



**FH AACHEN**  
UNIVERSITY OF APPLIED SCIENCES

Skript zur Vorlesung

# Angewandte Wahrscheinlichkeitstheorie

Georg Hoever

Fachbereich Elektrotechnik und  
Informationstechnik  
FH Aachen

---

## Inhaltsverzeichnis

<b>1. Grundlagen</b>	<b>1</b>
1.1. Wahrscheinlichkeiten bei diskreten Ereignissen . . . . .	1
1.2. Wahrscheinlichkeitsdichte . . . . .	8
1.3. Zufallsvariablen . . . . .	13
1.4. Verteilungsfunktion . . . . .	15
1.5. Erzeugung von Zufallszahlen . . . . .	18
<b>2. Weiterführende Begriffe</b>	<b>21</b>
2.1. Bedingte Wahrscheinlichkeit und Unabhängigkeit . . . . .	21
2.2. Erwartungswert und Varianz . . . . .	27
2.3. Korrelation . . . . .	30
2.4. Zentraler Grenzwertsatz . . . . .	32
<b>3. Statistik</b>	<b>35</b>
3.1. Grundlagen . . . . .	35
3.2. Schätzer . . . . .	38
3.3. Konfidenzbereiche . . . . .	43
3.3.1. Grundlagen . . . . .	43
3.3.2. Konfidenzbereiche für Erwartungswerte bei Normalverteilungen . .	44
3.3.3. Konfidenzbereiche für Bernoulli-Experimente . . . . .	47
3.4. Hypothesentests . . . . .	48
3.4.1. Grundlagen . . . . .	48
3.4.2. Tests für Erwartungswerte und Erfolgswahrscheinlichkeiten . . . .	50
3.4.3. Der $p$ -Wert . . . . .	54
3.4.4. Hypothesentests und Konfidenzbereiche . . . . .	55
3.4.5. $\chi^2$ -Verteilung und $\chi^2$ -Tests . . . . .	57
<b>4. Stochastische Prozesse und Markov-Ketten</b>	<b>60</b>
<b>5. Warteschlangen</b>	<b>65</b>
5.1. Einführung . . . . .	65
5.2. Die Poisson-Verteilung . . . . .	67
5.3. Eigenschaften von Warteschlangen . . . . .	69
<b>A. Binomialkoeffizient</b>	<b>72</b>
<b>B. Tabelle: Standardnormalverteilung</b>	<b>74</b>
<b>C. Tabelle: Quantile zur <math>t</math>-Verteilung</b>	<b>75</b>
<b>D. Tabelle: Quantile zur <math>\chi^2</math>-Verteilung</b>	<b>76</b>

---

---

<b>E. Übersicht: Verteilungen</b>	<b>77</b>
E.1. Diskrete Verteilungen . . . . .	77
E.2. Stetige Verteilungen . . . . .	79
<b>Index</b>	<b>80</b>

---

## 1. Grundlagen

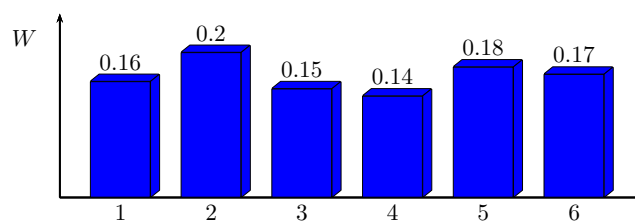
### 1.1. Wahrscheinlichkeiten bei diskreten Ereignissen

#### Beispiel 1:

1. Bei einem Münzwurf gibt es die Ergebnisse „Kopf“ oder „Zahl“. Bei einer „fairen“ Münze kommt jedes Ergebnis mit der Wahrscheinlichkeit  $\frac{1}{2}$  vor.
2. Bei einem „fairen“ Würfel kommt jede Augenzahl (1,2,...,6) gleich wahrscheinlich vor. Die Wahrscheinlichkeit ist jeweils  $\frac{1}{6}$ .
3. Bei einem ungleichmäßigen Würfel könnte es die folgenden Auftrittswahrscheinlichkeiten geben:

Augenzahl	1	2	3	4	5	6
Wahrscheinlichkeit	0.16	0.2	0.15	0.14	0.18	0.17

Als Balkendiagramm:



#### Definition 1.1

Bei einem *Zufallsexperiment* mit endlich vielen *Ergebnissen*  $\omega_1, \omega_2, \dots, \omega_N$  kann man jedem Ergebnis  $\omega_k$  eine *Wahrscheinlichkeit*  $p(\omega_k) = p_k$  zuordnen.

Es gilt dann  $\sum_{k=1}^N p_k = 1$ .

Ist jedes Ergebnis gleich wahrscheinlich, so heißt das Zufallsexperiment *Laplace-Experiment*. Für jedes  $k$  gilt dann  $p_k = \frac{1}{N}$ .

#### Bemerkungen:

1. Die Menge aller möglichen Ergebnisse wird üblicherweise mit  $\Omega$  bezeichnet. Die Wahrscheinlichkeit, dass ein Ergebnis in einer Teilmenge  $A \subseteq \Omega$  liegt, wird mit  $P(A)$  bezeichnet.
2. Tritt bei  $n$  Versuchen ein Ergebnis  $a$ -mal auf, so nennt man  $\frac{a}{n}$  die *relative Häufigkeit*. Diese nähert sich mit steigendem  $n$  der Wahrscheinlichkeit  $p$  des Ergebnisses.

Kennt man die exakte Wahrscheinlichkeit nicht, so ist die relative Häufigkeit eine Schätzung dieser Wahrscheinlichkeit. *Monte-Carlo Simulation* bedeutet die gezielte

Durchführung von Versuchen, um so die Wahrscheinlichkeit approximativ zu bestimmen.

3. Ein Zufallsexperiment mit nur zwei möglichen Ergebnissen  $\omega_1, \omega_2$  heißt *Bernoulli-Experiment*. Ist  $p = p(\omega_1)$ , so ist  $p(\omega_2) = 1 - p$ .

Statt  $\{\omega_1, \omega_2\}$  nimmt man häufig  $\{1, 0\}$ , also  $p(1) = p$ ,  $p(0) = 1 - p$ .

### Beispiel 2:

- Der Wurf mit einem fairen Würfel ist ein Laplace-Experiment. Betrachtet man nur die Fälle „Der Wurf ergibt eine 6“ und „Der Wurf ergibt keine 6“, so ist dies ein Bernoulli-Experiment mit  $p = \frac{1}{6}$ .
- Bei einem Wurf mit zwei fairen Würfeln sind alle Ergebnisse

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

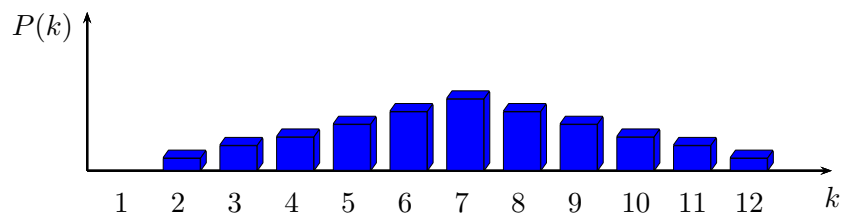
gleich wahrscheinlich, d.h. jedes Tupel (Augenzahl 1.Würfel, Augenzahl 2.Würfel) hat die Wahrscheinlichkeit  $\frac{1}{36}$ .

Bei einem Wurf mit zwei Würfeln erhält man 5 als Augensumme bei den Ergebnissen (1, 4), (2, 3), (3, 2) und (4, 1). Die Wahrscheinlichkeit für „Augensumme = 5“ ist somit  $4 \cdot \frac{1}{36} = \frac{1}{9}$ . Ähnliche Überlegungen für andere Summenwerte ergeben die folgende Tabelle:

$k$	1	2	3	4	5	6	7	8	9	10	11	12
$P(\text{Augensumme} = k)$	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Die Summe der Werte ist

$$0 + \frac{1}{36} + \frac{1}{18} + \frac{1}{12} + \frac{1}{9} + \frac{5}{36} + \frac{1}{6} + \frac{5}{36} + \frac{1}{9} + \frac{1}{12} + \frac{1}{18} + \frac{1}{36} = 1.$$



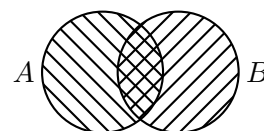
**Satz 1.2**

1. Bei der Kombination zweier unabhängiger Zufallsexperimente ergeben sich die Wahrscheinlichkeiten der Ergebnistupel als Produkt der entsprechenden Einzelwahrscheinlichkeiten.
2. Die Wahrscheinlichkeit für eine diskrete Ergebnismenge ergibt sich als Summe der Einzelwahrscheinlichkeiten:

$$\text{Für } A \subseteq \Omega \text{ ist } P(A) = \sum_{\omega_k \in A} p(\omega_k).$$

3. Für Ergebnismengen  $A$  und  $B$  gilt

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



Sind  $A$  und  $B$  disjunkt, d.h.  $A \cap B = \{\}$ , so gilt speziell

$$P(A \cup B) = P(A) + P(B).$$

**Bemerkungen:**

1. Aus Definition 1.1 folgt  $P(\Omega) = 1$ .
2. Bei einem Laplace-Experiment gilt für  $A \subseteq \Omega$ :

$$P(A) = \frac{\text{Anzahl der Elemente von } A}{\text{Anzahl der Elemente von } \Omega}$$

**Beispiel 3** (vgl. Beispiel 2):

Beim Würfeln mit 2 Würfeln sind alle Ergebnistupel gleich wahrscheinlich. Es ist also ein Laplace-Experiment mit

$$\Omega = \{(1, 1), \dots, (1, 6), \dots, (6, 1), \dots, (6, 6)\}.$$

Für

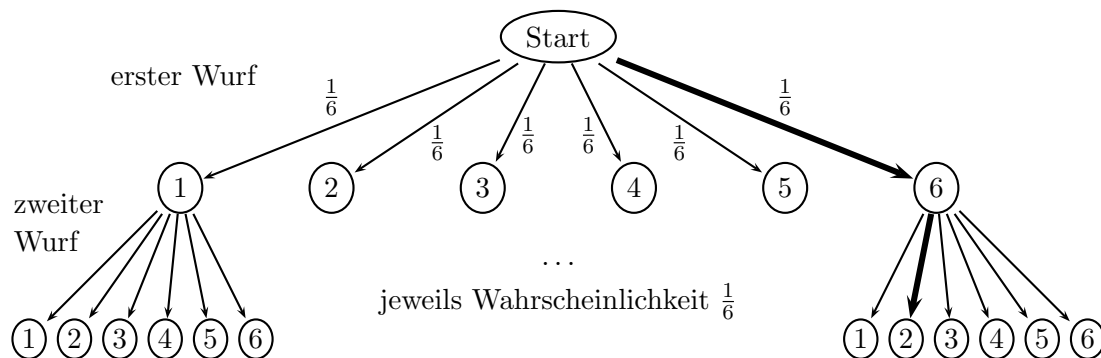
$$\begin{aligned} A &:= \{\text{Tupel} \mid \text{Summe der Augen} = 5\} \\ &= \{(1, 4), (2, 3), (3, 2), (4, 1)\} \end{aligned}$$

$$\text{ist } P(A) = \frac{4}{36} = \frac{1}{9}.$$

3. Mehrstufige Experimente beschreibt man häufig auch durch ein Baumdiagramm. Entsprechend Satz 1.2.1. ergibt sich die Wahrscheinlichkeit der Blätter als Produkt der Wahrscheinlichkeiten entlang des Pfades.

**Beispiel 4:**

Der Wurf mit zwei fairen Würfeln kann beschrieben werden durch den folgenden Baum.



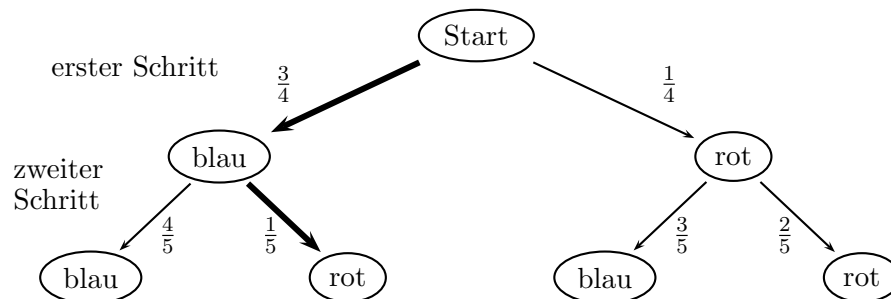
Fett markiert ist der Pfad zum Ereignis „Erster Wurf gleich 6, zweiter Wurf gleich 2“, also zu (6, 2) mit der Wahrscheinlichkeit  $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$ .

Entsprechendes gilt auch bei ungleichen Wahrscheinlichkeiten, und auch, wenn die Wahrscheinlichkeiten des zweiten Experiments vom Ergebnis des ersten Experiments abhängen. Man nennt die Wahrscheinlichkeiten in der zweiten Stufe dann auch Übergangswahrscheinlichkeiten.

### Beispiel 5:

In einer Urne befinden sich drei blaue und eine rote Kugel. Im ersten Schritt wird zufällig eine Kugel gezogen. Diese wird dann zusammen mit einer Kugel gleicher Farbe in die Urne zurückgelegt. Im zweiten Schritt wird erneut eine Kugel gezogen.

Dies führt zu folgenden Baum:



Nach der Ziehung von blau im ersten Schritt wird diese Kugel und eine neue blaue in die Urne gelegt, d.h., dort sind nun vier blaue und eine rote Kugeln. Die Übergangswahrscheinlichkeiten nach dieser Ziehung sind also  $\frac{4}{5}$  für eine blaue und  $\frac{1}{5}$  für eine rote Kugel. Die Wahrscheinlichkeit des Ereignisses „erste Kugel blau, zweite Kugel rot“ ergibt sich damit als  $\frac{3}{4} \cdot \frac{1}{5} = \frac{3}{20}$ .

Als Summe von Ereignissen ergibt sich nun beispielsweise

$$\begin{aligned}
 & P(\text{zweite Kugel rot}) \\
 = & P(\text{erste Kugel blau, zweite rot} \quad \text{oder} \quad \text{erste und zweite Kugel rot}) \\
 = & P(\text{erste Kugel blau, zweite rot}) \quad + \quad P(\text{erste und zweite Kugel rot}) \\
 = & \frac{3}{4} \cdot \frac{1}{5} \quad + \quad \frac{1}{4} \cdot \frac{2}{5} \\
 = & \frac{3}{20} \quad + \quad \frac{2}{20} \\
 = & \frac{5}{20} = \frac{1}{4}.
 \end{aligned}$$

Als Komplementärereignis ist dann

$$P(\text{zweite Kugel blau}) = 1 - P(\text{zweite Kugel rot}) = 1 - \frac{1}{4} = \frac{3}{4}.$$

### Beispiel 6:

Bei einer Glühbirnen-Produktion sei die Wahrscheinlichkeit, dass eine Glühbirne defekt ist,  $p_D = \frac{1}{15}$ . Wie groß ist die Wahrscheinlichkeit, bei 10 Birnen genau eine defekte zu haben?

Modellierung als 10-faches Bernoulli-Experiment mit

$$\begin{aligned}
 P(\text{„Birne defekt“}) &= p_D, \\
 P(\text{„Birne intakt“}) &= 1 - p_D.
 \end{aligned}$$

Die Wahrscheinlichkeit des Ereignisses „dritte Birne defekt, Rest intakt“ ist:

$$\begin{aligned}
 \text{Birne} & \qquad \qquad \qquad 1 \qquad \qquad 2 \qquad \qquad 3 \qquad \qquad 4 \qquad \qquad \qquad 10 \\
 \text{Wahrscheinlichkeit} &= (1 - p_D) \cdot (1 - p_D) \cdot p_D \cdot (1 - p_D) \cdot \dots \cdot (1 - p_D) \\
 &= p_D \cdot (1 - p_D)^9.
 \end{aligned}$$

Genauso erhält man für das Ereignis „sechste Birne defekt, Rest intakt“ die Wahrscheinlichkeit  $p_D \cdot (1 - p_D)^9$ .

Das Ereignis „genau eine von 10 Birnen defekt“ ergibt sich als Summe der disjunkten Ereignisse „ $k$ -te Birne defekt, Rest intakt“, also

$$\begin{aligned}
 & P(\text{„genau eine von 10 Birnen defekt“}) \\
 = & \sum_{k=1}^{10} P(\text{„}k\text{-te Birne defekt, Rest intakt“}) \\
 = & \sum_{k=1}^{10} p_D \cdot (1 - p_D)^9 \\
 = & 10 \cdot p_D \cdot (1 - p_D)^9 = 10 \cdot \frac{1}{15} \left(\frac{14}{15}\right)^9 \approx 0.36.
 \end{aligned}$$



**Bemerkung (Binomialverteilung):**

Beispiel 6 kann man verallgemeinern:

Betrachtet wird die  $n$ -fache Ausführung eines Bernoulli-Experiments mit den Ergebnissen  $\omega_0$  und  $\omega_1$ ,  $p = p(\omega_0)$ ,  $p(\omega_1) = 1 - p$ . Dann gilt:

$$P(\text{„genau } k\text{-mal } \omega_0\text{“}) = \binom{n}{k} \cdot p^k (1-p)^{n-k}.$$

Anzahl der Möglichkeiten  $k$ -mal  $\omega_0$  auf  $n$  Plätze zu verteilen ( $\hat{=}$  Ziehung  $k$  aus  $n$ )

Wahrscheinlichkeit einer bestimmten Kombination mit  $k$ -mal  $\omega_0$  und  $(n-k)$ -mal  $\omega_1$

Eine solche Wahrscheinlichkeitsverteilung heißt *Binomialverteilung*. Sinnvoll ist hier  $k = 0, 1, \dots, n$ . Tatsächlich ist nach dem Binomialsatz

$$\begin{aligned} \sum_{k=0}^n p_k &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \\ &= (p + (1-p))^n = 1^n = 1. \end{aligned}$$

**Bemerkungen (Geometrische Verteilung):**

1. Man kann ähnlich auch Zufallsexperimente mit abzählbar unendlich vielen Ergebnissen betrachten.

**Beispiel 7:**

Wie oft muss man würfeln um zum ersten Mal eine 6 zu erhalten? Bzw. wie groß ist die Wahrscheinlichkeit  $p_k$  nach  $k$  Würfeln zum ersten Mal eine 6 zu würfeln?

Die Wahrscheinlichkeit eine 6 zu würfeln ist  $p = \frac{1}{6}$ ; die Wahrscheinlichkeit keine 6 zu würfeln ist dann  $1 - p = \frac{5}{6}$ . Damit gilt

$$\begin{aligned} P_k &= P(\text{„nach } k \text{ Würfeln zum ersten Mal eine } 6\text{“}) \\ &= P(\text{„die ersten } k-1 \text{ Würfeln keine } 6\text{“}) \cdot P(\text{„}k\text{-ter Wurf ist eine } 6\text{“}) \\ &= (1-p)^{k-1} \cdot p. \end{aligned}$$

2. Eine solche Wahrscheinlichkeitsverteilung mit  $p_k = (1-p)^{k-1} \cdot p$  mit einem  $p \in [0, 1]$  und  $k = 1, 2, 3, \dots$  heißt *geometrische Verteilung*. Es gilt:

$$\begin{aligned}\sum_{k=1}^{\infty} p_k &= \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p \\ &= p \cdot \sum_{l=0}^{\infty} (1-p)^l \\ &= p \cdot \frac{1}{1-(1-p)} = 1.\end{aligned}$$

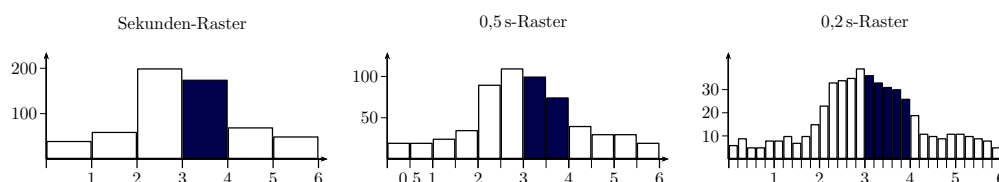
3. Manchmal betrachtet man in Situationen wie in Beispiel 7 die Zahl der Fehlversuche bis zum ersten Erfolg (an Stelle der Versuche bis zum ersten Erfolg einschließlich), also jeweils eine um 1 erniedrigte Zahl.

Man hat dann  $p_l = (1-p)^l \cdot p$  mit  $l = 0, 1, 2, \dots$

---

## 1.2. Wahrscheinlichkeitsdichte

Betrachtet man den zeitlichen Abstand zwischen Autos, die an einer bestimmten Stelle vorbei fahren, ist eine entsprechende Zuordnung (z.B. „Wahrscheinlichkeit, dass der Abstand genau 3,415 Sekunden ist“) nicht sinnvoll. Man kann allerdings Intervallen eine Wahrscheinlichkeit zuordnen, z.B. dass der Abstand zwischen 1s und 2s ist oder zwischen 2s und 3s u.s.w.. So erhält man beispielsweise folgende Histogramme:



### Definition 1.3

Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}^{\geq 0}$  mit  $\int_{-\infty}^{\infty} f(x) dx = 1$  beschreibt eine *Wahrscheinlichkeitsdichte* auf  $\Omega = \mathbb{R}$ .

Die Wahrscheinlichkeit eines Ereignisses  $A \subseteq \mathbb{R}$  ergibt sich durch

$$P(A) = \int_A f(x) dx.$$

### Bemerkungen:

1.  $f$  heißt auch *Verteilungsdichtefunktion*.
2. Bei einem Zufallsexperiment mit einer Wahrscheinlichkeitsdichte  $f$  ist die Wahrscheinlichkeit für ein einzelnes Ergebnis gleich 0:

$$P(\{x_0\}) = \int_{x_0}^{x_0} f(x) dx = 0.$$

Daher ist es bei Intervallen irrelevant, ob die Grenzen eingeschlossen sind oder nicht:

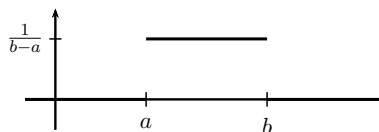
$$P([c, d]) = P(]c, d]) = P(]c, d[).$$

3. Bei einer stetigen Dichtefunktion  $f$  und kleinem  $\Delta x$  ist

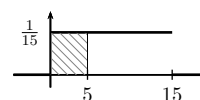
$$P([x_0, x_0 + \Delta x]) = \int_{x_0}^{x_0 + \Delta x} f(x) dx \approx \int_{x_0}^{x_0 + \Delta x} f(x_0) dx = \Delta x \cdot f(x_0).$$

**Spezielle Verteilungen:**1. Gleichverteilung auf  $[a, b]$ :

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

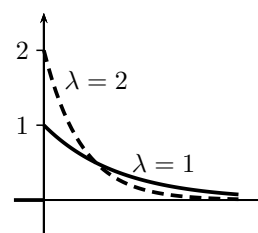
**Beispiel 1:**

Ein Bus fährt alle 15 Minuten. Wenn man nicht weiß, wann er genau fährt, und irgendwann zur Haltestelle geht, ist die Wartezeit gleichverteilt auf  $[0, 15]$  (in Minuten). Die Wahrscheinlichkeit, dass man weniger als 5 Minuten warten muss, ist  $\int_0^5 \frac{1}{15} dx = \frac{1}{3}$ .

