Ruprecht-Karls-Universität Heidelberg
Institut für Angewandte Mathematik
Professor Dr. Jan JOHANNES

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Skript zur Vorlesung

EINFÜHRUNG IN DIE WAHRSCHEINLICHKEITSTHEORIE UND STATISTIK

Wintersemester 2020/21

Fassung Stand 27. Oktober 2020

Falls Sie **Fehler im Skript** finden, teilen Sie mir diese bitte per eMail an johannes@math.uni-heidelberg.de mit.

MATHEMATIKON, Im Neuenheimer Feld 205, 69120 Heidelberg
Telefon: +49 6221 54.14.190 – Fax: +49 6221 54.14.101
eMail: johannes@math.uni-heidelberg.de
Webseite: sip.math.uni-heidelberg.de

Inhaltsverzeichnis

Kapitel 1

Prolog

Beispiel I: Geschlecht von Konsumierenden

Einem amerikanischen Produzenten eines *kalorienarmen, koffeinhaltigen Erfrischungsgetränkes* wurde das Gerücht zugetragen, dass es mehr weibliche als männliche Konsumierende des Getränkes gibt. Um diesen Rumor zu überprüfen, hat der firmeneigene Verbraucherservice das Geschlecht (*{M,W}*) von 1000 Konsumierenden des Getränkes erhoben. Mit folgendem Resultat:

W W W W M W W W M W M W W M W W W M W M W W M W M M W W W W W W W W W W W M
W W M W W W M W W W W W W M W W M W W M W W W W M M M W W W W W W M W M M W
W W M W W M W W M W W M W M W W W W W M W M W W W M M W W W W W W W M W W W
M W W W W W W W M W W W M W W M M W W W M W M M W W W W W M W W M W M W W W
W W W W W M W W W W W M W W M M W W M W W M M M W W W W W M W M W W W M M M
W W M W W W W W W W W W W W M W M W W W W W W W W M M W W M W W W W W M W W
W W M M W M W M W M W M W W W W M M W W W W M M W M W W W W W M W W W M W M
W W W W W W W W M W W W M W M W W W W W M M M M W M M W W W M W W W W W W M
W W W W W W W W M W W W W W M W M W W W W M W W W M M W W M W M W M W W W M
W W M W M W W W W M M W M W M W W M W W W W W W W M W W W M W W M W W W M W
M W M W W M M W W W W M W M W M W M M M W W M W M M W W W W W W W W W W W W
M W W W W W M W W W W M M W W M W W W W W M M M W W W M M W W W W W W M W W
M W M M W W W W W W M W W W M W M W W W W W M M W W W M W W M W M W W W W W
W M M M W W W M W W M W W M W W W W M M W W W W M W W W W M W M W W W W M W
W W M M W W M W M M W W W M W M M W M W W W W M W M M M M W W M W W W W W M
M W M W W W M W W M W M M M M M M W W W W W W W W M W W W W M W W W W W W W
W M W M M W W W M W M W W W W M M W M M W M W W M M W M M W W W W M W W W W
M W W W W M M W M M W M M W M M W W W M W W M W W W M W M W M W W W W M W M
M W W M W M W W W W W M W W W W M W W W W M W W W M W W W W W W M W W W M M
W W W M W W W W W W W M W W W W W W W M W W W W M W W W W M W W W M W W W M
M W W W W W W W M M M W W W W W M M M W W W W M M W M W W W W W W W W M W W
W W M W W W W W W W W M W W W W M W W W W W W M W W M W W M W W W W W W W W
W W W M M M W M W W W W W M W W W W M M M W W W W W W M M M M M W W W M W M
W M W M M W W M M M W W W W M W W W W M M W W W W M W W W W W W W W W W W W
W W W W W M M W W W M W W M W M W W M W M M M W M W W M W M M W W M W M M M
W W W W W W M W W M W M W W W W M W W M W W W W W M W M M W W W W W W W W W
M W W W W M W M W W W W

(Anzahl W=699)

Der amerikanische Produzent hat daraufhin entschieden, seine Produktpalette um ein koffeinhaltiges Erfrischungsgetränk ohne Zucker zu erweitern. Dabei war das Ziel, eine *neue Marke speziell für Männer* zu kreieren. Zur Kontrolle des Ziels hat der firmeneigene Verbraucherservice wieder das Geschlecht von 1000 Konsumierenden des Getränkes erhoben. Mit folgendem Resultat:

M M M M M M M M W M W M W W M W M M M M M W M M W W W M M M M W W M M M M M
M M M W M W W M M M M M W W M M W M W W W W W W M M W W M M W W W W W M M M
M M M M M M M M M M M W M M M M M M W M M W M M W W W M W W M W M W W M M W
M W W W W M W M M M M W M M W W W M M M M M M W M M W M W M M M W W W M W M
W W W W M W W M M W W M M W M W M M M M M M W M M M M M M M M M W M M W W W
M W M M M W M M M W W M M M W M M M M W M M W M M M M M M M W M M M W M M M
W M M M M W M M M W W W W M M W M M W M M M W W M W M M M M M M W W M W M M
W M W W M M M W W M W W M M M W M M M M M M W M M W M W W M M W M W M M W M
W W W W M M W M M W W M W M W M M W W M M M W W M M M M W M M W M W M M W M
M M W W M M M W M W M M M M M M M W M M M M M M M M W M M W W M M M M M W W
W M M M M M M W W M W M M M W W W M M M W W W M M M W W W M M M W M M M M W
M M M W M W M M M M W M W W M W W M M M W W M W W W M M W W M W M M M M M M
W W W W M M M M W M M W M W M W M M W M M M M M W W W M M W M M M M M W W M
W M M M M W W M M M M W M M M M W M M W M M M W W M M M W M M M M W M W W M
M M M W M M W W M M W M M W M M W M W W W W M M W W M M M M W M W W M M M M
W M W M W M M M W M W W W W M W M M M W M W W W W M M M M W M W M M M M W M
M M M W M M W M M W M W W M M W M W M W M W W M M W W W M M W W M M M M M M
W M W M M W M W W W M M M W M M M W W M M W M M M W M M M M W W W M W M W M
M M M W W W M M M M W W W M W M W W M M M M M M W M W M W M M W W W W M W W
M W W M M M M W W M M W W M W M M M M W M M M M M W M M W W M M M W M M M M
M M M W M M W W M M M W W M M W W M M M M M M M M M W M M W M W M W W M M M
M W W W M M W M M M M W M W M M M M W M M M M W M W M M M W W W W W W W W W
M M W M M W M M M W M W M W M M M M M M M M M W M W M W M W W M W W W W M M
M M M M M M M M M W W W M W W W M M W M M M W M W W M M M W M M M M M M W W
W W M M M M W W W M W W W M M W M M W W W W W M W W W M M W W M W M W M W W
M M W W W M M W W M M M W M W M M W M M M M M W W M W M W M W M M M W W M M
W W W M M W M W W M M M

(Anzahl M=611)

Beispiel II: Qualitätskontrolle

Eine Schraubenherstellerin hat einen Vertrag mit einem Klienten abgeschlossen, in dem sie sich verpflichtet, dass in der nächsten Lieferung *höchstens 1 von 100 Schrauben beschädigt* ist. Der Klient hat 10.000 Beutel a je 50 Schrauben bestellt. Zur Kontrolle der Qualität ihrer Lieferung, zählte die Herstellerin in 100 Beuteln der Lieferung die beschädigten Schrauben. Mit dem folgenden Resultat:

0 2 0 0 0 1 0 1 0 1 0 0 2 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 2 0 1 0 0 0 0 0 0 0 2 0 0 0 0 2 1 0 0 0 1 1 2
0 0 0 1 1 2 0 0 0 0 1 0 4 0 0 1 0 0 0 0 2 1 0 0 1 0 0 0 0 1 0 0 1 0 0 2 0 0 1 0 1 1 0 1

(Summe=46)

Beispiel III: Anzahl

Ein Unternehmen bietet seinen Klienten eine Telefonhotline von Montag bis Freitag zwischen 8:00 Uhr und 18:00 Uhr an. Um den Kundenservice zu verbessern, zählt das Unternehmen eine Woche lang *die Anrufe innerhalb einer Viertelstunde*. Mit dem folgenden Resultat:

18 17 16 8 19 8 12 15 14 11 15 11 13 8 12 14 15 13 8 17 16 15 12 13 10 15 15 13 13 16 13 15 13 16 8 9 11 10 14 15 8 8 15 13 8 15 10 18 8 11 7 16 13 8 15 9 14 9 7 13 10 8 16 15 13 9 14 15 14 18 6 10 6 14 9 7 14 6 14 10 11 16 13 9 14 11 13 11 12 12 11 11 9 15 11 7 20 10 8 12 24 14 11 8 14 13 11 10 12 7 13 11 10 13 14 10 7 16 19 12 6 13 9 15 10 22 12 15 14 16 13 12 14 6 11 13 10 6 11 13 5 15 9 14 12 12 11 9 7 13 13 9 12 9 11 12 19 13 9 14 10 10 11 10 9 11 12 10 10 13 10 11 21 14 12 10 8 11 16 16 18 13 12 11 13 13 14 17 11 9 13 11 11 9 15 15 7 15 20 5 18 14 10 11 20 19 12 10 14 7 9 9 12 7 8 8 6 14 12 13 9 12 12 14 9 10 7 11 14 18 10 12 11 12 13 11 16 14 10 11 15 10 10 14 9 14 15 12 10 13 15 15 8 13 11 10 12 11 11 12 9 3 5 13 13 13 17 13 16 13 9 15 13 17 12 9 7 13 9

(Summe=3348)

Beispiel IV: Wartezeit

Ein Wissenschaftler fährt jeden Tag mit der U-Bahn in Berlin. Während drei Monaten misst er seine tägliche *Wartezeit an der U-Bahn Haltestelle*. Mit dem folgenden Resultat (in Minuten):

