

# Mathematik III

## Wahrscheinlichkeitsrechnung

---

### Inhaltsverzeichnis

0 Vorbemerkungen .....	2
1 Zufällige Ereignisse und Wahrscheinlichkeiten .....	2
1.1 Elementare Wahrscheinlichkeitsrechnung .....	2
1.2 Berechnung von Wahrscheinlichkeiten durch kombinatorische Überlegungen .....	3
1.2.1 Spezialisierung des Fundamentalprinzips .....	4
1.2.2 Schema zur elementaren Kombinatorik .....	7
1.3 Endliche Wahrscheinlichkeiten .....	7
1.4 Geometrische Wahrscheinlichkeit .....	15
1.5 Bedingte Wahrscheinlichkeit, Unabhängige Ereignisse .....	18
1.5.1 Unabhängigkeit von Ereignissen .....	22
1.6 Axiomatische Begründung der Wahrscheinlichkeitstheorie .....	26
2 Zufallsgrößen und deren Verteilung .....	28
2.1 Zufallsgrößen, Verteilungen .....	33
2.2 Erwartungswert .....	41
2.2.1 Eigenschaften des Erwartungswertes .....	42
2.2.2 Eigenschaften der Streuung .....	42
2.3 Faltung zweier Verteilungen .....	44
2.3.1 Momente einer Zufallsgröße .....	46
2.4 Charakteristische Funktionen .....	46
2.5 Einführung der Kovarianz (Abweichungen) .....	50
2.6 Folgen von Zufallsgrößen .....	50
2.6.1 Tschebyscheffsche Ungleichung .....	51
2.7 Gesetze der großen Zahlen .....	52
2.7.1 Bernoullisches Gesetz der großen Zahlen: .....	53

## 0 Vorbemerkungen

- Mitte des 17. Jh. Chevalier de Mèrè,

Blaise Pascal

Spiel mit 3 Würfeln Augensumme 11 öfter als Augensumme 12

Auffassung, daß Ereignisse „11“ oder „12“ gleichwahrscheinlich sind  
Überlegung: 6 Möglichkeiten, eine 11 zu würfeln

6-4-1          6-3-2          5-5-1          5-4-2          5-3-3          4-4-3

aber auch 6 Möglichkeiten für die 12:

6-5-1          6-4-2          6-3-3          5-5-2          5-4-3          4-4-4

Briefwechsel zwischen Pascal und Fermat:          Anfang der Wahrscheinlichkeitstheorie

- Moderne Theorie

Stochastik → Wahrscheinlichkeitstheorie und Statistik  
seit Anfang dieses Jahrhunderts: Anstrengungen zur Formalisierung

## 1 Zufällige Ereignisse und Wahrscheinlichkeiten

### 1.1 Elementare Wahrscheinlichkeitsrechnung

**Beispiel:** Münzwurf

deterministischer Vorgang, Aufgrund unscharfer Anfangsbedingungen ist Ergebnis nicht exakt vorhersagbar. Zwei einander ausschließende Ergebnisse: Kopf oder Zahl. Jedes Ergebnis ist gleichberechtigt → Kopf oder Zahl treten jeweils mit Wahrscheinlichkeit  $\frac{1}{2}$  auf.

**Fazit:** Experiment, Versuch mit ungewissem Ausgang, zumindest prinzipiell oder gedanklich unter gleichen Versuchsbedingungen beliebig oft wiederholbar. Beobachtung: Bei unabhängiger Wiederholung derartiger Versuche sind Gesetzmäßigkeiten erkennbar.

**Spezialfall:** Alle endlich vielen Versuchsausgänge schließen einander aus und sind gleichberechtigt. Versuch wird „auf gut Glück“ durchgeführt.

Unter diesen Voraussetzungen sei A ein Ereignis, das aus verschiedenen Versuchsausgängen zusammengesetzt sein kann.

Wahrscheinlichkeit (von Laplace)

$$P(A) = \frac{\text{Anzahl der für A günstigen Fälle}}{\text{Anzahl aller möglichen Fälle}}$$

**Beispiel:** Münze  
Spielwürfel - 6 Ausgänge, A = „6“, A = „gerade Zahl“  
zwei Spielwürfel - 36 Ausgänge, A = „gleiche Augenzahl“

**Praxis:** lange Beobachtungsreihe: mehrmalige unabhängige Durchführung ein und desselben Experiments, Ereignis A det. Bei n Versuchen trete n(A) - mal das Ereignis A ein. Dann zeigt sich, daß die relative Häufigkeit

$$\frac{n(A)}{n} \quad (n = 1, 2, 3, \dots)$$

für wachsendes n stabil ist. Sie schwankt mehr oder weniger um einen gewissen Wert, nämlich die Wahrscheinlichkeit P(A).

→ statistischer Wahrscheinlichkeitsbegriff, begründet Kontakt der (noch zu entwickelnden) Theorie mit der Realität

Problem beim Laplace'schen Wahrscheinlichkeitsbegriff: gleichwahrscheinliche Ereignisse finden, mit deren Hilfe das interessierende Ereignis zusammengesetzt werden kann.

## 1.2 Berechnung von Wahrscheinlichkeiten durch Kombinatorische Überlegungen

K Mengen  $A_1, A_2, \dots, A_k$   
vom Umfang  $n_1, n_2, \dots, n_k$  Elemente

Kombination von Elementen bilden  $(a_1, a_2, \dots, a_k)$  mit  $a_i \in A_i$   
Wie viele solcher Kombinationen sind möglich ?

### Fundamentalprinzip der Kombinatorik:

Anzahl verschiedener geordneter k-Tupel  $N = n_1 \cdot n_2 \cdot \dots \cdot n_k$

**Beispiel:** Wie groß ist die Wahrscheinlichkeit für das Ereignis A, beim Spielen mit 3 Würfeln die maximale Augenzahl zu erreichen ?

$A_i$  (i = 1, 2, 3) gleich, sechselementig

→  $N = 6^3$  Möglichkeiten

→ 1 günstig →  $P(A) = 1/216$

## 1.2.1 Spezialisierung des Fundamentalprinzips

- geordnete Probe mit Wiederholung

Elemente können mehrfach berücksichtigt werden, „Auswahl mit Zurücklegen“. Eine Menge  $A$  mit  $n$  Elementen,  $k$  Elemente wählen

→ Zurückführung auf Fundamentalprinzip:

$$A_1 = A_2 = \dots = A_k = A \quad \rightarrow \quad N = n^k$$

- geordnete Probe ohne Wiederholung

**Eine** Menge  $A$  mit  $n$  Elementen,  $k \leq n$  Elemente wählen, dabei Zurücklegen verboten

→ Zurückführung auf Fundamentalprinzip:

$A_1 = A$	Element $a_1$ wählen	$n$	Elemente
$A_2 = A \setminus \{a_1\}$	Element $a_2$ wählen	$n - 1$	Elemente
$A_3 = A \setminus \{a_1, a_2\}$	Element $a_3$ wählen	$n - 2$	Elemente
$A_k = A \setminus \{a_1, a_2, \dots, a_{k-1}\}$	Element $a_k$ wählen	$n - k + 1$	Elemente

$$\rightarrow N = n(n-1)(n-2) \dots (n-k+1) = \frac{n!}{(n-k)!}$$

Noch spezieller:

$k = n$  setzen → Anzahl der Permutationen von  $n$  Elementen ohne Wiederholung

$$N = n!$$

### Beispiel:

Gruppe von  $k$  Studenten sitzt in einem Zug mit  $n \geq k$  Wagen. Jeder Student habe seinen Wagen unabhängig und „auf gut Glück“ gewählt. Gesucht ist die Wahrscheinlichkeit, daß alle Studenten in verschiedenen Wagen sitzen.

Zunächst **alle** Möglichkeiten:

$k$  Studenten auf  $n$  Wagen aufteilen, Problem mit Wiederholung und mit Berücksichtigung der Reihenfolge → insgesamt  $N = n^k$  Möglichkeiten, Studenten auf die Wagen aufzuteilen.

Ereignis  $A$ : „höchstens ein Student pro Wagen“  
für  $A$  günstige Fälle: Wiederholung verboten, Zahl der Kombinationen (mit Berücksichtigung der Reihenfolge, ohne Wiederholung)

$$N(A) = \frac{n!}{(n-k)!} \text{ Möglichkeiten}$$

$$\rightarrow \frac{n!}{n^k \cdot (n-k)!} = \frac{n \cdot (n-1) \cdot (n-2) \dots (n-k+1)}{n^k}$$

Jetzt die Berücksichtigung der Reihenfolge aufgeben.

Menge A vom Umfang n, k-elementige Teilmengen  $\{a_1, a_2, \dots, a_k\}$  bilden. Also ohne Rücklegen, ohne Wiederholung. Gesuchte Anzahl: N

Wir bekommen alle möglichen **geordneten** Proben aus A vom Umfang k ohne Wiederholung und jede nur einmal, indem wir zunächst eine beliebige k-elementige Teilmenge von A wählen und dann alle ihre Permutationen bilden.

Nach dem Fundamentalprinzip ergibt sich:

$$N \cdot k! = \frac{n!}{(n-k)!}$$

Teilmengen · Permutation = Proben (m. R., o. W.)  
k-elementig

$$\rightarrow N = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k} \quad \text{Binominalkoeffizient}$$

Zerlegung der Menge A in Teilmengen  $B_1, B_2, \dots, B_s$  vom Umfang  $k_1, k_2, \dots$ , bzw.  $k_s$

$$k_1 + k_2 + \dots + k_s = n$$

→ Permutation mit Wiederholung

$$N = \frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_s!} \quad \text{Polynominalkoeffizient}$$

**Beispiel:**

Qualitätskontrolle, Los von 100 Teilen, 10 wurden „auf gut Glück“ gewählt und kontrolliert. Falls kein Ausschuß, Annahme des Loses. Gesucht ist die Wahrscheinlichkeit des Ereignisses A, daß ein Los mit 100 Ausschußteilen nicht beanstandet wird ?

- Anzahl der Möglichkeiten 10 Teile aus 100 zu wählen, Kombinationen ohne Berücksichtigung der Reihenfolge, ohne Wiederholung

$$N = \binom{100}{10} = \frac{100!}{10! \cdot 90!}$$

- Anzahl der für das Ereignis A günstigen Fälle: 90 gute Teile, **davon** 10 heraus greifen

$$N = \binom{90}{10} = \frac{90!}{10! \cdot 80!}$$

$$P(A) = \frac{90! \cdot 90!}{100! \cdot 80!} = \frac{81 \cdot 82 \cdot \dots \cdot 90}{91 \cdot 92 \cdot \dots \cdot 100} \approx \left(\frac{9}{10}\right)^{10} = \left(1 - \frac{1}{10}\right)^{10} \approx \frac{1}{e} \approx 0,36$$

Noch offen:

Aus n Elementen k herauszugreifen, mit Wiederholung (zurücklegen erlaubt), ohne Berücksichtigung der Reihenfolge.

Anzahl der möglichen Kombinationen

$$N = \binom{n+k-1}{k}$$

**Beispiel:**  
aufteilen

Auf wieviele Weisen lassen sich k Markstücke auf n Personen

A Menge von n Personen, numerieren

Problem: ohne Berücksichtigung der Reihenfolge, mit Wiederholung

Dies gibt Anlaß zu alternativen Interpretationen von Kombinationen:

Gegeben sind n Zellen, auf die k Teilchen aufgeteilt werden sollen.

- Teilchen sind unterscheidbar oder nicht  
Problem mit oder ohne Berücksichtigung der Reihenfolge
- Mehrfachbelegung einer Zelle möglich oder nicht  
Problem mit oder ohne Wiederholung

**Beispiel:**

Es sind n unterschiedliche Dinge auf N Schubfächer zu verteilen. Jede Möglichkeit der Verteilung sei gleich wahrscheinlich. Mit welcher Wahrscheinlichkeit gelangen bei der Verteilung in ein Schubfach k Dinge ?

Jedes Ding läßt sich in eines der N Schubfächer legen, N Möglichkeiten ein Ding zu verteilen.

Fundamentalprinzip

→ Anzahl aller Möglichkeiten bei der Verteilung ist  $N^n$

Anzahl der günstigen Fälle:

→ k von n Dingen auf ein Schubfach verteilen (ohne Wiederholung, ohne Berücksichtigung der Reihenfolge)

$$\binom{n}{k} \text{ Möglichkeiten dafür}$$

→ die übrigen n - k Dinge in N - 1 Schubfächer (egal wie) verteilen

$$(N - 1)^{n-k}$$

$$\rightarrow P = \frac{\binom{n}{k} \cdot (N - 1)^{n-k}}{N^n} = \binom{n}{k} \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{n-k}$$

### 1.2.2 Schema zur elementaren Kombinatorik

Proben vom Umfang k	ohne Berücksichtigung der Reihenfolge	mit Berücksichtigung der Reihenfolge	
ohne Wiederholung	$\binom{n}{k}$	$\frac{n!}{(n-k)!}$	keine Mehrfachbelegung
mit Wiederholung	$\binom{n+k-1}{k}$	$n^k$	mit Mehrfachbelegung
	Teilchen sind nicht unterscheidbar	Teilchen sind unterscheidbar	Verteilung von k Teilchen auf n Zellen

### 1.3 Endliche Wahrscheinlichkeiten

→ nächster Schritt zur Formalisierung des Wahrscheinlichkeitsbegriffes, durch Identifizierung zufälliger Ereignisse mit **Mengen**

Zufälliger Versuch habe Ausgang  $\omega$  : Realisierung, Stichprobe oder Elementarereignis

**Elementarereignis:** alle Elementarereignisse (alle Versuchsausgänge) in einer Menge  $\Omega$  zusammenfassen: Raum aller Elementarereignisse. (gewisse) Teilmengen von  $\Omega$  bilden dann die zufälligen Ereignisse,  $\Omega$  und  $\emptyset$  seien Ereignisse.

**Beispiel:** Würfelexperiment  
 $\Omega = \{1, 2, 3, 4, 5, 6\}$   
 Ereignis A: „gerade Zahl würfeln“  
 $A = \{2, 4, 6\} \subseteq \Omega$





$\Omega$  das „sichere Ereignis „ tritt stets ein  
 $\emptyset$  das „unmögliche Ereignis“ tritt niemals ein

→ Das Eintreten des Elementarereignisses  $\omega$  hat alle Ereignisse  $A$  mit  $\omega \in A$  zur Folge. Die Ereignisse  $A$  sind Teilmengen des Raumes der Elementarereignisse

$$\omega \in A \subseteq \Omega$$

Damit ist eine sichere Basis für alle näheren Betrachtungen gelegt.  
Raum der Elementarereignisse entsprechend den Regeln der Mengenlehre strukturieren.

