwaches Gesetz der großen Zahl; Anwendungen

### 4. Grundbegriffe der Schätztheorie

Motivation; allgemeiner Rahmen; Schätzer; Maximum-Likelihood Schätzer; Erwartungstreue; Mittelwert; mittlerer quadratischer Fehler

### 5. Approximationen der Binomialverteilung

Stirlingsche Formel; Dichte der Standard-Normalverteilung; Satz von Moivre-Laplace; Verteilungsfunktion der Standard-Normalverteilung; Anwendungen; Poissonapproximation; Poissonverteilung

### 6. Tests

Motivation; Hypothese; Annehmen und Verwerfen einer Hypothese; Verwerfungsbereich oder kritischer Bereich; Fehler erster und zweiter Art; Teststatistik; kritischer Wert; Gütefunktion

## XII Stochastik II: Allgemeine Modelle

### 1. Wahrscheinlichkeitsmaße mit Dichten

Ergebnismengen;  $\sigma$ -Algebren; Borelmengen; Wahrscheinlichkeitsmaße; Wahrscheinlichkeitsräume; Dichten; Gleichverteilung auf einem Intervall; Exponentialverteilung; Normalverteilung; Produktdichten

## 2. Zufallsvariable und ihre Momente

Messbare Funktionen; Zufallsvariable; Verteilungsfunktion; Dichte; Unabhängigkeit; Faltung; Faltung von Normalverteilungen; Erwartungswert; Eigenschaften des Erwartungswertes; Varianz; Eigenschaften der Varianz

## 3. Schätzverfahren

Maximum-Likelihood Schätzung; Methode der kleinsten Quadrate; Regressionsgerade; Median

## 4. Tests

Likelihood-Quotienten; Likelihood-Quotienten Test;  $t$ -Test;  $\chi^2$ -Test; Anwendungen

## XIII Fourier-Analysis

### 1. Trigonometrische Polynome und Reihen

Periodische Funktionen; direkte periodische Fortsetzung; gerade periodische Fortsetzung; ungerade periodische Fortsetzung; trigonometrische Polynome; trigonometrische Reihen

### 2. Fourier-Reihen

Fourier-Koeffizienten; Fourier-Reihe; Rechteckschwingung; Rechenregeln; Bessel-Ungleichung; Riemann-Lemma; Konvergenz der Fourier-Reihe einer Funktion; Parseval-Gleichung; Anwendung auf gewöhnliche Differentialgleichungen

### 3. Die Fourier-Transformation

Fourier-Transformation; Fourier-Transformierte einer Funktion; Rechteckimpuls; Rechenregeln; Fourier-Transformierte einiger wichtiger Funktionen; Existenz und Eindeutigkeit der Fourier-Transformierten; Parseval-Gleichung

## XIV Partielle Differentialgleichungen

### 1. Einführung

Beispiele: Membrangleichung, Plattengleichung, Gasgleichung, Wärmeleitungsgleichung, Grundwasserströmung, Wellengleichung; Differentialoperatoren; Hauptteil; quasilinear; Koeffizienten; Differentialgleichung; Typen: elliptisch, parabolisch, hyperbolisch; Randbedingungen: Dirichlet, Neumann, gemischt

### 2. Die Wärmeleitungsgleichung

Kompatibilitätsbedingung; Fortsetzung der Randwerte; Lösung der homogenen Gleichung mit homogenen Randbedingungen; Separationsansatz; Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen; Integraldarstellung; Wärmeleitungskern; höhere Raumdimensionen; Zusammenhang mit dem Eigenwertproblem für den Laplace-Operator

### 3. Die Wellengleichung

Lösung der homogenen Gleichung mit homogenen Randbedingungen; Separationsansatz; Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen; Integraldarstellung; Wellengleichungskern; höhere Raumdimensionen; Dämpfung

### 4. Die Poissongleichung

Lösung des Eigenwertproblems im Rechteck; Lösung der Poissongleichung im Rechteck; Lösung des Eigenwertproblems im Kreis; Lösung der Poissongleichung im Kreis



## Index

- $*$ , 56
- $[\cdot]$ , 32
- $\text{Cov}(X, Y)$ , 26
- $EX$ , 24
- $E(X)$ , 24, 59
- $\mathcal{F}(f)$ , 85
- $\mathcal{F}^{-1}(F)$ , 85
- $\Omega$ , 7
- $P(A|B)$ , 15
- $\Phi$ , 41
- $\text{Var}(X)$ , 26
- $b_{n,p}$ , 20
- $\mathcal{P}$ , 7
- $\hat{f}$ , 85
- $\hat{f}_k$ , 78
- $h(s; n, N, S)$ , 12
- $\omega$ , 7
- $\varphi$ , 40
- $\rho_{XY}$ , 26
- $\sigma_X$ , 26
  
- Alternative, 47
- Anfangsbedingung, 93
- Annahme der Hypothese, 47
  
- Bayessche Formel, 17
- bedingte Wahrscheinlichkeit, 15
- Bernoulli-Experiment, 20
- Bernoulli-Verteilung, 20
- Bessel-Ungleichung, 81
- Bias, 34
- biharmonische Gleichung, 91
- Binomialverteilung, 20
- Borelmengen, 52
  
- $\chi^2$ -Test, 72
- $\chi^2$ -Verteilung, 71
  
- Dichte, 53, 55
- Dichte der
  - Standard-Normalverteilung, 40
  - Differentialoperator  $m$ -ter Ordnung, 95
- Diffusivität, 94
- Dirichlet-Randbedingung, 91
  