7.44 6.21 6.03 8.08 7.07 6.33 9.00 1.28 5.92 2.38 5.89 0.09 1.59 5.57 8.56 7.01 1.54 1.21 0.99 2.30 1.00 7.82 8.53 0.18 2.33 7.81 7.77 9.05 6.86 3.82 1.36 7.39 2.06 0.96 2.28 3.54 4.01 7.59 7.06 7.58 6.59 1.35 8.64 1.01 8.88 2.27 5.96 1.25 6.45 2.45 9.02 5.83 3.80 1.95 9.28 0.29 0.73 3.67 2.74 0.17 9.32 0.02 4.76 7.49 8.49 2.36 2.15 2.06 3.73 2.86 5.35 1.27 6.08 9.50 4.70 5.44 4.11 9.20 2.59 5.98 1.17 5.98 8.40 2.97 2.66 5.25 0.97 3.74 2.79 8.81

(Maximum=9.5)

Beispiel V: Lebensdauer

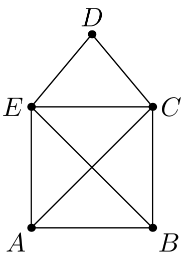
Eine Stadtverwaltung benutzt in ihren Gebäuden 35.000 Glühlampen. Der Hersteller der Glühlampen garantiert, dass die Lebensdauer der Glühlampe mindestens 1000 Stunden beträgt. Um diese Aussage zu überprüfen, führt die Einkaufsabteilung eine Studie mit 100 Glühlampen durch. Mit dem folgenden Resultat (in Stunden):

2109.35 1074.86 28.83 1279.98 156.99 5019.57 1996.70 478.79 999.25 2253.01 465.35 530.50 624.27 4167.61 721.24 134.58 1292.09 174.07 504.60 83.62 2824.01 881.01 42.30 227.83 156.20 13.34 1740.43 23.56 908.03 271.18 23.28 2191.34 357.66 1917.95 357.95 1281.17 183.21 98.11 700.71 820.09 739.14 23.79 923.54 17.19 108.47 467.33 761.10 15.48 5635.45 2714.75 457.12 271.27 155.30 1396.21 644.90 393.85 382.58 1087.93 4547.50 1241.29 807.20 2291.24 2027.31 150.71 5031.31 811.09 1049.09 988.20 1003.60 1264.92 1488.36 1603.13 1923.11 204.41 1765.73 224.21 1011.45 587.16 888.23 1274.92 1222.33 295.25 631.28 48.22 109.15 692.58 11.14 1855.80 377.98 200.43 731.42 823.03 106.78 2233.35 409.38 115.18 1542.88 112.48 1278.54 2345.72

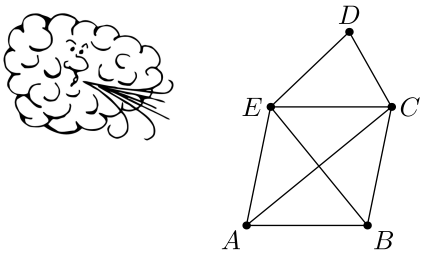
(Mittelwert=1026.37)

Beispiel VI: Messfehler

Nach einem starken Sturm fürchtet der alte Fürst G. Őmetry dass sein Schloss (siehe Bild unten) beschädigt ist. Genauer, er hat den Eindruck dass die Fassade, die vor dem Sturm ein Rechteck war, sich in ein nicht-rechteckiges Parallelogram verformt hat.



intaktes Schloss



beschädigtes Schloss

Um diesen Eindruck objektiv zu überprüfen, entschließt sich der Fürst G. Őmetry die Segmente $\overrightarrow{AC}$ und $\overrightarrow{EB}$ 31 mal zu messen, mit folgendem Resultat für das Segmente $\overrightarrow{AC}$ (in m):

9.60 9.35 9.57 8.65 10.11 10.05 10.82 9.17 11.02 9.02 9.48 9.34 9.27 9.96 9.50 9.31 9.21 10.11 9.96 9.00 7.77 9.44 9.08 9.97 9.68 10.55 8.66 9.14 9.06 9.43 9.56

(Mittelwert=9.51)

und für das Segment $\overrightarrow{EB}$ (in m)

10.38 11.32 9.01 9.51 10.19 10.94 10.26 10.12 10.57 9.93 12.17 11.88 10.30 11.33 10.95 10.23 8.72 10.22 11.42 11.84 10.29 11.55 8.43 9.06 9.45 9.99 10.40 11.00 12.68 9.73 11.36

(Mittelwert=10.49)

Die belgische Abbaye de Rochefort ist berühmt für ihr selbst gebrautes Bier. Die Mönche möchten in einen neuen Abfüllautomaten für 33cl Bierflaschen investieren. Der Hersteller des Automaten gibt eine Genauigkeit der Abfüllung von 0.8 an. Um diese Herstellerangabe zu überprüfen, messen die Mönche das Volumen von 42 zufällig ausgewählten 33cl Bierflaschen, mit folgendem Resultat (in Zentiliter):

32.68 32.19 33.50 33.78 33.95 33.04 33.17 30.62 33.92 32.99 32.13 32.99 33.64 34.41 33.88 32.42 33.08 32.48 32.49 33.52 33.86 33.95 32.80 32.66 32.57 33.93 31.02 32.71 31.89 34.28 33.76 32.79 31.74 33.03 33.57 33.72 31.55 32.57 32.87 32.03 32.47 32.87

(Standardabweichung=0.86)