**Definition:** Das Ereignis  $A$  aus  $\Omega$ , ziehe das Ereignis  $B$  aus  $\Omega$  nach sich,  $A \subseteq B$ , falls  $\omega \in A \Rightarrow \omega \in B$

Stets gilt:  $\emptyset \subseteq A, A \subseteq \Omega$

**Definition:** Zieht  $A \subseteq \Omega$  das Ereignis  $B \subseteq \Omega$  sowie  $B$  das Ereignis  $A$  nach sich, so heißen die Ereignisse  $A$  und  $B$  gleich.

$$A = B \Leftrightarrow A \subseteq B \wedge B \subseteq A$$

**Definition:** Die Summe (Vereinigung)  $A \cup B$  der Ereignisse  $A, B \subseteq \Omega$ , tritt genau dann ein, wenn wenigstens eines der Ereignisse  $A$  oder  $B$  eintritt.

$$\omega \in A \cup B \Leftrightarrow \omega \in A \vee \omega \in B$$

$$\begin{array}{ll} \text{Stets gilt:} & A \cup \emptyset = A & A \subseteq A \cup B \\ & A \cup A = A & B \subseteq A \cup B \\ & A \cup \Omega = \Omega & \end{array}$$

$$\begin{array}{ll} \text{Kommutativität:} & A \cup B = B \cup A \\ \text{Assoziativität:} & A \cup (B \cup C) = (A \cup B) \cup C \end{array}$$

**Allgemein:** endliche Summe von Ereignissen

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n$$

$(A_i)_{i=1,2,\dots}$  Folge von Ereignissen,  $A_i \subseteq \Omega$ . Dann bedeutet  $\bigcup_{i=1}^{\infty} A_i$  das Ereignis, das genau dann eintritt, wenn mindestens eines der  $A_i$  eintritt.

**Definition:** Das Produkt (der Durchschnitt)  $A \cap B$  der Ereignisse  $A, B \in \Omega$  tritt genau dann ein, wenn sowohl  $A$  als auch  $B$  eintreten.

$$\omega \in A \cap B \Leftrightarrow \omega \in A \wedge \omega \in B$$

Stets gilt:

$A \cap \emptyset = \emptyset$	$A \cap B \subseteq A$
$A \cap A = A$	$A \cap B \subseteq B$
$A \cap \Omega = A$	

Außerdem gilt das Kommutativ- und das Assoziativgesetz bezüglich  $\cap$ .

**Allgemein:** endliches Produkt von Ereignissen

$$A_i \subseteq \Omega$$

$$\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap \dots \cap A_n \quad \text{jedes der } A_i \text{ tritt ein}$$

$$\bigcap_{i=1}^{\infty} A_i$$

**Definition:** Das zu  $A \subseteq \Omega$  komplementäre Ereignis  $\bar{A}$  tritt genau dann ein, wenn A nicht eintritt. Es gilt

$$A \cap \bar{A} = \emptyset \quad A \cup \bar{A} = \Omega$$

, d.h. A und  $\bar{A}$  zerlegen den Raum der Elementarereignisse

$$\bar{\Omega} = \emptyset \quad \bar{\emptyset} = \Omega \quad A \subseteq B \Rightarrow \bar{B} \subseteq \bar{A}$$

**Definition:** Sind A und B zufällige Ereignisse, so bezeichnen wir das Ereignis das genau dann eintritt, wenn A aber nicht B entritt, mit  $A \setminus B$ .

Es gilt:  $A \setminus B = A \cap \bar{B}$

**Bemerkung:** Aufgrund der von uns gewählten Konstruktion gelten für Ereignisse grundsätzlich die Regeln der Mengenlehre, etwa die Distributivgesetze:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

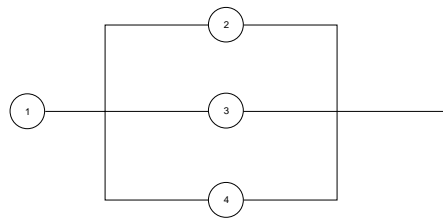
oder die DeMorgan'sche Regel:

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \quad \text{allgemein} \quad \overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \bar{A}_i$$

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \text{allgemein} \quad \overline{\bigcup_{i=1}^{\infty} A_i} = \bigcap_{i=1}^{\infty} \bar{A}_i$$

**Beispiel:**

4 Geräte seien in der folgenden Weise geschaltet:



$A_i$  bezeichne das zufällige Ereignis: „Das Gerät  $i$  fällt aus“  
( $i = 1, 2, 3, 4$ )

Ereignis A: „Das System fällt aus“  $A = A_1 \cup (A_2 \cap A_3 \cap A_4)$

Ereignis B: „Das System fällt nicht aus“  $B = A_1 \cap (\overline{A_2} \cup \overline{A_3} \cup \overline{A_4})$

**Definition:**

Eine Menge  $M$  zufälliger Ereignisse heißt **Ereignisalgebra**, wenn gilt:

E1:  $\Omega \in M$

E2:  $A \in M, B \in M \rightarrow A \cup B \in M$  ( $M$  ist  $\cup$ -stabil)

E3:  $A \in M \rightarrow \overline{A} \in M$

Enthält  $M$  unendlich viele Elemente, so habe  $M$  überdies die

Eigenschaft E4 :  $A_i \in M (i = 1, 2, \dots) \rightarrow \bigcup_{i=1}^{\infty} A_i \in M$

In diesem Fall ist  $M$  eine  $\sigma$ -Algebra.

**Folgerungen:**

Eine Ereignisalgebra besitzt die Eigenschaften:

1.  $\emptyset \in M$

2.  $A, B \in M \rightarrow A \cap B \in M, A \setminus B \in M$  ( $M$  ist  $\cap$ -stabil)

3.  $A_i \in M (i = 1, 2, \dots) \rightarrow \bigcap_{i=1}^{\infty} A_i \in M$

**Beispiel:**

1.  $\Omega$  gegeben, sowie  $A \subseteq \Omega ; A \neq \Omega ; A \neq \emptyset ;$   
Ereignisalgebra  $M$  mit  $A \in M$  konstruieren

$$M = \{A, \overline{A}, \emptyset, \Omega\}$$

2. Kleinste Algebra:  $M = \{\emptyset, \Omega\}$

3. Größte Algebra:  $M = \wp(\Omega)$  (Potenzm. aller Teilmeng.)

Sei  $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$  endlich mit  $N$  Elementen. Wieviele Elemente enthält dann die Potenzmenge?

Mengen  $A \subseteq \Omega$  konstruieren: Mengenbildungsprinzip

$$\omega_i \in A \text{ oder } \omega_i \notin A$$

Also  $N$  „Teilchen“ auf **zwei** Zellen aufteilen.

→  $\wp(\Omega)$  enthält  $2^N$  Elemente.

**Beispiel:** Spiel mit zwei Würfeln:  $6 \cdot 6 = 36$  Elementarereignisse  
 $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\}$

$\wp(\Omega)$  enthält  $2^{36} = 68.719.476.736$  Elemente

$\wp(\Omega)$  enthält qualitativ mehr Elemente als  $\Omega$ .

$\Omega$  abzählbar  $\Omega = \{\omega_1, \omega_2, \dots\}$

$\wp(\Omega)$  überabzählbar

In diesem Falle lassen sich „gerade noch“ Wahrscheinlichkeiten auf allen Ereignissen  $A \subset \wp(\Omega)$  definieren. Bei überabzählbaren  $\Omega$ , z.B.  $\Omega = \mathbb{R}^1$  gilt das nicht mehr.

### Borel-Mengen

Sei  $\Omega = \mathbb{R}^1$  die reelle Achse. Nach dem Ereignis  $\{\omega : \omega = x\}$  ( $x \in \mathbb{R}$ ) zu fragen ist oft nicht sinnvoll, dagegen nach  $\{\omega : \omega \in \text{Intervall}\}$  sehr.

→ spezielle Teilmengen von  $\Omega$  betrachten:

$$A = (a, b] ; -\infty < a < b < \infty$$

Bilden diese Mengen eine Algebra ?

$$(a, b] \cup (b, c] = (a, c] \quad , \quad \bigcup_{i=-\infty}^{\infty} (i, i+1] = \mathbb{R}$$

Wir müssen Menge hinzunehmen.

**Definition:** Sei  $\Omega = \mathbb{R}^1$ . Mit  $\mathcal{L}^1$  bezeichnen wir die kleinste  $\sigma$ -Algebra (Existenz gesichert), die alle betroffenen Intervalle  $(a, b]$ ,  $-\infty < a < b < \infty$  enthält.

$\mathcal{L}^1$  heißt  $\sigma$ -Algebra der Borel-Mengen in  $\mathbb{R}^1$ . Mittels geeigneter Parallelepipede (Rechteck, Quader, ...) definiert man die  $\sigma$ -Algebra  $\mathcal{L}^n$  der Borel-Mengen im  $\mathbb{R}^n$ .

Im folgenden setzen wir nun stets voraus, daß der jeweils betrachtete Raum der Elementarereignisse  $\Omega$  durch eine geeignete Ereignisalgebra strukturiert sei. Nur die Elemente  $A \in \mathcal{M}$  wollen wir künftig als Ereignisse zulassen.

**Definition:** Es sei  $\Omega$  ein Elementarereignisraum versehen mit der Ereignisalgebra

**M.** Zwei Ereignisse  $A, B \in \mathcal{M}$  heißen **unvereinbar** oder **disjunkt**, falls ihr gemeinsames Eintreten unmöglich ist.

$$A \cap B = \emptyset$$

Die Ereignisse A und B schließen einander aus.

**Beispiel:** 1 Würfel

Ereignis A = „ungerade Zahl würfeln“

Ereignis B = „eine Zahl größer 5 würfeln“

**Definition:** Eine Menge nicht unmöglich zufälliger Ereignisse  $\{A_1, A_2, \dots, A_n, \dots\}$ , ( $\emptyset \neq A \in \mathcal{M}$ ) heißt vollständiges Ereignissystem wenn gilt:

V1:  $A_i$  ( $i = 1, 2, \dots$ ) paarweise unvereinbar

$$A_i \cap A_k = \emptyset \quad (i \neq k)$$

V2: Vollständigkeit

$$A_1 \cup A_2 \cup \dots \cup A_n \cup \dots = \Omega$$

Vollständiges Ereignissystem **zerlegt** den Raum  $\Omega$ .

Nächstes Ziel ist die Weiterentwicklung des Wahrscheinlichkeitsbegriffes.

Spezialisierung: Raum der Elementarereignisse sei höchstens abzählbar:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \text{ oder } \Omega = \{\omega_1, \omega_2, \dots\}$$

Ereignisalgebra  $\mathcal{M}$  enthalte alle einpunktigen Ereignisse  $A_i = \{\omega_i\} \in \mathcal{M}$  ( $i = 1, 2, \dots$ ).  $\{A_1, A_2, \dots, A_n\}$  bildet also vollständiges Ereignissystem.

→ Für alle  $A \in \mathcal{M}$  Wahrscheinlichkeit  $P(A)$  definieren:

$$A \rightarrow P(A) \quad , \quad A \in \mathcal{M}$$

Zuerst Funktion  $P : \mathcal{M} \rightarrow [0, 1]$  auf dem vollständigen Ereignissystem definieren:

$$A_i \rightarrow P(A_i) = P(\{\omega_i\}) = p_i \quad i = 1, 2, \dots$$

Dabei muß für die Folge  $(p_i)_{i=1,2,\dots}$  natürlich gelten  $p_i > 0$  und  $\sum_i p_i = 1$  für  $i = 1, 2, \dots$

Für ein beliebiges Ereignis  $A \in \mathcal{M}$  setzen wir  $P(A) := \sum_{\omega_i \in A} p_i$ .

**Beispiel:**  $\Omega = \{0, 1, 2, \dots\}$  Menge der natürlichen Zahlen.

$\lambda > 0$  feste reelle Zahl

$$p_k = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$\rightarrow p_k > 0 \quad \sum_{k=0}^{\infty} p_k = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1$$

*Eigenschaften der Wahrscheinlichkeit P:*

1.  $0 \leq P(A) \leq 1$  , mit  $P(A) = \sum_{\omega_i \in A} p_i$

2.  $P(\Omega) = \sum_{\omega_i \in \Omega} p_i = 1$

3.  $P(\bar{A}) = \sum_{\omega_i \in \bar{A}} p_i = \sum_{\omega_i \in \Omega} p_i - \sum_{\omega_i \in A} p_i = 1 - P(A)$

4.  $P(\emptyset) = 1 - P(\Omega)$

5. Additionstheorem: A und B unvereinbar  $\rightarrow P(A \cup B) = P(A) + P(B)$

6. Verallgemeinerung:  $A_i, i = 1, 2, \dots$  höchstens abzählbar viele paarweise unvereinbare Ereignisse

$$A_i \cap A_k = \emptyset \quad (i \neq k; i, k = 1, 2, \dots)$$

$$\rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

7. Unvereinbarkeit fallen lassen:  $A, B \in \mathcal{M}$  beliebig

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

8. Isotonie der Wahrscheinlichkeit:

$$A \leq B \rightarrow P(A) \leq P(B)$$

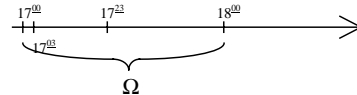
## 1.4 Geometrische Wahrscheinlichkeit

Zufällige Experimente mit überabzählbar vielen Ausgängen können mit elementaren Methoden behandelt werden.

**Beispiel:** Eine Dame verspricht, zwischen 17.00 und 18.00 Uhr zu einem Rendezvous zu erscheinen, nähere Angaben macht sie nicht.

Wahrscheinlichkeit für das Ereignis, daß sie zwischen 17.03 und 17.23 Uhr eintrifft ist gesucht.

Anschauung und Intuition:

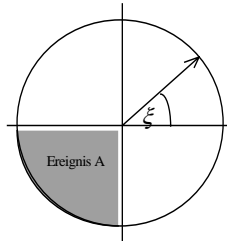


$$P(A) = \frac{20 \text{ min}}{60 \text{ min}} = \frac{1}{3}$$

Abstraktion: „Inhalt“ der Menge  $A = \{t : 17.03 \leq t \leq 17.23\}$   
 „Inhalt“ der Menge  $\Omega = \{t : 17.00 \leq t \leq 18.00\}$

$$P(A) = \frac{\text{Inhalt (A)}}{\text{Inhalt } \Omega}$$

**Beispiel:** Glücksrad mit Zeiger; zufälliger Versuch: Rad drehen, wo bleibt Zeiger stehen, jede Zeigerstellung gleichberechtigt.



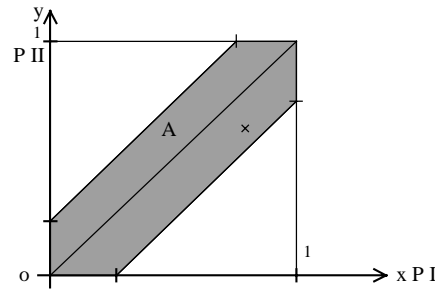
$$\Omega = \{\gamma : 0 \leq \gamma \leq 2 \cdot \pi\}$$

für Ereignis A günstige Zeigerstellungen identifizieren

$$A = \{ \xi : \pi \leq \xi \leq \frac{3}{2} \pi \}$$

$$P(A) = \frac{\text{Inhalt (A)}}{\text{Inhalt } (\Omega)} = \frac{\frac{\pi}{2}}{2\pi} = \frac{1}{4}$$

**Beispiel:** Zwei Personen vereinbaren, sich zwischen 12 und 13 an einem bestimmten Ort zu treffen. Jeder wartet auf den anderen nötigenfalls 15 Minuten, danach geht er. Wie groß ist die Wahrscheinlichkeit für ein Treffen ?