2. Exponentialverteilung mit Parameter  $\lambda$ :

$$f(x) = \begin{cases} 0 & \text{für } x < 0, \\ \lambda e^{-\lambda x} & \text{für } x \geq 0. \end{cases}$$

Die Exponentialverteilung spielt bei Warte- oder Ausfallprozessen eine Rolle.

**Beispiel 2:**

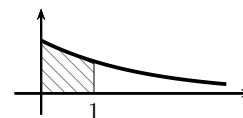
In einem Call-Center ist der zeitliche Abstand zwischen zwei Anrufen exponentialverteilt. Bei einem frequentierten Call-Center ist  $\lambda$  groß, bei wenig nachgefragtem Call-Center ist  $\lambda$  klein.

**Beispiel 3:**

Ein Zufallsexperiment sei exponentialverteilt mit  $\lambda = \frac{1}{2}$ .

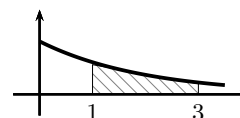
Dann ergibt sich als Wahrscheinlichkeit, dass ein Ergebnis kleiner 1 ist:

$$\begin{aligned} P([0, 1]) &= \int_0^1 \frac{1}{2} e^{-\frac{1}{2}x} dx = -e^{-\frac{1}{2}x} \Big|_0^1 \\ &= -e^{-\frac{1}{2}} + 1 \approx 0.39. \end{aligned}$$



Die Wahrscheinlichkeit für ein Ergebnis zwischen 1 und 3 ist

$$\begin{aligned} P([1, 3]) &= \int_1^3 \frac{1}{2} \cdot e^{-\frac{1}{2}x} dx = -e^{-\frac{1}{2}x} \Big|_1^3 \\ &= -e^{-1.5} + e^{-0.5} \approx 0.38. \end{aligned}$$



3. Normalverteilung mit Parameter  $\mu$  und  $\sigma^2$  :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

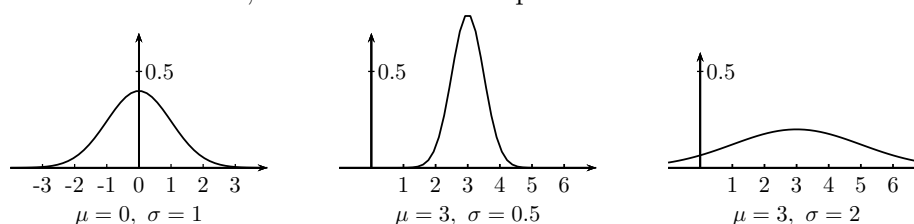
$\mu$  heißt *Mittelwert* oder *Erwartungswert*,  $\sigma$  heißt *Standardabweichung* und  $\sigma^2$  *Varianz*.

Der Spezialfall  $\mu = 0$  und  $\sigma^2 = 1$ , also

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

heißt *Standardnormalverteilung*.

Die Dichte  $f$  ist symmetrisch um  $\mu$  und hat dort ihr Maximum; ein großes  $\sigma$  entspricht einer flachen Form, ein kleines  $\sigma$  einer spitzen Form.



$\mu \pm \sigma$  sind genau die Wendestellen der Dichtefunktion.

Die Normalverteilung entsteht, wenn sich viele unabhängige Zufallsprozesse überlagern.

#### Beispiel 4:

1. Durch viele verschiedene Störeinflüssen ergeben Messungen, z.B. die Messung eines Widerstandswertes, zufällig Ergebnisse um den wahren Wert. Eine Messung wird oft interpretiert als ein normalverteiltes Zufallsexperiment mit  $\mu =$  „wahrer Wert“ und  $\sigma$  klein bei genauer und groß bei ungenauer Messung.
2. Bei der Serienproduktion, z.B von 100  $\Omega$ -Widerständen, unterscheiden sich die Produkte auf Grund von Toleranzen. Die Werte können als normalverteilt modelliert werden.

#### Bemerkung:

Durch entsprechende Umskalierung kann man Wahrscheinlichkeitswerte von beliebigen Normalverteilungen auf solche von Standardnormalverteilungen zurückführen.

#### Beispiel 5:

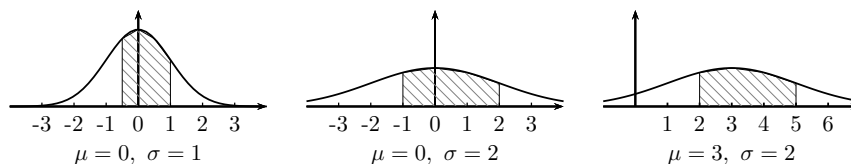
Betrachtet werden Wahrscheinlichkeitswerte zu drei Normalverteilungen:

- $P_{0,1}$  zur Standardnormalverteilung ( $\mu = 0, \sigma = 1$ ),

- $P_{0,2}$  zu einer „gestreckten“ Normalverteilung mit  $\mu = 0$  und  $\sigma = 2$ ,
- $P_{3,2}$  zu einer zusätzlich verschobenen Normalverteilung mit  $\mu = 3$  und  $\sigma = 2$ .

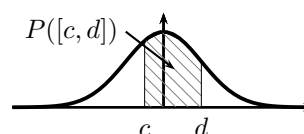
Dann gilt für entsprechende Wahrscheinlichkeitswerte beispielsweise

$$P_{0,1}([-0.5, 1]) = P_{0,2}([-1, 2]) = P_{3,2}([2, 5]).$$



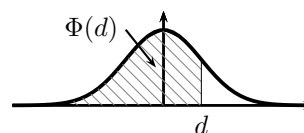
Die Wahrscheinlichkeitswerte

$$P_{0,1}([c, d]) = \int_c^d \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



zur Standardnormalverteilung sind nicht elementar integrierbar. Die Werte von

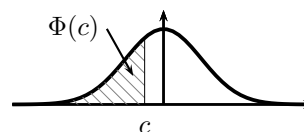
$$\Phi(x) := P_{0,1}([-\infty, x]) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



sind tabelliert (siehe Anhang B).

Damit erhält man

$$P_{0,1}([c, d]) = \Phi(d) - \Phi(c).$$



#### Satz 1.4

Bei einem mit  $(\mu, \sigma^2)$ -normalverteiltem Zufallsexperiment ist

$$P([a, b]) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right),$$

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

#### Beispiel 6:

(Vgl. Beispiel 5.) Für eine Normalverteilung mit  $\mu = 3$  und  $\sigma = 2$  ist

$$P([2, 5]) = \Phi\left(\frac{5 - 3}{2}\right) - \Phi\left(\frac{2 - 3}{2}\right) = \Phi(1) - \Phi(-0.5).$$

Entsprechend der Tabelle in Anhang B ist  $\Phi(1) \approx 0,84134$  und

$$\Phi(-0,5) = 1 - \Phi(0,5) \approx 1 - 0,69146 = 0,30854,$$

also

$$P([2, 5]) \approx 0,84134 - 0,30854 = 0,5328.$$

**Bemerkung:**

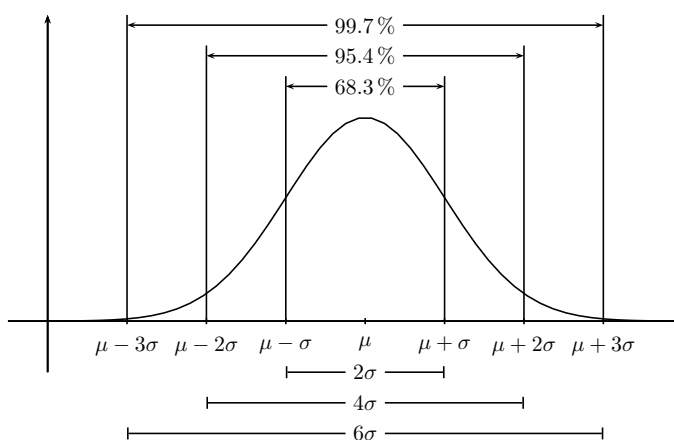
Bei einer  $(\mu, \sigma^2)$ -Normalverteilung ist die Wahrscheinlichkeit eines Ereignisses im Bereich einer  $k$ -fachen Standardabweichung um den Mittelwert

$$\begin{aligned} P([\mu - k\sigma, \mu + k\sigma]) &= \Phi\left(\frac{\mu + k\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - k\sigma - \mu}{\sigma}\right) \\ &= \Phi(k) - \Phi(-k) \\ &= \Phi(k) - (1 - \Phi(k)) \\ &= 2 \cdot \Phi(k) - 1, \end{aligned}$$

also nur von  $k$  abhängig. Im Bereich  $\pm 1\sigma$  ist die Wahrscheinlichkeit also z.B.

$$P([\mu - \sigma, \mu + \sigma]) = 2 \cdot \Phi(1) - 1 \approx 2 \cdot 0,84134 - 1 = 0,68268.$$

Für weitere Werte ergibt sich:



Genau 95 % erhält man bei  $\mu \pm 1.96\sigma$ , genau 99 % bei  $\mu \pm 2.58\sigma$ .

**Beispiel 7:**

1% -Widerstände sind solche bei denen die Fertigungstoleranzen so sind, dass maximal 5% der Widerstände um mehr als 1% vom angegebenen Widerstandswert abweichen.

Bei 1% -100  $\Omega$ -Widerständen ist also  $1.96\sigma = 1 \Omega$ .

### 1.3. Zufallsvariablen

**Definition 1.5**

Beschreibt  $X \in \mathbb{R}$  das Ergebnis eines Zufallsexperiments, so nennt man  $X$  auch *Zufallsvariable*. Ein konkretes Ergebnis nennt man *Realisierung*. Kann  $X$  nur endlich oder abzählbar viele Werte annehmen, so heißt  $X$  *diskrete Zufallsvariable*, ansonsten *stetige Zufallsvariable*.

**Beispiel 1:**

Die Summe der Augen bei einem Wurf mit 2 Würfeln ist eine diskrete Zufallsvariable.

Entsprechend der auftretenden Wahrscheinlichkeiten sagt man z.B. „ $X$  ist exponentialverteilt mit Parameter  $\lambda = 3$ “ oder „ $X$  ist auf  $M$  gleichverteilt“.

**Beispiel 2:**

Sei  $X$  standardnormalverteilt (also  $\mu = 0$  und  $\sigma = 1$ ). Dann ist  $P(X \in [-1, 1]) \approx 0.683$  (vgl. die Bemerkung Seite 12).

**Bemerkung:**

Mit Zufallsvariablen kann man auch rechnen; z.B. ist mit  $X$  auch  $2 \cdot X$  eine Zufallsvariable und mit  $X_1$  und  $X_2$  ist auch  $X_1 + X_2$  eine Zufallsvariable.

Speziell gilt:

1. Ist  $X$  standardnormalverteilt, so ist  $Y = \sigma \cdot X + \mu$  eine  $(\mu, \sigma^2)$ -normalverteilte Zufallsvariable.
2. Sind die Zufallsvariablen  $X_1, \dots, X_n$  bernoulli-verteilt, also  $P(X_i = 1) = p$  und  $P(X_i = 0) = 1 - p$ , und voneinander unabhängig (zur genauen Definition s. Abschnitt 2.1), so ist  $X = \sum_{i=1}^n X_i$  binomialverteilt.

Dies ist klar, denn das Ereignis „ $X = k$ “ kommt genau dann zustande, wenn  $k$  der  $X_i$  gleich Eins sind.

3. Sind  $X_1$  und  $X_2$  unabhängig und normalverteilt mit Erwartungswert  $\mu_1$  bzw.  $\mu_2$  und Varianz  $\sigma_1^2$  bzw.  $\sigma_2^2$ , so ist  $X_1 + X_2$  normalverteilt mit Erwartungswert  $\mu_1 + \mu_2$  und Varianz  $\sigma_1^2 + \sigma_2^2$  (s. auch Satz 2.6).

**Satz 1.6**

Sind  $X$  und  $Y$  unabhängige stetige Zufallsvariablen mit Dichten  $f$  und  $g$ , so besitzt  $W = X + Y$  die Dichte

$$h(w) = f * g(w) = \int_{-\infty}^{\infty} f(w - s)g(s) ds.$$



**Bemerkung:**

$f * g$  heißt auch Faltung von  $f$  und  $g$ . Durch die Substitution  $t = w - s$ , also  $s = w - t$  sieht man:  $f * g = g * f$ .

---

### 1.4. Verteilungsfunktion

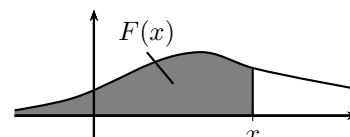
**Definition 1.7**

Zu einer Zufallsvariable  $X$  heißt  $F : \mathbb{R} \rightarrow \mathbb{R}^{\geq 0}$ ,  $F(x) = P(X \leq x)$  *Verteilungsfunktion*.

**Bemerkung:**

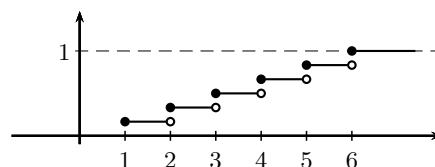
Ist  $X$  eine stetige Zufallsvariable mit der Wahrscheinlichkeitsdichte  $f$ , so ist

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{und} \quad F'(x) = f(x).$$



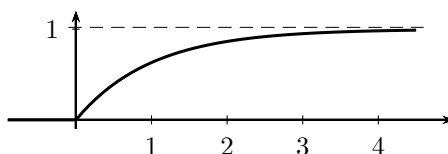
**Beispiel 1:**

1. Ist  $X$  gleichverteilt auf  $\{1, 2, 3, 4, 5, 6\}$ , so ist die Verteilungsfunktion die nebenstehende Sprungfunktion.



2. Ist  $X$  exponentialverteilt mit Parameter  $\lambda$ , so ist  $F(x) = 0$  für  $x < 0$  und für  $x \geq 0$  gilt

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = -e^{-\lambda x} - (-e^{-\lambda \cdot 0}) = 1 - e^{-\lambda x}.$$



**Bemerkung:**

Das Ereignis  $X \in ]a, b]$  kann man zerlegen in

$$(X \leq b) \quad \text{und} \quad (\text{nicht } X \leq a).$$

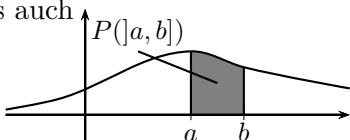
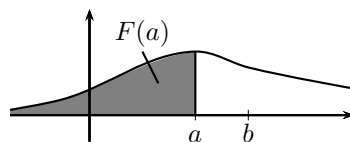
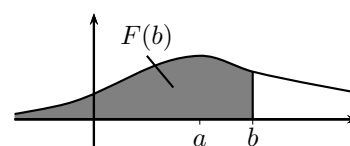
Entsprechend ist

$$\begin{aligned} P(X \in ]a, b]) &= P(X \leq b \text{ und nicht } X \leq a) \\ &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a). \end{aligned}$$

Bei stetigen Zufallsvariablen ist  $P(X = a) = 0$ , so dass auch

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

gilt.



**Satz 1.8**

Ist  $F$  die Verteilungsfunktion der Zufallsvariablen  $X$ , so gilt

$$P(X \in ]a, b]) = F(b) - F(a).$$

**Beispiel 2:**

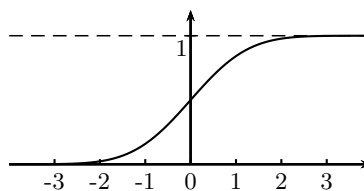
1. Sei  $X$  exponentialverteilt mit  $\lambda = \frac{1}{2}$ . Dann ist

$$\begin{aligned} P(X \in [1, 3]) &= F(3) - F(1) \\ &= (1 - e^{-\frac{3}{2}}) - (1 - e^{-\frac{1}{2}}) = e^{-\frac{1}{2}} - e^{-\frac{3}{2}} \approx 0.38. \end{aligned}$$

(Vgl. Beispiel 3 in Abschnitt 1.2.)

2. Die Verteilungsfunktion zur Standardnormalverteilung ist

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



Ist  $X$  standard-normalverteilt, so ist  $P(x \in [a, b]) = \Phi(b) - \Phi(a)$  (vgl. Satz 1.4).

**Satz 1.9**

Für eine Verteilungsfunktion  $F$  gilt  $\lim_{x \rightarrow -\infty} F(x) = 0$  und  $\lim_{x \rightarrow \infty} F(x) = 1$ .

**Definition 1.10**

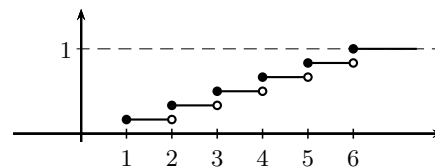
Das  $p$ -Quantil  $x_p$  ( $p \in ]0, 1[$ ) zu einer Zufallsvariablen  $X$  ist der kleinste Wert  $x$  mit  $P(X \leq x) \geq p$ .

**Bemerkungen:**

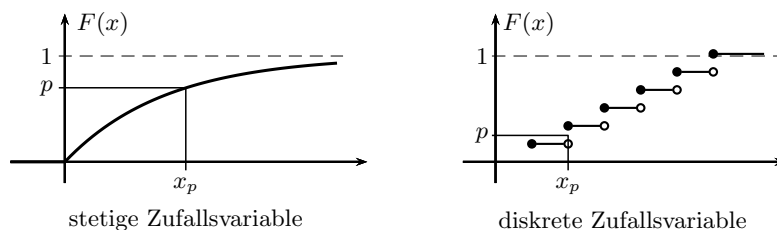
- Mit der Verteilungsfunktion  $F$  kann man genauso sagen: Das  $p$ -Quantil ist der kleinste Wert  $x$  mit  $F(x) \geq p$ .
- Bei einer stetigen Zufallsvariablen ist  $F$  stetig; jeder Wert zwischen 0 und 1 wird dann angenommen, so dass gilt: Das  $p$ -Quantil ist der kleinste Wert  $x$  mit  $F(x) = p$ . Bei einer diskreten Zufallsvariablen muss das nicht so sein.

**Beispiel 3:**

Ist  $X$  gleichverteilt auf  $\{1, 2, 3, 4, 5, 6\}$ , so ist  $x_{1/3} = x_{0.3} = 2$ , denn es ist  $F(2) = \frac{1}{3} \geq 0.3$  und für  $x < 2$  gilt  $F(x) < \frac{1}{3}$  und auch  $F(x) < 0.3$ .

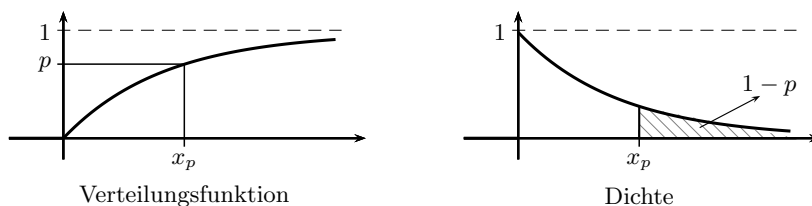


3. Die Quantil-Funktion ist quasi die Umkehrfunktion zur Verteilungsfunktion:



4. Bei einem  $p$ -Quantil  $x_p$  ist  $P(X \leq x_p) \geq p$ , also  $P(X > x_p) \leq 1 - p$ .

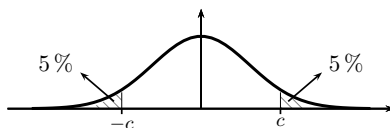
Bei einer stetigen Zufallsvariablen gilt sogar  $P(X > x_p) = 1 - p$



#### Beispiel 4:

$X$  sei standardnormalverteilt. In welchem Intervall  $[-c, c]$  liegen 90% der Werte?

$c$  ist gleich dem 95% -Quantil (auch 0.95-Quantil genannt). Mit der Tabelle in Anhang B findet man für das  $c$  mit  $\Phi(c) = 0.95$  den Wert  $c = 1.64$ .



#### Bemerkungen:

1. Die Quantile  $x_p$  zur Standardnormalverteilung kann man aus der Tabelle der Verteilungsfunktion zur Standardnormalverteilung (s. Anhang B) ablesen: Man sucht den Wert  $x_p$  mit  $\Phi(x_p) = p$ .
2. Ist  $X$  normalverteilt mit  $\mu$  und  $\sigma^2$ , so ist das  $p$ -Quantil zu  $X$  gleich  $\mu + \sigma \cdot x_p$  wobei  $x_p$  das  $p$ -Quantil zur Standardnormalverteilung ist.

### 1.5. Erzeugung von Zufallszahlen

Häufig stellen Programmiersprachen nur einen Zufallszahlengenerator zur Verfügung, der in  $[0,1]$  gleichverteilte Zahlen erzeugt. Wie kann man daraus einen Generator für andere Verteilungen erzeugen?

Sei im folgenden  $U$  eine Zufallsvariable, die auf  $[0,1]$  gleichverteilt ist.

#### Beispiel 1:

Eine diskrete Zufallsvariable  $X$  soll simuliert werden, die

mit Wahrscheinlichkeit  $\frac{1}{4}$  den Wert 1

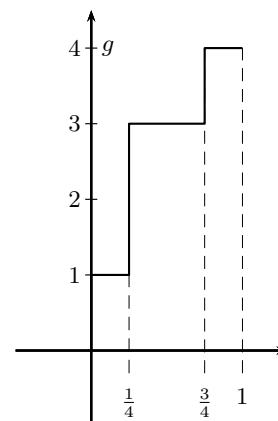
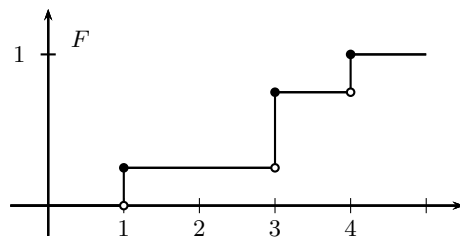
mit Wahrscheinlichkeit  $\frac{1}{2}$  den Wert 3

mit Wahrscheinlichkeit  $\frac{1}{4}$  den Wert 4

annimmt. Offensichtlich gelingt dies mit  $g(U)$  durch

$$g(x) = \begin{cases} 1, & \text{falls } x \leq \frac{1}{4}, \\ 3, & \text{falls } \frac{1}{4} < x \leq \frac{3}{4}, \\ 4, & \text{falls } x > \frac{3}{4}. \end{cases}$$

Der Graf von  $g$  entspricht der Umkehrfunktion der Verteilungsfunktion  $F$  zu  $X$ .



#### Satz 1.11

Sei  $X$  eine Zufallsvariable mit Verteilungsfunktion  $F$ . Ist  $U$  gleichverteilt auf  $[0, 1]$  und  $g$  die Umkehrfunktion zu  $F$ , so besitzt  $g(U)$  die gleiche Verteilung wie  $X$ .

#### Bemerkung:

Im strengen Sinn besitzt die Verteilungsfunktion  $F$  von  $X$  aus Beispiel 1 keine Umkehrfunktion. Es ist dann

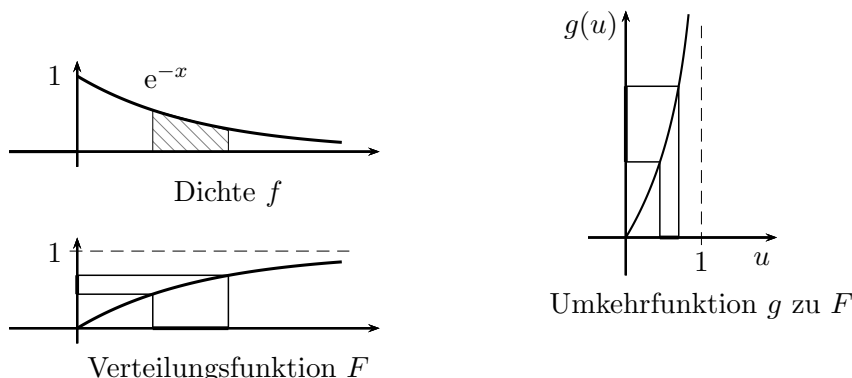
$$g(u) = \min\{x \mid F(x) \geq u\}$$

zu nehmen, was genau der Quantil-Funktion entspricht.