Laut Hersteller kann der Abfüllautomat mit derselben Genauigkeit auch 75cl Bierflaschen abfüllen. Die Mönche messen zusätzlich das Volumen von 42 zufällig ausgewählten 75cl Bierflaschen, mit folgendem Resultat (in Zentiliter):

75.71 76.28 75.33 75.37 74.84 73.13 76.34 75.61 75.15 74.56 74.30 74.35 75.43 73.70 74.49 75.18 74.23 76.62 73.29 74.76 75.12 74.12 75.57 75.89 75.67 75.68 74.66 73.93 76.08 75.50 75.52 75.59 75.36 74.95 72.84 76.13 75.37 75.53 73.80 74.56 74.44 73.98

(Standardabweichung=0.9)

Kapitel 2

Wahrscheinlichkeitsraum

§01 Stichprobenraum

Ein zufälliges Experiment ist ein Experiment, in dem das Ergebnis nicht mit Sicherheit vorhergesagt werden kann. Um ein zufälliges Experiment mathematisch zu untersuchen, ist es notwendig alle Versuchsausgänge beschreiben zu können.

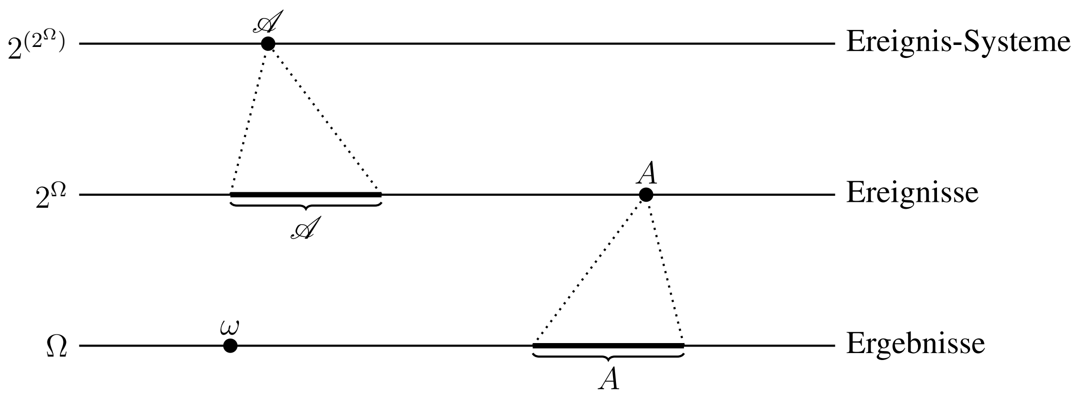
§01.01 **Definition**. Eine nicht-leere Menge Ω aller möglichen Ausgänge eines zufälligen Experiments wird *Ergebnismenge* (*Grundmenge* oder *Stichprobenraum*) genannt. Ein möglicher Versuchsausgang ω des zufälligen Experiments, also ein Element von Ω, kurz $\omega \in \Omega$ heißt *Ergebnis* (*Stichprobe*). Im Folgenden sei stets Ω eine nicht-leere Menge. □

Ein Ereignis ist typischerweise in Form einer Frage gegeben, deren Antwort nur vom Versuchsausgang des zufälligen Experiments abhängt. Von Interesse ist dabei insbesondere, ob das Ereignis eingetreten ist (sich realisiert hat) oder eben nicht.

§01.02 **Definition**. Ein *Ereignis* ist eine Teilmenge der Grundmenge Ω, also ein Element der Potenzmenge 2^Ω von Ω. Für einen Versuchsausgang $\omega \in \Omega$ ist ein Ereignis $A \in 2^\Omega$ eingetreten, wenn $\omega \in A$ gilt. Wir bezeichnen mit $A^c := \Omega \setminus A \in 2^\Omega$ das Komplement von A in Ω. □

§01.03 **Sprechweise**. Die Mengen $\emptyset$, Ω und $\{\omega\}$, $\omega \in \Omega$, heißen das (absolut) *unmögliche Ereignis*, das (absolut) *sichere Ereignis* bzw. ein *Elementarereignis*; Ereignisse A und B mit $A \cap B = \emptyset$ nennt man *unvereinbar* oder *unverträglich*. □

§01.04 **Skizze**. Die folgende Abbildung stellt drei begriffliche Ebenen der Stochastik dar.



Die Abbildung wurde auf der Grundlage von Georgii [2015, Abb.1.1, S.11] erstellt. □

§01.05 **Definition**. Ein Teilmengensystem $\mathscr{A} \subseteq 2^\Omega$ heißt *σ-Algebra*, falls gilt: (a) $\Omega \in \mathscr{A}$; (b) $A^c \in \mathscr{A}$ für alle $A \in \mathscr{A}$; (c) $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$ für alle $A_n \in \mathscr{A}$, $n \in \mathbb{N}$. Das Paar $(\Omega, \mathscr{A})$ wird *messbarer Raum* genannt und $\mathscr{A}$ wird als *Menge der interessierenden Ereignisse* bezeichnet. □

§01.06 **Lemma**. *Es sei $\mathscr{E} \subseteq 2^\Omega$ ein System von Teilmengen von Ω. Dann ist*

$$\boxed{\sigma(\mathscr{E})} := \bigcap \{\mathscr{A} \mid \mathscr{A} \text{ ist } \sigma\text{-Algebra auf } \Omega \text{ und } \mathscr{E} \subseteq \mathscr{A}\}$$

die kleinste σ-Algebra auf Ω, die $\mathscr{E}$ enthält. $\mathscr{E}$ heißt Erzeuger *von $\sigma(\mathscr{E})$ und $\sigma(\mathscr{E})$ die von $\mathscr{E}$* erzeugte σ-Algebra *auf Ω.*

§01.07 **Beweis** von Lemma §01.06. Übungsaufgabe. □

§01.08 **Beispiel**.