Punkt  $(x, y)$  repräsentiert  
Ankunftszeiten für beide  
Personen I und II.

$$\Omega = [0, 1]^2$$

$$A = \{(x, y) \in \Omega : |x - y| \leq \frac{1}{4}\}$$

$$\rightarrow y = x + \frac{1}{4} \text{ und } y = x - \frac{1}{4}$$

$$P(A) = \frac{\text{Inhalt}(A)}{\text{Inhalt}(\Omega)} = 1 - P(\bar{A}) = 1 - \left(\frac{3}{4}\right)^2 = \frac{7}{16}$$

**Allgemeines Modell:**  
in

Versuch lasse sich als zufälliges Werfen eines Punktes

einem beschränkten Grundbereich  $\Omega$  des n-dimensionalen euklidischen Raumes zu interpretieren. Dabei gelte:

- Der geworfene Punkt kann auf jeden beliebigen Punkt  $\omega \in \Omega$  fallen
- Inhaltsgleiche Teilmengen von  $\Omega$  (= Ereignisse) haben die gleiche Wahrscheinlichkeit
- Das sichere Ereignis entspricht dem Grundbereich  $\Omega$

Dann berechnet sich die Wahrscheinlichkeit des Ereignisses  $A \subseteq \Omega$  nach der Formel:

$$P(A) = \frac{\text{Inhalt}(A)}{\text{Inhalt}(\Omega)}$$

**Bemerkungen:**

→ Wahrscheinlichkeit eines Ereignisses ist also unabhängig von der speziellen Lage und Gestalt in  $\Omega$

→ Die Analogie zum klassischen Laplace'schen Wahrscheinlichkeitsbegriff ist offensichtlich.

→ Die derartig definierte geometrische Wahrscheinlichkeit hat viel Anlaß zu Mißverständnissen und Einwänden gegeben. Grund: Paradoxon von Bertrand.

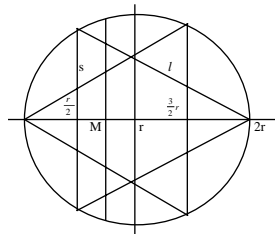
**Aufgabe:**

In einem Kreis wird zufällig eine Sehne gezogen. Wie groß ist die Wahrscheinlichkeit dafür, daß deren Länge die Seite eines im Kreis eingeschriebenen gleichseitigen Dreiecks übertrifft ?



### 1. Auffassung:

Aus Symmetriegründen o.B.d.A. Richtung der Sehne festhalten. Dann senkrecht Durchmesser des Kreises betrachten.



Wann gilt  $s > l$  ?

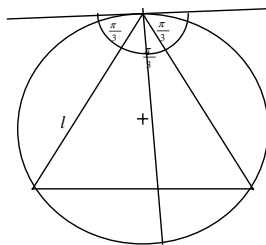
Mittelpunkt M der Sehne zwischen  $\frac{r}{2}$  und

$$\frac{3}{2}r$$

$$\rightarrow P = \frac{r}{2r} = \frac{1}{2}$$

### 2. Auffassung:

Spitzen des gleichseitigen Dreiecks in einem Endpunkt der Sehne, Winkel der Sehne mit Tangente.

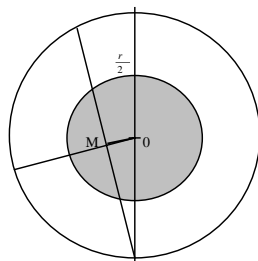


günstiger Bereich:  $\frac{\pi}{3} \dots \frac{2}{3}\pi$

$$\rightarrow \frac{\frac{\pi}{3}}{\pi} = \frac{1}{3}$$

### 3. Auffassung:

Mittelpunkt M der Sehne zufällig im Kreisinnern wählen. Außerdem muß Abstand M zu 0 kleiner als  $\frac{r}{2}$  sein.



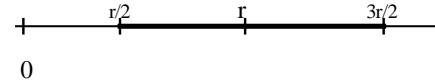
$$P = \frac{\text{Inhalt (Kreis mit } \frac{r}{2})}{\text{Inhalt (Kreis mit } r)}$$

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

Lösung der Aufgabe offenbar von Lösungsweg abhängig.

Auflösung des Paradoxon: Es wurden jeweils verschiedene Aufgaben formuliert.

1. „zufällig“ bedeutet hier, den Punkt M „auf gut Glück“ auf  $(0, 2r)$  zu wählen
  2. Winkel  $\xi$  „auf gut Glück“ in  $(0, \pi)$  wählen
  3. Punkt „auf gut Glück“ in der inneren Kreisfläche wählen
- damit ganz verschiedene Zufallsmechanismen vorausgesetzt
- unterschiedliche Lösungen



### 1.5 Bedingte Wahrscheinlichkeit, Unabhängige Ereignisse

- Gegeben:**
- Wahrscheinlichkeitsraum  $(\Omega, M, P)$
  - Modell für (realen) Bedingungskomplex eines zufälligen Experiments
  - Ereignis  $B \in M, P(B) > 0$

zusätzliche Hypothese: „Das Ereignis B tritt ein“

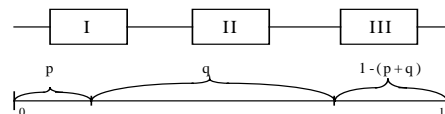
Durch Hinzunahme dieser Hypothese wird der Bedingungskomplex geändert. Folglich werden sich i.a. auch die Wahrscheinlichkeiten der Ereignisse  $A \in M$  ändern.

**Definition:** Unter den obigen Voraussetzungen heißt:

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

die bedingte Wahrscheinlichkeit des Ereignisses A unter der Bedingung B.

**Beispiel:** Maschinensystem in Reihe



Wahrscheinlichkeit für Ursache eines Ausfalls des Systems bei:

- Maschine I:  $p$
- Maschine II:  $q$
- Maschine III:  $1 - (p + q)$

System sei ausgefallen, bereits vergeblich nach einem Fehler in Maschine I gesucht. Wie ist die Wahrscheinlichkeit dafür, daß dann die Ursache in Maschine II liegt?

Ereignis A: „Ursache des Ausfalls liegt genau an Maschine II“

Ereignis B: „Ursache des Ausfalls liegt nicht an Maschine II“

Gesucht:  $P(A | B)$

Es gilt hier  $A \subseteq B$ , sowie  $P(A) = q$

$P(B) = 1 - P(A) = 1 - p$  und weiter

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \stackrel{A \subseteq B}{=} \frac{P(A)}{P(B)} = \frac{q}{1-p}$$

### Eigenschaften der bedingten Wahrscheinlichkeit

1.  $0 \leq P(A | B) \leq 1$

2. B unvereinbar  $\rightarrow P(A | B) = \frac{\overbrace{P(A \cap B)}^{\emptyset}}{P(B)} = 0$

3.  $B \subseteq A \rightarrow P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

4.  $[X1] | \Omega = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A)$

5. Die bedingte Wahrscheinlichkeit  $P(A | B)$  kann kleiner, größer oder gleich der unbedingten Wahrscheinlichkeit  $P(A)$  sein.

### Beispiel:

1 Spielwürfel

B = „gerade Augenzahl“  $\rightarrow P(B) = \frac{1}{2}$

a) A = „Augenzahl nicht größer als 3“  $\rightarrow P(A) = \frac{1}{2}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} < P(A) = \frac{1}{2}$$

b) A = „Augenzahl gleich 2, 3 oder 4“  $\rightarrow P(A) = \frac{1}{2}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} > P(A) = \frac{1}{2}$$

c) A = „Augenzahl gleich 1 oder 2“  $\rightarrow P(A) = \frac{1}{3}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} = P(A)$$

**Hypothese:** B,  $P(B) > 0$  festhalten  
 Funktion  $A \rightarrow P_B(A) := P(A | B)$ ,  $A \in M$  betrachten.

Bei festem B besitzt  $P(B)$  alle Eigenschaften einer Wahrscheinlichkeit;  
 ist also eine neue Wahrscheinlichkeit auf M. - ohne Beweis -

A, B seien Ereignisse mit  $P(A) > 0$ ,  $P(B) > 0 \rightarrow P(A \cap B) = ?$   
 Zuweilen  $P(A | B)$  oder  $P(B | A)$  bekannt oder leichter zu ermitteln. Definition der bedingten Wahrscheinlichkeit

**Multiplikationssatz:** Es seien A und B Ereignisse positiver Wahrscheinlichkeit.  
 Dann gilt:

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

**Folgerung:**  $\frac{P(A|B)}{P(A)} = \frac{P(B|A)}{P(B)}$

**Beispiel:** 10 Bauelemente sind in einer Kiste, 4 davon sind defekt.  
 2 Elemente werden nacheinander „auf gut Glück“ entnommen.  
 Mit welcher Wahrscheinlichkeit sind beide Elemente intakt ?  
 (Ereignis A)

$A_i =$  „i-tes Element intakt“  $(i = 1, 2)$

$$P(A_1) = \frac{6}{10}, P(A_2 | A_1) = \frac{5}{9}$$

$$\rightarrow P(A) = P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 | A_1) = \frac{6}{10} \cdot \frac{5}{9} = \frac{1}{3}$$

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

**Verallgemeinerung:** Es seien  $A_1, A_2, \dots, A_n$  zufällige Ereignisse mit  
 $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$ , dann gilt:

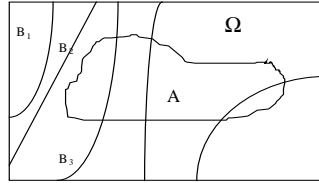
$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdot \dots \cdot P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

**Gegeben:** Ereignis A, vollständiges Ereignissystem  $\{B_1, B_2, \dots, B_n\}$   
 $B_i$  paarweise disjunkt  $\bigcup_i B_i = \Omega$ ,  $P(B_i) > 0$

**Standpunkt:**  $P(A)$  gesucht,  $P(A | B_i)$  dagegen bekannt.

Formel der totalen Wahrscheinlichkeit

$$P(A) = \sum_{i=1}^n P(A|B_i) * P(B_i)$$



**Beweis:**

$$A = A \cap \Omega = A \cap \left( \bigcup_i B_i \right) = \bigcup_i \underbrace{(A \cap B_i)}_{\text{paarweise unvereinbar}} \Rightarrow$$

$$P(A) = P\left(\bigcup_i (A \cap B_i)\right) = \sum_i P(A \cap B_i) = \sum_i \frac{P(A \cap B_i)}{P(B_i)} * P(B_i)$$

**Beispiel:** „Ruin des Spielers“

Spieler nimmt an Spielrunden teil: Erraten des Resultats eines Münzwurfs. Es wird um eine Mark gespielt.

Münze erraten - richtig  $\rightarrow$  + 1 DM - falsch  $\rightarrow$  - 1 DM

Anfangskapital: x DM ; x = 0, 1, ... (0 = keine Spielrunde möglich)

Strategie des Spielers: Solange spielen, bis Summe a Mark erreicht ist ( $a \geq x$ )

**Problem:** Mit welcher Wahrscheinlichkeit verliert der Spieler sein Kapital?

$P(x)$  = Wahrscheinlichkeit, daß sich der Spieler mit Anfangskapital von x DM ruiniert. ( $x = 0, 1, \dots, a$ )

Ereignis  $B_1$  = „Spieler gewinnt in der ersten Runde“  $x \rightarrow x + 1$

Ereignis  $B_2$  = „Spieler verliert in der ersten Runde“  $x \rightarrow x - 1$

Ereignis A = „Der Spieler wird ruiniert“

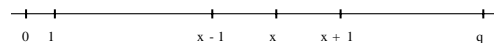
$B_1, B_2$  : vollständiges Ereignissystem:  $B_1 \cup B_2 = \Omega$  und  $B_1 \cap B_2 = \emptyset$

$$P(B_1) = \frac{1}{2} \text{ und } P(B_2) = \frac{1}{2}$$

$P(A)$  nicht erkennbar, aber Beziehung für bedingte Wahrscheinlichkeiten.

$$P(A|B_2) = p(x - 1)$$

$$P(A|B_1) = p(x + 1)$$



$\rightarrow$  Anwenden der Formel der totalen Wahrscheinlichkeit:

$$P(x) = P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)$$

$$= \frac{p(x+1)}{2} + \frac{p(x-1)}{2} \rightarrow \text{Differenzengleichung} \rightarrow P(x) = \frac{1}{2}(p(x+1) + p(x-1))$$

$$P(0) = 1, P(a) = 0$$

allgemeine Lösung der Differenzengleichung:  $P(x) = c_1 + c_2 \cdot x$   
 Einsetzen in die Randbedingungen:

$$1 = P(0) = c_1 \quad 0 = p(a) = c_1 (= 0) + c_2 \cdot a \rightarrow c_2 = -\frac{1}{a}$$

$$\text{Lösung: } P(x) = 1 - \frac{x}{a} \quad (x = 0, 1, \dots, a)$$

### 1.5.1 Unabhängigkeit von Ereignissen

A, B seien Ereignisse mit  $P(A), P(B) > 0$

möglicher Spezialfall:

$$(*) P(A | B) = P(A)$$

$$\rightarrow P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A \cap B)}{P(B)} \cdot \frac{P(B)}{P(A)} = \overbrace{\frac{1}{P(A|B)}}^{(*)} \cdot \frac{P(B)}{P(A)} = P(B)$$

Damit ist  $\frac{P(A \cap B)}{P(A)} = P(B) \Rightarrow P(A \cap B) = P(A) \cdot P(B)$  falls (\*) gilt.

**Definition:** Zufällige Ereignisse A und B heißen unabhängig, falls gilt:

$$P(A \cap B) = P(A) \cdot P(B)$$

**Bemerkung:** Voraussetzungen  $P(A), P(B) > 0$  haben wir vernachlässigt. Hat wenigstens eines der Ereignisse A oder B die Wahrscheinlichkeit Null, so sind A und B unabhängig.

Vergleich mit dem Multiplikationssatz:

$$\forall A, B \in M, P(A), P(B) > 0$$

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

Man unterscheide die Begriffe „unabhängig“ und „unvereinbar“ für Ereignisse.

*Unvereinbarkeit von A und B:*

$$A \cap B = \emptyset \Rightarrow P(A \cap B) = 0$$

Folglich sind zwei unvereinbare Ereignisse A, B positiver Wahrscheinlichkeit nicht unabhängig:

$$\underbrace{P(A)}_{>0} \cdot \underbrace{P(B)}_{>0} + P(A \cap B) = 0$$

**Beispiel:** „Skatblatt“  
eine Karte „auf gut Glück“ ziehen

Ereignis  $A_1$  = „Farbe ist Pik“

Ereignis  $A_2$  = „Karte ist Dame“

Sind  $A_1$  und  $A_2$  unabhängig?

*Formal:*  $P(A_1) = \frac{1}{4}, P(A_2) = \frac{1}{8}$

$$P(A_1 \cap A_2) = \frac{1}{32}$$

Es gilt:  $P(A_1) \cdot P(A_2) = \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{32} = P(A_1 \cap A_2)$

→ Ereignisse sind unabhängig

$$P(A_1 | A_2) = \frac{1}{4} = P(A_1), P(A_2 | A_1) = \frac{1}{8} = P(A_2)$$

**Bemerkung:** Unabhängigkeit von A und B drückt aus, daß A und B wahrscheinlichkeits-theoretisch in dem Sinne keinen Einfluß aufeinander haben, daß die Information „B tritt ein“ - wenn sie überhaupt positive Wahrscheinlichkeit hat - nichts an der Wahrscheinlichkeit von A ändert.