**Beispiel 2:**

Ist  $X$  exponentialverteilt mit  $\lambda = 1$  so ist die Verteilungsfunktion  $F(x) = 1 - e^{-x}$ .  
Für die Umkehrfunktion  $g(u)$  gilt:

$$\begin{aligned} u &= 1 - e^{-g(u)} \\ \Leftrightarrow e^{-g(u)} &= 1 - u \\ \Leftrightarrow g(u) &= -\ln(1 - u). \end{aligned}$$



Ist  $U$  gleichverteilt in  $[0, 1]$ , so ist auch  $1 - U$  gleichverteilt in  $[0, 1]$ . Daher kann man statt  $g(u) = -\ln(1 - u)$  auch  $\tilde{g}(u) = -\ln u$  nehmen.

Da die Verteilungsfunktion  $\Phi$  zur Standardnormalverteilung keine elementare Funktion ist, ist das Vorgehen aus Satz 1.11 für die Normalverteilung nicht praktikabel.

**Satz 1.12**

Sind  $U_1$  und  $U_2$  unabhängige, auf  $[0, 1]$  gleichverteilte Zufallsvariablen, so sind

$$\begin{aligned} X_1 &= \sqrt{-2 \ln U_1} \cdot \sin(2\pi \cdot U_2) \\ \text{und } X_2 &= \sqrt{-2 \ln U_1} \cdot \cos(2\pi \cdot U_2) \end{aligned}$$

(unabhängige) standardnormalverteilte Zufallsvariablen.

**Herleitung:**

Zwei unabhängige standardnormalverteilte Zufallsvariablen  $X_1, X_2$  besitzen die gemeinsame Dichte

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}} = \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}} = \frac{1}{2\pi} e^{-\frac{r^2}{2}}$$

mit  $r = \sqrt{x_1^2 + x_2^2}$ . Daraus folgt, dass die gemeinsame Verteilung rotationssymmetrisch ist; sie kann also auch dargestellt werden durch einen in  $[0, 2\pi[$  gleichverteilten

Winkel und einen Radius mit Verteilungsfunktion

$$\begin{aligned} F(r) &= P\left(\sqrt{x_1^2 + x_2^2} \leq r\right) = \int_{\tilde{K}_r} f(x_1, x_2) d(x_1, x_2) \\ &= \int_0^r \frac{1}{2\pi} e^{-\frac{r^2}{2}} \cdot 2\pi r \, dr \\ &= -e^{-\frac{r^2}{2}} \Big|_0^r \\ &= 1 - e^{-\frac{r^2}{2}}. \end{aligned}$$

Für die Umkehrfunktion  $g$  gilt

$$\begin{aligned} u &= 1 - e^{-\frac{g(u)^2}{2}} \\ \Leftrightarrow e^{-\frac{g(u)^2}{2}} &= 1 - u \\ \Leftrightarrow -\frac{g(u)^2}{2} &= \ln(1 - u) \\ \Leftrightarrow g(u) &= \sqrt{-2 \ln(1 - u)}. \end{aligned}$$

Damit lässt sich die Verteilung des Radius simulieren, wobei  $1 - U$  auch durch  $U$  ersetzt werden kann.

**Bemerkung:**

Ist  $X$  standardnormalverteilt, so ist  $\mu + \sigma \cdot X$  normalverteilt mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ .

---

## 2. Weiterführende Begriffe

### 2.1. Bedingte Wahrscheinlichkeit und Unabhängigkeit

#### Beispiel 1:

Es wird zwei Mal gewürfelt und nur verraten, dass die Augensumme größer als sieben ist. Wie groß ist die Wahrscheinlichkeit, dass der erste Wurf eine 6 ergab?

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Die Aussage „Summe  $> 7$ “ schränkt die möglichen Ereignisse auf die eingerahmten 15 Paare ein. Von denen erfüllen 5 die Eigenschaft „erster Wurf gleich 6“ (letzte Zeile), die gesuchte Wahrscheinlichkeit ist also  $\frac{5}{15} = \frac{1}{3}$ .

#### Definition 2.1

Die *bedingte Wahrscheinlichkeit*  $P(A|B)$  von  $A$  unter (der Bedingung)  $B$ , d.h. die Wahrscheinlichkeit eines Ergebnisses in  $A$  unter der Bedingung, dass es in  $B$  liegt, ist:

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

#### Beispiel 2 (Fortsetzung von Beispiel 1):

Es ist  $A = \{\text{„erster Wurf} = 6\}$  und  $B = \{\text{„Augensumme} > 7\}$ , also

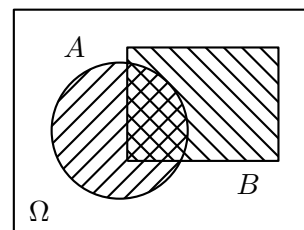
$$P(B) = \frac{15}{36},$$

$$P(A \cap B) = P(\text{„erster Wurf} = 6 \text{ und Augensumme} > 7\text{“}) = \frac{5}{36}$$

$$\Rightarrow P(A|B) = \frac{\frac{5}{36}}{\frac{15}{36}} = \frac{5}{15} = \frac{1}{3}.$$

#### Bemerkungen:

- Bei einer bedingten Wahrscheinlichkeit wird „unnormiert“: Statt des gesamten Wahrscheinlichkeitsraums  $\Omega$  wird nur noch  $B$  als Grundmenge genommen und jeweils die Schnitte  $A \cap B$  betrachtet. Die entsprechenden Wahrscheinlichkeiten müssen allerdings noch unnormiert werden, um  $P(B|B) = 1$  zu erreichen; dies geschieht durch die Division durch  $P(B)$ .





2. Es folgt

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

und analog

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

### Beispiel 3:

1. Ein Zufallsexperiment liefere die möglichen Ergebnisse 1, 2, 3, 4 und 5 jeweils mit den Wahrscheinlichkeiten

$$p_1 = \frac{1}{9}, \quad p_2 = \frac{2}{9}, \quad p_3 = \frac{1}{9}, \quad p_4 = \frac{3}{9}, \quad p_5 = \frac{2}{9}.$$

Wie ist die Verteilung unter der Bedingung, dass das Ergebnis  $\geq 3$  ist?

Die Bedingung ist

$$B = \{\text{Ergebnis} \geq 3\} = \{\text{Ergebnis} = 3, 4 \text{ oder } 5\}$$

mit  $P(B) = p_3 + p_4 + p_5 = \frac{6}{9} = \frac{2}{3}$ .

Die bedingten Verteilungswerte sind dann

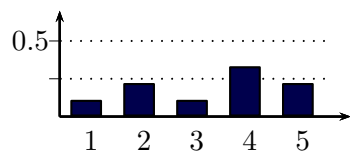
$$P(1|B) = \frac{P(\{1\} \cap \{3, 4, 5\})}{P(B)} = \frac{P(\{1\})}{P(B)} = \frac{0}{\frac{2}{3}} = 0,$$

$$P(2|B) = \frac{P(\{2\} \cap \{3, 4, 5\})}{P(B)} = \frac{P(\{2\})}{P(B)} = \frac{0}{\frac{2}{3}} = 0,$$

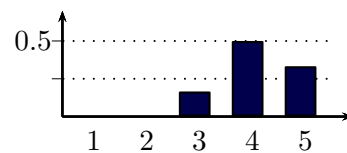
$$P(3|B) = \frac{P(\{3\} \cap \{3, 4, 5\})}{P(B)} = \frac{P(\{3\})}{P(B)} = \frac{p_3}{\frac{2}{3}} = \frac{\frac{1}{9}}{\frac{2}{3}} = \frac{1}{6},$$

$$P(4|B) = \frac{P(\{4\} \cap \{3, 4, 5\})}{P(B)} = \frac{P(\{4\})}{P(B)} = \frac{p_4}{\frac{2}{3}} = \frac{\frac{3}{9}}{\frac{2}{3}} = \frac{3}{6},$$

$$P(5|B) = \frac{P(\{5\} \cap \{3, 4, 5\})}{P(B)} = \frac{P(\{5\})}{P(B)} = \frac{p_5}{\frac{2}{3}} = \frac{\frac{2}{9}}{\frac{2}{3}} = \frac{2}{6}.$$

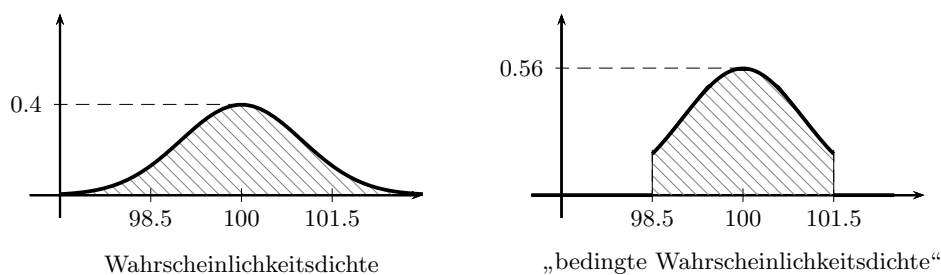


unbedingte Wahrscheinlichkeiten



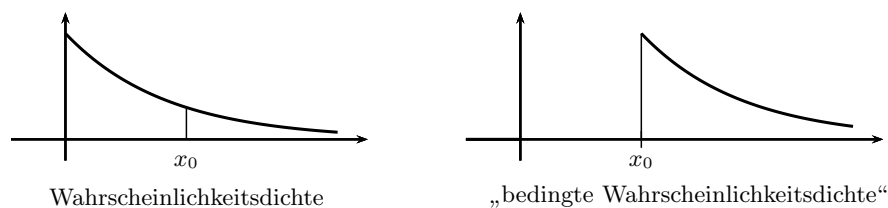
bedingte Wahrscheinlichkeiten

2. Eine Produktion liefert Widerstände mit Erwartungswert  $\mu = 100 \Omega$  und Standardabweichung  $\sigma = 1 \Omega$ . Widerstände kleiner als  $98.5 \Omega$  und größer als  $101.5 \Omega$  werden aussortiert. Die Verteilung der übrigen Widerstände entspricht dem „abgeschnittenen Berg“ der Normalverteilung mit entsprechender Umnormierung:



3. Die Wahrscheinlichkeitsdichte  $f_{x_0}$  einer Exponentialverteilung mit Parameter  $\lambda$  unter der Bedingung, dass bis zum Zeitpunkt  $x_0$  das Ereignis nicht eingetreten ist, ist gleich 0 für  $x < x_0$ ; für  $x \geq x_0$  gilt:

$$\begin{aligned} f_{x_0}(x) &= \frac{1}{P([x_0, \infty[)} \cdot f(x) = \frac{1}{\int_{x_0}^{\infty} \lambda e^{-\lambda s} ds} \cdot \lambda e^{-\lambda x} = \frac{1}{-e^{-\lambda s} \Big|_{x_0}^{\infty}} \cdot \lambda e^{-\lambda x} \\ &= \frac{1}{e^{-\lambda x_0}} \cdot \lambda e^{-\lambda x} = \lambda e^{-\lambda(x-x_0)}. \end{aligned}$$



Die bedingte Wahrscheinlichkeitsdichte ist also identisch mit der verschobenen ursprünglichen Verteilung. Die Exponentialverteilung ist *gedächtnislos*.

4. Wie groß ist die Wahrscheinlichkeit noch  $k$ -mal bis zur ersten 6 würfeln zu müssen, wenn man schon  $l$  mal vergeblich versuchte, eine 6 zu würfeln?

$$A = \{\text{„Erste 6 nach } (k+l)\text{-mal würfeln“}\},$$

$$B = \{\text{„}l\text{-mal keine 6“}\}.$$

Es ist

$$P(A \cap B) = P(A) = (1-p)^{k+l-1} \cdot p \quad \text{mit} \quad p = \frac{1}{6}$$

(geometrische Verteilung siehe Seite 6) und  $P(B) = (1-p)^l$ , also

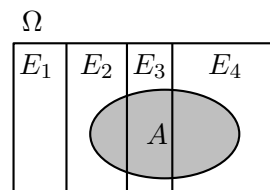
$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{(1-p)^{k+l-1} \cdot p}{(1-p)^l} = (1-p)^{k-1} \cdot p \\ &= P(\{\text{„Erste 6 nach } k\text{-mal würfeln“}\}). \end{aligned}$$

Die geometrische Verteilung ist also *gedächtnislos*.

**Satz 2.2** (Formel von der totalen Wahrscheinlichkeit)

Bilden  $E_1, \dots, E_n$  eine disjunkte Zerlegung von  $\Omega$ , so gilt

$$\begin{aligned} P(A) &= \sum_{k=1}^n P(A \cap E_k) \\ &= \sum_{k=1}^n P(A|E_k) \cdot P(E_k). \end{aligned}$$



**Beispiel 4:**

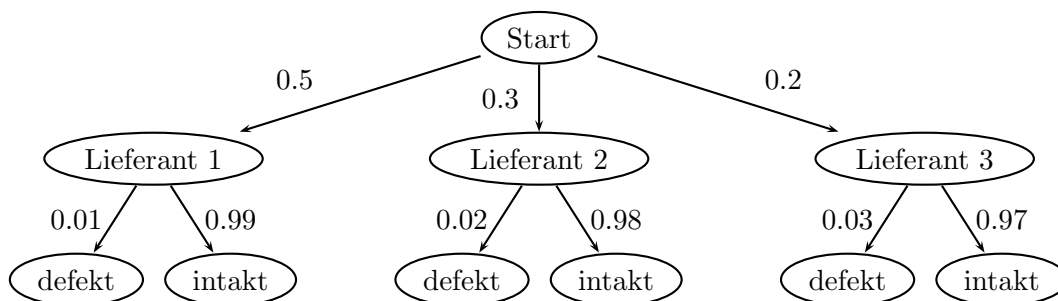
Ein Betrieb hat drei Zulieferer, die unterschiedliche Fehlerraten haben:

Lieferant	1	2	3
Anteil	50%	30%	20%
Fehlerhafte Teile	1%	2%	3%

Dann ist

$$\begin{aligned} P(\text{Teil defekt}) &= \sum_{k=1}^3 P(\text{Teil defekt} \mid \text{Teil kommt von Lieferant } k) \\ &\quad \cdot P(\text{Teil kommt von Lieferant } k) \\ &= 0.01 \cdot 0.5 + 0.02 \cdot 0.3 + 0.03 \cdot 0.2 = 0.017. \end{aligned}$$

Die Situation kann man auch mit Hilfe eines Baumdiagramms beschreiben. Die Übergangswahrscheinlichkeiten entsprechen dabei genau den bedingten Wahrscheinlichkeiten:



Man kann nun auch umgekehrt fragen: Man beobachtet ein defektes Teil. Wie wahrscheinlich ist, dass dieses von Lieferant 3 kommt?

Gefragt ist nach

$$P(\text{Lieferant 3} \mid \text{Teil defekt}) = \frac{P(\text{Lieferant 3 und Teil defekt})}{P(\text{Teil defekt})}.$$

Wegen  $P(\text{Lieferant 3 und Teil defekt}) = 0.2 \cdot 0.03 = 0.006$  ergibt sich

$$P(\text{Lieferant 3} \mid \text{Teil defekt}) = \frac{0.006}{0.017} \approx 0.353.$$

Die Üblegungen zum Ende des vorigen Beispiels führen zu folgendem Satz:

**Satz 2.3** (Formel von Bayes)

Bilden  $E_1, \dots, E_n$  eine disjunkte Zerlegung von  $\Omega$ , so gilt

$$P(E_l|A) = \frac{P(E_l \cap A)}{P(A)} = \frac{P(A|E_l) \cdot P(E_l)}{\sum_{k=1}^n P(A|E_k) \cdot P(E_k)}.$$

**Beispiel 5** (Fortsetzung von Beispiel 4):

Die Anwendung der Formel sieht man gut bei der Darstellung

$$P(\text{Lieferant 3}|\text{Teil defekt}) = \frac{0.03 \cdot 0.2}{0.01 \cdot 0.5 + 0.02 \cdot 0.3 + 0.03 \cdot 0.2}.$$

**Definition 2.4**

1. Zwei Ergebnismengen  $A$  und  $B$ , die sich gegenseitig nicht bedingen, heißen *unabhängig*, d.h. es gilt eine (und damit alle) der Eigenschaften:

- $P(A|B) = P(A)$ ,
- $P(B|A) = P(B)$ ,
- $P(A \cap B) = P(A) \cdot P(B)$ .

2. Zwei Zufallsvariablen  $X$  und  $Y$  heißen *unabhängig*  $:\Leftrightarrow$  für alle  $A, B \subseteq \mathbb{R}$  gilt

$$P(X \in A \text{ und } Y \in B) = P(X \in A) \cdot P(Y \in B).$$

**Bemerkung:**

Aus  $P(A|B) = P(A)$  folgt

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

(vgl. Satz 1.2 (1.)) und daraus

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

**Beispiel 6:**

1. (Vgl. Beispiel 2, Abschnitt 1.1). Der Wurf mit zwei Würfeln kann durch zwei unabhängige Zufallsvariablen  $X$  und  $Y$ , die jeweils einen einzelnen Wurf darstellen, modelliert werden. Dann ergibt sich beispielsweise die Wahrscheinlichkeit des Ergebnistupels (3,5) als

$$\begin{aligned} P(X = 3 \text{ und } Y = 5) &= P(X = 3) \cdot P(Y = 5) \\ &= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}. \end{aligned}$$

2. Seien  $X_1$ ,  $X_2$  und  $X_3$  drei unabhängige Bernoulli-Zufallsvariablen, die jeweils die Werte 0 und 1 mit der Wahrscheinlichkeit  $\frac{1}{2}$  annehmen. Aus diesen Zufallsvariablen werden zwei neue Zufallsvariablen gemacht:

$$Y := X_1 + X_2 \quad \text{und} \quad Z := X_1 + X_3.$$

$Y$  und  $Z$  sind nicht unabhängig, denn es ist

$$\begin{aligned} P(Y = 2) &= P(X_1 = 1 \text{ und } X_2 = 1) \\ &= P(X_1 = 1) \cdot P(X_2 = 1) \\ &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \end{aligned}$$

entsprechend auch  $P(Z = 2) = \frac{1}{4}$ , aber

$$\begin{aligned} P(Y = 2 \text{ und } Z = 2) &= P(X_1 = 1 \text{ und } X_2 = 1 \text{ und } X_3 = 1) \\ &= P(X_1 = 1) \cdot P(X_2 = 1) \cdot P(X_3 = 1) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \\ &\neq \frac{1}{16} = \frac{1}{4} \cdot \frac{1}{4} = P(Y = 2) \cdot P(Z = 2). \end{aligned}$$

### Bemerkung:

Oft treten Folgen  $(X_i)_{i \in \mathbb{N}}$  von Zufallsvariablen auf.

Sind die einzelnen Zufallsvariablen  $X_i$  identisch verteilt und unabhängig voneinander so spricht man auch von einer *iid*-Folge (independent and identically distributed).

### Beispiel 7:

1. Wiederholtes Würfeln kann man als Folge  $X_i$  von Zufallsvariablen auffassen. Dabei ist  $X_i$  unabhängig von  $X_{i-1}$ .
  2. Bei der Bitübertragung kann man ein gesendetes Bit als  $b = \pm 1$  interpretieren, das bei der Übertragung zu  $b + x$  gestört wird, wobei  $x$  normalverteilt um 0 ist. Die Störungen  $x_1, x_2, \dots$  aufeinander folgender Bits kann man als Folge von Zufallsvariablen  $X_i$  betrachten. In der Regel sind die  $X_i$  nicht unabhängig, denn wenn zum Zeitpunkt  $i$  der Kanal stark gestört wird, ist es wahrscheinlich, dass er auch zur Zeit  $i - 1$  oder  $i + 1$  gestört wird. Man nimmt oft dennoch Unabhängigkeit an (*AWGN-Kanal: additive white gaussian noise*), was in der Praxis durch Interleaving erreicht wird, d.h. eigentlich codierungsmäßig zusammengehörige Bits werden zu auseinanderliegenden Zeiten gesendet.
-

## 2.2. Erwartungswert und Varianz

### Beispiel 1:

Bei einem Gewinnspiel würfelt der Teilnehmer und erhält die Augenzahl in Euro. Dies kann man als diskrete Zufallsvariable mit möglichen Werten 1, 2, 3, 4, 5 und 6 betrachten. Wieviel erhält der Spieler im Durchschnitt? Wie ist es bei einem Würfel mit anderen aufgedruckten Zahlen?

Bei einem normalen Würfel ist der Durchschnitt offensichtlich 3,5. Bei einem mit vier Zweien und zwei Sechsen beschriebenen Würfel erhält man bei vielen Durchführungen in ca. zwei Drittel der Fälle 2€ und in einem Drittel der Fälle 6€. Bei  $N$  Durchführungen ist also

$$\begin{aligned} \text{mittlerer Gewinn} &= \frac{1}{N} \cdot \left( 2\text{€} \cdot \frac{2}{3}N + 6\text{€} \cdot \frac{1}{3}N \right) \\ &= 2\text{€} \cdot \frac{2}{3} + 6\text{€} \cdot \frac{1}{3} = \frac{10}{3}\text{€} = 3,33\dots\text{€}. \end{aligned}$$

Die Werte  $\frac{2}{3}$  bzw.  $\frac{1}{3}$  entsprechen dabei genau den Wahrscheinlichkeiten der Ergebnisse 2€ bzw. 6€.

### Definition 2.5

Zu einer Zufallsvariable  $X$  definiert man den *Erwartungswert*  $E(X)$  und die *Varianz*  $V(X)$  durch

- falls  $X$  eine diskrete Zufallsvariable mit Werten  $x_1, x_2, \dots$  und Wahrscheinlichkeiten  $p(x_1), p(x_2), \dots$  ist:

$$E(X) = \sum_{x_k} x_k \cdot p(x_k), \quad V(X) = \sum_{x_k} (x_k - E(X))^2 \cdot p(x_k).$$

- falls  $X$  eine stetige Zufallsvariable mit Dichte  $f$  ist:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx, \quad V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f(x) dx.$$

$\sqrt{V(X)}$  heißt *Standardabweichung* von  $X$ .

### Bemerkung:

Die Standardabweichung ist ein Maß für die Streuung der Werte. Die Varianz hat keine anschauliche Bedeutung. Mit ihr lässt sich aber besser rechnen.

### Beispiel 2:

1. Eine Zufallsvariable  $X$  nehme die Werte 1, 2 oder 4 an mit den Wahrscheinlichkeiten

$$P(X = 1) = 0.2, \quad P(X = 2) = 0.5 \quad \text{und} \quad P(X = 4) = 0.3.$$

Dann ist

$$E(X) = 1 \cdot 0.2 + 2 \cdot 0.5 + 4 \cdot 0.3 = 2.4,$$

$$V(X) = (1 - 2.4)^2 \cdot 0.2 + (2 - 2.4)^2 \cdot 0.5 + (4 - 2.4)^2 \cdot 0.3 = 1.24.$$

Die Standardabweichung ist dann  $\sqrt{V(X)} = \sqrt{1.24} \approx 1.11$ .