(a) Auf jeder nicht-leeren Ergebnismenge Ω existieren die *triviale σ-Algebra* $\{\emptyset, \Omega\}$ sowie die Potenzmenge 2^Ω als σ-Algebren.

(b) Für jedes nicht-leeres $A \subseteq \Omega$ gilt $\sigma(\{A\}) = \{\emptyset, \Omega, A, A^c\}$.

(c) Für eine höchstens abzählbar unendliche (d.h. endlich oder abzählbar unendliche), kurz *abzählbare*, Indexmenge $\mathcal{I}$ sei $\mathscr{E} = \Big\{A_i \in 2^\Omega, i \in \mathcal{I}, \text{ paarweise disjunkt und } \biguplus_{i\in\mathcal{I}} A_i = \Omega\Big\}$ eine Partition von Ω. Dann ist $\sigma(\mathscr{E}) = \Big\{\biguplus_{j\in\mathcal{J}} A_j \mid \mathcal{J} \subseteq \mathcal{I}\Big\}$. □

§01.09 **Bemerkung**. Wie im letzten Beispiel §01.08, bezeichnen wir zur optischen Verdeutlichung die Vereinigung paarweiser disjunkter Mengen mit dem Symbol $\biguplus$. □

§01.10 **Lemma**. *Für jede σ-Algebra $\mathscr{A}$ auf Ω gilt:*

(i) *$\emptyset \in \mathscr{A}$;*

(ii) *Für $A, B \in \mathscr{A}$ sind $A \cup B$, $A \cap B$, $A \setminus B := A \cap B^c$ und $A\Delta B := (A \setminus B) \cup (B \setminus A)$ Elemente von $\mathscr{A}$;*

(iii) *Für eine abzählbare Indexmenge $\mathcal{I}$ und $\{A_i : i \in \mathcal{I}\} \subseteq \mathscr{A}$ gilt $\bigcap_{i\in\mathcal{I}} A_i \in \mathscr{A}$.*

§01.11 **Beweis** von Lemma §01.10. Übungsaufgabe. □

§01.12 **Erinnerung.**. Ein Folge $(x_n)_{n\in\mathbb{N}}$ aus $\mathbb{R}$ heißt monoton wachsend (bzw. fallend), wenn $x_n \leqslant x_{n+1}$ (bzw. $x_{n+1} \leqslant x_n$) für alle $n \in \mathbb{N}$ gilt. Ist eine monotone wachsende (bzw. fallende) Folge konvergent, etwa $x = \lim_{n\to\infty} x_n$, so schreiben wir kurz $x_n \uparrow x$ (bzw. $x_n \downarrow x$). □

§01.13 **Definition**. Sei $A_n \subseteq \Omega$ für $n \in \mathbb{N}$. Die Folge $(A_n)_{n\in\mathbb{N}}$ heißt *monoton wachsend* (bzw. *fallend*), wenn $A_n \subseteq A_{n+1}$ (bzw. $A_{n+1} \subseteq A_n$) für alle $n \in \mathbb{N}$ gilt. Weiterhin heißen

$$\boxed{A_\star := \liminf_{n\to\infty} A_n} := \bigcup_{n\geqslant 1}\bigcap_{m\geqslant n} A_m := \bigcup\Big\{\bigcap\{A_m \mid m \in \mathbb{N} \wedge m \geqslant n\} \mid n \in \mathbb{N}\Big\} \quad \text{und}$$

$$\boxed{A^\star := \limsup_{n\to\infty} A_n} := \bigcap_{n\geqslant 1}\bigcup_{m\geqslant n} A_m.$$

Limes inferior bzw. *Limes superior* der Folge $(A_n)_{n\in\mathbb{N}}$. Die Folge $(A_n)_{j\in\mathbb{N}}$ heißt *konvergent*, wenn $A_\star = A^\star =: A$ gilt. In diesem Fall schreiben wir kurz $\boxed{\lim_{n\to\infty} A_n} = A$. □

§01.14 **Bemerkung**.

(i) Jede *monoton wachsende* (bzw. *fallende*) Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen von Ω ist konvergent mit $A := \lim_{n\to\infty} A_n = \bigcup_{n\in\mathbb{N}} A_n$ (bzw. $A := \lim_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} A_n$). In diesem Fall schreiben wir kurz $A_n \uparrow A$ (bzw. $A_n \downarrow A$).