**Satz:** Sind die Ereignisse A und B unabhängig, so sind es auch die Ereignisse  $\bar{A}$  und B, A und  $\bar{B}$ , wie auch  $\bar{A}$  und  $\bar{B}$ .

**Beweis:** Seien A und B unabhängige Ereignisse:

$$P(A \cap B) = P(A) \cdot P(B)$$

Es genügt zu zeigen: Dann sind auch  $\bar{A}$  und B unabhängig.  
Ereignis B disjunkt zerlegen:

$$B = \Omega \cap B = (A \cup \bar{A}) \cap B = (A \cap B) \cup (\bar{A} \cap B)$$

$$(A \cap B) \cap (\bar{A} \cap B) = \emptyset$$

*Additivität:*

$$\begin{aligned}
P(B) &= P(A \cap B) \cup P(\bar{A} \cap B) \\
&= P(A) \cdot P(B) \cup P(\bar{A} \cap B) \Rightarrow P(\bar{A} \cap B) = P(B) - P(A) \cdot P(B) \\
&= [1 - P(A)] \cdot P(B) \\
&= P(\bar{A}) \cdot P(B)
\end{aligned}$$

**Beispiel:**

zwei verschiedene Würfel werfen

Ereignis A = „Würfel 1 zeigt ungerade Augenzahl“  
 B = „Würfel 2 zeigt ungerade Augenzahl“  
 C = „Die Augensumme ist ungerade“

A und B offenbar unabhängig  $P(A) = \frac{1}{2}$ ,  $P(B) = \frac{1}{2}$

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

bedingte Wahrscheinlichkeiten:

$$P(C|A) = \frac{1}{2}, P(C|B) = \frac{1}{2} \text{ aber auch } P(C) = \frac{1}{2}$$

→ A, C sowie B, C jeweils unabhängig

anders ausgedrückt:  $P(A) = P(B) = P(C) = \frac{1}{2}$

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

Ereignisse A, B, C paarweise unabhängig.

**Aber:**  $P(C | A \cap B) = 0 \rightarrow C$  nicht unabhängig von  $\{A, B\}$   
 es besteht hier offenbar Abhängigkeit „zu dritt“ → Definition

**Definition:**  
unabhängig,

Die zufälligen Ereignisse  $A_1, A_2, \dots, A_n$  heißen vollständig

wenn für beliebige  $k = 2, 3, \dots, n$  und beliebige natürliche Zahlen  $i_1, \dots, i_k$   
 mit  $1 \leq i_1 < i_2 < \dots < i_k \leq n$  gilt:

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$$

Die zufälligen Ereignisse einer unendlichen Folge  $A_1, A_2, \dots$  heißen vollständig abhängig, wenn für jedes natürliche  $n = 2, 3, \dots$  die Ereignisse  $A_1, \dots, A_n$  vollständig sind.

**Folgerung:**

Sind die Ereignisse  $A_1, \dots, A_n$  vollständig unabhängig, so sind sie es auch paarweise.

**Bemerkungen:**

Die Umkehrung obiger Folgerungen gilt nicht (siehe Beispiel).  
 Aus der Beziehung  $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$  mit  $P(C) > 0$   
 folgt i.a. nicht  $P(A \cap B) = P(A) \cdot P(B)$

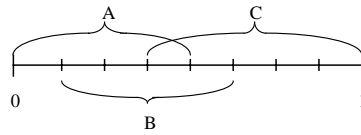


**Beispiel:**

$$\Omega = \{\omega : 0 < \omega < 1\}$$

$$A = \{\omega : 0 < \omega < \frac{4}{8}\}, B = \{\omega : \frac{1}{8} < \omega < \frac{5}{8}\}, C = \{\omega : \frac{3}{8} < \omega < \frac{7}{8}\}$$

P(D) - Inhalt von D (Länge)



Es gilt:  $P(A) = P(B) = P(C) = \frac{1}{2}$ , sowie

$$P(A \cap B \cap C) = \frac{1}{8} = P(A) \cdot P(B) \cdot P(C)$$

$$\textbf{Aber: } P(A \cap B) = \frac{3}{8} \neq \frac{2}{8} = P(A) \cdot P(B)$$

Dies begründet die komplizierte Definition der vollständigen Unabhängigkeit.

**Bayessche Formel:**

Formel der totalen Wahrscheinlichkeit:  $P(A) = \sum P(A | B_i) \cdot P(B_i)$

$\{B_1, B_2, \dots, B_n\}$  vollständiges Ereignissystem

Andere Fragestellung: Ereignis positiver Wahrscheinlichkeiten:  $P(A) > 0$

bekannt:  $P(B_i), P(A | B_i)$  ( $i = 1, 2, \dots, n$ )

gesucht:  $P(B_k | A) = ?$

**Satz:** (Bayessche Formel)

Unter den obigen Voraussetzungen gilt:

$$P(B_k | A) = \frac{P(A | B_k) \cdot P(B_k)}{\sum_{i=1}^n P(A | B_i) \cdot P(B_i)} \quad (k = 1, 2, \dots, n)$$

**Beweis:**

nach Definition der bedingten Wahrscheinlichkeit:

$$P(B_k | A) = \frac{P(B_k \cap A)}{P(A)}, \quad P(A | B_k) = \frac{P(B_k \cap A)}{P(B_k)}$$

$$P(B_k | A) \cdot P(A) = P(A | B_k) \cdot P(B_k)$$

$$\rightarrow P(B_k | A) = \frac{P(A | B_k) \cdot P(B_k)}{P(A)} \quad (k = 1, 2, \dots, n)$$

Voraussetzungen zur Anwendung der Formel der totalen Wahrscheinlichkeit für  $P(A)$  erfüllt.

$$\rightarrow P(B_k | A) = \frac{P(A | B_k) \cdot P(B_k)}{\sum_{i=1}^n P(A | B_i) \cdot P(B_i)} \quad (k = 1, 2, \dots, n)$$

**Sprechweise:**  $P(B_k)$  a priori Wahrscheinlichkeiten  
 $P(B_k | A)$  a posteriori Wahrscheinlichkeiten

**Physikalisches Analogon der Bayesschen Formel:**

In  $n$  Gefäßen seien Lösungen ein und desselben Stoffes in unterschiedlichen Konzentrationen enthalten. Das Gesamtvolumen der Lösungen sei 1 Liter.

$P(B_k)$  - Volumen der Lösungen im  $k$ -ten Gefäß ( $k = 1, 2, \dots, n$ )

$$P(B_k | A) = \frac{P(A | B_k) \cdot P(B_k)}{\sum_{i=1}^n P(A | B_i) \cdot P(B_i)} \quad k = 1, \dots, n$$

Anteil der Gesamtstoffmengen im  $k$ -ten Gefäß.

**1.6 Axiomatische Begründung der Wahrscheinlichkeitstheorie**

Kolmogorov 1933

**Vorgegeben:**  $\Omega$  - Raum der Elementarereignis  
 $\mathcal{M}$  - System von Teilmengen von  $\Omega$ :  $\sigma$ -Algebra

S1.  $\Omega \in \mathcal{M}$   
S2.  $A \in \mathcal{M} \Rightarrow \bar{A} \in \mathcal{M}$   
S3.  $A_i \in \mathcal{M} (i = 1, 2, \dots) \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$

$P$  reellwertig auf  $\mathcal{M}$ ,  $P$  erfüllt die Axiome

W1.  $P(A) \geq 0$   
W2.  $P(\Omega) = 1$   
W3.  $A_n \in \mathcal{M} (n = 1, 2, \dots)$  paarweise unvereinbar

$$\rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Die Funktion  $P$  heißt Wahrscheinlichkeit oder auch Wahrscheinlichkeitsmaß auf  $\mathcal{M}$ . Das Tripel  $(\Omega, \mathcal{M}, P)$  heißt Wahrscheinlichkeitsraum.

**Bemerkung:** Die nunmehr endgültige Fassung des Wahrscheinlichkeitsbegriffs verallgemeinert unsere bisherigen Konstruktionen. Viele der im speziellen Fall gezeigten Eigenschaften sind allgemeingültig.

**Eigenschaften:** (Auswahl)

1.  $P(\emptyset) = 0$

*Beweis:*  $P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \cup \dots) = P(\Omega) + P(\emptyset) + P(\emptyset) + \dots \rightarrow P(\emptyset) = 0$

2. endliche Additivität

$A, B$  unvereinbare Ereignisse  $\rightarrow P(A \cup B) = P(A) + P(B)$

*Beweis:*

$A_1 = A, A_2 = B, A_i = \emptyset (i = 3, 4, \dots)$

$$P(A \cup B) = P\left(\bigcup_i A_i\right) = \sum_i P(A_i) = P(A) + P(B) + \underbrace{\sum_{i=1}^{\infty} P(\emptyset)}_{=0}$$

3.  $P(\bar{A}) = 1 - P(A), A \in M$

*Beweis:*

$A \in M \rightarrow \bar{A} \in M, A \cup \bar{A} = \Omega, A \cap \bar{A} = \emptyset$

$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$

4. Monotonie

$A \subseteq B \Rightarrow P(A) \leq P(B); A, B \in M$

*Beweis:*

Ereignis  $B$  disjunkt zerlegen:

$B = (A \cup (B \cap \bar{A})) \Rightarrow A \cap (B \cap \bar{A}) = \emptyset$

$P(B) = P(A \cup (B \cap \bar{A})) = P(A) + \underbrace{P(B \cap \bar{A})}_{\geq 0} \geq P(A)$

5. Subadditivität

$A_n, (n = 1, 2, \dots)$  Folge von Ereignissen aus  $M \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$

*Beweis:*

Aus  $(A_n)$  eine Folge paarweiser disjunkter Ereignisse  $(B_n)$  konstruieren:

$B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \cup A_2), \dots, B_n = A_n \setminus (A_1 \cup A_2 \cup \dots \cup A_{n-1})$

$$= A_n \setminus \left(\bigcup_{k=1}^{n-1} A_k\right) \Rightarrow \bigcup_n A_n = \underbrace{A_1 \cup (A_2 \setminus A_1)}_{A_1 \cup A_2} \cup \underbrace{(A_3 \setminus (A_1 \cup A_2)) \cup \dots}_{A_1 \cup A_2 \cup A_3} = \bigcup_n B_n$$

Weiterhin gilt:  $B_k \subseteq A_k \xRightarrow{\text{Monotonie}} P(B_k) \leq P(A_k) (k = 1, 2, \dots)$

$$\rightarrow P\left(\bigcup_n A_n\right) = P\left(\bigcup_n B_n\right) = \sum_n P(B_n) \leq \sum_n P(A_n)$$

## 6. Stetigkeit

a) **isotone** Folge von Ereignissen  $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$

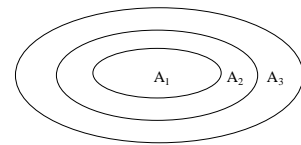
$\exists$  Limes der Mengenfolge:  $\bigcup_n A_n$

$$\rightarrow P\left(\bigcup_n A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

*Beweis:*

Wieder die Glieder der Folge  $(A_n)$  disjunkt machen:

$$B_1 = A_1, B_2 = A_2 \setminus B_1, \dots, B_n = A_n \setminus \left(\bigcup_{k=1}^{n-1} B_k\right)$$



Folge  $(B_n)$  paarweise disjunkter Ereignisse,  $A_n = \bigcup_{k=1}^n B_k$

$$\rightarrow P\left(\bigcup_n A_n\right) = P\left(\bigcup_n B_n\right) = \sum_n P(B_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n P(B_k) = \lim_{n \rightarrow \infty} P\left(\bigcup_{k=1}^n B_k\right) = \lim_{n \rightarrow \infty} P(A_n)$$

b) **antitone** Folge von Ereignissen  $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$  hier gilt  $\exists$  Limes  $\bigcap_n A_n$

$$\rightarrow P\left(\bigcap_n A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

*Beweis:*

durch Übergang zu den komplementären Ereignissen:  $\overline{A_1} \subseteq \overline{A_2} \subseteq \dots$  monoton wachsende Folge

$$P\left(\bigcap_n A_n\right) = 1 - P\left(\overline{\bigcap_n A_n}\right) = 1 - P\left(\bigcup_n \overline{A_n}\right) \stackrel{a)}{=} 1 - \lim_{n \rightarrow \infty} P(\overline{A_n}) = \lim_{n \rightarrow \infty} [1 - P(\overline{A_n})] = \lim_{n \rightarrow \infty} P(A_n)$$

## 2 Zufallsgrößen und deren Verteilung

**Einführung:** Bernoulli-Schema:  $(\Omega, \mathcal{M}, P)$  Wahrscheinlichkeitsraum  
zufälliges Ereignis  $A$ ,  $P(A) = p$ ,  $0 < p < 1$   
Experiment wird  $n$ -mal unabhängig voneinander ausgeführt

$A$  tritt ein: „Erfolg“,  $\overline{A}$  - Alternative

Indikator  $1_A^{(k)}$  bewertet, ob der k-te Versuch erfolgreich ist, d.h. ob A eintritt.

$$1_A^{(k)} = \begin{cases} 1 & \text{A tritt ein} \\ 0 & \text{\bar{A} tritt ein} \end{cases}$$

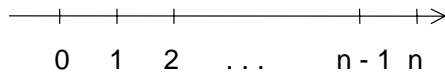
→ Folge „unabhängiger“ vom Zufall abhängiger Größen  $1_A^{(k)}$  ( $k = 1, \dots, n$ )



Die Größe X zählt die Anzahl der Erfolge bei Durchführung von n Versuchen:

$$X = \sum_{k=1}^n 1_A^{(k)}$$

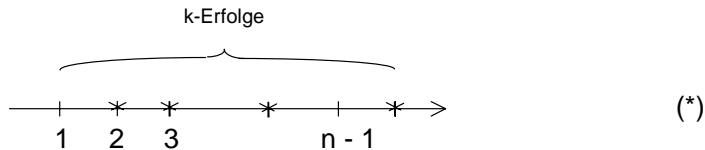
Welche Werte kann X annehmen ?



Mit welchen Wahrscheinlichkeiten ?

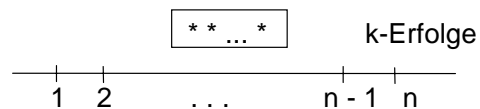
→ Verteilung auf  $\{0, 1, \dots, n\}$

Ereignis  $\{X = k\}$ : A tritt in der Versuchsserie k-mal und  $n - k$ -mal nicht auf.



Jede konkrete Folge (\*) hat die Wahrscheinlichkeit  $p^k(1-p)^{n-k}$ .

Wieviele solcher Folgen gibt es ?  
k-mal „Erfolg“ auf n Plätze verteilen.