- Eigenfunktion, 112
- Eigenwert, 112
- Eigenwertproblem, 112
- Eigenwertproblem für den Laplace-Operator, 105
- Elementarereignis, 7
- elliptische pDgl, 97
- empirischen Median, 67
- Ereignis, 7, 52
- Erfolgswahrscheinlichkeit, 20
- Ergebnis, 7
- Ergebnismenge, 7
- erwartungstreu, 34
- erwartungstreuer Schätzer, 34
- Erwartungswert, 24, 59
- Erwartungswert der
  - Binomialverteilung, 25
  - hypergeometrischen Verteilung, 26
- erzeugte  $\sigma$ -Algebra, 51
- Exponentialverteilung, 54
  
- Faltung, 56
- Fehler erster Art, 47
- Fehler zweiter Art, 47
- Formel von Bayes, 17
- Formel von der totalen Wahrscheinlichkeit, 17
- Fortsetzung der Randwerte, 99
- Fourier-Koeffizienten, 78
- Fourier-Reihe, 78
- Fourier-transformierbar, 85
- Fourier-Transformierte, 85
  
- Gasgleichung, 92

- gemeinsame Verteilungsfunktion, 22
- geometrische Verteilung, 21
- Gleichverteilung, 9, 54
- Grad, 76
- Grundwasserströmung, 94
- Gütefunktion, 48
- Hauptteil eines Differentialoperators, 96
- homogene Wärmeleitungsgleichung
  - mit homogenen Randbedingungen, 99
- hyperbolische pDgl, 97
- hypergeometrische Verteilung, 12
- Hypothese, 47
- inhomogene Wärmeleitungsgleichung
  - mit homogenen Rand- und Anfangsbedingungen, 99
- inverse Fourier-Transformation, 85
- Koeffizienten eines Differentialoperators, 96
- Kompabilitätsbedingung, 98, 106
- komplexe Fourier-Koeffizienten, 78
- Konduktivität, 94
- Korrelationskoeffizient, 26
- Kovarianz, 26
- kritischer Bereich, 47
- kritischer Wert, 47
- Laplace-Experiment, 9
- Laplacescher
  - Wahrscheinlichkeitsraum, 9
- Likelihood-Funktion, 33
- Likelihood-Quotient, 67
- Likelihood-Quotienten Test, 68
- linearer Differentialoperator, 96
- linearer Differentialoperator mit konstanten Koeffizienten, 96
- Maximum-Likelihood Ansatz, 32
- Maximum-Likelihood Schätzung, 34
- Median, 67
- Membrangleichung, 91
- messbar, 55
- Methode der kleinsten Quadrate, 65
- Mittelwert, 35
- mittlerer quadratischer Fehler, 36
- Multinomialkoeffizient, 13
- negative Binomialverteilung, 21
- Neumann-Randbedingung, 91
- Normalverteilung, 54
- normierte Form, 41
- parabolische pDgl, 97
- Parseval-Gleichung, 82, 88
- Partialsomme, 77
- partielle Differentialgleichung  $m$ -ter Ordnung, 96
- partielle Differentialgleichung zweiter Ordnung, 91
- Pascal-Verteilung, 21
- pDgl, 91
- pDgl  $m$ -ter Ordnung, 96
- Periode, 75
- periodisch, 75
- Plattengleichung, 91
- Poisson-Approximation, 44
- Poisson-verteilt, 43
- Poissongleichung, 91
- Potenzmenge, 7
- Produkt von
  - Wahrscheinlichkeitsräumen, 19
- Produktformel für
  - Wahrscheinlichkeiten, 16
- Quanti, 70
- quasilinearer Differentialoperator, 95
- Randbedingung, 91
- Rechteckschwingung, 79
- reelle Fourier-Koeffizienten, 79
- Regressionsgerade, 65, 66
- Riemann-Lemma, 82
- Satz von Moivre-Laplace, 41
- Schätzer, 33
- schwaches Gesetz der großen Zahlen, 30
- Separationsansatz, 100
- sicheres Ereignis, 7
- $\sigma$ -Algebra, 51
- Stabdiagramm, 22
- Standardabweichung, 26, 60
- standardisierte Form, 41
- Stichprobenraum, 33
- Stirlingsche Formel, 38
- Streuung, 26
- Superposition, 99, 106
- $T$ -periodisch, 75
- $t$ -Test, 70
- $t$ -Verteilung, 70
- Test, 47
- Teststatistik, 47
- $t_{n-1}$ -Verteilung, 70

- trigonometrische Reihe, 77
- trigonometrisches Polynom, 76
  
- unabhängig, 18, 24, 56
- unbiased, 34
- unkorreliert, 26
- unmögliches Ereignis, 7
  
- Varianz, 26, 60
- Varianz der Binomialverteilung, 28
- Varianz der hypergeometrischen  
Verteilung, 29
- Verteilungsfunktion, 22, 53, 55
- Verteilungsfunktion der  
Standard-Normalverteilung, 41
- Verwerfen der Hypothese, 47
- Verwerfungsbereich, 47
  
- Wärmeleitungsgleichung, 92
- Wärmeleitungskern, 105
- Wahrscheinlichkeit, 7
- Wahrscheinlichkeitsfunktion, 8
- Wahrscheinlichkeitsmaß, 7
- Wahrscheinlichkeitsmaß, 52
- Wahrscheinlichkeitsraum, 7, 52
- Wahrscheinlichkeitsverteilung, 7
- Wellengleichung, 94
- Wellengleichungskern, 111
  
- Zufallsvariable, 21, 55