2. Zu einer Bernoulli-Zufallsvariable  $X$  mit den Werten 1, 0 und  $P(X = 1) = p$ , also  $P(X = 0) = 1 - p$ , ist

$$E(X) = 1 \cdot p + 0 \cdot (1 - p) = p,$$

$$\begin{aligned} V(X) &= (1 - p)^2 \cdot p + (0 - p)^2 \cdot (1 - p) = (1 - p) \cdot p \cdot ((1 - p) + p) \\ &= (1 - p) \cdot p. \end{aligned}$$

3. Ist  $X$  gleichverteilt auf  $[a, b]$ , so ist

$$E(X) = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{2} x^2 \Big|_a^b = \frac{1}{2} \cdot \frac{b^2 - a^2}{b-a} = \frac{1}{2}(a+b).$$

4. Ist  $X$   $(\mu, \sigma^2)$ -normalverteilt, so ist  $E(X) = \mu$  und  $V(X) = \sigma^2$ .
5. In Anhang E sind Erwartungswerte und Varianzen der verschiedenen Verteilungen aufgeführt.

**Satz 2.6**

Für Zufallsvariablen  $X, Y$  und  $a, b \in \mathbb{R}$  gilt:

1.  $V(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2.$

2.  $E(aX + b) = a \cdot E(X) + b, \quad V(aX + b) = a^2 \cdot V(X).$

3.  $E(X + Y) = E(X) + E(Y).$

4. Sind  $X$  und  $Y$  unabhängig, so gilt

4.1  $V(X + Y) = V(X) + V(Y),$

4.2  $E(X \cdot Y) = E(X) \cdot E(Y).$

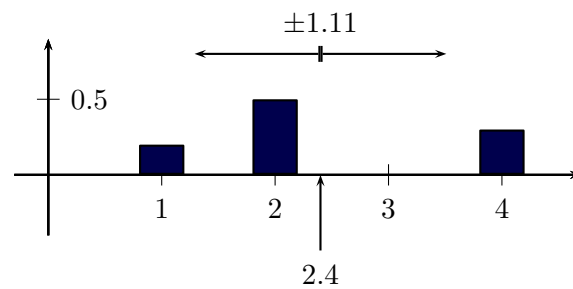
**Bemerkung:**

Der Erwartungswert gibt den Schwerpunkt an von

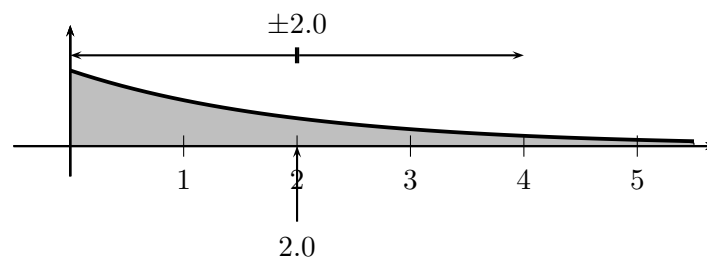
- Gewichten  $p(x_k)$ , die an den Stellen  $x_k$  positioniert sind (bei einer diskreten Zufallsvariablen),
- der Wahrscheinlichkeitsdichte (bei einer stetigen Zufallsvariablen).

**Beispiel 3:**

1. Bei den Werten aus Beispiel 2.1 ergibt sich bildlich mit Erwartungswert 2.4 und Standardabweichung 1.11:



2. Bei einer Exponentialverteilung mit  $\lambda = 0.5$  ist der Erwartungswert gleich 2.0 und auch die Standardabweichung gleich 2.0.





### 2.3. Korrelation

**Definition 2.7**

Zu zwei Zufallsvariablen  $X$  und  $Y$  heißt

$$C(X, Y) = E((X - E(X)) \cdot (Y - E(Y)))$$

Kovarianz und

$$\rho_{X,Y} = \frac{C(X, Y)}{\sqrt{V(X) \cdot V(Y)}}$$

Korrelationskoeffizient.

Die Kovarianz und der Korrelationskoeffizient beschreiben, wie eng zwei Zufallsvariablen zusammenhängen.

Sind  $X$  und  $Y$  unabhängig, so gilt wegen  $E(X - E(X)) = E(X) - E(X) = 0$ , dass

$$C(X, Y) = E((X - E(X)) \cdot (Y - E(Y))) = E(X - E(X)) \cdot E(Y - E(Y)) = 0$$

ist. Hängen  $X$  und  $Y$  derart voneinander ab, dass  $X$  tendenziell in die gleiche Richtung von  $E(X)$  abweicht, wie  $Y$  von  $E(Y)$ , so wird  $C(X, Y)$  positiv sein. Im Extremfall, dass  $X$  und  $Y$  immer dieselben Werte liefern, also  $X = Y$ , ist

$$C(X, Y) = E((X - E(X))^2) = V(X).$$

**Satz 2.8**

Es gilt

$$C(X, Y)^2 \leq V(X) \cdot V(Y) \quad \text{und} \quad \rho_{X,Y} \in [-1, 1].$$

**Beispiel 1:**

Sei  $X$  standardnormalverteilt und  $Y = a \cdot X + b \cdot Z$ , wobei  $Z$  eine von  $X$  unabhängige Standardnormalverteilung ist.

Dann ist  $E(X) = 0 = E(Y)$  und

$$\begin{aligned} C(X, Y) &= E(X \cdot (a \cdot X + b \cdot Z)) \\ &= E(a \cdot X^2 + b \cdot X \cdot Z) \\ &= a \cdot E(X^2) + b \cdot E(X \cdot Z) \\ &= a \cdot V(X) + b \cdot E(X) \cdot E(Z) \\ &= a \cdot 1 + b \cdot 0 \cdot 0 = a \end{aligned}$$

und wegen  $V(X) = 1$  und

$$\begin{aligned}V(Y) &= V(a \cdot X + b \cdot Z) \\&= V(a \cdot X) + V(b \cdot Z) \\&= a^2 \cdot V(X) + b^2 \cdot V(Z) \\&= a^2 \cdot 1 + b^2 \cdot 1 = a^2 + b^2\end{aligned}$$

ist

$$\rho_{X,Y} = \frac{a}{\sqrt{a^2 + b^2}}.$$

Beispielsweise ist

$$\text{bei } a = 0: \rho_{X,Y} = 0,$$

$$\text{bei } b = 0: \rho_{X,Y} = \frac{a}{|a|} = \begin{cases} 1, & \text{falls } a > 0, \\ -1, & \text{falls } a < 0, \end{cases}$$

$$\text{bei } a = 1, b = 1: \rho_{X,Y} = \frac{1}{\sqrt{2}} \approx 0.7,$$

$$\text{bei } a = 1, b = 2: \rho_{X,Y} = \frac{1}{\sqrt{5}} \approx 0.48.$$

**Bemerkung:**

Die Kovarianz bzw. der Korrelationskoeffizient beschreiben die Stärke des Zusammenhangs zweier Zufallsvariablen. Es gilt:

Sind  $X$  und  $Y$  unabhängig, so ist  $C(X, Y) = 0 = \rho_{X,Y}$ ,

Ist  $Y = a \cdot X$ , so ist  $\rho_{X,Y} = \begin{cases} 1, & \text{falls } a > 0, \\ -1, & \text{falls } a < 0. \end{cases}$

Allerdings folgt aus  $C(X, Y) = 0$  nicht notwendigerweise, dass  $X$  und  $Y$  unabhängig im Sinne von Definition 2.4 sind.

---

## 2.4. Zentraler Grenzwertsatz

Was passiert bei der Überlagerung vieler gleichartiger zufälliger Effekte?

Sei  $Z_n = X_1 + \dots + X_n$ , wobei die  $X_k$  unabhängige Zufallsvariablen mit einheitlichem Erwartungswert  $\mu = E(X_k)$  und einheitlicher Varianz  $\sigma^2 = V(X_k)$  sind. Dann gilt nach Satz 2.6:

$$\begin{aligned} E(Z_n) &= E(X_1 + \dots + X_n) \\ &= E(X_1) + \dots + E(X_n) \\ &= \mu + \dots + \mu \\ &= n \cdot \mu \\ \text{und } V(Z_n) &= V(X_1 + \dots + X_n) \\ &= V(X_1) + \dots + V(X_n) \\ &= \sigma^2 + \dots + \sigma^2 \\ &= n \cdot \sigma^2. \end{aligned}$$

Die verschobene Zufallsvariable  $Z_n - n\mu$  hat den Erwartungswert 0 und die Varianz  $n \cdot \sigma^2$ . Die normalisierte Zufallsvariable  $X_n^* = \frac{1}{\sqrt{n}\sigma}(Z_n - n\mu)$  hat dann den Erwartungswert 0 und die Varianz 1.

### Satz 2.9 (Zentraler Grenzwertsatz)

*Sind  $X_1, X_2, \dots$  unabhängige identisch verteilte Zufallsvariablen mit einheitlichem Erwartungswert  $\mu = E(X_k)$  und einheitlicher Varianz  $\sigma^2 = V(X_k)$ , so konvergiert die Zufallsvariable*

$$X_n^* = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}$$

*gegen die Standardnormalverteilung, d.h. für  $x \in \mathbb{R}$  gilt  $\lim_{n \rightarrow \infty} P(X_n^* \leq x) = \Phi(x)$ .*

### Bemerkungen:

1. Der Satz gilt auch unter deutlich schwächeren Bedingungen. Insbesondere brauchen die  $X_i$  nicht alle die gleiche Verteilung zu besitzen, sondern es reicht, dass große Werte gleichmäßig selten vorkommen.
2. Vereinfacht besagt der Satz: Die Summe unabhängiger zufälliger Einflüsse ist ungefähr normalverteilt.  
Dies findet beispielsweise in der Messtechnik seine Anwendung: Gemessen wird der „wahre Wert“, überlagert mit vielen zufälligen Störeinflüssen. Eine Messung kann man daher betrachten als Realisierung einer Zufallsvariablen, die normalverteilt um den „wahren Wert“ ist.

3. Mit dem arithmetischen Mittel  $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$  ist

$$X_n^* = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} = n \cdot \frac{\frac{1}{n} \sum_{k=1}^n X_k - \mu}{\sqrt{n}\sigma} = \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma}.$$

### Anwendungsbeispiel:

Die Wahrscheinlichkeit für die Produktion einer fehlerhaften LED sei  $p$ . Wie groß ist die Wahrscheinlichkeit, bei  $n$  LEDs höchstens  $k$  defekte zu haben?

Modellierung als Bernoulli-Experiment:

Jeder LED entspricht eine Zufallsvariable  $X_k$  mit Wert 1, falls sie defekt ist (Wahrscheinlichkeit  $p$ ) und Wert 0, wenn sie in Ordnung ist. Die Anzahl defekter LEDs unter  $n$  Stück ist  $Z_n = \sum_{k=1}^n X_k$ . Gesucht ist  $P(Z_n \leq k)$ . Nach Satz 2.9 gilt für große  $n$ :

$$P\left(\frac{Z_n - n\mu}{\sqrt{n}\sigma} \leq x\right) \approx \Phi(x)$$

Dabei ist  $\mu = E(X_k) = p$  und  $\sigma^2 = V(X_k) = (1-p) \cdot p$ , also  $\sigma = \sqrt{(1-p) \cdot p}$ . Es gilt

$$Z_n \leq k \Leftrightarrow \frac{Z_n - n\mu}{\sqrt{n}\sigma} \leq \frac{k - n\mu}{\sqrt{n}\sigma},$$

also

$$P(Z_n \leq k) \approx \Phi\left(\frac{k - n\mu}{\sqrt{n}\sigma}\right).$$

### Beispiel 1:

$p = 1\% = 0.01$ ,  $n = 1000$ ,  $k = 15$ :

$$P(Z_{1000} \leq 15) \approx \Phi\left(\frac{15 - 1000 \cdot 0.01}{\sqrt{1000} \cdot \sqrt{0.99 \cdot 0.01}}\right) = \Phi(1.59) = 0.9441.$$

Eine bessere Approximation erhält man, wenn die Werte  $0, 1, 2, \dots, k$  von  $Z_n$  in der stetigen Version durch die Intervalle

$$[-0.5; 0.5] \cup [0.5; 1.5] \cup \dots \cup [k - 0.5; k + 0.5] = [-0.5; k + 0.5]$$

ersetzt werden, also

$$\begin{aligned} P(0 \leq Z_n \leq k) &\approx P\left(\text{Standardnormalverteilung in } \frac{[-0.5; k + 0.5] - n\mu}{\sqrt{n}\sigma}\right) \\ &= P\left(\text{Standardnormalverteilung in } \left[\frac{-0.5 - n\mu}{\sqrt{n}\sigma}; \frac{k + 0.5 - n\mu}{\sqrt{n}\sigma}\right]\right) \\ &= \Phi\left(\frac{k + 0.5 - n\mu}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{-0.5 - n\mu}{\sqrt{n}\sigma}\right). \end{aligned}$$

**Beispiel 2:**

$p = 0.01$ ,  $n = 1000$ ,  $k = 15$  ergibt

$$\begin{aligned} P(Z_{1000} \leq 15) &\approx \Phi(1.748) - \Phi(-3.33) \\ &= 0.9598 - 0.0004 = 0.9594. \end{aligned}$$

(Der exakte Wert mit Hilfe der Binomialverteilung ist 0.9521.)

**Satz 2.10**

*Ist  $X$  ein Bernoulli-Experiment mit Erfolgswahrscheinlichkeit  $p$  und  $Z_n$  die Anzahl der Erfolge bei  $n$ -maliger Durchführung von  $X$ , so gilt für  $a, b \in \{0, 1, \dots, n\}$*

$$P(a \leq Z_n \leq b) \approx \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right).$$

**Bemerkung:**

Die Näherung ist gut für  $n$  mit  $n \cdot p(1-p) > 9$ .

### 3. Statistik

#### 3.1. Grundlagen

Oft hat man Situationen, in denen man die Parameter der zugrunde liegenden Verteilung nicht kennt. Man kann dann Beobachtungen machen.

**Definition 3.1**

Zu einer *Stichprobe* bestehend aus  $n$  Beobachtungen  $x_1, \dots, x_n \in \mathbb{R}$  heißt

1.  $\bar{x} := \frac{1}{n} \sum_{k=1}^n x_k$  *Mittelwert der Stichprobe*,
2.  $s^2 := \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$  *Varianz der Stichprobe*,
3.  $s = \sqrt{s^2}$  *Standardabweichung der Stichprobe*.

**Bemerkungen:**

1. Die Standardabweichung ist ein Maß für die Streuung der Werte. Die Varianz hat keine anschauliche Bedeutung. Mit ihr lässt sich aber besser rechnen.
2. Durch elementare Umformungen erhält man als alternative Darstellung der Varianz:

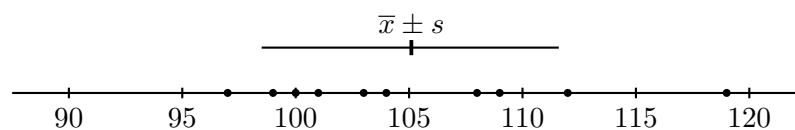
$$s^2 = \frac{1}{n-1} \sum_{k=1}^n x_k^2 - \frac{n}{n-1} (\bar{x})^2.$$

Diese Formel ist nützlich, wenn sukzessive neue Beobachtungen hinzukommen.

**Beispiel 1:**

Bei Biegeversuchen wird getestet, nach wie vielen Hin- und Herbiegungen ein Material bricht. Man erhält folgende Zahlen: 100, 118, 97, 104, 112, 99, 109, 108, 103, 101.

Der Mittelwert ist  $\bar{x} = \frac{1}{10}(100 + 118 + 97 + \dots + 101) = 105.1$ . Für die Varianz ergibt sich  $s^2 = 43.2$  und damit die Standardabweichung  $s = 6.6$ .


**Bemerkung:**

Manchmal betrachtet man auch den *Median*; das ist der Wert, so dass 50% der Stichprobenwerte unterhalb und 50% oberhalb liegen.

In Beispiel 1 liegt der Median zwischen 103 und 104.

Der Median ist unempfindlich gegenüber Ausreißern.

Kennt man die Art der Verteilung, aber nicht den Wert / die Werte der Parameter, kann man Mittelwert, Varianz oder ähnliches zur Parameterschätzung nutzen.

**Satz 3.2**

*Bei immer größer werdendem Stichprobenumfang zu einer Zufallsvariablen  $X$  gilt:*

1. *Der Stichproben-Mittelwert konvergiert in gewissem Sinn gegen  $E(X)$ .*
2. *Die Stichproben-Varianz konvergiert in gewissem Sinn gegen  $V(X)$ .*

**Bemerkung:**

Hat man eine diskrete Zufallsvariable mit den möglichen Ereignissen  $x_1, \dots, x_N$  und kennt deren wahre Wahrscheinlichkeiten  $p(x_k)$ , so berechnet sich der Erwartungswert  $E(X)$  durch  $E(X) = \sum_{k=1}^N x_k \cdot p(x_k)$ . Dabei bezeichnet  $N$  die Anzahl der möglichen Ereignisse und jedes Ereignis kommt in der Summe nur ein Mal vor.

Bei der Berechnung eines Stichproben-Mittelwerts  $\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k$  bezeichnet  $n$  die Anzahl der Stichproben. Unter den  $x_k$  können nun auch Wiederholungen auftreten.

**Beispiel 2:**

Eine Zufallsvariable  $X$  nehme die Werte 1, 2 oder 4 an mit den Wahrscheinlichkeiten  $P(X = 1) = 0.2$ ,  $P(X = 2) = 0.5$  und  $P(X = 4) = 0.3$ . Dann ist

$$E(X) = 1 \cdot 0.2 + 2 \cdot 0.5 + 4 \cdot 0.3 = 2.4.$$

Eine Stichprobe könnte beispielsweise folgende zehn Werte liefern:

1, 4, 2, 2, 1, 2, 4, 1, 2, 4.

Der Stichproben-Mittelwert ist dann

$$\bar{x} = \frac{1}{10} \cdot (1 + 4 + 2 + 2 + 1 + 2 + 4 + 1 + 2 + 4) = 2.3.$$

2. Der Faktor  $\frac{1}{n-1}$  (statt des vielleicht erwarteten  $\frac{1}{n}$ ) bei der Definition von  $s^2$  verhindert eine Unterschätzung der Varianz, denn der Mittelwert  $\bar{x}$  passt sich der Stichprobe an, so dass die bei der Definition betrachteten quadratischen Abstände im Mittel kleiner als die Abstände zum wahren Erwartungswert  $E(X)$  sind.

Bei Kenntnis von  $E(X)$  ist  $\frac{1}{n} \sum_{k=1}^n (x_k - E(X))^2$  ein Schätzer, der gegen  $V(X)$  konvergiert.

**Beispiel 3:**

1. Eine Produktionstraße soll 100  $\Omega$ -Widerstände herstellen. Die Produktion ist  $(\mu, \sigma^2)$ -normalverteilt. Durch Ziehen einer Stichprobe kann man  $\mu \approx \bar{x}$  und  $\sigma^2 \approx s^2$  schätzen.
2. Bei der Serienfertigung von LEDs kann man die Frage, wie wahrscheinlich eine produzierte LED defekt ist, als Frage nach der Wahrscheinlichkeit  $p$  eines Bernoulli-Experiments mit  $P(\text{„LED defekt“}) = p$  und  $P(\text{„LED intakt“}) = 1 - p$  auffassen. Bei Ziehen einer Stichprobe ist

$$p \approx \frac{\text{Anzahl defekter LEDs}}{\text{Gesamtzahl}} \quad (\text{relative H\u00e4ufigkeit}).$$

---



## 3.2. Schätzer

### Situation:

Man kennt die Verteilungsart einer Zufallsvariablen  $X$ , nicht aber den konkreten Parameter  $\vartheta$  („theta“).

### Ziel:

Aufgrund von  $n$  Beobachtungen  $x_1, \dots, x_n$  schätzt man den Parameter durch eine Funktion  $T(x_1, \dots, x_n)$ .

Da die  $x_k$  Realisierungen von Zufallsvariablen  $X_k$  sind, ist  $T$  selbst eine Zufallsvariable:  $T = T(X_1, \dots, X_n)$ .

### Beispiel 1:

1. Der Mittelwert einer Stichprobe,  $T(x_1, \dots, x_n) = \frac{1}{n} \cdot \sum_{k=1}^n x_k$ , ist ein Schätzer für den Erwartungswert  $E(X)$ .
2. Ist  $X$  ein Bernoulli-Experiment mit Erfolgswahrscheinlichkeit  $p$ , also

$$P(X = 1) = p, \quad P(X = 0) = 1 - p,$$

so ist  $E(X) = p$ , also  $T$  aus 1. ein Schätzer für den Parameter  $p$ .

### Beispiel 2:

Die 8-fache Wiederholung eines Bernoulli-Experiments ergab

$$(x_1, \dots, x_n) = (0, 1, 0, 0, 1, 0, 1, 0),$$

also 3 Erfolge und 5 Misserfolge. Damit ist

$$T(0, 1, 0, 0, 1, 0, 1, 0) = \frac{1}{8} \cdot 3 = \frac{3}{8}$$

ein Schätzer für  $p$ .

Formal kann man die Stichproben als unabhängige Zufallsvariablen  $X_k$  und  $T = T(X_1, \dots, X_n)$  als Zufallsvariable betrachten. Die obige Realisierung ergab dann  $\frac{3}{8}$  als Realisierung von  $T$ .

### Bemerkung:

Der Parameter  $\vartheta$  ist (durch die Wirklichkeit) festgelegt, die Ergebnisse des Schätzers sind Zufallswerte.

---

**Definition 3.3**

Ein Schätzer  $T$  für einen Parameter  $\vartheta$  heißt *erwartungstreu*

$:\Leftrightarrow$  der Erwartungswert von  $T$  ist  $\vartheta$ .

**Bemerkung:**

1. Erwartungstreue bedeutet anschaulich, dass bei wiederholten Stichproben der Mittelwert der Schätzwerte gegen den entsprechenden Parameter konvergiert.
2. Sind  $X_1, X_2, \dots, X_n$  unabhängig und verteilt wie  $X$  so gilt für den Erwartungswert des Schätzers  $T(X_1, \dots, X_n) = \frac{1}{n} \cdot \sum_{k=1}^n X_k$ :

$$\begin{aligned} E(T(X_1, \dots, X_n)) &= E\left(\frac{1}{n} \sum_{k=1}^n X_k\right) \\ &\stackrel{\text{Satz 2.6}}{=} \frac{1}{n} \cdot \sum_{k=1}^n E(X_k) \\ &= \frac{1}{n} \cdot \sum_{k=1}^n E(X) \\ &= \frac{1}{n} \cdot (n \cdot E(X)) \\ &= E(X), \end{aligned}$$

d.h.  $T$  ist ein erwartungstreuer Schätzer für  $E(X)$ .

3. Man kann zeigen, dass die Schätzer

$$T_1(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$$

und

$$T_2(x_1, \dots, x_n) = \frac{1}{n} \sum_{k=1}^n (x_k - E(X))^2$$

erwartungstreu für  $V(X)$  sind.