(ii) Für den Limes inferior bzw. superior einer Folge $(A_n)_{n\in\mathbb{N}}$ gilt

$$A_\star = \liminf_{n\to\infty} A_n = \{\omega \in \Omega : |\{n \in \mathbb{N} : \omega \notin A_n\}| < \infty\} \quad \text{bzw.}$$
$$A^\star = \limsup_{n\to\infty} A_n = \{\omega \in \Omega : |\{n \in \mathbb{N} : \omega \in A_n\}| = \infty\}.$$

In anderen Worten, $A_\star$ ist also das Ereignis, dass *schließlich alle der A_n eintreten*. $A^\star$ ist hingegen das Ereignis, dass *unendlich viele der A_n eintreten*. Insbesondere gilt $A_\star = \liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n = A^\star$. □

§01.15 **Lemma**. *Sei $(\Omega, \mathscr{A})$ ein messbarer Raum und $(A_n)_{n\in\mathbb{N}}$ eine Folge von Teilmengen aus $\mathscr{A}$. Dann gilt:*

(i) $\liminf_{n\to\infty} A_n \in \mathscr{A}$ *und* $\limsup_{n\to\infty} A_n \in \mathscr{A}$*;*

(ii) *Falls $A := \lim_{n\to\infty} A_n$ existiert, dann ist $A \in \mathscr{A}$.*

§01.16 **Beweis** von Lemma §01.15. Übungsaufgabe. □

§01.17 **Definition**. Es sei $\mathcal{S}$ ein metrischer (oder topologischer) Raum und $\mathscr{O}$ das System der offenen Teilmengen von $\mathcal{S}$. Dann heißt die von den offen Mengen $\mathscr{O}$ erzeugte σ-Algebra $\mathscr{B}_S := \sigma(\mathscr{O})$ die *Borel-σ-Algebra* über S. Die Elemente von $\mathscr{B}_S$ heißen *Borel-Mengen*. □

§01.18 **Bemerkung**. Häufig sind wir an der Borel-σ-Algebra $\mathscr{B}^n := \mathscr{B}_{\mathbb{R}^n}$ über $\mathbb{R}^n$ interessiert, wobei $\mathbb{R}^n$ versehen ist mit dem Euklidischen Abstand $d(x,y) = \|x-y\| = \sqrt{\sum_{i\in[\![n]\!]}(x_i-y_i)^2}$ für $x = (x_i)_{i\in[\![n]\!]}$, $y = (y_i)_{i\in[\![n]\!]} \in \mathbb{R}^n$ mit $[\![n]\!] := [1,n] \cap \mathbb{N}$. □.

§01.19 **Schreibweise**. Wir schreiben $[\![a,b]\!] := [a,b] \cap \mathbb{Z}$ für $a,b \in \mathbb{R}$ mit $a \leqslant b$. Wir setzen $\mathbb{R}^+ := [0,\infty)$, $\mathbb{Q}^+ := [0,\infty) \cap \mathbb{Q}$, $\mathbb{R}^+_{\setminus 0} := (0,\infty)$, $\mathbb{Q}^+_{\setminus 0} := (0,\infty) \cap \mathbb{Q}$, $\mathbb{R}_{\setminus 0} := \mathbb{R} \setminus \{0\}$, sowie $\mathbb{Q}_{\setminus 0} := \mathbb{Q} \setminus \{0\}$. Für $a, b \in \mathbb{R}^n$ schreiben wir $a < b$, wenn $a_i < b_i$ für alle $i \in [\![n]\!]$ gilt. Für $a < b$, definieren wir den offenen *Quader* als das Kartesische Produkt $(a,b) := \times_{i\in[\![n]\!]}(a_i,b_i) := (a_1,b_1) \times (a_2,b_2) \times \cdots \times (a_n,b_n)$. Analog, sind $[a,b]$, $(a,b]$ sowie $[a,b)$ definiert. Weiterhin, sei $(-\infty,b) := \times_{i\in[\![n]\!]}(-\infty,b_i)$ und analog $(-\infty,b]$ definiert. □.

§01.20 **Bemerkung**. Wie in Schreibweise §01.19 bezeichnet $\times_{i\in\mathcal{I}} \mathcal{S}_i$ das Kartesische Produkt der Mengen $\mathcal{S}_i$, $i \in \mathcal{I}$, und für $s \in \times_{i\in\mathcal{I}} \mathcal{S}_i$ bezeichnen s_i, $i \in \mathcal{I}$, stets die Komponenten von $s = (s_i)_{i\in\mathcal{I}}$. □

§01.21 **Satz**. *Die Borel-σ-Algebra $\mathscr{B}^n$ über $\mathbb{R}^n$ wird auch erzeugt von folgenden Mengensystemen:* (i) $\mathscr{E}_1 := \{A \subseteq \mathbb{R}^n \,|\, A \text{ ist abgeschlossen}\}$; (ii) $\mathscr{E}_2 := \{A \subseteq \mathbb{R}^n \,|\, A \text{ ist kompakt}\}$; (iii) $\mathscr{E}_3 := \{(a,b) \,|\, a,b \in \mathbb{Q}^n, a < b\}$; (iv) $\mathscr{E}_4 := \{[a,b] \,|\, a,b \in \mathbb{Q}^n, a < b\}$; (v) $\mathscr{E}_5 := \{(a,b] \,|\, a,b \in \mathbb{Q}^n, a < b\}$; (vi) $\mathscr{E}_6 := \{[a,b) \,|\, a,b \in \mathbb{Q}^n, a < b\}$; (vii) $\mathscr{E}_7 := \{(-\infty,b] \,|\, b \in \mathbb{Q}^n\}$; (viii) $\mathscr{E}_8 := \{(-\infty,b) \,|\, b \in \mathbb{Q}^n\}$; (ix) $\mathscr{E}_9 := \{(a,\infty) \,|\, a \in \mathbb{Q}^n\}$ *und* (x) $\mathscr{E}_{10} := \{[a,\infty) \,|\, a \in \mathbb{Q}^n\}$.