### Kombinatorische Standardaufgabe:

Verteilung von k Teilchen auf n Zellen

- Teilchen nicht unterscheidbar
- keine Mehrfachbelegungen

→  $\binom{n}{k}$  verschiedene Kombinationen

$$\rightarrow P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k} \quad (k = 0, 1, \dots, n)$$

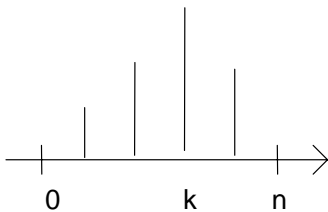
Das ist die Verteilung der Anzahl der Erfolge:  
„Binominalverteilung mit den Parametern n und p“

$$\rightarrow \sum_{k=0}^n P\{X = k\} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \stackrel{\text{Binomischer Satz}}{=} [p + (1-p)]^n = 1$$

Versuchsserie: „Bernoulli-Schema“  
„unabhängige Verteilung ein- und desselben Versuches“

**Bemerkung:** die Größen  $1_A^{(k)}$  und X hängen vom Zufall ab.  
ausführlich:  $1_A^{(k)}(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in \bar{A} \end{cases}$ ,  $X(\omega) = \sum_{k=1}^n 1_A^{(k)}(\omega)$   
Spätere Definition solcher Größen als „Zufallsgrößen“.

Bewertung des Punktes k entsprechend der Binominalverteilung:



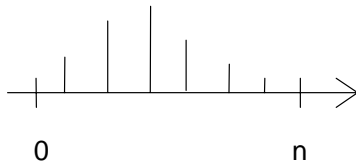
$$\begin{aligned} \sum_{k=0}^{\infty} k * P\{X = k\} &= \sum_{k=0}^{\infty} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^{\infty} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} = np \sum_{k=0}^{n-1} \frac{(n-1)!}{k![n-(k+1)]!} p^k (1-p)^{n-(k+1)} \\ &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k![(n-1)-k]!} p^k (1-p)^{(n-1)-k} = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \stackrel{\text{bin. Satz}}{=} np[p + (1-p)]^{n-1} = np \end{aligned}$$

→ Parameter der Lage der Verteilung

**anderer Parameter:** charakterisiert die Streuung der Verteilung  
mittlere quadratische Abweichung

$$\text{Mittelwert } \underbrace{(X - np)^2}_{\text{Zentrieren auf Null}} = \overset{\text{ähnliche Betr.}}{\dots} = np(1-p)$$

**Konstruktion X:**

$$X = \sum_{k=1}^n 1_A^{(k)} =: X_n$$


Was passiert bei wachsendem n ?

**Trick:** Größe  $X_n$  zentrieren und normieren, d.h.  $X_n^* = \frac{X - np}{\sqrt{np(1-p)}}$ .

Dann gilt der Grenzwertsatz von de MOIRE-LAPLACE

$$\lim_{n \rightarrow \infty} P\{x_1 \leq X_n^* \leq x_2\} = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx, \quad x_1 < x_2 \in \mathbf{R}$$

**praktische Bedeutung:**

n groß, nicht Einzelwahrscheinlichkeiten  $P\{X_n = k\}$ , sondern  $P\{a \leq X_n \leq b\}$  seien interessant ( $a, b \in \mathbf{R}$ )

$$\begin{aligned} P\{a \leq X_n \leq b\} &= P\left\{ \frac{a - np}{\sqrt{np(1-p)}} \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq \frac{b - np}{\sqrt{np(1-p)}} \right\} \\ &= P\left\{ \frac{a - np}{\sqrt{np(1-p)}} \leq X_n^* \leq \frac{b - np}{\sqrt{np(1-p)}} \right\} \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{a - np}{\sqrt{np(1-p)}}}^{\frac{b - np}{\sqrt{np(1-p)}}} e^{-\frac{x^2}{2}} dx \Rightarrow \text{einfacher darstellen} \end{aligned}$$

Uneigentliches Integral:

$$\Phi := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{x^2}{2}}, \quad x \in \mathbf{R}$$

Integral interessiert, es gilt sogar  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

Funktion  $\Phi$  vertafelt oder im Computer

$$P\{a \leq X \leq b\} \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right)$$

**Beispiel:** Produktion von Glühlämpchen

Kartons zu je 1000 Stück  
Erfahrungstatsache: Ausschuß im Mittel 3% → (vage) Erwartung,  
pro Karton 30 Lämpchen defekt.

*Realisierbare Frage:* Wahrscheinlichkeit dafür, daß 20 bis 40  
Lämpchen defekt sind.

Modell: X (zufällige) Anzahl defekter Lämpchen in einem  
„auf gut Glück“ gewähltem Karton mit 1000 Lämpchen.  
X ist binominalverteilt mit den Parametern  $n = 1000$  und  
 $p = 0,03$

→ Mittel:  $n \cdot p = 1000 \cdot 0.03 = 30$

*mittlere quadratische Abweichung:*  
 $n \cdot p \cdot q = 1000 \cdot 0.03 \cdot (1 - 0.03) = 29.1$

gesuchte Wahrscheinlichkeit:

$$P\{20 \leq X \leq 40\} = \sum_{k=20}^{40} P\{X = k\} = \sum_{k=20}^{40} \binom{1000}{k} 0.03^k (1 - 0.03)^{1000-k}$$

→ zu kompliziert

→ Näherung

$$P\{20 \leq X \leq 40\} \approx \Phi\left(\frac{40 - 1000 \cdot 0.03}{\sqrt{1000 \cdot 0.03 \cdot 0.97}}\right) - \Phi\left(\frac{20 - 1000 \cdot 0.03}{\sqrt{1000 \cdot 0.03 \cdot 0.97}}\right)$$

$$= \Phi\left(\frac{10}{\sqrt{29.1}}\right) - \Phi\left(\frac{-10}{\sqrt{29.1}}\right) = 2\Phi(1.85) - 1 \approx 0.94$$

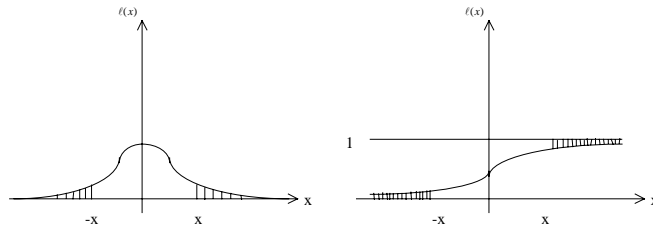
Symmetrie des Integranden  $e^{-\frac{x^2}{2}}$  in  $\Phi$



2. Das Integral  $\Phi$  besitzt eine Dichte:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy = \int_{-\infty}^x \ell(x) dx, \quad x \in \mathbb{R}$$

$$\ell(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$



$$\rightarrow \Phi(x) + \Phi(-x) = 1$$

$$\rightarrow \Phi(x) - \Phi(-x) = \Phi(x) - (1 - \Phi(x)) = 2\Phi(x) - 1$$

Die Funktion definiert die Verteilung einer (stetigen) zufälligen Größe  $Y$ , die Werte aus ganz  $\mathbb{R}$  annehmen kann, es gilt:

$$P\{Y < x\} = \Phi(x), \quad x \in \mathbb{R}$$

es gilt außerdem:  $P\{a \leq Y < b\} = \Phi(b) - \Phi(a)$ ,  $a < b$

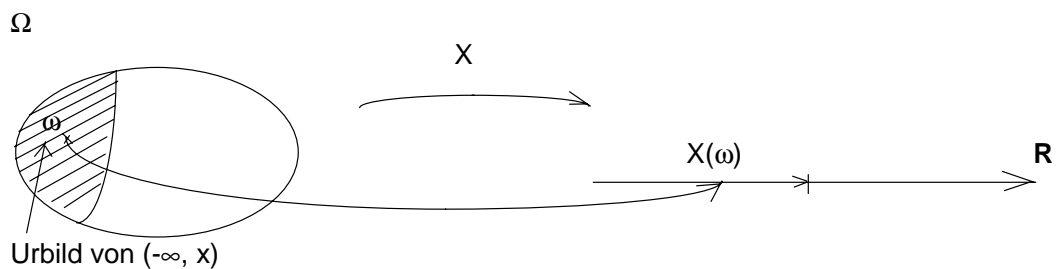
Die Größe  $Y$  heißt normalverteilt, genauer standard-normal-verteilt nach  $N(0, 1)$ .

## 2.1 Zufallsgrößen, Verteilungen

**vorgegeben:**      Wahrscheinlichkeitsraum  $(\Omega, \mathcal{M}, P)$

**Definition:**      Eine reellwertige Funktion  $X$  auf  $\Omega$   $\omega \rightarrow X(\omega)$  ( $\omega \in \Omega$ )  
heißt Zufallsgröße, falls für jede reelle Zahl  $x$  gilt:

$$\{\omega \in \Omega : X(\omega) < x\} \in \mathcal{M} \quad (*)$$



**Bemerkungen:**

1. Den Regeln der  $\sigma$ -Algebra folgend sind dann auch Mengen der Gestalt:

$$\{\omega: x_1 \leq X(\omega) \leq x_2\} = \{\omega: X(\omega) < x_2\} \setminus \{\omega: X(\omega) < x_1\}$$

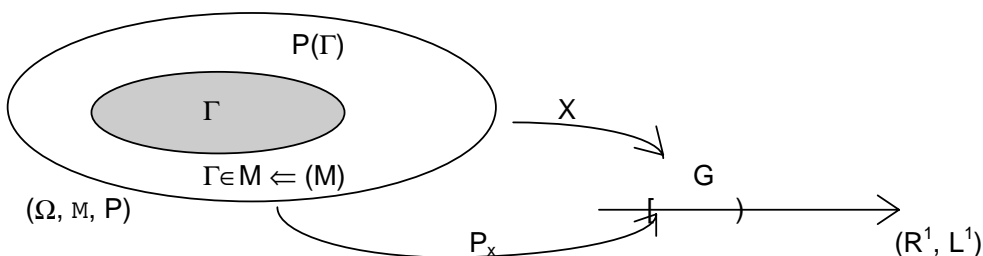
2. Die Bedingung (\*) genannt „Meßbarkeit von X“ ist technischer Natur; sie sichert, daß alle Mengen  $\{\omega: X(\omega) < x\}$ ,  $\{\omega: x_1 \leq X < x_2\}$ , ... Ereignisse sind, für die folgende Wahrscheinlichkeiten gebildet werden können:

$$\omega: X(\omega) < x, P\{\omega: x_1 \leq X < x_2\}$$

3. üblicherweise schreibt man kurz:

$$\{X < x\}, \{x_1 \leq X < x_2\} \text{ und } P\{X < x\}, P\{x_1 \leq X \leq x_2\}, \dots$$

$\{\omega: X(\omega) < x\}, \{\omega: x_1 \leq X(\omega) < x_2\} \quad (M)$   
 $P\{\omega: X(\omega) < x\}$   
 $\{X < x\}$   
 $P\{X < x\}$



Mittels der Abbildung  $X$  wird in der  $\sigma$ -Algebra  $L^1$  der Borel-Mengen von  $\mathbb{R}^1$  das Bildmaß  $P_x$  von  $P$  definiert:

$$P_x(G) := P\{\omega : X(\omega) \in G\}$$

**Definition:** Das Maß  $P_x$  heißt Verteilung, auch Wahrscheinlichkeitsverteilung von  $X$  auf  $(\mathbb{R}^1, L^1)$ .

Dieses Maß regelt, mit welcher Wahrscheinlichkeit die Zufallsgröße  $X$  bei einer Realisierung einen Wert aus einer gegebenen Borel-Menge annimmt:

$$P_x([x_1, x_2)) = P\{x_1 \leq X < x_2\}$$

Wir unterscheiden 2 Spezialfälle:

**Definition:** Die Zufallsgröße  $X$  heißt *diskret*, wenn sie nur endlich oder abzählbar unendlich viele Werte annehmen kann. Es gilt dann:

$$P\{X = x\} = P\{\omega : X(\omega) = x\} \quad (x = \underbrace{x_1, x_2, \dots}_{\text{feste Werte}})$$

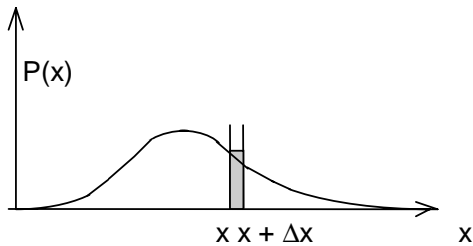
$$\sum_{x_i} P\{X = x_i\} = 1 \quad \text{---|---|---|---|---|---> \mathbb{R}^1$$

**Beispiel:** Binominalverteilte Zufallsgröße mit Parametern  $(n, p)$  nimmt ausschließlich Werte in  $\{0, 1, \dots, n\}$  an.

**Definition:** Die Zufallsgröße  $X$  heißt stetig, wenn ihre Wahrscheinlichkeitsverteilung eine Dichte  $p \geq 0$  besitzt, so daß

$$P\{x_1 \leq X < x_2\} = \int_{x_1}^{x_2} p(x) dx \quad (x_1 < x_2 \in \mathbb{R}) \text{ gilt.}$$

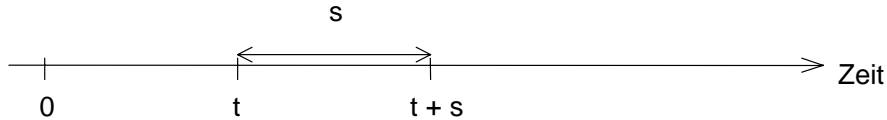
**Folgerung:**  $P\{\omega: X(\omega) \in \mathbb{R}\} = P\{-\infty < x < \infty\} = \int_{-\infty}^{\infty} p(x) dx = 1$



$$P\{X \in [x, x + \Delta x]\} = p(x) \cdot \Delta x + o(\Delta x)$$

**Beispiel:** standardnormalverteilte Zufallsgröße  $X$ , andere Verteilung konstruieren: Zufallsgröße  $X \geq 0$

**Beispiel:**  $X$  Zeit des Eintretens irgendeines Ereignisses; (Lebensdauer, Zerfallszeit)



Angenommen  $\{X > t\}$  **Frage:**  $\{X > t + s\}$  ?

Zufallsgröße  $X$  sei „gedächtnislos“, falls  $P\{X > t + s \mid X > t\} = P\{X > s\}$  ( $s, t \geq 0$ ) gilt.

$$\rightarrow \frac{P\{x > t + s, X > t\}}{P\{X > t\}} = P(X > s)$$

$$\rightarrow \frac{P\{X > t + s\}}{P\{X > t\}} > P(X > s) \Rightarrow P\{X > t + s\} = P\{X > t\} * P\{x > s\}$$

**Funktionalitätsgleichung:**  $[\Psi(t + s) = \Psi(t) \cdot \Psi(s)]$

gesucht: beschränkte Lösung  $\rightarrow P\{X > t\} = e^{-\lambda t}$  ( $t \geq 0$ )

Komplement bilden:

$$P\{X < t\} = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Zufallsgröße mit dieser Verteilung heißt **exponentiell verteilt** mit dem Parameter  $\lambda > 0$ .

$$P\{X < \infty\} = \lim_{t \rightarrow \infty} (1 - e^{-\lambda t}) = 1$$

**Zufallsgröße X hat eine Dichte:**

$$\frac{dP\{X < t\}}{dt} = \frac{d[1 - e^{-\lambda t}]}{dt} = \lambda e^{-\lambda t}, \quad t \geq 0$$

$$\rightarrow p(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$$\text{Also: } P\{X < t\} = \int_{-\infty}^t p(s) ds, \quad t \in \mathbb{R}$$

Sei  $X_n$  binomialverteilte Zufallsgröße mit den Parametern  $n$  und  $p$ :

$$P\{X_n = k\} = \binom{n}{k} p^k (1-p)^{n-k} \quad (k = 0, 1, \dots, n)$$

**Problem:**  $p$  werde klein  $\rightarrow$  mittlere Zahl der Erfolge klein, lässt man simultan  $n$  wachsen, dann ist der Mittelwert trotzdem bedeutend.