**Beispiel 3:**

1. Die Busse der ASEAG sind durchnummeriert. Sie fahren drei Mal und nutzen dabei die Busse Nr. 37, 78 und 56. Wieviel Busse hat die ASEAG?

Das Beispiel kann folgendermaßen modelliert werden: Man zieht  $n$  mal aus einer Urne mit von 1 bis  $N$  durchnummerierten Kugeln. Auf Grundlage der Beobachtungen  $x_1, \dots, x_n$  soll ein Schätzer für den unbekannt Parameter  $N$  gebildet werden. (Dieses Problem ist auch als *Taxiproblem* bekannt.)

Man kann zeigen, dass, falls ohne Zurücklegen gezogen wird, der Schätzer

$$T(x_1, \dots, x_n) = \frac{n+1}{n} \cdot \max\{x_1, \dots, x_n\} - 1$$

ein erwartungstreuer Schätzer für  $N$  ist.

Bei den drei Beobachtungen 37, 78, 56 erhält man damit als Schätzwert

$$T(37, 78, 56) = \frac{4}{3} \cdot 78 - 1 = 103.$$

2. Ist  $X$  exponentialverteilt mit Parameter  $\lambda$ , so ist

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

Es ist daher plausibel,  $\frac{1}{T}$  mit  $T(x_1, \dots, x_n) = \frac{1}{n} \cdot \sum_{k=1}^n x_k$  als Schätzer für  $\lambda$  zu benutzen. Dieser Schätzer ist allerdings nicht erwartungstreu!

Das folgende Beispiel soll eine Möglichkeit verdeutlichen, einen Schätzer zu konstruieren:

#### Beispiel 4:

Bei den  $n$ -fachen Durchführung eines Bernoulli-Experiments mit unbekanntem  $p$  wurden  $k$  Erfolge verzeichnet. Nach der Binomialverteilung ist

$$P(k \text{ Erfolge bei } n \text{ Durchführungen}) = \binom{n}{k} \cdot p^k (1-p)^{n-k}. \quad (*)$$

Man kann als Schätzwert das  $p$  nehmen, bei dem dieser Ausdruck maximal ist, die Beobachtung also maximale Wahrscheinlichkeit (engl: „Maximum Likelihood“) hat. Statt (\*) zu maximieren, bietet es sich an,

$$\begin{aligned} f(p) &= \ln(p^k \cdot (1-p)^{n-k}) \\ &= k \cdot \ln p + (n-k) \ln(1-p) \end{aligned}$$

zu maximieren:

$$\begin{aligned} f'(p) &= k \cdot \frac{1}{p} + (n-k) \cdot \frac{-1}{1-p} \stackrel{!}{=} 0 \\ \Leftrightarrow \frac{k}{p} &= \frac{n-k}{1-p} \\ \Leftrightarrow (1-p) \cdot k &= (n-k) \cdot p \\ \Leftrightarrow p &= \frac{k}{n}. \end{aligned}$$

Dies ist tatsächlich die Maximalstelle.

**Definition 3.4**

Ein Schätzer, der den Parameterwert ergibt, unter dem die Beobachtung die größte Wahrscheinlichkeit hat (bei stetigen Zufallsvariablen, die größte Wahrscheinlichkeitsdichte), heißt *Maximum-Likelihood Schätzer*.

**Bemerkungen:**

1. Bei mehreren unabhängigen Beobachtungen ergibt sich die Wahrscheinlichkeit der Gesamtbeobachtungen als Produkt der einzelnen Wahrscheinlichkeiten.

Bei stetigen Zufallsvariablen nimmt man statt der Wahrscheinlichkeit den Dichtewert bzw. bei mehreren Beobachtungen das Produkt der Dichtewerte.

2. Die Funktion, die die Wahrscheinlichkeit(sdichte) einer Beobachtung in Abhängigkeit des Parameters angibt, heißt *Likelihood-Funktion*.

**Beispiel 5:**

1. (Fortsetzung von Beispiel 3.1)

Beim Taxiproblem mit einer Gesamtzahl  $N$  ist die Wahrscheinlichkeit einer Beobachtung von  $x_1, \dots, x_n$  gleich Null, falls das Maximum der  $x_k$  größer als  $N$  ist, ansonsten gleich  $(\frac{1}{N})^n$  (bei mit Zurücklegen), da alle Beobachtungen gleiche Wahrscheinlichkeit besitzen.

Die Likelihood-Funktion  $f(N)$  zu einer Beobachtung  $x_1, \dots, x_n$  ist also

$$f(N) = \begin{cases} 0, & \text{falls } \max\{x_1, \dots, x_n\} > N, \\ (\frac{1}{N})^n, & \text{sonst.} \end{cases}$$

Diese Funktion wird maximal, wenn  $N$  minimal im zweiten Fall ist, also für  $N = \max\{x_1, \dots, x_n\}$ . Damit wird allerdings die Gesamtzahl offensichtlich unterschätzt.

2. Bei normalverteiltem  $X$  wurde  $x_1, \dots, x_n$  beobachtet. Bei Parametern  $\mu$  und  $\sigma^2$  ergeben sich als einzelne Dichtewerte

$$f_i = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}.$$

Die Likelihood-Funktion ist das Produkt

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

Eine Maximierung dieser Funktion bzgl.  $\mu$  ist äquivalent zur Minimierung von  $g(\mu) = \sum_{i=1}^n (x_i - \mu)^2$ . Es ist

$$\begin{aligned} g'(\mu) &= \sum_{i=1}^n -2(x_i - \mu) \stackrel{!}{=} 0 \\ \Leftrightarrow \sum_{i=1}^n x_i &= n \cdot \mu \\ \Leftrightarrow \mu &= \frac{1}{n} \sum_{i=1}^n x_i. \end{aligned}$$

Also ist  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  Maximum-Likelihood-Schätzer für  $\mu$ .

---

### 3.3. Konfidenzbereiche

#### 3.3.1. Grundlagen

**Situation:**

Im vorigen Abschnitt wurden sogenannte Punktschätzer betrachtet, z.B. war bei einer  $n$ -fachen Realisierung eines Bernoulli-Experiments, bei der  $k$  Erfolge beobachtet wurden, der Wert  $t = \frac{k}{n}$  ein Schätzer für die Erfolgswahrscheinlichkeit  $p$ .

**Beispiel 1:**

Bei einer Stichprobe von  $n = 100$  Teilen sind  $k = 10$  Teile defekt. Damit ist die Wahrscheinlichkeit defekter Teile ungefähr 0.1.

Der wahre Wert des Parameters wird aber von dem geschätzten Wert abweichen. Kann man für den Parameter verlässliche Grenzen angeben?

Statt *eines* Wertes soll nun ein *Intervall* geschätzt werden, das den wahren Parameter mit einer vorgegebenen Wahrscheinlichkeit bzw. bis auf eine Irrtumswahrscheinlichkeit  $\alpha$  überdeckt.

**Definition 3.5**

$X$  besitze eine Verteilung mit einem unbekanntem Parameter  $\vartheta \in \mathbb{R}$ .

Ein *Konfidenzbereich*  $I$  mit Irrtumswahrscheinlichkeit  $\alpha$  zu einer Stichprobe  $x = (x_1, \dots, x_n)$  ist ein Bereich  $I(x) = I(x_1, \dots, x_n) \subset \mathbb{R}$ , so dass für jedes  $\vartheta$  gilt

$$P_{\vartheta}(\vartheta \in I(x)) \geq 1 - \alpha \quad (\alpha \text{ klein}).$$

$1 - \alpha$  heißt *Konfidenzniveau*.

**Bemerkung:**

1. Der Parameter  $\vartheta$  hängt nicht vom Zufall ab, sondern der Bereich  $I(x_1, \dots, x_n)$ .

Das wesentliche beim Konfidenzbereich ist die festgelegte Funktionsvorschrift  $I(\dots)$ . Die Stichprobenergebnisse  $x_i$  hängen vom Zufall ab, so dass der resultierende Bereich  $I(x_1, \dots, x_n)$  vom Zufall abhängt.

Dieser zufällige Bereich soll – egal welchen Wert  $\vartheta$  hat – den wahren Parameter in den meisten Fällen (genauer: mit Wahrscheinlichkeit  $1 - \alpha$ ) überdecken.

Die Schreibweise  $P_{\vartheta}(\dots)$  verdeutlicht dies; gemeint ist die Wahrscheinlichkeit, die sich beim Parameterwert  $\vartheta$  ergibt.

2. Übliche Werte für die Irrtumswahrscheinlichkeit sind  $\alpha = 0.1$ ,  $\alpha = 0.05$  oder  $\alpha = 0.01$ .
-

**Beispiel 2:**

Ein Zufallsexperiment liefert gleichverteilt Werte im Intervall  $[\mu - 1; \mu + 1]$ , wobei der Parameter  $\mu$  unbekannt ist.

Welche der folgenden Aussagen beschreibt einen Konfidenzbereich zum Konfidenzniveau 90% für den (wahren) Wert  $\mu$  in Abhängigkeit von einer Stichprobe  $x$ ?

- a) Falls  $x < 0$ :  $\mu \leq 0.8$ , falls  $x \geq 0$ :  $\mu \geq -0.8$ .  
 b) Falls  $x < 0$ :  $\mu$  liegt *nicht* in  $]0; 0.2[$ , falls  $x \geq 0$ :  $\mu$  liegt *nicht* in  $] - 0.2; 0[$ .

Man muss jeweils für jedes mögliche  $\mu$  testen, ob die Aussage mit Wahrscheinlichkeit 90% stimmt.

- a) Wenn  $\mu < -1$  ist, ist  $x$  garantiert kleiner 0, und die Aussage stimmt.

Wenn  $\mu \in [-1; -0.8[$  ist, ist die Wahrscheinlichkeit, dass  $x \geq 0$  ist, kleiner als 10%, also zu 90%  $x < 0$ , und die Aussage stimmt.

Wenn  $\mu \in [-0.8; 0.8]$  ist, stimmt die Aussage in jedem Fall, egal, ob  $x$  positiv oder negativ ist.

Die verbleibenden Fälle  $\mu \in ]0.8; 1]$  und  $\mu > 1$  sind entsprechend.

Die Aussagen beschreibt also einen Konfidenzbereich zum Konfidenzniveau 90%.

- b) Wenn  $\mu = 0.1$  ist, ist  $x$  mit Wahrscheinlichkeit 45% negativ. Die Aussage stimmt dann also zu 45% nicht, d.h., es liegt kein Konfidenzbereich zum Konfidenzniveau 90% vor.

**3.3.2. Konfidenzbereiche für Erwartungswerte bei Normalverteilungen**

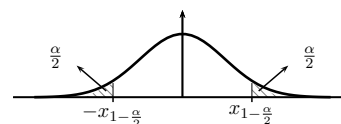
Sei  $X_1, \dots, X_n$  normalverteilt mit Erwartungswert  $\mu$  und Standardabweichung  $\sigma$ .

Dann ist

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$$

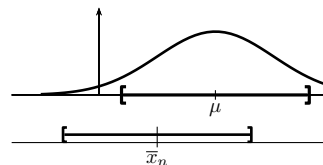
normalverteilt mit Erwartungswert  $\mu$  und Standardabweichung  $\frac{1}{\sqrt{n}}\sigma$ .

Betrachtet man nun zu einer vorgegebenen Irrtumswahrscheinlichkeit  $\alpha$  das  $(1 - \frac{\alpha}{2})$ -Quantil  $c = x_{1 - \frac{\alpha}{2}}$ , so



gilt (unabhängig vom wahren  $\mu$ ):

$$\begin{aligned} 1 - \alpha &= P\left(\bar{X}_n \in \left[\mu - \frac{\sigma}{\sqrt{n}} \cdot c; \mu + \frac{\sigma}{\sqrt{n}} \cdot c\right]\right) \\ &= P\left(\mu \in \left[\bar{X}_n - \frac{\sigma}{\sqrt{n}} \cdot c; \bar{X}_n + \frac{\sigma}{\sqrt{n}} \cdot c\right]\right). \end{aligned}$$



### Satz 3.6

Ist  $X$  normalverteilt mit Standardabweichung  $\sigma$ , so erhält man für den Erwartungswert  $\mu$  einen Konfidenzbereich zum Konfidenzniveau  $1 - \alpha$  zu  $n$  Beobachtungen  $x_1, \dots, x_n$  durch

$$\left[\bar{x}_n - \frac{\sigma}{\sqrt{n}} \cdot c; \bar{x}_n + \frac{\sigma}{\sqrt{n}} \cdot c\right]$$

mit  $\bar{x}_n = \frac{1}{n} \sum_{k=1}^n x_k$  und  $c = x_{1-\frac{\alpha}{2}}$ , dem  $1 - \frac{\alpha}{2}$ -Quantil zur Standardnormalverteilung.

### Beispiel 1:

Ein Merkmal ist normalverteilt mit  $\sigma = 1$ . Eine Stichprobe aus 10 Beobachtungen ergibt den Mittelwert  $\bar{x} = 12$ .

Zu  $\alpha = 0.1$  ist  $x_{1-\frac{\alpha}{2}} = x_{0.95} = 1.64$ . Der Konfidenzbereich für  $\mu$  ist also

$$\left[12 - \frac{1}{\sqrt{10}} \cdot 1.64; 12 + \frac{1}{\sqrt{10}} \cdot 1.64\right] = [11.48; 12.52].$$

Oft ist es nicht realistisch, dass man die Standardabweichung  $\sigma$  kennt. Statt dessen nutzt man den Schätzwert

$$S_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2}.$$

Satz 3.6 beruht im Prinzip darauf, dass die normalisierte Zufallsvariable  $\frac{\bar{X}_n - \mu}{\frac{1}{\sqrt{n}} \sigma}$  standardnormalverteilt ist. Ersetzt man nun  $\sigma$  durch den Schätzwert  $S_n$ , erhält man die Zufallsvariable

$$t_{n-1} = \frac{\bar{X}_n - \mu}{\frac{1}{\sqrt{n}} S_n},$$

wobei allgemein  $t_m$  die  $t$ -Verteilung (auch Student-Verteilung) mit  $m$  Freiheitsgraden ist.



Mit den entsprechenden Quantilen  $t_{m,p}$  zur  $t$ -Verteilung erhält man als Konfidenzbereich

$$\left[ \bar{X}_n - \frac{S_n}{\sqrt{n}} \cdot c; \bar{X}_n + \frac{S_n}{\sqrt{n}} \cdot c \right], \quad c = t_{n-1, 1-\frac{\alpha}{2}}.$$

**Satz 3.7**

Ist  $X$  normalverteilt mit unbekannter Standardabweichung, so erhält man für den Erwartungswert  $\mu$  einen Konfidenzbereich zum Konfidenzniveau  $1 - \alpha$  zu  $n$  Beobachtungen  $x_1, \dots, x_n$  durch

$$\left[ \bar{x}_n - \frac{s_n}{\sqrt{n}} \cdot c; \bar{x}_n + \frac{s_n}{\sqrt{n}} \cdot c \right]$$

mit  $\bar{x}_n = \frac{1}{n} \sum_{k=1}^n x_k$ ,  $s_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x}_n)^2}$  und  $c = t_{n-1, 1-\frac{\alpha}{2}}$ , dem  $1 - \frac{\alpha}{2}$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden.

**Beispiel 2:**

Ist die Standardabweichung  $s = 1$  aus Beispiel 1 nur geschätzt, ergibt sich mit  $t_{n-1, 1-\frac{\alpha}{2}} = t_{9, 0.95} = 1.833$  als Konfidenzbereich

$$\left[ 12 - \frac{1}{\sqrt{10}} \cdot 1.833; 12 + \frac{1}{\sqrt{10}} \cdot 1.833 \right] = [11.42; 12.58],$$

ein wegen der Unsicherheit der Schätzung der Standardabweichung größerer Bereich als bei Beispiel 1.

**Bemerkung:**

Man kann auch an nur *einseitigen Konfidenzbereichen* interessiert sein, z.B. „der Erwartungswert  $\mu$  ist mindestens 15“. Ähnlich wie oben kann man dann einseitige Begrenzungen analysieren, wobei man dann jeweils das  $1 - \frac{\alpha}{2}$ -Quantil durch das  $1 - \alpha$ -Quantil ersetzt.

**Beispiel 3:**

Will man in Beispiel 1 die Lage von  $\mu$  mit  $\alpha = 0.1$  nur einseitig abschätzen, so erhält man mit  $x_{1-\alpha} = x_{0.9} = 1.28$  einen einseitigen Konfidenzbereich für  $\mu$  durch

$$\mu \leq 12 + \frac{1}{\sqrt{10}} \cdot 1.28 = 12.40.$$

Ist die Standardabweichung  $s = 1$  nur geschätzt, ergibt sich mit  $t_{n-1, 1-\alpha} = t_{9, 0.9} = 1.383$  als Konfidenzbereich

$$\mu \leq 12 + \frac{1}{\sqrt{10}} \cdot 1.383 = 12.437.$$

### 3.3.3. Konfidenzbereiche für Bernoulli-Experimente

Man kann ähnliche Überlegungen auch zur Bestimmung von Konfidenzbereichen zu Bernoulli-Experimenten nutzen:

Sind  $X_1, X_2, \dots$  unabhängig und Bernoulli-verteilt mit Erfolgswahrscheinlichkeit  $p$  (also  $E(X) = p$  und  $V(X) = p(1-p)$ ), so ist nach dem zentralen Grenzwertsatz (Satz 2.9)

$$X_n^* := \frac{\sum_{i=1}^n X_i - n \cdot p}{\sqrt{n \cdot p(1-p)}} = \frac{k - np}{\sqrt{np(1-p)}}$$

(mit  $k$ =Anzahl der Erfolge) für große  $n$  ungefähr standardnormalverteilt, also

$$P\left(-x_{1-\frac{\alpha}{2}} \leq X_n^* \leq x_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha.$$

Löst man hier den Bereich nach  $p$  auf, erhält man einen Konfidenzbereich für  $p$ :

#### Satz 3.8

Einen (approximativen) Konfidenzbereich zum Konfidenzniveau  $1 - \alpha$  für die Erfolgswahrscheinlichkeit  $p$  bei einem Bernoulli-Experiment erhält man bei  $n$  Versuchen mit  $k$  Erfolgen durch  $p \in [p_1, p_2]$  mit

$$p_{1,2} = \frac{1}{c^2 + n} \cdot \left( k + \frac{c^2}{2} \pm c \cdot \sqrt{\frac{k(n-k)}{n} + \frac{c^2}{4}} \right)$$

mit  $c = x_{1-\frac{\alpha}{2}}$ , dem  $1 - \frac{\alpha}{2}$ -Quantil zur Standardnormalverteilung.

(Zur Anwendung der Formel sollte  $n \geq 20$  und  $\frac{k}{n}$  nicht zu nahe bei 0 oder 1 sein.)

#### Bemerkung:

Für große  $n$  und  $k$  ( $k \geq 50$  und  $n - k \geq 50$ ) kann man die Grenzen weiter approximieren durch

$$p_{1,2} = \frac{1}{n} \cdot \left( k \pm c \cdot \sqrt{\frac{k(n-k)}{n}} \right) = \frac{k}{n} \pm \frac{c}{\sqrt{n}} \cdot \sqrt{\frac{k}{n} \cdot \left(1 - \frac{k}{n}\right)}.$$

#### Beispiel 1:

Eine Wahlumfrage ergibt eine Prognose von 40% für Partei A. Befragt wurden 1000 Wähler. Wie groß ist der Konfidenzbereich zu einem Niveau  $1 - \alpha = 0.95$ ?

Mit  $x_{1-\frac{\alpha}{2}} = x_{0.975} = 1.96$ ,  $n = 1000$  und  $k = 400$  ergibt sich als Konfidenzbereich  $[0.370; 0.431]$ .

#### Bemerkung:

Man kann auch hier an nur *einseitigen Konfidenzbereichen* interessiert sein, z.B. „die Fehlerrate  $p$  ist kleiner 1%“. Ähnlich wie oben kann man dann einseitige Begrenzungen mit dem  $1 - \alpha$ -Quantil  $x_{1-\alpha}$  nutzen.

### 3.4. Hypothesentests

#### 3.4.1. Grundlagen

##### Beispiel 1:

Das Nachwiegen von 10 Schokoladentafeln, die eigentlich 100g wiegen sollen, ergibt Gewichte (in Gramm) von

99.2, 99.8, 100.2, 98.9, 99.9, 99.8, 100.4, 99.8, 100.0 und 99.1

mit dem Mittelwert  $\bar{x} = 99.71\text{g}$ .

Kann man dem Hersteller Betrug vorwerfen, d.h., spricht die Beobachtung signifikant gegen die Hypothese  $\mu = 100\text{g}$ ?

Man kann die Hypothese ablehnen, wenn das Experiment ein sehr unwahrscheinliches Ergebnis liefert.

##### Definition 3.9

Bei einem *Hypothesentest* zur Überprüfung einer *Hypothese*  $H_0$  zu einem *Signifikanzniveau*  $1 - \alpha$  ( $\alpha$  klein) legt man einen *Akzeptanzbereich*  $A$  so fest, dass, falls die Hypothese gilt, für die zugrunde liegende Zufallsvariable  $X$  gilt:

$$P(X \in A) \geq 1 - \alpha.$$

Bei dem (unter der Hypothese unwahrscheinlichen) Fall, dass für eine Realisierung  $x \notin A$  gilt, lehnt man die Hypothese ab und entscheidet sich für die *Alternative*.

##### Beispiel 2 (Fortsetzung von Beispiel 1):

Im Beispiel 1 ist die Hypothese  $\mu = 100\text{g}$ , die Alternative also  $\mu \neq 100\text{g}$ , und die betrachtete Zufallsvariable der Mittelwert  $\bar{X}_{10}$  über 10 Beobachtungen.

Bei vorgegebenem Signifikanzniveau  $1 - \alpha$  kann die Hypothese abgelehnt werden, wenn  $\bar{X}_{10}$  einen Wert in den unwahrscheinlichen Randbereichen annimmt.



##### Bemerkungen:

1. Mit einem Hypothesentest kann man nicht eine Hypothese „beweisen“, sondern nur eine Hypothese signifikant widerlegen.

Will man also eine Aussage signifikant beweisen, muss man übergehen zum Gegenteil und dies als Hypothese formulieren.

2. Folgende Szenarien sind möglich:

	Entscheidung	
	Hypothese beibehalten	Hypothese verwerfen
Hypothese gilt	OK	Fehler 1.Art ( $\alpha$ -Fehler)
Hypothese gilt nicht	Fehler 2.Art ( $\beta$ -Fehler)	OK

Die Wahrscheinlichkeit für einen Fehler erster Art (Hypothese fälschlich verwerfen) steuert man mit dem Signifikanzniveau; er ist höchstens  $\alpha$ .

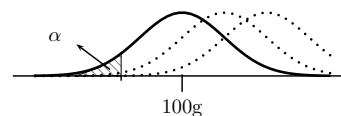
Die Wahrscheinlichkeit für einen Fehler zweiter Art ist schwierig zu bestimmen. Durch größeren Stichprobenumfang kann dieser Fehler verkleinert werden.

3. Es gibt sehr viele verschiedene Tests und verschiedene Prüfgrößen  $X$ , z.B. Tests auf Gleichheit von Erwartungswerten bei Normalverteilungen oder von Erfolgswahrscheinlichkeiten, auf Verteilung entsprechend einer Normalverteilung, ...

Man kann zur gleichen Situation auch unterschiedliche Tests betrachten.