§01.22 **Beweis** von Satz §01.21. In der Vorlesung Wahrscheinlichkeitstheorie 1. □

§01.23 **Lemma**. *Seien $(\Omega, \mathscr{A})$ ein messbarer Raum und $B \subseteq \Omega$ eine nicht-leere Teilmenge. Dann ist $\mathscr{A}\big|_B := \mathscr{A} \cap B := \{A \cap B \,|\, A \in \mathscr{A}\}$ eine σ-Algebra über B, die sogenannte* Spur *oder* Einschränkung *von $\mathscr{A}$ auf B. Des Weiteren, für $\mathscr{E} \in 2^\Omega$ gilt $\sigma(\mathscr{E})\big|_B = \sigma(\mathscr{E}\big|_B)$.*

§01.24 **Beweis** von Lemma §01.23. Übungsaufgabe. □

§01.25 **Schreibweisen.** Wir bezeichnen mit $\overline{\mathscr{B}} := \mathscr{B}_{\overline{\mathbb{R}}}$ die Borel-σ-Algebra über der kompaktifizierten Zahlengerade $\overline{\mathbb{R}} := [-\infty, \infty]$, wobei die Mengen $\{-\infty\}$, $\{\infty\}$ und $\mathbb{R}$ in $\overline{\mathbb{R}}$ abgeschlossen bzw. offen, also Borel-Mengen sind. Insbesondere, ist $\mathscr{B} := \mathscr{B}_{\mathbb{R}} = \overline{\mathscr{B}} \cap \mathbb{R}$ die Borel-σ-Algebra über $\mathbb{R}$. Weiterhin schreiben wir $\overline{\mathscr{B}}^+ := \overline{\mathscr{B}} \cap \overline{\mathbb{R}}^+$, $\mathscr{B}^+ := \mathscr{B} \cap \mathbb{R}^+$ und $\mathscr{B}^+_{\setminus 0} := \mathscr{B} \cap \mathbb{R}^+_{\setminus 0}$. □.

§02 Wahrscheinlichkeit

In einem zufälligem Experiment, das durch einen messbaren Raum $(\Omega, \mathscr{A})$ beschrieben wird, möchten wir den interessierenden Ereignissen eine Wahrscheinlichkeit zu ordnen. Die folgende Definition wurde 1933 von dem russischen Mathematiker A.N. Kolmogorov eingeführt.

§02.01 **Definition.** Sei $(\Omega, \mathscr{A})$ ein messbarer Raum. Eine Abbildung $\mathbb{P} : \mathscr{A} \to [0, 1]$ heißt *Wahrscheinlichkeitsmaß* (*Wahrscheinlichkeitsverteilung* oder kurz *Verteilung*) auf $\mathscr{A}$, wenn sie folgenden Bedingungen genügt:

(a) $\mathbb{P}(\Omega) = 1$; (*normiert*)

(b) $\mathbb{P}(\biguplus_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$ für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Ereignisse aus $\mathscr{A}$, d.h. $A_n \cap A_m = \emptyset$ für alle $n, m \in \mathbb{N}$ mit $n \neq m$. (*σ-additiv*) □

Wir bezeichnen mit $\mathcal{W} := \mathcal{W}(\mathscr{A})$ die Menge aller Wahrscheinlichkeitsmaße auf $\mathscr{A}$. Ein Tripel $(\Omega, \mathscr{A}, \mathbb{P})$ bestehend aus einer Ergebnismenge Ω, einer σ-Algebra $\mathscr{A}$ interessierender Ereignisse sowie einem Wahrscheinlichkeitsmaß $\mathbb{P} \in \mathcal{W}$ wird *Wahrscheinlichkeitsraum* genannt. □

§02.02 **Sprechweise.** Ereignisse $A, B \in \mathscr{A}$ mit $\mathbb{P}(A) = 0$ und $\mathbb{P}(B) = 1$ werden *Nullmenge* bzw. *Einsmenge* genannt. □

§02.03 **Beispiel.** Sei $(\Omega, \mathscr{A})$ ein messbarer Raum. Für $A \subseteq \Omega$ bezeichnet $\mathbb{1}_A : \Omega \to \{0, 1\}$ mit $\mathbb{1}_A^{-1}(\{1\}) = A$ und $\mathbb{1}_A^{-1}(\{0\}) = A^c$ die *Indikatorfunktion* auf A. Für jedes $\omega \in \Omega$ ist das *Einpunkt*-oder *Diracmaß* $\delta_\omega : \mathscr{A} \to \{0, 1\}$ mit $\delta_\omega(A) := \mathbb{1}_A(\omega)$ für alle $A \in \mathscr{A}$ ein Wahrscheinlichkeitsmaß aus $\mathcal{W}(\mathscr{A})$. Ist $(\mathbb{P}_n)_{n\in\mathbb{N}}$ eine Folge von Wahrscheinlichkeitsmaße in $\mathcal{W}$, so ist auch jede Konvexkombination $\sum_{n\in\mathbb{N}} w_n \mathbb{P}_n$ mit $w_n \geqslant 0$, $n \in \mathbb{N}$, und $\sum_{n\in\mathbb{N}} w_n = 1$ in $\mathcal{W}$. Die Diracmaße bilden Extremalpunkte der konvexen Menge aller Wahrscheinlichkeitsmaße. □