**Ansatz:**  $n \cdot p = \lambda$ ;  $\lambda > 0$  konstant.

Was passiert für  $n \rightarrow \infty$ ?

$k = \text{konstant}$ :

$$P\{X_{n\#} = k\} = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{p^k}{k!} (1-p)^n (1-p)^{-k} *$$

$$\frac{n!}{(n-k)!} = \frac{(p^n)^k}{k!} \left(1 - \frac{p^n}{n}\right)^n \left(1 - \frac{p^n}{n}\right)^{-k} \overbrace{\frac{n(n-1) \dots (n-k+1)}{n^k}}^{k \text{ Faktoren}} =$$

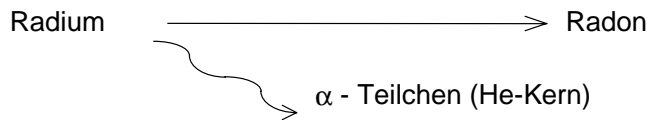
$$= \frac{\lambda^k}{k!} \left(\underbrace{1 - \frac{\lambda}{n}}_{e^{-\lambda}}\right)^n \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_1 \underbrace{\frac{n}{n}}_1 * \underbrace{\frac{n-1}{n}}_1 * \dots * \underbrace{\frac{n-k+1}{n}}_1$$

$$\Rightarrow \lim_{n \rightarrow \infty} P\{X_n = k\} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (\text{Grenzwertsatz})$$

Wegen  $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1$  ist damit eine Wahrscheinlichkeitsverteilung auf  $\{0, 1, 2, \dots\}$  erklärt.

**Definition:** Eine Zufallsgröße  $X$  mit  $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$  ( $k = 0, 1, \dots$ ) heißt **POISSON-verteilt** mit dem Parameter  $\lambda > 0$ .

**Modell:** radioaktiver Zerfall



Im Zeitintervall der Länge  $t$  zerfällt das Radiumatom mit Wahrscheinlichkeit  $p(t)$ .

$n$  Radiumabnahme

Abstand sehr groß, Zerfall eines Kerns erfolgt unabhängig von allen anderen.

Mittlere Zahl der ausgesandten  $\alpha$  - Teilchen während  $t$ :  $a(t) = n \cdot p(t)$

Experimentelle Erfahrungen (Messungen) für  $t = 1\text{s}$  und  $n = 10^{22}$  (= 19 Radium):  
 $a(t) \approx 10^{10} \rightarrow p(t) \approx 10^{-12}$  (also sehr klein)

*Versuch:* Beobachtungen eines dieser Atome  
*Erfolg:* Zerfall während 1s

- $\rightarrow$  gleichzeitig laufen also  $10^{22}$  solcher Versuche ab.
- $\rightarrow$  Voraussetzungen des Bernoullischema erfüllt.

*Zerfallsgröße*  $X(t)$ : Anzahl der während  $t$  ausgesandten  $\alpha$ -Teilchen.

$n$  sehr groß,  $p$  sehr klein  $\rightarrow$  annähernd POISSON-verteilt mit Parameter  $\lambda = n \cdot p$ .

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

### **Verallgemeinerung des Begriffes der Zufallsgröße:**

**Definition:** Es seien  $x_1, x_2, \dots, x_n$  Zufallsgrößen auf ein- und demselben Wahrscheinlichkeitsraum  $(\Omega, \mathcal{M}, P)$ .

Dann heißt das  $n$ -Tupel  $\vec{X} = (X_1, X_2, \dots, X_n)$  ein **Zufallsvektor**.

**Bemerkungen:** zufällige Vektoren sind Vektoren, deren Komponenten Zufallsgrößen sind. Man kann  $X$  als (meßbare) Abbildung von  $\Omega$  in den Raum  $\mathbf{R}^n$  auffassen.

Die Verteilung von  $\vec{X}$  ist das Bildmaß  $P_{\vec{X}}$  von  $P$  in der  $\sigma$ -Algebra der Borelmengen  $L^n$  von  $\mathbf{R}^n$ .

$$P_{\vec{X}}(B) = P\{\vec{X} \in B\} = P\{\omega : X_1(\omega), X_2(\omega), \dots, X_n(\omega) \in B\} \quad B \in L^n$$

### Beispiele für Zufallsvektoren:

simultane Messung verschiedener Größen bei einem Experiment. (p-V-Diagramm, Körpergröße eines Menschen, Dimension eines Werkstückes)

**Definition:** Ein Zufallsvektor  $\vec{X} = (X_1, X_2, \dots, X_n)$  heißt **diskret**, wenn er höchstens abzählbar viele verschiedene Werte  $\vec{X} = (X_1, X_2, \dots, X_n) \in \mathbf{R}^n$  annehmen kann.

**Definition:** Der Zufallsvektor  $\vec{X}$  heißt **stetig**, wenn seine Verteilung eine Dichte  $p(x_1, x_2, \dots, x_n) \geq 0$  besitzt:

$$P\{X \in B\} = \int_B \dots \int p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad B \in L^n$$

$X_1, X_2, \dots, X_n, \dots$  höchstens abzählbar viele Zufallsgrößen auf  $(\Omega, M, P)$

**Definition:** Die Zufallsgrößen  $X_1, X_2, \dots, X_n, \dots$  heißen **unabhängig**, falls für beliebige Zahlen  $x_k' \leq x_k''$  die Ereignisse

$$\{x_k' \leq X_k < x_k''\} \quad k = 1, 2, \dots, n, \dots$$

vollständig unabhängig sind.

### Folgerungen:

1. Bei diskreten unabhängigen Zufallsgrößen  $X_1, X_2, \dots, X_n$  gilt für die (gemeinsame)

Verteilung von  $\vec{X} = (X_1, X_2, \dots, X_n)$ :

$$P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = P\{X_1 = x_1\} \cdot P\{X_2 = x_2\} \cdot \dots \cdot P\{X_n = x_n\}, (x_1, x_2, \dots, x_n) \in \mathbf{R}^n.$$

2. Bei stetigen unabhängigen Zufallsgrößen  $X_1, X_2, \dots, X_n$  mit den Dichten  $p_1, p_2, \dots, p_n$  gilt für die (gemeinsame) Dichte des zufälligen Vektors  $\vec{X} = (X_1, X_2, \dots, X_n)$ :

$$p(x_1, x_2, \dots, x_n) = p_1(x_1) \cdot p_2(x_2) \cdot \dots \cdot p_n(x_n) \text{ mit } (x_1, x_2, \dots, x_n) \in \mathbf{R}^n.$$

### Funktionen von Zufallsgrößen

$X$  sei Zufallsgröße auf  $(\Omega, M, P)$ ,  $g$  ist reelle Funktion  $g: \underset{\text{aus}}{\mathbf{R}^1} \rightarrow \underset{\text{in}}{\mathbf{R}^1}$ .

Unter ganz schwachen Voraussetzungen (Meßbarkeit muß gesichert sein) ist dann auch die durch  $\omega \rightarrow g(X(\omega))$ ,  $\omega \in \Omega$  definierte Abbildung eine Zufallsgröße.

Analoges gilt für einen Zufallsvektor  $(X_1, X_2, \dots, X_n)$  und die Funktion  $h: \mathbf{R}^n \rightarrow \mathbf{R}^1$ .

$Y = h(X_1, X_2, \dots, X_n)$  ist dann Zufallsgröße auf  $(\Omega, M, P)$ .

## Charakterisierung von Zufallsgrößen und Vektoren

X sei Zufallsgröße auf  $(\Omega, \mathcal{M}, P)$

wichtigste Charakteristik von X: Verteilung  $P_X$  (aber oft „unhandlich“)

**Definition:** Die durch  $F_X(x) = P\{X \leq x\}$ ,  $x \in \mathbf{R}$  definierte Funktion heißt **Verteilungsfunktion** der Zufallsgröße X.

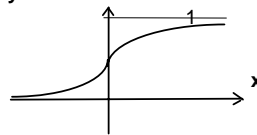
**Bemerkung:** Die Verteilungsfunktion der Zufallsgröße X charakterisiert die Verteilung einer Zufallsgröße vollständig, es gilt:

$$P\{x_1 < X < x_2\} = F_X(x_2) - F_X(x_1)$$

**Beispiel:**

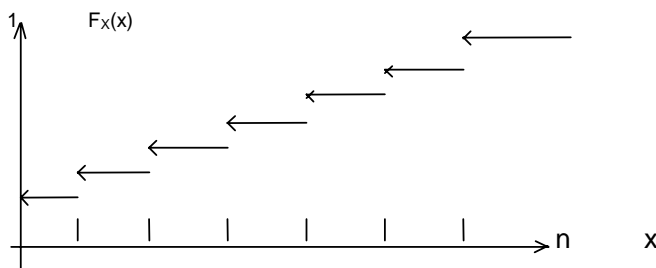
Standard-Normalverteilung  $N(0, 1)$

$$F_X = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy, \quad x \in \mathbf{R}$$



Binominalverteilung mit Parametern  $n, p$

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ \sum_{0 \leq k \leq x} \binom{n}{k} p^k (1-p)^{n-k}, & 0 < x \leq n \\ 1, & x > n \end{cases}$$



## Eigenschaften einer Verteilungsfunktion $F_X$

1.  $0 \leq F_X(x) \leq 1$ ,  $x \in \mathbf{R}$
2.  $F_X$  monoton wachsend  $x_1 \leq x_2 \rightarrow F(x_1) \leq F(x_2)$
3.  $F_X$  linksseitig stetig  $\lim_{y \rightarrow x-0} F_X(y) = F_X(x)$ ,  $x \in \mathbf{R}$
4.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$ ,  $\lim_{x \rightarrow \infty} F_X(x) = 1$

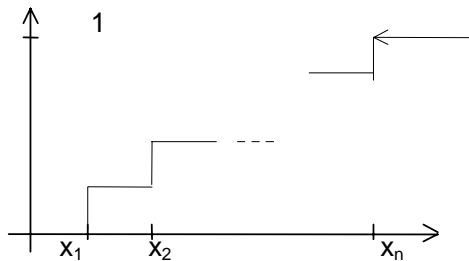
**Bemerkung:** Diese Eigenschaften sind sogar charakterisierend, d.h. jede Funktion  $F$

mit diesen Eigenschaften 1 - 4 ist Verteilungsfunktion einer gewissen Zufallsgröße X.

Verteilungsfunktion enthält volle Information über die Verteilung, aber kompliziert und schwierig zu bestimmen (bei Zufallsgrößen mit a priori unbekannter Verteilung)  
 → Wunsch nach Informationsverdichtung

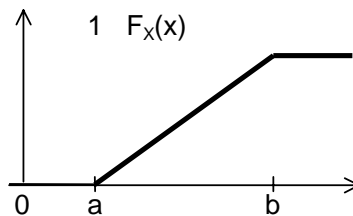
**Beispiel:** (Gleichverteilung)  
 Die diskrete Zufallsgröße X mit den Werten  $x_1, x_2, \dots, x_n$  heißt **gleichverteilt**, falls gilt:

$$P\{X = x_i\} = \frac{1}{n} \quad (i = 1, 2, \dots, n)$$



Analog gilt für die Dichte einer stetigen Gleichverteilung auf dem Intervall  $[a, b] \subset \mathbb{R}$ :

$$p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{sonst} \end{cases}$$



**Mittelmarke für X:**

X **diskret**: arithmetisches Mittel =  $\frac{1}{n} \sum_{k=1}^n x_k$

X **stetig**: Intervallmitte =  $\frac{a+b}{2}$

Verallgemeinerung für diese Mittelungen für beliebige diskrete und stetige Zufallsgrößen X?



## 2.2 Erwartungswert

**Definition:** Es sei  $X$  eine diskrete Zufallsgröße mit den Werten  $x_k$ ,  $k = 1, 2, \dots$ . Dann heißt die durch

$$EX = \sum_k x_k P\{X = x_k\}$$

definierte Zahl der **Erwartungswert** der Zufallsgröße  $X$ . Dabei wird der Eindeutigkeit wegen die absolute Konvergenz obiger Reihe vorausgesetzt:

$$\sum_k x_k P\{X = x_k\} < \infty$$

**Definition:** Für eine stetige Zufallsgröße  $X$  definieren wir  $EX = \int_{-\infty}^{\infty} X p(x) dx$  als

Erwartungswert. Es wird wieder die absolute Konvergenz

vorausgesetzt:  $\int_{-\infty}^{\infty} |x| p(x) dx < \infty$ .

$$\begin{array}{l} EX = \sum_k x_k P\{X=x_k\} \\ EX = \int_{-\infty}^{\infty} x p(x) dx \end{array} \quad \left. \begin{array}{l} \text{(diskret)} \\ \text{(stetig)} \end{array} \right\} \text{ absolut konvergent}$$

**Bemerkungen:** Der Begriff des Erwartungswertes kommt unserer anschaulichen Vorstellung eines Mittels (Mittlung) sehr nahe. Deutet man eine diskrete oder stetige Wahrscheinlichkeitsverteilung als Massenverteilung eines diskret oder stetig verteilten mechanischen Systems auf der Achse, so ist  $EX$  der Schwerpunkt des Systems (im physikalischen Sinne).

Es gibt aber auch andere Methoden zu Mitteln (Median, Zentralwert)

**Beispiele:**

- Poisson-Verteilung  $EX = \sum_{k=0}^{\infty} k P\{X = k\} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$   
 $= \lambda e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^{\lambda}} = \lambda$

Parameter der Poisson-Verteilung ist in diesem Falle = Erwartungswert derselben, vorteilhaft für statistische Untersuchungen.

- Exponentialverteilung  $EX = \int_{-\infty}^{\infty} xp(x)dx = \int_0^{\infty} x\lambda e^{-\lambda x} dx \stackrel{\text{partielle Integr.}}{=} \dots = \frac{1}{\lambda}$

### 2.2.1 Eigenschaften des Erwartungswertes

1. Linearität  $E(aX + bY) = aEX + bEY$
2. Funktion einer Zufallsgröße  $g(X)$

$$X \text{ **diskret**: } \quad Eg(X) = \sum_k g(x_k) P\{X = x_k\}$$

$$X \text{ **stetig**: } \quad Eg(X) = \int_{-\infty}^{\infty} g(x) \underbrace{p(x)}_{\text{Dichte von X}} dx$$

absolute Konvergenz der rechten Seite jeweils vorausgesetzt.

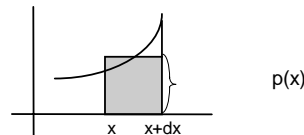
Der Erwartungswert charakterisiert die Lage des Zentrums einer Verteilung (Lageparameter); über die Stärke möglicher Abweichungen der Zufallsgröße vom Zentrum gibt er keine Auskunft.

→ Wunsch nach einem Maß für die „Streuung“

**Definition:** Es sei  $X$  eine Zufallsgröße mit dem Erwartungswert  $EX$ . Dann wird im Falle der Existenz durch  $D^2X = E(X - EX)^2$  die **Streuung** (oder Varianz) der Zufallsgröße  $X$  definiert.