**Beispiel 3** (Fortsetzung von Beispiel 1):

Verbraucherschützer interessiert nur, ob die Schokoladentafeln signifikant weniger als 100g wiegen. Die Hypothese ist daher  $\mu \geq 100\text{g}$ , die Alternative  $\mu < 100\text{g}$ . Bei diesem einseitigen Test ist der Ablehnungsbereich einseitig: Bei zu geringem Durchschnittsgewicht wird die Hypothese signifikant abgelehnt.



### 3.4.2. Tests für Erwartungswerte und Erfolgswahrscheinlichkeiten

#### 1. Erwartungswerte bei Normalverteilungen mit bekannter Standardabweichung

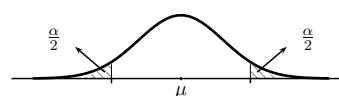
##### Ziel:

Bestimmung eines Akzeptanzbereichs zum Test der Hypothese „Erwartungswert =  $\mu$ “ bei einer normalverteilten Zufallsvariable mit Standardabweichung  $\sigma$  auf Basis von  $n$  Beobachtungen  $X_1, \dots, X_n$ .

Bei  $(\mu, \sigma^2)$ -normalverteiltem  $X$  ist  $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$  normalverteilt um  $\mu$  mit Standardabweichung  $\frac{1}{\sqrt{n}}\sigma$ .

Bei einem Signifikanzniveau  $1-\alpha$  erhält man daher mit dem  $1-\frac{\alpha}{2}$ -Quantil  $x_{1-\frac{\alpha}{2}}$  zur Standardnormalverteilung den Akzeptanzbereich

$$\left[ \mu - \frac{\sigma}{\sqrt{n}} x_{1-\frac{\alpha}{2}}, \mu + \frac{\sigma}{\sqrt{n}} x_{1-\frac{\alpha}{2}} \right].$$



##### Beispiel 1 (Fortsetzung von Beispiel 1 des vorigen Abschnitts):

Das Nachwiegen von 10 Schokoladentafeln ergab ein Durchschnittsgewicht von  $\bar{x} = 99.71\text{g}$ .

Der Hersteller habe einen normalverteilten Prozess mit Standardabweichung von  $\sigma = 0.5\text{g}$ . Ist die Hypothese „Erwartungswert  $\mu = 100\text{g}$ “ zum Signifikanzniveau 0.95, also  $\alpha = 0.05$ , verträglich mit den 10 Beobachtungen?

Wegen

$$x_{1-\frac{\alpha}{2}} = x_{1-\frac{0.05}{2}} = x_{0.975} \approx 1.96$$

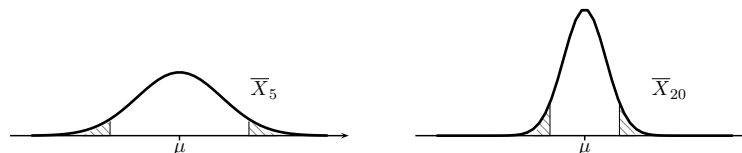
ist der Akzeptanzbereich

$$\left[ 100\text{g} - \frac{0.5\text{g}}{\sqrt{10}} \cdot 1.96, 100\text{g} + \frac{0.5\text{g}}{\sqrt{10}} \cdot 1.96 \right] = [99.69\text{g}, 100.31\text{g}].$$

Die Beobachtung liefert  $\bar{x} = 99.71\text{g}$ , d.h. die Hypothese ist konform zu den Beobachtungen.

##### Bemerkung:

Bei wachsendem Stichprobenumfang  $n$  schrumpft die Breite des Akzeptanzbereichs für  $\bar{x}_n$  also mit  $\frac{1}{\sqrt{n}}$ , da sich die Dichte von  $\bar{X}_n$  entsprechend um  $\mu$  konzentriert.



**Bemerkungen:**

1. Statt einen Akzeptanzbereich für den Mittelwert, also für die Testgröße  $\bar{X}_n$ , zu bestimmen, kann man auch die Testgröße

$$X_n^* = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} = \frac{n \cdot \left(\frac{1}{n} \sum_{k=1}^n X_k - \mu\right)}{\sqrt{n}\sigma} = \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma}$$

nutzen, die standardnormalverteilt ist (s. Seite 32).

Damit erhält man direkt den Akzeptanzbereich

$$|X_n^*| \leq x_{1-\frac{\alpha}{2}}.$$

2. Bei einem einseitigen Test kann man die ganze Irrtumswahrscheinlichkeit  $\alpha$  auf eine Seite nehmen und einen entsprechenden einseitigen Test mit  $x_{1-\alpha}$  machen.

2. *Erwartungswerte bei Normalverteilungen mit unbekannter Standardabweichung*

**Ziel:**

Bestimmung eines Akzeptanzbereichs zum Test der Hypothese „Erwartungswert =  $\mu$ “ bei einer normalverteilten Zufallsvariable mit unbekannter Standardabweichung auf Basis von  $n$  Beobachtungen  $X_1, \dots, X_n$ .

Wie bei den Konfidenzbereichen kann man mit dem Schätzwert

$$S_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x}_n)^2}$$

die Testgröße

$$\hat{X}_n = \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{S_n} \quad \left(\text{mit } \bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k\right)$$

bilden, die  $t$ -verteilt mit  $n-1$  Freiheitsgraden ist. Zur Bestimmung der Akzeptanzgrenzen verwendet man dann die Quantile  $t_{m,p}$  zur  $t$ -Verteilung (s. Anhang C).

Bei Vorgabe eines Signifikanzniveaus  $1 - \alpha$  hat man den Akzeptanzbereich

$$|\hat{X}_n| \leq t_{n-1, 1-\frac{\alpha}{2}}.$$

**Beispiel 2** (Fortsetzung von Beispiel 1):

Bei unbekanntem  $\sigma$  erhält man mit den Werten aus Beispiel 1  $S_n = 0.49\text{g}$  und man hat

$$\hat{x}_n = \sqrt{10} \cdot \frac{\bar{x}_n - 100\text{g}}{0.49\text{g}} \approx -1.87$$

Das  $(1 - \frac{0.05}{2})$ -Quantil zur  $t$ -Verteilung mit 9 Freiheitsgraden ist

$$c = t_{9,0.975} = 2.262.$$

Wegen  $|\hat{x}_n| < c$  ist die Hypothese  $\mu = 100$  zum Signifikanzniveau 0.95 also verträglich mit der Beobachtung.

Bei einem Signifikanzniveau von nur 0.9, also  $\alpha = 0.1$ , erhält man

$$t_{9,1-\frac{\alpha}{2}} = t_{9,0.95} \approx 1.83,$$

so dass man bei diesem Signifikanzniveau die Hypothese ablehnen muss.

**Bemerkung:**

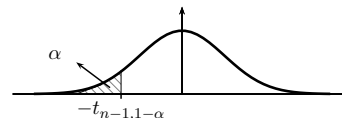
Auch hier sind natürlich einseitige Tests mit  $t_{n-1,1-\alpha}$  möglich.

**Beispiel 3** (Fortsetzung von Beispiel 1 und 2):

Verbraucherschützer interessiert nur, ob die Schokoladentafeln signifikant weniger als 100g wiegen. Die Hypothese ist dann  $\mu \geq 100$ , die Alternative  $\mu < 100$ .

Der Ablehnungsbereich ist einseitig: Bei zu geringem Durchschnittsgewicht wird die Hypothese signifikant abgelehnt, also bei

$$\hat{X}_n < -t_{n-1,1-\alpha}.$$



Zum Signifikanzniveau 0.95 ergibt sich das entsprechende Quantil  $t_{9,0.95} = 1.833$ .

Wegen  $\hat{x}_n = -1.87$  wird die Hypothese  $\mu \geq 100$  zum Signifikanzniveau 95% abgelehnt; man kann also begründet sagen, dass die Schokoladen zu leicht sind.

### 3. Erfolgswahrscheinlichkeiten bei Bernoulli-Experimenten

**Ziel:**

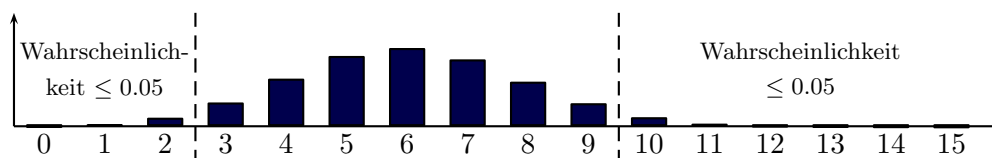
Bestimmung eines Akzeptanzbereichs zum Test der Hypothese „Erfolgswahrscheinlichkeit =  $p$ “ bei einem Bernoulli-Experiment auf Basis von  $n$  Durchführungen.

Unter der Hypothese „Erfolgswahrscheinlichkeit =  $p$ “ ist die Anzahl der Erfolge bei  $n$  Durchführungen binomial-verteilt. Bei dieser Verteilung kann man die unwahrscheinlichen Randbereiche abschneiden.

**Beispiel 4:**

Betrachtet werden 15 Durchführungen eines Bernoulliexperimentes mit der Hypothese  $p = 0.4$ .

Will man einen Akzeptanzbereich zum Signifikanzniveau 0.9 bestimmen, kann man für die Anzahl der Erfolge an beiden Seiten des Wahrscheinlichkeits-Histogramms eine Wahrscheinlichkeit von 0.05 streichen:



Die konkreten Werte für die kumulierten Wahrscheinlichkeiten sind (unter der Hypothese  $p = 0.4$ )

$$\begin{array}{ll}
 P(0 \text{ Erfolge}) \approx 0.0005, & P(15 \text{ Erfolge}) \approx 1.1 \cdot 10^{-6}, \\
 P(\leq 1 \text{ Erfolg}) \approx 0.0052, & P(\geq 14 \text{ Erfolge}) \approx 2.5 \cdot 10^{-5}, \\
 P(\leq 2 \text{ Erfolg}) \approx 0.0271, & P(\geq 13 \text{ Erfolge}) \approx 2.8 \cdot 10^{-4}, \\
 P(\leq 3 \text{ Erfolg}) \approx 0.0905, & P(\geq 12 \text{ Erfolge}) \approx 0.0019, \\
 P(\leq 4 \text{ Erfolg}) \approx 0.217, & P(\geq 11 \text{ Erfolge}) \approx 0.0093, \\
 P(\leq 5 \text{ Erfolg}) \approx 0.403, & P(\geq 10 \text{ Erfolge}) \approx 0.0338, \\
 P(\leq 6 \text{ Erfolg}) \approx 0.609, & P(\geq 9 \text{ Erfolge}) \approx 0.095.
 \end{array}$$

Damit erhält man den Akzeptanzbereich  $[3; 9]$ .

Bei großem Stichprobenumfang kann man auch die Näherung mittels der Normalverteilung (s. Satz 2.10) zur Bestimmung der Akzeptanzgrenzen nutzen:

Für die Anzahl  $Z_n$  der Erfolge bei  $n$ -maliger Durchführung gilt

$$P(Z_n \leq b) \approx \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) \quad \text{und} \quad P(Z_n \geq a) \approx 1 - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right).$$

Mit Hilfe der Quantile zur Standardnormalverteilung kann man damit Grenzen eines Akzeptanzbereichs bestimmen und erhält als beidseitigen Akzeptanzbereich für die Anzahl der Erfolge  $Z_n$ :

$$Z_n \in [k_1; k_2] \quad \text{mit} \quad k_{1,2} = np \pm (\sqrt{np(1-p)} \cdot x_{1-\frac{\alpha}{2}} + 0.5).$$

### Beispiel 5:

100 Würfe mit einer Münze ergeben 58 mal Zahl und nur 42 mal Kopf. Ist die Münze unfair, also  $p \neq 0.5$  bei einer entsprechenden Bernoulli-Modellierung (Kopf  $\hat{=} 0$ , Zahl  $\hat{=} 1$ )?

Die zu testende Hypothese ist  $p = 0.5$ . Zum Signifikanzniveau 0.95, also  $\alpha = 0.05$ , ist  $x_{1-\frac{\alpha}{2}} = 1.96$ . Damit wird der Akzeptanzbereich begrenzt durch

$$100 \cdot 0.5 \pm (\sqrt{100 \cdot 0.5(1-0.5)} \cdot 1.96 + 0.5) = 50 \pm 10.3.$$

Damit liegt das Ergebnis noch im Rahmen.



### 3.4.3. Der $p$ -Wert

Statt bei einem Hypothesentest das Signifikanzniveau  $1 - \alpha$  festzulegen, kann man auch den sogenannten  $p$ -Wert einer Beobachtung betrachten:

#### Definition 3.10

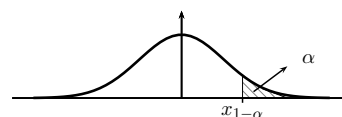
Bei einem Hypothesentest ist der  $p$ -Wert zu einer Beobachtung das kleinste  $\alpha$ , so dass die Beobachtung bei einem Signifikanzniveau von  $1 - \alpha$  zur Ablehnung von  $H_0$  führt.

#### Beispiel

Bei einem normalverteilten Merkmal mit  $\sigma = 2$  soll getestet werden, ob für den Erwartungswert  $\mu \leq 4$  gilt. Beobachtet wurden fünf Werte mit Mittelwert  $\bar{x} = 5.62$ .

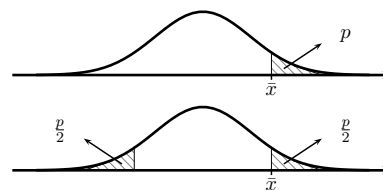
Bei einem Test zum Niveau  $1 - \alpha$  lehnt man die Hypothese ab bei  $\bar{x} \geq 4 + \frac{2}{\sqrt{5}} \cdot x_{1-\alpha}$ . Für das kleinst-mögliche  $\alpha$ , das zur Ablehnung führt, gilt also

$$\begin{aligned} 5.62 &= 4 + \frac{2}{\sqrt{5}} \cdot x_{1-\alpha} \\ \Leftrightarrow x_{1-\alpha} &= 1.62 \cdot \frac{\sqrt{5}}{2} \approx 1.811 \\ \Leftrightarrow 1 - \alpha &\approx 0.965 \quad \Leftrightarrow \quad \alpha \approx 0.035. \end{aligned}$$



Der  $p$ -Wert zu der Beobachtung ist also 0.035.

Anschaulich beschreibt der  $p$ -Wert die Wahrscheinlichkeit des entsprechenden Restbereichs jenseits des beobachteten Werts bzw. bei einem zweiseitigen Test der beiden Restbereiche.



Legt man ein Signifikanzniveau  $1 - \alpha$  fest, so gilt:

$$\begin{aligned} p\text{-Wert} \leq \alpha &\Rightarrow \text{ Lehne die Hypothese ab!} \\ p\text{-Wert} > \alpha &\Rightarrow \text{ Die Beobachtung ist „im Rahmen“.} \end{aligned}$$

Mit dem  $p$ -Wert sieht man auch, wie weit man von der Signifikanz-Grenze entfernt ist.

(Unter <http://www.aerzteblatt.de/archiv/64533/> gibt es einen schönen Artikel, der den  $p$ -Wert mit Konfidenzintervallen vergleicht.)

### 3.4.4. Hypothesentests und Konfidenzbereiche

Konfidenzbereiche sind eng verwandt mit Hypothesentests:

Bei einem Konfidenz-Intervall bestimmt man ein Intervall  $I(x)$  zu einer Stichprobe so, dass gilt:

$$P_{\vartheta}(\vartheta \in I(x)) \geq 1 - \alpha.$$

Daraus kann man einen Hypothesentest für  $\vartheta$  machen, indem die Hypothese bei einer Stichprobe  $x$  akzeptiert wird, wenn  $\vartheta \in I(x)$  ist.

Hat man umgekehrt eine Vorschrift für einen Hypothesentest, so kann man damit Konfidenzintervalle bestimmen:

Bei einem Hypothesentest zu „Parameter  $\vartheta$ “ bestimmt man einen Akzeptanzbereich  $A(\vartheta)$  für eine Testgröße  $T(X)$  zu einer Stichprobe  $X$ , meist indem man unwahrscheinliche Randbereiche ausschließt. Dabei gilt

$$P_{\vartheta}(T(X) \in A(\vartheta)) \geq 1 - \alpha$$

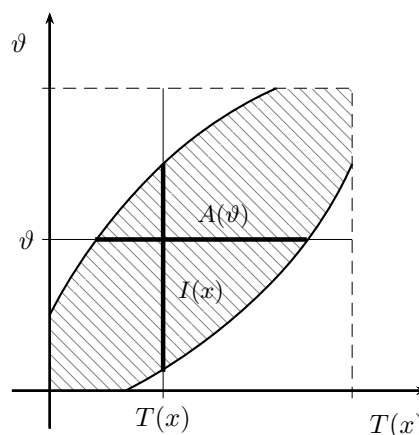
zu einem Signifikanzniveau  $1 - \alpha$ .

Man erhält nun einen Konfidenzbereich durch

$$I(x) = \{\vartheta | T(x) \in A(\vartheta)\},$$

denn dann ist

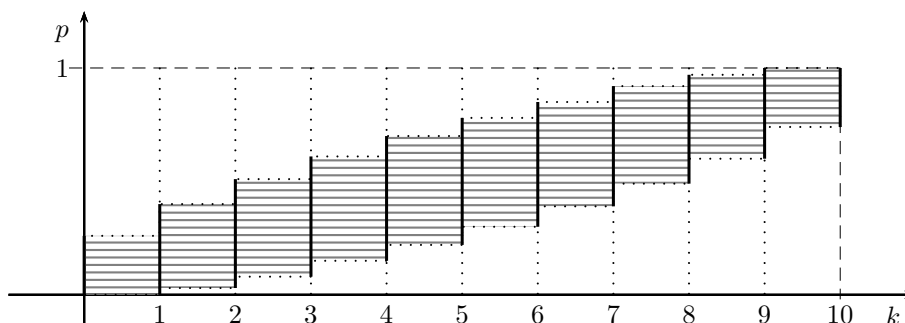
$$P_{\vartheta}(\vartheta \in I(x)) = P_{\vartheta}(T(x) \in A(\vartheta)) \geq 1 - \alpha.$$



#### Beispiel 1:

Bei einem 10-fachen Bernoulliexperiment kann man zur Hypothese „Erfolgswahrscheinlichkeit =  $p$ “ (zu einem Signifikanzniveau  $1 - \alpha$ ) durch Abschneiden der unwahrscheinlichen Ränder Akzeptanzbereiche  $A(p)$  für die Anzahl der Erfolge  $k$  bestimmen, also einen Bereich  $A(p)$  mit

$$P_p(k \in A(p)) \geq 1 - \alpha.$$



In Abhängigkeit von den Erfolgen  $k$  erhält man damit Konfidenzbereiche für  $p$ :

$$I(k) = \{p | k \in A(p)\}.$$

Dies ist tatsächlich ein Konfidenzbereich, denn wegen  $p \in I(k) \Leftrightarrow k \in A(p)$  ist

$$P_p(p \in I(k)) = P_p(k \in A(p)) \geq 1 - \alpha.$$

---

### 3.4.5. $\chi^2$ -Verteilung und $\chi^2$ -Tests

Die  $\chi^2$ -Verteilung (lies „Chiquadrat“) mit  $m$  Freiheitsgraden entsteht als Summe der Quadrate von  $m$  standardnormalverteilten Zufallsvariablen:

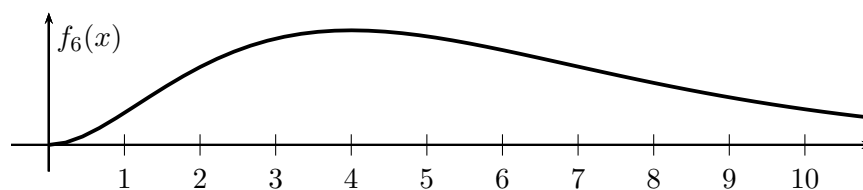
**Definition 3.11**

Sind  $X_1, \dots, X_m$  unabhängige standardnormalverteilte Zufallsvariablen, so heißt die Verteilung von  $\sum_{k=1}^m X_k^2$   $\chi^2$ -Verteilung mit  $m$  Freiheitsgraden

Die  $\chi^2$ -Verteilung  $\chi_m^2$  mit  $m$  Freiheitsgraden besitzt die Dichte

$$f_m(x) = c_m \cdot x^{\frac{m}{2}-1} \cdot e^{-\frac{x}{2}} \quad \text{für } x > 0$$

mit einer Normierungskonstanten  $c_m$  und  $f(x) = 0$  für  $x \leq 0$ .



Für Erwartungswert und Varianz gilt

$$E(\chi_m^2) = m \quad \text{und} \quad V(\chi_m^2) = 2m.$$

Die  $p$ -Quantile  $\chi_{p,m}^2$  zur  $\chi^2$ -Verteilung mit  $m$  Freiheitsgraden sind zu einigen  $p$ -Werten in Anhang D aufgelistet.

**Anwendungen:**

1. Die Verteilung der Stichprobenvarianz

$$s^2 := \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$$

ist eng verwandt mit der  $\chi^2$ -Verteilung:

Sind die zugrunde liegenden Zufallsprozesse normalverteilt mit Varianz  $\sigma^2$ , so ist  $\frac{n-1}{\sigma^2} \cdot s^2$  genau wie  $\chi_{n-1}^2$  verteilt. (Bei anderen Zufallsprozessen erhält man aufgrund des zentralen Grenzwertsatzes eine entsprechende approximative Verteilung.)

Damit gilt mit der Wahrscheinlichkeit  $1 - \alpha$ :

$$\begin{aligned} \chi_{\frac{\alpha}{2}, n-1}^2 &\leq \frac{n-1}{\sigma^2} \cdot s^2 \leq \chi_{1-\frac{\alpha}{2}, n-1}^2 \\ \Leftrightarrow \sigma^2 &\leq \frac{n-1}{\chi_{\frac{\alpha}{2}, n-1}^2} \cdot s^2 \quad \text{und} \quad \frac{n-1}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \cdot s^2 \leq \sigma^2 \\ \Leftrightarrow \sigma^2 &\in \left[ \frac{n-1}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \cdot s^2; \frac{n-1}{\chi_{\frac{\alpha}{2}, n-1}^2} \cdot s^2 \right]. \end{aligned}$$

Damit hat man ein Konfidenzintervall zum Niveau  $1 - \alpha$  für die tatsächliche Varianz.

2. Die  $\chi^2$ -Verteilung dient als Grundlager vieler Hypothesentests ( $\chi^2$ -Tests).

Eine Anwendung ist beispielsweise der Test, ob eine beobachtete Verteilung einer hypothetischen Verteilung entspricht:

*Situation 1: Diskrete Verteilung*

Eine Zufallsvariable  $X$  kann  $N$  verschiedene Werte  $x_1, \dots, x_N$  annehmen. Zu testen ist die Hypothese  $P(X = x_k) = p_k$  zu vorgegebenen  $p_k$ .

Dazu werden  $n$  Realisierungen von  $X$  durchgeführt und die relativen Anteile  $r_k$  des Auftretens von  $x_k$  bestimmt.

Die Testgröße

$$T = n \cdot \sum_{k=1}^N \frac{(r_k - p_k)^2}{p_k}$$

ist dann (unter der Hypothese) für große  $n$  annähernd  $\chi^2$ -verteilt mit  $N - 1$  Freiheitsgraden.