**Speziell gilt:**

$$X \text{ **diskret**: } \quad D^2X = \sum_k (x_k - EX)^2 P\{X=x_k\}$$



$$X \text{ **stetig**: } \quad D^2X = \int_0^{\infty} (x - EX)^2 p(x) dx$$

denn  $D^2X$  wird über die Funktion  $g(X) = (X - EX)^2$  von  $X$  bestimmt.

**Bemerkung:** Auch die Streuung gestattet eine mechanische Interpretation. Veranschaulicht man sich die Verteilung als diskretes oder stetiges Massensystem auf der Achse mit dem Schwerpunkt  $EX$ , so entspricht  $D^2X$  dem Trägheitsmoment dieses Systems bezüglich einer Achse durch den Schwerpunkt. (Existenz jeweils vorausgesetzt).

### 2.2.2 Eigenschaften der Streuung

1.  $D^2(aX + b) = a^2 D^2X$

$$\begin{aligned} \text{Beweis: } \quad D^2(aX + b) &= E[(aX + b) - E(aX + b)]^2 = E[a(X - EX) + (b - Eb)]^2 \\ &= E[a^2 (X - EX)^2] = a^2 D^2X \end{aligned}$$

Folgerung: •  $D^2(-X) = D^2X$   
 •  $D^2\left(\frac{X}{\sqrt{D^2 X}}\right) = 1$  (Normieren, Standardisieren von X)

2. Annahme: X einpunktverteilt:

$P\{X = c\} = 1$  c feste reelle Zahl

→ X = c fast sicher, EX = c

Die Streuung einer Zufallsgröße X ist genau dann Null, wenn X einpunktverteilt ist.

3.  $D^2X = EX^2 - (EX)^2$

Beweis:  $D^2X = E(X - EX)^2 = E[X^2 - 2X EX + (EX)^2] = EX^2 - 2EX (EX) + (EX)^2 = EX^2 - (EX)^2$

Diese Aussage ist mit dem Steinerschen Satz äquivalent:

$$\underbrace{EX^2}_{\text{Trägheitsmoment bzgl. d. Achse durch 0}} = \underbrace{D^2X}_{\text{Trägheitsmoment bzgl. d. Achse durch Schwerpunkt}} + \underbrace{(EX)^2}_{\text{Abstandsquadrat: Schwerpunkt-Nullpunkt}}$$

4.  $D^2(X+Y) = D^2X + D^2Y + 2[E(XY) - (EX)(EY)]$

Beispiele:

1. X ist **Poisson-verteilt** mit Parameter  $\lambda > 0$

$$\begin{aligned} EX^2 &= \sum_{g(x)=x^2} Eg(X) = \sum_{k=0}^{\infty} g(x_k) P\{X = x_k\} = \sum_{k=0}^{\infty} k^2 P\{X = k\} = \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \left[ \sum_{k=0}^{\infty} k P\{X = k\} + \sum_{k=0}^{\infty} P\{X = k\} \right] = \lambda[EX + 1] \\ &= \lambda(\lambda + 1) \end{aligned}$$

Steinerscher Satz →  $D^2X = EX^2 - (EX)^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda$

2. X ist **exponentialverteilt** mit Parameter  $\lambda > 0$

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 p(x) dx = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = \dots \text{2 mal partiell integriert} \dots = \frac{2}{\lambda^2} \\ \rightarrow D^2X &= EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}$$

3. **Normalverteilung**

Y sei eine standardnormalverteilte Zufallsgröße

→  $EY = 0, D^2Y = 1$  (Streuung)

kurz: N(0, 1)

$X = \sigma Y + \mu$  - lineare Transformation von  $Y$ ;  $\sigma > 0$ ,  $\mu \in \mathbf{R}$

$$F_X(x) = P\{X < x\} = P\{\sigma Y + \mu < x\} = P\left\{Y < \frac{x - \mu}{\sigma}\right\}$$

$$= F_Y\left(\frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right) = \int_{-\infty}^{\frac{x - \mu}{\sigma}} \ell(u) du = \frac{1}{\sigma} \int_{-\infty}^x \ell\left(\frac{t - \mu}{\sigma}\right) dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t - \mu)^2}{2\sigma^2}} dt$$

**Definition:** Eine Zufallsgröße  $X$  mit dieser Verteilungsfunktion  $F_X$  heißt **normalverteilt** mit den Parametern  $\sigma > 0$ ,  $\mu \in \mathbf{R}$  oder  $N(\mu, \sigma^2)$  verteilt.

**Bedeutung der Parameter:**

$$EX = E(\sigma Y + \mu) = \sigma \underbrace{EY}_{=\mu} + \mu = \mu$$

$$D^2X = D^2(\sigma Y + \mu) = \sigma^2 \underbrace{D^2Y}_{=1} = \sigma^2$$

### 2.3 Faltung zweier Verteilungen

$X, Y$  seien Zufallsgrößen auf  $(\Omega, \mathcal{M}, P)$

$g: \mathbf{R}^2 \rightarrow \mathbf{R}$  Funktion, (Meßbarkeit sei gesichert)  $\rightarrow$  neue Zufallsgröße:  $Z = g(X, Y)$

Verteilung von  $Z$  ?

*speziell:*  $X, Y$  seien disjunkt mit Werten  $x_i, y_k$ ;  $(i, k = 0, 1, \dots)$

$$P\{Z = z\} = \sum_{g(x_i, y_k) = z} P\{X = x_i, Y = y_k\} \quad z \in \mathbf{R}$$

Summiert wird also über disjunkte Indizes  $i, k$  für die  $g(x_i, y_k) = z$ .  
Existieren keine solchen Werte  $x_i, y_k$ , so ist die Summe gleich Null.  
Zur Berechnung von  $P\{X = z\}$  muß man also i. a. die gemeinsame Verteilung von  $X$  und  $Y$  kennen.

*noch spezieller:* Summe von  $X$  und  $Y$

$$P\{Z = z\} = \sum_{x_i + y_k = z} P\{X = x_i, Y = y_k\}$$

Übung:  $X, Y$  Poisson-verteilt  $\rightarrow$  Speziell  $X, Y$  mit Werten  $0, 1, 2, \dots$

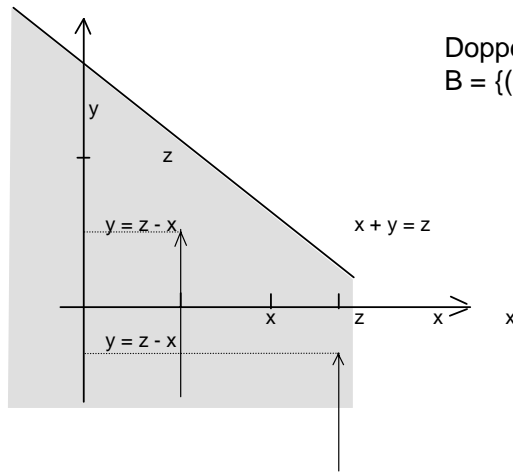
$$P\{Z = n\} = \sum_{k=0}^n P\{X = k; Y = n - k\} = (X, Y \text{ unabhängig})$$

$$= \sum_{k=0}^n P\{X = k\} P\{Y = n - k\}$$

nennt man „**Faltung**“ der Verteilung von  $X$  und  $Y$  !!“  
**Ziel:** Übertragung der Methode auf stetige Zufallsgrößen  $X, Y$  mit der gemeinsamen Dichte  $f_{X,Y}$

gesucht:  $Z = X + Y$  stetig? , Dichte  $f_Z$  ?

$$z \in \mathbf{R}: f_Z(z) = P\{Z < z\} = P\{X + Y < z\} = \underbrace{\iint_{X+Y < z} f_{X,Y}(x,y) dx dy}$$



Doppelintegral, Integrationsgebiet  
 $B = \{(x,y): -\infty < x < \infty, -\infty < y < z-x\}$

$$= (\text{Integral iterieren}) = \int_{x=-\infty}^{\infty} \left( \int_{y=-\infty}^{z-x} f_{X,Y}(x,y) dy \right) dx = (\text{Substitution im inneren Integral})$$

Substitution:  $\zeta = x + y, \quad y = \zeta - x$   
 $dy = d\zeta$   
 $y = -\infty \rightarrow \zeta = -\infty$   
 $y = z - x \rightarrow \zeta = x + y = x + (z - x) = z$

$$= \int_{x=-\infty}^{\infty} \left( \int_{z=-\infty}^z f_{X,Y}(x, z-x) dz \right) dx$$

**Ziel:** Funktion von  $z \rightarrow$  Integralreihenfolgetausch

$$= \int_{\zeta=-\infty}^z \left( \int_{x=-\infty}^z f_{X,Y}(x, \zeta - x) dx \right) d\zeta$$

Dichte für  $f_Z(z) : f_Z(\zeta)$

$\rightarrow$  Dichte für  $Z = X + Y$

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x) dy; \quad z \in \mathbf{R}$$

$\rightarrow$  Damit ist  $Z = X + Y$  eine stetige Zufallsgröße

*speziell:* Annahme:  $X$  und  $Y$  unabhängig

$$\rightarrow f_{X,Y}(x,y) = f_X(x) * f_Y(y) \quad (x, y \in \mathbf{R})$$

$$\rightarrow f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x) dx = \int_{-\infty}^{\infty} f_X(x) * f_Y(z-x) dx$$

„Faltungsformel“ für die Dichte  $f_X$  und  $f_Y$  (bei Unabhängigkeit)

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) * f_Y(z-x) dx \quad ; \quad z \in \mathbf{R}$$

**Bemerkung:** Durch Vertauschen der Rolle von  $x$  und  $y$  in den obigen Überlegungen beweist man die Formel:

$$f_Z(z) = \int_{y=-\infty}^{\infty} f_X(z-y) * f_Y(y) dy \quad ; \quad z \in \mathbf{R}$$

weitere Aussagen über unabhängige Zufallsgrößen  $X$  und  $Y$ :

- $EXY = EX EY$
- $f, g$  reelle Funktionen,  $X, Y$  unabhängig  $\Rightarrow f(X), g(X)$  unabhängig

### 2.3.1 Momente einer Zufallsgröße

$X$  - Zufallsgröße

**Definition:** Im Falle der Existenz heißt  $\mu_k = EX^k$  ( $k = 1, 2, \dots$ ) das  $k$ -te Moment der Zufallsgröße  $X$ .

allgemein:  $\mu_k = E(x - c)^k$  ( $k = 1, 2, \dots$ )  
heißt  $k$ -tes Moment von  $X$  bezüglich  $c \in \mathbf{R}$

Momente sind sowohl theoretisch (Momentenproblem) als auch praktisch (Statistik) bedeutsam.

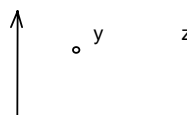
### 2.4 Charakteristische Funktionen

**Vorbemerkung:** komplexe Zufallsgröße  
 $X, Y$  (reelle) Zufallsgrößen auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{M}, P)$   
 $Z = X + iY$  komplexe Zufallsgröße auf  $(\Omega, \mathcal{M}, P)$

- Meßbarkeit überträgt sich
- Verteilung von  $Z$  kann durch die gemeinsame Verteilung von  $X, Y$  charakterisiert werden
- Erwartungswert von  $Z$ : formal summieren  $EZ = EX + iEY$
- $Z_j = X_j + iY_j$  ( $j = 1, 2$ )  
 $Z_1, Z_2$  unabhängig: ( $X_1, Y_1$  unabhängig von  $X_2, Y_2$ )

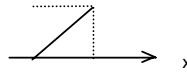
- Betrag  $\underbrace{|EZ|}_{\text{komplexe Zahl}} \leq E \underbrace{|Z|}_{\text{reelle Zufallsgröße}}$  Anmerkung:  $|z| = \sqrt{x^2 + y^2}$

**Darstellung komplexer Zahlen (Wdh.)**



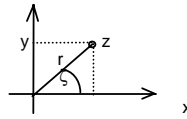
- *algebraische Form:*

$$z = x + iy$$



- *trigonometrische Form:*

$$z = r(\cos \zeta + i \sin \zeta)$$



- *Exponentialform:*

$$z = r e^{i\zeta}$$

**Definition:**

Sei  $X$  eine (reelle) Zufallsgröße auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{M}, P)$ . Dann bezeichnen wir die Abbildung  $t \rightarrow \zeta_X(t) := E(e^{itX})$ ,  $-\infty < t < \infty$  als charakteristische Funktion der Zufallsgröße  $X$  oder der Verteilung von  $X$ .

**Bemerkung:**

Aufgrund von  $|e^{itx}| = 1$  ist obiger Ausdruck absolut und gleichmäßig in  $t$  konvergent. Zu jeder Zufallsgröße  $X$  existiert also die charakteristische Funktion  $\zeta_X(t)$ .

$X$  diskrete Zufallsgröße:

$$\zeta(t) = E(e^{itX}) \stackrel{[g(x)=e^{itx}]}{=} \sum_{k=0}^{\infty} e^{itx_k} * P\{X = x_k\}, t \in \mathbf{R}$$

(absolute Konvergenz)

**Beispiel:**

$X$  sei zweipunktverteilt

$$P\{X = -1\} = P\{X = 1\} = \frac{1}{2}$$

$$\rightarrow \zeta(t) = e^{-it} * \frac{1}{2} + e^{it} = \frac{e^{-it} + e^{it}}{2} = \cos t$$

$X$  stetig:

Verteilung von  $X$  besitzt Dichte  $f_X$

$$\zeta(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx, t \in \mathbf{R}$$

$$\text{Konvergenz: } \int_{-\infty}^{\infty} \underbrace{|e^{itx}|}_{=1} f_X(x) dx = \int_{\mathbf{R}} f_X(x) dx = 1$$

(absolut konvergent und gleichmäßig in  $t$ )

**X stetig:** Verteilung von X besteht aus Dichte  $f_X$

$$\varphi(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx, \quad t \in \mathbf{R}$$

$g(x) = e^{itx}$

**Konvergenz:**  $\int_{-\infty}^{\infty} |e^{itx}| f_X(x) dx = \int_{-\infty}^{\infty} f_X(x) dx = 1$

**Beispiel:** X auf [0, 1] gleichmäßig verteilt:

$$f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

$$\rightarrow \varphi(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx = \int_{x=0}^1 e^{itx} * 1 dx = \left[ \frac{e^{itx}}{it} \right]_{x=0}^{x=1} = \frac{e^{it} - 1}{it}, \quad t \in \mathbf{R}$$

**Eigenschaften charakteristischer Funktionen**

$$\begin{aligned} \varphi(0) &= E(e^{itX})|_{t=0} = E(e^0) = e^0 = 1 \\ |\varphi(t)| &= |E(e^{itX})| \leq |E(e^{itX})| = E^*1 = 1 && \text{(qualitative Aussage)} \\ \varphi(-t) &= \overline{\varphi(t)} \quad t \in \mathbf{R} && \text{(konjugiert komplex)} \end{aligned}$$

[jede charakteristische Funktion erfüllt (notwendigerweise) diese Bedingungen; sie sind indessen nicht hinreichend]

**lineare Transformation von X:**

Zufallsgröße  $Y = aX + b$  charakteristische Funktion ?

$$\varphi_Y(t) = E(e^{itY}) = E(e^{it(aX+b)}) = E(e^{itb} e^{itaX}) = e^{itb} E(e^{itaX}) = e^{itb} \varphi_X(at), \quad t \in \mathbf{R}^1$$

**Multiplikationssatz:**

X, Y seien zwei unabhängige Zufallsgrößen  
 charakteristische Funktion der Summe  $Z = X + Y$  ?  
 $\varphi_Z(t) = E(e^{itZ}) = E(e^{it(X+Y)}) = E(e^{itX} e^{itY}) =$  (unter der Voraussetzung, daß X und Y unabhängig sind)  $= E(e^{itX}) E(e^{itY}) = \varphi_X(t) * \varphi_Y(t), \quad t \in \mathbf{R}$

**Satz:** Die charakteristische Funktion der Summe endlich vieler vollständig unabhängiger Zufallsgrößen ist gleich dem Produkt der charakteristischen Funktionen dieser Zufallsgrößen.