Dies kann man sich wie folgt plausibilisieren:

Den relativen Anteil  $R_k$  kann man als Zufallsvariable entsprechend des Mittelwerts von  $n$  Bernoulliexperimenten mit Wert 1, falls  $x_k$  auftritt, und Wert 0 sonst, ansehen. Unter der Hypothese  $P(X = x_k) = p_k$  ist das also ein Bernoulliexperiment mit Erfolgswahrscheinlichkeit  $p_k$  und Varianz  $p_k \cdot (1 - p_k)$ .

Nach dem zentralen Grenzwertsatz ist dann

$$X_{n,k}^* = \sqrt{n} \cdot \frac{R_k - p_k}{\sqrt{p_k \cdot (1 - p_k)}}$$

annähernd standardnormalverteilt. Die Summe der Quadrate ist dann

$$\sum_{k=1}^N (X_{n,k}^*)^2 = \sum_{k=1}^N n \cdot \frac{(R_k - p_k)^2}{p_k \cdot (1 - p_k)} = n \cdot \sum_{k=1}^N \frac{(R_k - p_k)^2}{p_k} \cdot \frac{1}{1 - p_k}.$$

Sind die einzelnen  $p_k$  klein, ist  $\frac{1}{1 - p_k} \approx 1$  vernachlässigbar. Da die Summe aller  $R_k$  gleich 1 ist, sind die  $X_{n,k}$  nicht unabhängig, wodurch sich die  $N - 1$  (statt  $N$ ) Freiheitsgrade der  $\chi^2$ -Verteilung ergibt.

Bei großen Abweichungen, also großem  $T$ , sind die Beobachtungen nicht mehr konsistent zur Hypothese, d.h. zu einem Signifikanzniveau  $1 - \alpha$  wird die Hypothese abgelehnt, wenn  $T$  größer als das  $1 - \alpha$ -Quantil  $\chi_{1-\alpha; N-1}^2$  der  $\chi^2$ -Verteilung mit  $N - 1$  Freiheitsgraden ist.

*Situation 2: Stetige Verteilung*

Um zu testen, ob eine Zufallsvariable einer bestimmten stetigen Verteilung genügt (z.B. ob sie standardnormalverteilt ist), kann man einzelne Klassen definieren (z.B. einzelne Teilintervalle) und einerseits die theoretischen Wahrscheinlichkeiten  $p_k$  entsprechend der hypothetischen Verteilung und andererseits die relativen Häufigkeiten  $r_k$  zu den einzelnen Klassen bei  $n$  Realisierungen bestimmen.

Durch diese Klassenbildung erhält man eine diskrete Verteilung und kann nun wie oben vorgehen:

Für große  $n$  ist die oben definierte Testgröße  $T$  bei  $N$  Klassen (unter der Hypothese) annähernd  $\chi^2$ -verteilt mit  $N - 1$  Freiheitsgraden, und man lehnt die Hypothese bei  $T > \chi_{1-\alpha; N-1}$  ab.

---

## 4. Stochastische Prozesse und Markov-Ketten

### Definition 4.1

Sei  $T$  eine Indexmenge und zu jedem  $t \in T$  sei  $X_t$  eine Zufallsvariable. Dann heißt die Familie  $(X_t)_{t \in T}$  *stochastischer Prozess*.

### Bemerkung:

Meist ist  $T = \mathbb{R}^{\geq 0}$  (zeitstetiger Prozess), z.B. wenn  $X_t$  einen Aktienkurs beschreibt, oder  $T = \mathbb{N}$  oder  $T = \mathbb{N}_0$  (zeitdiskreter Prozess), z.B. wenn  $X_n$  das Vermögen eines Glücksspielers nach dem  $n$ -ten Spiel beschreibt.

### Beispiel 1 (Symmetrische Irrfahrt (random walk) in $\mathbb{Z}$ ):

$T = \mathbb{N}_0$ ,  $X_0 = 0$ . Abhängig von  $X_n$  ist  $X_{n+1} = X_n + 1$  oder  $X_{n+1} = X_n - 1$  jeweils mit Wahrscheinlichkeit  $\frac{1}{2}$ .

### Beispiel 2 (Symmetrische Irrfahrt in $\mathbb{Z}^d$ ):

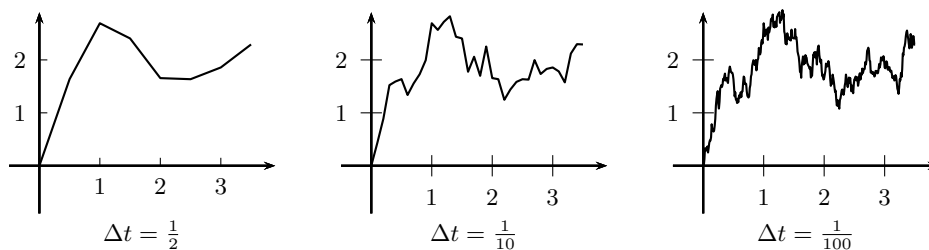
$T = \mathbb{N}_0$ ,  $X_n \in \mathbb{Z}^d$ .  $X_0 = 0$  und  $X_{n+1}$  ergibt sich aus  $X_n$  durch Addition/Subtraktion eines zufällig gewählten Vektors  $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{pmatrix}$  (gleichverteilt).

### Beispiel 3 (Glücksspiel):

$T = \mathbb{N}_0$ ,  $X_0 =$  Startguthaben. Bei  $X_n > 0$  ist  $X_{n+1} = X_n \pm 1$  jeweils mit Wahrscheinlichkeit  $\frac{1}{2}$ ; bei  $X_n = 0$  ist auch  $X_{n+1} = 0$ .

### Beispiel 4 (Brownsche Bewegung (Wiener-Prozess)):

$T = \mathbb{R}^{\geq 0}$ ,  $X_0 = 0$ . Die Zuwächse  $X_t - X_s$  ( $t > s$ ) sind  $(0, t - s)$ -normalverteilt. Eine Approximation geschieht durch Diskretisierung und  $(0, \Delta t)$ -normalverteilte Zuwächse.



### Bemerkung:

Man kann zeigen, dass bei der symmetrischen Irrfahrt in  $\mathbb{Z}$  und  $\mathbb{Z}^2$  jeder Punkt mit der Wahrscheinlichkeit 1 irgendwann erreicht wird. Dies bedeutet auch, dass man von jedem Punkt den Ausgangspunkt mit Wahrscheinlichkeit 1 wieder erreicht. In  $\mathbb{Z}^d$ ,  $d > 2$ , ist das nicht so. („Ein betrunkenen Mann findet immer heim, ein betrunkenen Vogel nicht“).

Damit folgt, dass man beim Glücksspiel mit Sicherheit irgendwann bankrott ist.

**Definition 4.2**

Eine *Markov-Kette* (erster Ordnung) ist ein stochastischer Prozess, mit  $T = \mathbb{R}^{\geq 0}$  (Markov-Prozess) oder  $T = \mathbb{N}$  (diskrete Markov-Kette), bei dem die Zukunft nur vom aktuellen Zeitpunkt und nicht von der Vorgeschichte abhängig ist.

**Bemerkung:**

Man sagt auch, eine Markov-Kette ist *gedächtnislos*, oder schreibt

$$\text{für } t > t_n > t_{n-1} > \dots > t_0 \text{ gilt: } P(X_t | X_{t_n}, X_{t_{n-1}}, \dots, X_{t_0}) = P(X_t | X_{t_n}).$$

**Beispiel 5:**

Die Beispiele 1 bis 4 sind Markov-Ketten.

**Beispiel 6:**

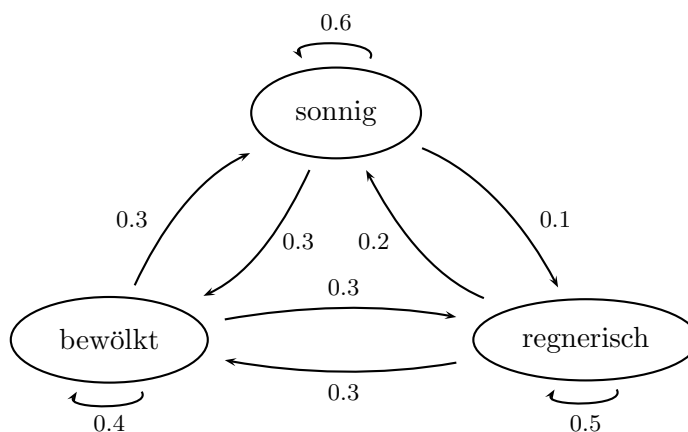
Sind  $X_1, X_2, \dots$  unabhängig, so ist  $(X_i)_{i \in \mathbb{N}}$  eine Markov-Kette.

**Beispiel 7:**

Betrachtet werden die Zustände „sonnig“, „bewölkt“ und „regnerisch“.

Wetterbeobachtungen zeigen folgende Zustandsübergänge:

		Heute		
		sonnig	bewölkt	regnerisch
Morgen	sonnig	0.6	0.3	0.2
	bewölkt	0.3	0.4	0.3
	regnerisch	0.1	0.3	0.5





**Definition 4.3**

Eine diskrete Markov-Kette  $(X_n)_{n \in \mathbb{N}_0}$  heißt *homogen*

$$:\Leftrightarrow \text{für alle } n_1 > n_0 \text{ und } k \in \mathbb{N} \text{ gilt } P(X_{n_1+k} | X_{n_0+k}) = P(X_{n_1} | X_{n_0})$$

(die Zustandsübergänge sind unabhängig von der Zeit).

**Bemerkung:**

Eine homogene Markov-Kette  $(X_n)_{n \in \mathbb{N}_0}$  ist eindeutig bestimmt durch die Verteilung von  $X_0$  und die Übergänge  $P(X_{n+1} | X_n) = P(X_1 | X_0)$ .

Bei einem diskreten Zustandsraum  $X_n \in \{\omega_1, \dots, \omega_N\}$  kann der Übergang durch eine *Übergangsmatrix*  $P = (p_{ij})_{1 \leq i, j \leq N}$  mit  $p_{ij} = P(X_1 = \omega_i | X_0 = \omega_j)$  beschrieben werden.

**Beispiel 8:**

In Beispiel 7 gibt es die Zustände  $\omega_1 =$  „sonnig“,  $\omega_2 =$  „bewölkt“, und  $\omega_3 =$  „regnerisch“. Die Übergangsmatrix ist

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.5 \end{pmatrix}.$$

Der Ausgangszustand sei „sonnig“, also beschrieben durch den Wahrscheinlichkeitsvektor  $x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . Dann ist der Wahrscheinlichkeitsvektor zu  $X_1$

$$x_1 = \begin{pmatrix} 0.6 \\ 0.3 \\ 0.1 \end{pmatrix} = P \cdot x_0.$$

Wie ist es bei  $X_2$ ?

Zu 60% gibt es die Verteilung  $\begin{pmatrix} 0.6 \\ 0.3 \\ 0.1 \end{pmatrix}$  zu 30%  $\begin{pmatrix} 0.3 \\ 0.4 \\ 0.3 \end{pmatrix}$  und zu 10%  $\begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix}$ :

$$\begin{aligned} x_2 &= 0.6 \cdot \begin{pmatrix} 0.6 \\ 0.3 \\ 0.1 \end{pmatrix} + 0.3 \cdot \begin{pmatrix} 0.3 \\ 0.4 \\ 0.3 \end{pmatrix} + 0.1 \cdot \begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix} \\ &= \begin{pmatrix} 0.6 & 0.3 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.5 \end{pmatrix} \cdot \begin{pmatrix} 0.6 \\ 0.3 \\ 0.1 \end{pmatrix} \\ &= P \cdot x_1 \\ &= P \cdot P \cdot x_0 = P^2 \cdot x_0. \end{aligned}$$

**Bemerkung:**

Besitzt eine homogene Markov-Kette die Übergangsmatrix  $P$ , so gilt für die Wahrscheinlichkeitsverteilungen  $x_n$

$$x_{n+1} = P \cdot x_n \quad \text{und} \quad x_n = P^n \cdot x_0.$$

**Beispiel 9:**

In Beispiel 8 ergibt sich

$$\begin{array}{cccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & \dots \\ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & \begin{pmatrix} 0.6 \\ 0.3 \\ 0.1 \end{pmatrix}, & \begin{pmatrix} 0.47 \\ 0.33 \\ 0.20 \end{pmatrix}, & \begin{pmatrix} 0.42 \\ 0.33 \\ 0.25 \end{pmatrix}, & \begin{pmatrix} 0.40 \\ 0.33 \\ 0.26 \end{pmatrix}, & \begin{pmatrix} 0.39 \\ 0.33 \\ 0.27 \end{pmatrix}, & \begin{pmatrix} 0.39 \\ 0.33 \\ 0.28 \end{pmatrix}, & \begin{pmatrix} 0.39 \\ 0.33 \\ 0.28 \end{pmatrix}, & \begin{pmatrix} 0.39 \\ 0.33 \\ 0.28 \end{pmatrix} & \dots \end{array}$$

**Definition 4.4**

Eine Wahrscheinlichkeitsverteilung  $\hat{x}$  heißt *stationär* zu einer homogenen Markov-Kette mit Übergangsmatrix  $P$   $:\Leftrightarrow$  Es gilt  $\hat{x} = P \cdot \hat{x}$ .

**Satz 4.5**

Zu jeder Übergangsmatrix  $P$  gibt es eine stationäre Verteilung.

**Bemerkung:**

Ein  $x \in \mathbb{R}^n, x \neq 0$ , heißt *Eigenvektor* mit *Eigenwert*  $\alpha$  zur Matrix  $A \in \mathbb{R}^{n \times n}$   $:\Leftrightarrow A \cdot x = \alpha \cdot x$ .

Eine stationäre Verteilung ist also ein Eigenvektor von  $P$  zum Eigenwert 1.

Eigenvektoren zu einem Eigenwert  $\alpha$  kann man wegen

$$Px = \alpha \cdot x \Leftrightarrow Px = \alpha \cdot I \cdot x \Leftrightarrow (P - \alpha \cdot I) \cdot x = 0$$

bestimmen als nichttriviale Lösungen des homogenen Gleichungssystems

$$(P - \alpha \cdot I) \cdot x = 0.$$

**Beispiel 10:**

Berechnung einer stationäre Verteilung zu  $P$  aus Beispiel 7 bzw. 8:

Suche einen Eigenvektor  $x$  zum Eigenwert 1 von  $P = \begin{pmatrix} 0.6 & 0.3 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.5 \end{pmatrix}$ , also eine

Lösung  $x$  zu

$$(P - I) \cdot x = 0.$$

Mit elementaren Zeilenoperationen (mit gedachter rechter Seite gleich Null) erhält man

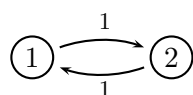
$$\begin{aligned}
 P - I &= \begin{pmatrix} -0.4 & 0.3 & 0.2 \\ 0.3 & -0.6 & 0.3 \\ 0.1 & 0.3 & -0.5 \end{pmatrix} \begin{array}{l} \text{alles} \cdot 10, \\ \text{tausche I und III} \end{array} \\
 &\rightarrow \begin{pmatrix} 1 & 3 & -5 \\ 3 & -6 & 3 \\ -4 & 3 & 2 \end{pmatrix} \begin{array}{l} -3 \cdot \text{I} \\ +4 \cdot \text{I} \end{array} \\
 &\rightarrow \begin{pmatrix} 1 & 3 & -5 \\ 0 & -15 & 18 \\ 0 & 15 & -18 \end{pmatrix} \begin{array}{l} +\frac{1}{5} \cdot \text{II} \\ +\text{II} \end{array} \\
 &\rightarrow \begin{pmatrix} 1 & 0 & -\frac{7}{5} \\ 0 & 1 & -\frac{6}{5} \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

Eigenvektoren sind also  $\lambda \cdot \begin{pmatrix} 7 \\ 6 \\ 5 \\ 1 \end{pmatrix}$ . Bei  $\lambda = \frac{5}{18}$  erhält man  $\hat{x} = \begin{pmatrix} \frac{7}{18} \\ \frac{6}{18} \\ \frac{5}{18} \\ 1 \end{pmatrix}$ , der eine Wahrscheinlichkeitsverteilung beschreibt.

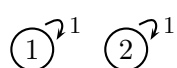
**Bemerkung:**

Häufig ist die stationäre Verteilung eindeutig und kann durch  $P^n \cdot x_0$ , ausgehend von einer beliebigen Anfangsverteilung  $x_0$  berechnet werden (z.B. wenn alle Einträge in  $P$  positiv sind).

**Beispiel 11:**



ist eine Markovkette mit Übergangsmatrix  $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Die stationäre Verteilung ist  $\hat{x} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$ , aber für  $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  gilt nicht  $P^n \cdot x_0 \xrightarrow{n \rightarrow \infty} \hat{x}$ .



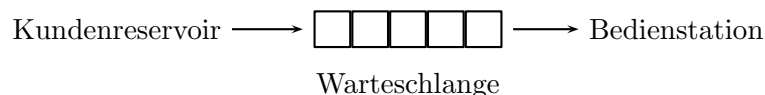
ist eine Markovkette mit Übergangsmatrix  $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Jede Wahrscheinlichkeitsverteilung ist dann eine stationäre Verteilung.

## 5. Warteschlangen

### 5.1. Einführung

#### Situation:

Aus einem (als unendlich angenommenen) Kundenreservoir kommen Kunden und reihen sich in eine Warteschlange ein, um an der Bedienstation bedient zu werden.



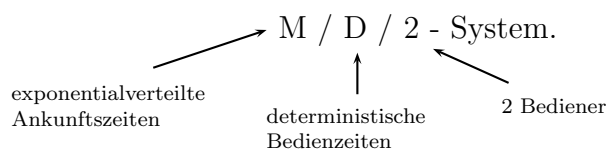
Die (Zwischen-) Ankunftszeiten sowie die Bedienzeiten der Kunden können

(M) exponentialverteilt sein („memoryless“),

(D) deterministisch sein,

(G) irgendwie anders verteilt sein („general“).

Entsprechend der Verteilung bezeichnet man ein Warteschlangensystem beispielsweise als



Ist die Warteschlangenlänge begrenzt, kann diese Länge hinten angehängt werden. z.B. M/D/2/10. (*Kendall-Notation*)

#### Beispiel 1:

Eine Supermarktkasse kann man als M/M/1-System modellieren, die Schlange an einem Geldautomaten angenähert als M/D/1.

Wichtige Kenngrößen sind

$$\text{Ankunftsrate} = \frac{1}{E(\text{Zwischenankunftszeit})},$$

$$\text{Bedienrate} = \frac{1}{E(\text{Bedienzeit})}.$$

#### Bemerkung:

Bei einer Warteschlange mit exponentialverteilten Zwischenankunftszeiten mit Parameter  $\lambda$  ist  $E(\text{Zwischenankunftszeit}) = \frac{1}{\lambda}$ , d.h., die Ankunftsrate ist gleich  $\lambda$ , entsprechend bei exponentialverteilten Bedienzeiten.

Es gibt unterschiedliche Bedienstrategien:

- FIFO bzw. FCFS  
First-in-first-out bzw. First-come-first-serve
- LIFO bzw. LCFS  
Last-in-first-out bzw. Last-come-first-serve
- RR oder Round Robin  
Ist die Bedienung nach einer festen Zeit noch nicht beendet, muss der Kunde (mit der restlich nötigen Bedienzeit) wieder hinten in die Warteschlange.  
Beispiel: Multitasking Betriebssysteme.

Bei unterschiedlichen Prioritäten der Aufgaben/Kunden gibt es zwei Strategien:

- Preemptive priority  
Sobald ein wichtiger Kunde kommt, wird die aktuelle Bearbeitung unterbrochen und der neue Kunde vorgelassen.
- Non preemptive priority  
Der wichtigere Kunde kommt direkt nach dem aktuellen dran; dessen Bedienung wird aber zunächst abgeschlossen.

## 5.2. Die Poisson-Verteilung

### Situation:

Bei einem Ankunftsprozess beobachtet man im Durchschnitt eine Ankunft in der Zeit  $T_0$ . Wie groß ist die Wahrscheinlichkeit

$$P(T, k) = W(\text{in } [0, T] \text{ gibt es genau } k \text{ Ankünfte})?$$

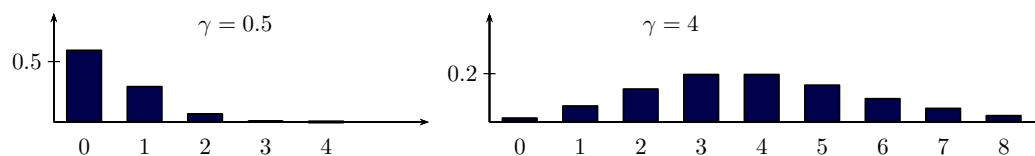
### Definition 5.1

Die diskrete Zufallsvariable  $X$  mit

$$P(X = k) = \frac{\gamma^k}{k!} e^{-\gamma} \quad (k = 0, 1, 2, \dots)$$

heißt *Poisson-verteilt* mit Parameter  $\gamma > 0$ .

Es ist  $E(X) = \gamma$  und  $V(X) = \gamma$ .



### Satz 5.2

Bei einem exponentialverteilten Ankunftsprozess mit der Ankunftsrate  $\alpha = \frac{1}{T_0}$  (im Durchschnitt eine Ankunft in  $T_0$ ) ist die Anzahl der Ankünfte in der Zeit  $T$  Poisson-verteilt mit Parameter  $\gamma = \alpha T = \frac{T}{T_0}$ .

### Bemerkungen:

1. Die Wahrscheinlichkeiten einer Poisson-Verteilung addieren sich tatsächlich zu 1:

$$\sum_{k=0}^{\infty} \frac{\gamma^k}{k!} e^{-\gamma} = e^{\gamma} \cdot e^{-\gamma} = 1.$$

2. Statt von einem Ankunftsprozess mit exponentialverteilten Zwischenankunftszeiten zu sprechen, spricht man auch von einem Poisson-verteilten Ankunftsprozess.

Entsprechendes gilt für den Bedienprozess.

3. Die Poisson-Verteilung approximiert die Binomialverteilung für  $\gamma = n \cdot p$  bei großem  $n$  und kleinem  $p$  (die Approximation ist gut für  $n > 50$  und  $p < 0.1$ ):

Die „Erfolgsrate“ ist  $p$  (durchschnittliche Anzahl Erfolge pro Versuch), und die Anzahl der Versuche  $n$  spielt die Rolle von  $T$ .

**Beispiel 1:**

$n = 60$ ,  $p = \frac{1}{30}$ , also  $\gamma = \frac{60}{30} = 2$ :

$$P_{\text{Binomial}}(2) = \binom{60}{2} \cdot \left(\frac{1}{30}\right)^2 \cdot \left(\frac{29}{30}\right)^{58} \approx 0.275$$

$$P_{\text{Binomial}}(3) = \binom{60}{3} \cdot \left(\frac{1}{30}\right)^3 \cdot \left(\frac{29}{30}\right)^{57} \approx 0.183$$

$$P_{\text{Poisson}}(2) = \frac{2^2}{2!} e^{-2} \approx 0.271$$

$$P_{\text{Poisson}}(3) = \frac{2^3}{3!} e^{-2} \approx 0.180$$

### 5.3. Eigenschaften von Warteschlangen

**Satz 5.3**

Betrachtet wird ein Poisson-verteilter Ankunfts- und Bedienprozess mit Ankunftsrate  $\alpha$  und Bedienrate  $\beta$ . Ist  $\alpha < \beta$ , so konvergiert

$$P_n(t) = P(\text{Zur Zeit } t \text{ sind } n \text{ Kunden im System})$$

gegen

$$P_n = \left(\frac{\alpha}{\beta}\right)^n \cdot \left(1 - \frac{\alpha}{\beta}\right).$$

**Bemerkung:**

Die Wahrscheinlichkeiten entsprechen denen einer geometrischen Verteilung (Variante 2 in Anhang E) mit Parameter  $p = 1 - \frac{\alpha}{\beta}$ .