**Beispiel:** X sei Poisson-verteilt mit Parameter  $\lambda > 0$

$$\begin{aligned} \varphi_X(t) &= E(e^{itX}) = \sum_{k=0}^{\infty} e^{itk} P\{X=k\} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \\ &= \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}, \quad t \in \mathbf{R} \end{aligned}$$



Y sei Poisson-verteilt mit Parameter  $\mu > 0$  und unabhängig von X.  
 Wie ist  $Z = X + Y$  verteilt ?

$$\varphi_Z(t) = \varphi_X(t) \varphi_Y(t) = e^{\lambda(e^{it}-1)} e^{\mu(e^{it}-1)} = \exp[\lambda(e^{it}-1) + \underbrace{\mu(e^{it}-1)}_{\text{charakteristische Funktion einer Poissonverteilung mit Parametern } \lambda + \mu}] = \exp [(\lambda+\mu) (e^{it}-1)]$$

**Bemerkung:** Falls  $\varphi_Z$  die Verteilung von Z, **eindeutig** charakterisiert, so ist Z Poisson-verteilt mit dem Parameter  $\lambda + \mu$ .

**Eindeutigkeitssatz:** Jede Verteilung ist durch ihre charakteristische Funktion eindeutig bestimmt.

### Erzeugung der Momente einer Zufallsgröße

Besitzt die Zufallsgröße X das Moment k-ter Ordnung  $m_k = EX^k$ , so existiert die k-te Ableitung von  $\varphi_X$  und es gilt:

$$m_k = \frac{1}{i^k} \left. \frac{d^k \varphi_X(t)}{dt^k} \right|_{t=0}$$

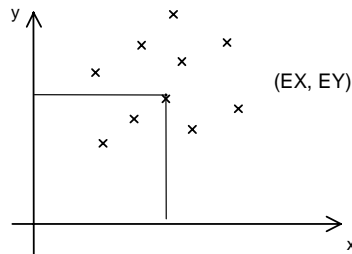
### Charakteristiken im 2-dimensionalen

Zufallsvektor (X, Y) auf dem Wahrscheinlichkeitsraum  $(\Omega, M, P)$

$$[P_{(X,Y)}(B) = P\{(X,Y) \in B\} = P\{\omega: (X(\omega), Y(\omega)) \in B\}, B \in L^2]$$

**Informationsverdichtung:** Erwartungswert, Streuung (Varianz) verallgemeinern.

**Erwartungsvektor:** (EX, EY)



### Varianz verallgemeinern:

zwei Zufallsgrößen X, Y aus  $D^2X, D^2Y$  aber auch Abhängigkeiten zwischen X und Y berücksichtigen.

## 2.5 Einführung der Kovarianz (Abweichungen)

$$\text{cov}(X, Y) = E(X - EX)(Y - EY) \rightarrow \text{cov}(X, Y) = D^2X$$

**Kovarianzmatrix:**

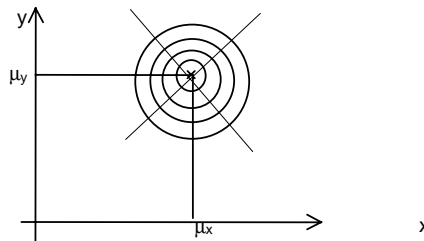
$$\begin{bmatrix} D^2X & \text{cov}(X, Y) \\ \text{cov}(X, Y) & D^2Y \end{bmatrix} \text{symmetrisch, positiv definierte Matrix}$$

**Kovarianz normieren:**  $\zeta(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D^2X} \sqrt{D^2Y}}$

**Aussage:** Komponenten X und Y unabhängig  $\rightarrow E(X \cdot Y) = (EX)(EY)$   
 $\text{cov}(X, Y) = 0$   
 $\zeta(X, Y) = 0$   
 $D^2(X + Y) = D^2X + D^2Y$   
 (Anmerkung:  $+2[E(XY) - EX EY] = \text{cov}(X, Y)$ )

**Beispiel:** 2-dimensionale Normalverteilung, angegeben durch die Dichte der Verteilung von (X, Y):

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\zeta^2}} \exp\left\{-\frac{1}{2(1-\zeta^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\zeta\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}$$



$$\text{cov}(X, Y) = \zeta \sigma_x \sigma_y, \quad \zeta(x, y) = \zeta$$

$$X \approx N(\mu_x, \sigma_x^2), \quad Y \approx N(\mu_y, \sigma_y^2)$$

## 2.6 Folgen von Zufallsgrößen

**Markovsche Ungleichungen:**

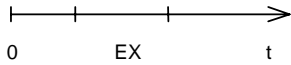
Es sei X eine fast sicher nicht negative Zufallsgröße:  $P\{X \geq 0\} = 1$ , für die der Erwartungswert EX existiert. Dann gilt für jede Zahl  $t > 0$  die Abkürzung:

$$P\{X \geq t\} \leq \frac{EX}{t}$$

**Beweis:**

$X$  ist diskret: Einzelwahrscheinlichkeiten  $P\{X = x_i\}$  ( $i = 0, 1, 2, \dots$ )

$$\begin{aligned} \rightarrow EX = \sum_i x_i P\{X = x_i\} &= \underbrace{\sum_{i: x_i < t} x_i P\{X = x_i\}}_{\geq 0} + \sum_{i: x_i > t} x_i P\{X = x_i\} \geq \sum_{i: x_i > t} x_i P\{X = x_i\} \\ &\geq t \sum_{i: x_i > t} P\{X = x_i\} = t P\{X \geq t\} \Rightarrow \text{Behauptung} \end{aligned}$$



$X$  sei fast sicher  $\geq 0$

$$P\{X \geq t\} \leq \frac{EX}{t}$$

$X$  sei stetig: Dichte  $f_X(x) > 0$

$$\begin{aligned} EX = \int_{-\infty}^{\infty} x f_X(x) dx &= \int_0^{\infty} x f_X(x) dx = \underbrace{\int_0^t x f_X(x) dx}_{\geq 0} + \int_t^{\infty} x f_X(x) dx \\ &\geq \int_t^{\infty} x f_X(x) dx \geq t \int_t^{\infty} f_X(x) dx = t P\{X \geq t\} \Rightarrow \text{Behauptung} \end{aligned}$$

**Bemerkung:** In dieser Ungleichung wird von der Verteilung von  $X$  lediglich der Erwartungswert  $EX$  benutzt, daher ist die obige Abschätzung oft recht grob. Zu besseren Resultaten kommt man, wenn man auch die Summe  $D^2X$  der Zufallsgrößen benutzen kann.

### 2.6.1 Tschebyscheffsche Ungleichung

Die Zufallsgröße  $X$  besitze Erwartungswert und Streuung. Dann gilt für  $t > 0$ :

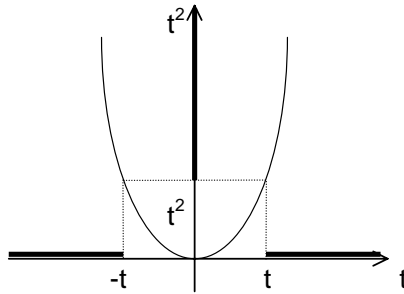
$$P\{|X - EX| \geq t\} \leq \frac{D^2X}{t^2}$$

**Beweis:** Unter Anwendung der Markovschen Ungleichung auf  $(X - EX)^2$ ,  $t$  wird durch  $t^2$  ersetzt:

$$P\{(X - EX)^2 \geq t^2\} \leq \frac{E(X - EX)^2}{t^2}$$

Das Ereignis  $\{(X - EX)^2 \geq t^2\}$  tritt genau dann ein,

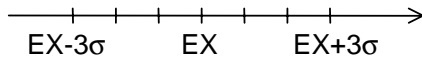
$$\text{wenn } \{|X - EX| \geq t\} \leq \frac{E(X - EX)^2}{t^2}$$



dick gekennzeichnete Bereiche auf der t-Achse entsprechen genau dem markierten Bereich der  $t^2$ -Achse

**Bemerkungen:** Für  $t^2 \leq D^2X$  wird die Aussage dieser Ungleichung trivial, denn dann ist  $\frac{D^2X}{t} \geq 1$ .

Im Falle  $t = 3\sigma = 3\sqrt{D^2X}$  (den „Drei-Sigma-Grenzen“) besagt die Tschebyscheffsche Ungleichung, daß für jede Zufallsgröße  $X$  - also für jede Verteilung - mit Existieren der Streuung gilt:



$$P\{|X - EX| < 3\sigma\} = 1 - P\{|X - EX| \geq 3\sigma\} \geq 1 - \frac{D^2X}{9\sigma^2} = 1 - \frac{\sigma^2}{9\sigma^2} = 1 - \frac{1}{9} = \underline{\underline{\frac{8}{9}}}$$

(quantitativ wichtige Aussage)

**Bemerkung:** Wenn die Wahrscheinlichkeit  $= \frac{8}{9}$  ist, liegt  $X$  im Intervall  $[EX - 3\sigma, EX + 3\sigma]$ .

## 2.7 Gesetze der großen Zahlen

**Bernoulli-Schema:** Serie von  $n$  unabhängigen Versuchen, Zufallsgröße  $X_n$  zählt das Eintreten des Ereignisses  $A$  in der Serie.

$P(A) = p \rightarrow X_n$  binomial verteilt mit den Parametern  $n, p$

$\rightarrow Ex_n = n \cdot p, D^2X_n = n \cdot p(1 - p)$

$X_n$ : absolute Häufigkeit des Eintretens von  $A$

relative Häufigkeit:  $h_n(A) := \frac{X_n}{n}$

**Naturbeobachtung:** Folge der relativen Häufigkeiten  $h_n(A)$  ( $n = 1, 2, \dots$ ) ist stabil.

*Erwartungswert und Streuung der relativen Häufigkeiten:*

$$E h_n(A) = E \frac{X_n}{n} = \frac{1}{n} Ex_n = \frac{np}{n} = p = P(A)$$

$\rightarrow E h_n(A) = P(A) \quad (n = 1, 2, \dots)$

[Erwartungswert d. relativen Häufigkeit von  $A$  ist gleich der Wahrscheinlichkeit von  $A$ ]

$$\rightarrow D^2 h_n(A) = \frac{p(1-p)}{n} \rightarrow 0 \text{ für } n \rightarrow \infty \quad (\text{deterministische Konvergenz})$$

Die Varianz und damit die Abweichung der relativen Häufigkeit  $h_n(A)$  von der dem Experiment zugrunde liegende Wahrscheinlichkeit  $P(A)$ , die wir als **objektiv gegeben** ansehen, wird immer kleiner. Mit Hilfe der Tschebyscheffschen Ungleichung läßt sich diese Abweichung sogar quantitativ erfassen:

$$\text{Für } \varepsilon > 0 \text{ gilt: } P\{|h_n(A) - P(A)| \geq \varepsilon\} = P\{|h_n(A) - E h_n(A)| \geq \varepsilon\} \leq \frac{D^2 h_n(A)}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}$$

### 2.7.1 Bernoullisches Gesetz der großen Zahlen:

Es sei  $A$  ein in einem zufälligen Versuch auftretendes Ereignis und  $h_n(A)$  die relative Häufigkeit des Eintretens von  $A$  in einer Serie von unabhängigen Wiederholungen dieses Versuches ( $n = 1, 2, \dots$ ). Dann gilt:

$$\lim_{n \rightarrow \infty} P\{|h_n(A) - P(A)| < \varepsilon\} = 1$$

Man sagt: „ $h_n(A)$  konvergiert stochastisch gegen die Wahrscheinlichkeit  $P(A)$ “.

#### Bemerkungen:

1. Neben seiner unmittelbaren Bedeutung lehrt dieses Gesetz auch, daß jedes Ereignis positiver Wahrscheinlichkeit - wie klein diese auch sein mag - in einer genügend langen Versuchsserie mit einer beliebig nahe an Eins gelegenen Wahrscheinlichkeit mindestens einmal vorkommt: „Gesetz der großen Zahlen“ - der Traum jeden Spielers.
2. Allgemein ist dieses Gesetz ein hervorragender Beleg für die adäquate Beschreibung realer zufälliger Phänomene durch die Wahrscheinlichkeitstheorie.

#### Verallgemeinerung:

$X_k$ ;  $k = 1, 2, \dots$  Folge von Zufallsgrößen  
arithmetisches Mittel der ersten  $n$  Zufallsgrößen  $X_k$ :

$$Y_n = \frac{1}{n} \sum_{k=1}^n X_k$$

#### Aussage:

$X_1, X_2, \dots, X_n, \dots$  Folge unabhängiger Zufallsgrößen mit gleichem

Erwartungswert  $E X_k = \mu$  ( $k = 1, 2, \dots$ ) und beschränkter Streuung  $D^2 X_k \leq M < \infty$ .

Für  $\varepsilon > 0$  gilt:  $(*) \lim_{n \rightarrow \infty} P\{|Y_n - \mu| \geq \varepsilon\} = 0$

#### Aussage:

$X_1, X_2, \dots, X_k, \dots$  paarweise unabhängige identisch verteilte Zufallsgrößen, deren Erwartungswert und Streuung existieren. Auch dann gilt das Gesetz der großen Zahlen (\*).

Wie wir schon beim Grenzwertsatz von de Moire-Laplace kennenlernten, sind auch Aussagen über Grenzverteilungen möglich.

Es sei  $(X_k)$  eine Folge unabhängiger, identisch verteilter Zufallsgrößen mit positiver Streuung.

Dann genügt die Folge  $(X_k)$  dem „zentralen Grenzwertsatz“:

$$EX_k = \mu \quad , \quad D^2X_k = \sigma^2 > 0 \quad (k = 1, 2, \dots)$$

$$Z_n = \frac{\sum_{k=1}^n (X_k - \mu)}{\sqrt{\sum_{k=1}^n \sigma^2}} = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n\sigma^2}} \quad (n = 1, 2, \dots)$$

$$P\left\{a < \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n\sigma}} < b\right\} \xrightarrow{n \rightarrow \infty} \Phi(b) - \Phi(a)$$

d.h. die Zufallsgröße  $Z_n = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n\sigma}}$  ( $n = 1, 2, \dots$ ) sind asymptotisch normalverteilt, genauer  $N(0, 1)$  verteilt.

**Bemerkung:** Der *Zentrale Grenzwertsatz* manifestiert die enorme theoretische und praktische Bedeutung der Normalverteilung. Ergibt sich eine Zufallsgröße aus der Überlagerung einer Vielzahl weitgehend unabhängiger zufälliger Effekte, so ist sie genähert normalverteilt. Zur Bestimmung der asymptotischen Verteilung ist dann nur die Kenntnis der Erwartungswerte und Streuungen erforderlich.