**Beispiel 1:**

An der Kasse eines Ladens werden ca. 25 Kunden pro Stunde bedient. Die durchschnittliche Bediendauer ist 2 Minuten. Also ist  $\alpha = 25$  (pro Stunde),  $\beta = 30$  (pro Stunde) und bei einer M/M/1-Modellierung ist

$$P_n = \left(\frac{25}{30}\right)^n \left(1 - \frac{25}{30}\right) = \left(\frac{5}{6}\right)^n \cdot \frac{1}{6}.$$

Wie viel Prozent ihrer Arbeitszeit ist die Kassiererin unbeschäftigt?

$$P_0 = \left(\frac{5}{6}\right)^0 \cdot \frac{1}{6} \approx 0.1667$$

$\Rightarrow$  in ca. 17% der Zeit ist kein Kunde an der Kasse.

**Bemerkungen:**

1. Die Wahrscheinlichkeit, dass die Bedienstation arbeitslos ist, ist  $P_0 = 1 - \frac{\alpha}{\beta}$ , daher bezeichnet man  $\frac{\alpha}{\beta}$  auch als *Auslastung*.

Dies ist auch intuitiv verständlich:  $N$  Kunden kommen ungefähr in einer Zeit  $N \cdot \frac{1}{\alpha}$ , die Bedienung dauert ungefähr  $N \cdot \frac{1}{\beta}$ , so dass sich als Auslastung  $\frac{N/\beta}{N/\alpha} = \frac{\alpha}{\beta}$  ergibt.

2. Als mittlere Anzahl Personen im System ergibt sich durch den Erwartungswert  $\frac{1-p}{p}$  der geometrischen Verteilung mit  $p = 1 - \frac{\alpha}{\beta}$ :

$$\text{mittlere Anzahl Personen} = \frac{1-p}{p} = \frac{1 - (1 - \frac{\alpha}{\beta})}{1 - \frac{\alpha}{\beta}} = \frac{\alpha}{\beta - \alpha}.$$



3. Manchmal interessiert man sich für die *Schlänglänge*, also die Anzahl der Personen abzüglich der gerade in Bedienung befindlichen Person. Es gilt

$$\text{mittlere Schlänglänge} = \frac{\alpha^2}{\beta(\beta - \alpha)}.$$

**Beispiel 2** (Fortsetzung von Beispiel 1):

Die Anzahl Personen im System ist  $\frac{25}{30-25} = 5$ .

**Bemerkung (Warteschlange als Markov-Kette):**

Die Anzahl der Personen im System kann man als Markov-Kette mit den möglichen Zuständen  $0, 1, 2, \dots$  betrachten. Die (unendliche) Übergangsmatrix wird beschrieben durch

nach\von	0	1	2	3	...
0	$1 - \alpha\Delta t$	$\beta\Delta t$	0	0	...
1	$\alpha\Delta t$	$1 - (\alpha + \beta)\Delta t$	$\beta\Delta t$	0	...
2	0	$\alpha\Delta t$	$1 - (\alpha + \beta)\Delta t$	$\beta\Delta t$	...
3	0	0	$\alpha\Delta t$	$1 - (\alpha + \beta)\Delta t$	$\ddots$
4	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$

Als (unendlichen) Eigenvektor zum Eigenwert 1 hat man

$$\begin{pmatrix} 1 - \frac{\alpha}{\beta} \\ \left(\frac{\alpha}{\beta}\right) \left(1 - \frac{\alpha}{\beta}\right) \\ \left(\frac{\alpha}{\beta}\right)^2 \left(1 - \frac{\alpha}{\beta}\right) \\ \vdots \end{pmatrix}.$$

**Satz 5.4**

*Bei einer FIFO-Warteschlange mit Poisson-verteilten Ankunfts- und Bedienprozessen mit Raten  $\alpha$  bzw.  $\beta$  sind die Aufenthaltszeiten exponentialverteilt mit Parameter  $\lambda = \beta - \alpha$ .*

**Bemerkung:**

Die mittlere Aufenthaltsdauer ist also  $\frac{1}{\lambda} = \frac{1}{\beta - \alpha}$ .

**Beispiel 3** (Fortsetzung von Beispiel 1):

Die mittlere Aufenthaltsdauer ist  $\frac{1}{30-25} = \frac{1}{5}$  (in Stunden), also 12 Minuten.

**Bemerkung:**

Manchmal unterscheidet man zwischen Aufenthaltsdauer und Wartezeit = Aufenthaltsdauer – Bedienzeit. Mit den Rechenregeln für Erwartungswerte gilt:

$$\begin{aligned}\text{mittlere Wartezeit} &= \text{mittlere Aufenthaltsdauer} - \text{mittlere Bedienzeit} \\ &= \frac{1}{\beta - \alpha} - \frac{1}{\beta} = \frac{\beta - (\beta - \alpha)}{(\beta - \alpha) \cdot \beta} = \frac{\alpha}{(\beta - \alpha) \cdot \beta}.\end{aligned}$$

## A. Binomialkoeffizient

Erinnerung:  $n! = 1 \cdot 2 \cdot \dots \cdot n$  ( $n$ -Fakultät), insbesondere  $0! = 1 = 1!$ .

### Definition A.1

Für  $n, k \in \mathbb{N}_0$ ,  $n \geq k$  heißt

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

*Binomialkoeffizient*, gelesen „ $n$  über  $k$ “.

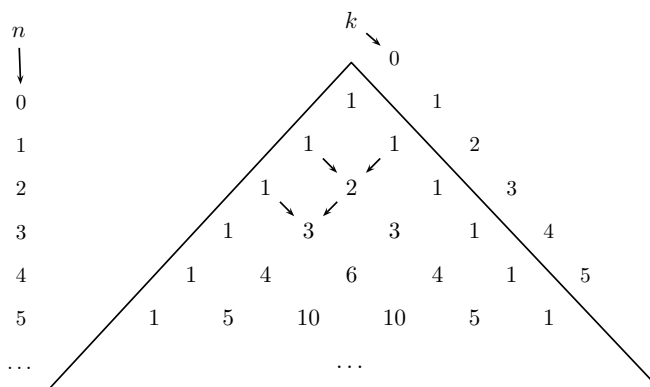
### Beispiel 1:

$$\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3} = \frac{7!}{3! \cdot 4!} = 35.$$

### Satz A.2

1. Für alle  $n, k \in \mathbb{N}_0$ ,  $n \geq k$  gilt  $\binom{n}{k} = \binom{n}{n-k}$ .
2. Für alle  $n, k \in \mathbb{N}$ ,  $n > k$  gilt  $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ .

Die Binomialkoeffizienten kann man sich gut in Form des sog. *Pascalschen Dreiecks* herleiten. Das Dreieck hat am Rand lauter Einsen. Innen ergeben sich die Zahlen durch die Summe der beiden darüber liegenden Werte. Dies entspricht dem Bildungsgesetz 2. aus Satz A.2.



### Satz 1.3 (Binomialsatz)

Für alle  $n \in \mathbb{N}$ ,  $a, b \in \mathbb{R}$  gilt  $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} \cdot b^k$ .

**Beispiel 2:**

$$\begin{aligned}
 (a+b)^2 &= \binom{2}{0} a^{2-0} \cdot b^0 + \binom{2}{1} a^{2-1} \cdot b^1 + \binom{2}{2} a^{2-2} \cdot b^2 \\
 &= a^2 + 2ab + b^2. \\
 (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3.
 \end{aligned}$$

**Bemerkungen:**

1. Der Satz begründet die Bezeichnung „Binomialkoeffizient“ für  $\binom{n}{k}$ , denn dies sind die Koeffizienten bei der ausmultiplizierten Form der Binomialausdrücke  $(a+b)^n$ .
2. Dass diese Koeffizienten genau denen entsprechen, die durch das Pascalsche Dreieck gebildet werden, kann man verstehen, wenn man sich überlegt, wie die Koeffizienten von  $(a+b)^{n+1} = (a+b) \cdot (a+b)^n$  aus denen von  $(a+b)^n$  entstehen.

**Beispiel 3:**

Beim Ausmultiplizieren von  $(a+b)^6$  treten Terme  $a^6, a^5b, a^4b^2, \dots, b^6$  auf, also

$$(a+b)^6 = c_0a^6 + c_1a^5b + c_2a^4b^2 + c_3a^3b^3 + c_4a^2b^4 + c_5ab^5 + c_6b^6.$$

Bei  $(a+b)^7$  gibt es einen Term  $a^4b^3$ . Beim Ausmultiplizieren von  $(a+b) \cdot (a+b)^6$  erhält man einerseits  $a \cdot c_3a^3b^3$  und andererseits  $b \cdot c_2a^4b^2$ , d.h. der Koeffizient von  $a^4b^3$  ist gleich  $c_2 + c_3$ .

Die Koeffizienten von  $(a+b)^{n+1}$  ergeben sich also aus der Summe der zwei entsprechenden Koeffizienten von  $(a+b)^n$  wie beim Pascalschen Dreieck.

**Satz A.4**

*Es gibt  $\binom{n}{k}$  Möglichkeiten aus  $n$  verschiedenen Elementen  $k$  Elemente auszuwählen.*

**Beispiel 4:**

Beim Lotto „6 aus 49“ gibt es  $\binom{49}{6} = 13983816$  verschiedene Ziehungsergebnisse.

## B. Tabelle: Standardnormalverteilung

Werte zur Verteilungsfunktion der Standardnormalverteilung

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

$x$	$\cdot,0$	$\cdot,1$	$\cdot,2$	$\cdot,3$	$\cdot,4$	$\cdot,5$	$\cdot,6$	$\cdot,7$	$\cdot,8$	$\cdot,9$
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99982	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

Ablesebeispiel:

Der Wert zu  $x = 1.96$  steht in der Zeile 1.9 und Spalte  $\cdot,6$ , also  $\Phi(1.96) \approx 0,97500$ .

Für  $x < 0$  gilt  $\Phi(x) = 1 - \Phi(|x|)$ .

### C. Tabelle: Quantile zur $t$ -Verteilung

$p$ -Quantile  $t_{m,p}$  zur  $t$ -Verteilung mit  $m$  Freiheitsgraden.

	$p = 0,9$	$p = 0,95$	$p = 0,975$	$p = 0,99$	$p = 0,995$
$m = 1$	3,0777	6,3137	12,7062	31,8210	63,6559
$m = 2$	1,8856	2,9200	4,3027	6,9645	9,9250
$m = 3$	1,6377	2,3534	3,1824	4,5407	5,8408
$m = 4$	1,5332	2,1318	2,7765	3,7469	4,6041
$m = 5$	1,4759	2,0150	2,5706	3,3649	4,0321
$m = 6$	1,4398	1,9432	2,4469	3,1427	3,7074
$m = 7$	1,4149	1,8946	2,3646	2,9979	3,4995
$m = 8$	1,3968	1,8595	2,3060	2,8965	3,3554
$m = 9$	1,3830	1,8331	2,2622	2,8214	3,2498
$m = 10$	1,3722	1,8125	2,2281	2,7638	3,1693
$m = 15$	1,3406	1,7531	2,1315	2,6025	2,9467
$m = 20$	1,3253	1,7247	2,0860	2,5280	2,8453
$m = 30$	1,3104	1,6973	2,0423	2,4573	2,7500
$m = 40$	1,3031	1,6839	2,0211	2,4233	2,7045
$m = 50$	1,2987	1,6759	2,0086	2,4033	2,6778
$m = 100$	1,2901	1,6602	1,9840	2,3642	2,6259
$m = 500$	1,2832	1,6479	1,9647	2,3338	2,5857
$m = 1000$	1,2824	1,6464	1,9623	2,3301	2,5807

Für große  $m$  ( $m > 40$ ) kann als Näherung

$$t_{m,p} \approx x_p \cdot \left(1 + \frac{1 + x_p^2}{4m}\right)$$

verwendet werden, wobei  $x_p$  das entsprechende Quantil zur Standardnormalverteilung ist.

## D. Tabelle: Quantile zur $\chi^2$ -Verteilung

$p$ -Quantile  $\chi_{p,m}^2$  zur  $\chi^2$ -Verteilung mit  $m$  Freiheitsgraden:

	$p = 0,025$	$p = 0,05$	$p = 0,9$	$p = 0,95$	$p = 0,975$	$p = 0,99$
$m = 1$	0.001	0.004	2.706	3.841	5.024	6.635
$m = 2$	0.051	0.103	4.605	5.991	7.378	9.210
$m = 3$	0.216	0.352	6.251	7.815	9.348	11.345
$m = 4$	0.484	0.711	7.779	9.488	11.143	13.277
$m = 5$	0.831	1.145	9.236	11.070	12.833	15.086
$m = 6$	1.237	1.635	10.645	12.592	14.449	16.812
$m = 7$	1.690	2.167	12.017	14.067	16.013	18.475
$m = 8$	2.180	2.733	13.362	15.507	17.535	20.090
$m = 9$	2.700	3.325	14.684	16.919	19.023	21.666
$m = 10$	3.247	3.940	15.987	18.307	20.483	23.209
$m = 20$	9.591	10.851	28.412	31.410	34.170	37.566
$m = 30$	16.791	18.493	40.256	43.773	46.979	50.892
$m = 40$	24.433	26.509	51.805	55.758	59.342	63.691
$m = 50$	32.357	34.764	63.167	67.505	71.420	76.154
$m = 100$	74.222	77.929	118.498	124.342	129.561	135.807

## E. Übersicht: Verteilungen

### E.1. Diskrete Verteilungen

(Diskrete) Gleichverteilung auf  $1, \dots, n$  (Laplace-Verteilung)

- Werte:  $1, \dots, n$ .
- $P(X = k) = \frac{1}{n}$ .
- $E(X) = \frac{n+1}{2}$ ,  $V(X) = \frac{n^2-1}{12}$ .

Bernoulli-Verteilung mit Parameter  $p$

- Werte: 0, 1 (oder „intakt“, „defekt“ o.ä.).
- $P(X = 1) = p$ ,  $P(X = 0) = 1 - p$ .
- $E(X) = p$ ,  $V(X) = p(1 - p)$ .
- Beschreibt ein Zufallsexperiment mit zwei möglichen Ereignissen.

Binomialverteilung mit Parametern  $p$  und  $n$

- Werte:  $0, 1, \dots, n$ .
- $P(X = k) = \binom{n}{k} \cdot p^k (1 - p)^{n-k}$ .
- $E(X) = np$ ,  $V(X) = np(1 - p)$ .
- Beschreibt die Wahrscheinlichkeit von  $k$  Erfolgen bei der  $n$ -fachen Wiederholung eines Bernoulli-Experiments mit Erfolgswahrscheinlichkeit  $p$ .

Geometrische Verteilung mit Parameter  $p$

- Variante 1:
  - Werte:  $1, 2, \dots$
  - $P(X = k) = (1 - p)^{k-1} \cdot p$ .
  - $E(X) = \frac{1}{p}$ ,  $V(X) = \frac{1-p^2}{p^2}$ .
  - Beschreibt die Wahrscheinlichkeit, dass bei der  $k$ -ten Wiederholung eines Bernoulli-Experiments mit Erfolgswahrscheinlichkeit  $p$  zum ersten Mal Erfolg eintritt.
- Variante 2:
  - Werte:  $0, 1, 2, \dots$



- $P(X = l) = (1 - p)^l \cdot p.$
- $E(X) = \frac{1}{p} - 1 = \frac{1-p}{p}, V(X) = \frac{1-p^2}{p}.$
- Beschreibt die Wahrscheinlichkeit, dass nach  $l$  Fehlversuchen eines Bernoulli-Experiments mit Erfolgswahrscheinlichkeit  $p$  zum ersten Mal Erfolg eintritt.

### Poisson-Verteilung mit Parameter $\lambda$

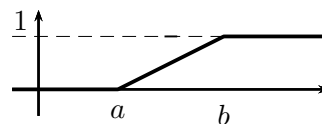
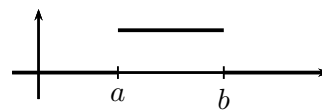
- Werte:  $0, 1, \dots$
- $P(X = k) = \frac{\gamma^k}{k!} e^{-\gamma}.$
- $E(X) = \gamma, V(X) = \gamma.$
- Bei einem exponentialverteilten Ankunftsprozess mit Ankunftsrate  $\alpha = \frac{1}{T_0}$  (im Durchschnitt eine Ankunft in  $T_0$ ) gibt die Poisson-Verteilung mit  $\gamma = \alpha T = \frac{T}{T_0}$  die Wahrscheinlichkeit für die Anzahl der Ankünfte in der Zeit  $T$  an.

Die Poisson-Verteilung approximiert die Binomialverteilung für  $\gamma = n \cdot p$  bei großem  $n$  und kleinem  $p$ .

## E.2. Stetige Verteilungen

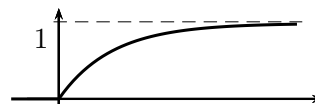
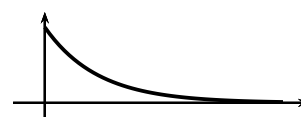
### (Stetige) Gleichverteilung auf $[a, b]$

- Werte:  $[a, b]$ .
- Dichte:  $f(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b], \\ 0 & \text{sonst.} \end{cases}$
- Verteilungsfunktion  $f(x) = \begin{cases} 0 & \text{für } x < a, \\ \frac{x-a}{b-a} & \text{für } x \in [a, b], \\ 1 & \text{für } x > b. \end{cases}$
- $E(X) = \frac{a+b}{2}$ ,  $V(X) = \frac{(b-a)^2}{12}$ .



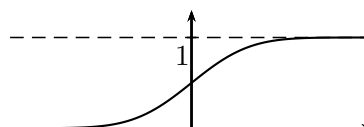
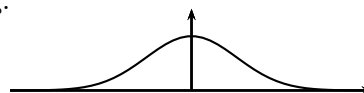
### Exponentialverteilung mit Parameter $\lambda$

- Werte:  $\mathbb{R}^{\geq 0}$ .
- Dichte:  $f(x) = \begin{cases} 0 & \text{für } x < 0, \\ \lambda e^{-\lambda x} & \text{für } x \geq 0. \end{cases}$
- Verteilungsfunktion  $f(x) = \begin{cases} 0 & \text{für } x < 0, \\ 1 - e^{-\lambda x} & \text{für } x \geq 0. \end{cases}$
- $E(X) = \frac{1}{\lambda}$ ,  $V(X) = \frac{1}{\lambda^2}$ .
- Beschreibt Warte- oder Ausfallprozesse.



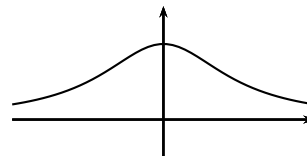
### Normalverteilung mit Parametern $\mu$ und $\sigma$

- Werte:  $\mathbb{R}$ .
- Für  $\mu = 0$  und  $\sigma = 1$ : Standardnormalverteilung.
- Dichte:  $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ .
- Verteilungsfunktion  $f(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ , wobei  $\Phi$  die Verteilungsfunktion zur Standardnormalverteilung ist.
- $E(X) = \mu$ ,  $V(X) = \sigma^2$ .



### Student-Verteilung mit $m$ Freiheitsgraden

- Werte:  $\mathbb{R}$ .
- Dichte:  $f(x) = c_m \cdot \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}$  mit einer Normierungskonstanten  $c_m$ .
- $E(X) = 0$ ,  $V(X) = \frac{n}{n-2}$ .
- Für  $m \rightarrow \infty$  konvergiert die Verteilung gegen die Standardnormalverteilung.



**Index**

<b>Symbols</b>	
$\chi^2$ -Verteilung.....	57, 76
<b>A</b>	
Ankunftsrate.....	65
<b>B</b>	
Bedienrate.....	65
bedingte Wahrscheinlichkeit.....	21
Bernoulli-Experiment.....	2
Bernoulli-Verteilung.....	2, 77
Binomialkoeffizient.....	72
Binomialverteilung.....	6, 77
Brownsche Bewegung.....	60
<b>E</b>	
erwartungstreuer Schätzer.....	39
Erwartungswert.....	10, 27
Exponentialverteilung.....	9, 79
<b>F</b>	
Fehler	
erster Art.....	49
zweiter Art.....	49
Formel	
der totalen Wahrscheinlichkeit.....	24
von Bayes.....	25
<b>G</b>	
gedächtnislose Verteilung.....	23
geometrische Verteilung.....	7, 77
Gleichverteilung.....	9, 77, 79
<b>H</b>	
Hypothesentest.....	48
<b>K</b>	
Kendall-Notation.....	65
Konfidenzbereich.....	43
bei Bernoulli-Experimenten.....	47
bei Normalverteilungen.....	44
Konfidenzniveau.....	43
Korrelationskoeffizient.....	30
Kovarianz.....	30
<b>L</b>	
Laplace-Experiment.....	1, 77
Likelihood-Funktion.....	41
<b>M</b>	
Markov-Kette.....	61
homogene.....	62
Maximum-Likelihood Schätzer.....	41
Median.....	36
Mittelwert.....	10
Stichproben-.....	35
Monte-Carlo Simulation.....	2
<b>N</b>	
Normalverteilung.....	10, 79
<b>P</b>	
Pascalsches Dreieck.....	72

Poisson-Verteilung .....	67, 78	Binomial- .....	6, 77
<b>Q</b>		Exponential- .....	9, 79
Quantil .....	16	gedächtnislos .....	23
<b>R</b>		geometrische .....	7, 77
Realisierung .....	13	Gleich- .....	77, 79
relative Häufigkeit .....	1	Normal- .....	10, 79
<b>S</b>		Poisson- .....	67, 78
Schätzer .....	38	stationäre .....	63
erwartungstreu .....	39	Verteilungsfunktion .....	15
Maximum-Likelihood .....	41	der Standardnormalverteilung .....	74
Signifikanzniveau .....	48	<b>W</b>	
Standardabweichung .....	10, 27	Wahrscheinlichkeit .....	1
Stichproben- .....	35	bedingte .....	21
Standardnormalverteilung .....	10, 74	Wahrscheinlichkeitsdichte .....	8
Stichprobe .....	35	Warteschlange .....	65
Stochastischer Prozess .....	60	FIFO .....	66
Symmetrische Irrfahrt .....	60	LIFO .....	66
<b>T</b>		Non preemptive priority .....	66
t-Verteilung .....	45, 75	Preemptive priority .....	66
Taxi-Problem .....	39, 41	Round Robin .....	66
<b>U</b>		Wiener-Prozess .....	60
Übergangsmatrix .....	63	<b>Z</b>	
unabhängige Ereignisse .....	25	Zentraler Grenzwertsatz .....	32
<b>V</b>		Zufallsvariable .....	13
Varianz .....	27	diskrete .....	13
Stichproben- .....	35	stetige .....	13
Verteilung .....		unabhängige .....	25
Bernoulli- .....	2, 77	Zufallszahl .....	
		Erzeugung .....	18