§02.04 **Definition.** Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{A})$ heißt ein Element $\omega \in \Omega$ mit $\{\omega\} \in \mathscr{A}$ und $\mathbb{P}(\{\omega\}) > 0$ *Atom*. Die Wahrscheinlichkeit $\mathbb{P}(\{\omega\})$ wird *Masse* des Atoms ω genannt.

§02.05 **Lemma.** *Ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{A})$ besitzt abzählbar viele Atome.*

§02.06 **Beweis** von Lemma §02.05. Da es höchstens n Atome mit Masse gleich oder größer $1/n$ gibt, kann die Anzahl der Atome höchstens abzählbar unendlich sein. □

§02.07 **Schreibweisen.** Für $a, b \in \mathbb{R}$ schreiben wir kurz $a \vee b := \max\{a, b\}$ sowie $a \wedge b := \min\{a, b\}$. □

§02.08 **Lemma.** *Für jedes Wahrscheinlichkeitsmaß $\mathbb{P} \in \mathcal{W}(\mathscr{A})$ gilt:*

(i) $\mathbb{P}(\emptyset) = 0$; (unmögliches Ereignis)

(ii) *Für jedes $A \in \mathscr{A}$ gilt $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;*

(iii) *Für alle $A, B \in \mathscr{A}$ gilt* *(Booles Regeln)*
$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$,
$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$,
$\mathbb{P}(A \Delta B) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B)$;
für disjunkte A und B, d.h. $A \cap B = \emptyset$, also $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$; *(additiv)*

(iv) *Für $n \in \mathbb{N}$ und $A_1, \dots, A_n \in \mathscr{A}$ gilt* *(Poincarés Formel)*
$\mathbb{P}\big(\bigcup_{i \in [\![n]\!]} A_i\big) = \sum_{\emptyset \neq \mathcal{I} \subseteq [\![n]\!]} (-1)^{|\mathcal{I}|-1} \mathbb{P}\big(\bigcap_{i \in \mathcal{I}} A_i\big)$.;

(v) *Für alle $A, B \in \mathscr{A}$ mit $A \subseteq B$ gilt $0 \leqslant \mathbb{P}(A) \leqslant \mathbb{P}(B) \leqslant 1$* *(Isotonie);*

(vi) *Für $A, B \in \mathscr{A}$ gilt:* *(Bonferroni Ungleichungen)*
$\mathbb{P}(A) \vee \mathbb{P}(B) \leqslant \mathbb{P}(A \cup B) \leqslant \{\mathbb{P}(A) + \mathbb{P}(B)\} \wedge 1$;
$\{\mathbb{P}(A) + \mathbb{P}(B) - 1\} \vee 0 \leqslant \mathbb{P}(A \cap B) \leqslant \mathbb{P}(A) \wedge \mathbb{P}(B)$;
$|\mathbb{P}(A) - \mathbb{P}(B)| \leqslant \mathbb{P}(A \Delta B)$.

§02.09 **Beweis** von Lemma §02.08. Übungsaufgabe. □

§02.10 **Proposition**. *Für jedes $\mathbb{P} \in \mathcal{W}(\mathscr{A})$ und jede Folge $(A_n)_{n \in \mathbb{N}}$ von Teilmengen aus $\mathscr{A}$ gilt:*

(i) $\mathbb{P}(\cup_{n \in \mathbb{N}} A_n) \leqslant \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$; *(subadditiv)*

(ii) *Falls $A_n \downarrow \emptyset$, dann gilt $\mathbb{P}(A_n) \downarrow 0$;* *($\sigma$-stetig)*
falls $A_n \uparrow A$, dann gilt $\mathbb{P}(A_n) \uparrow \mathbb{P}(A)$;

(iii) *Falls $\lim_{n \to \infty} A_n = A$, dann gilt $\lim_{n \to \infty} \mathbb{P}(A_n \Delta A) = 0$ und somit $\lim_{n \to \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.*

§02.11 **Beweis** von Proposition §02.10. (iii) in der Vorlesung und (i)-(ii) Übungsaufgabe. □

§02.12 **Lemma**. *Jede normierte, additive Mengenfunktion $\mathbb{P} : \mathscr{A} \to [0, 1]$, die σ-stetig ist, ist auch σ-additiv und damit ein Wahrscheinlichkeitsmaß, also $\mathbb{P} \in \mathcal{W}(\mathscr{A})$.*

§02.13 **Beweis** von Lemma §02.12. In der Vorlesung. □

§02.14 **Bemerkung**. Die letzte Aussage in Lemma §02.12 erlaubt eine alternative aber zu Definition §02.01 äquivalente Definition eines Wahrscheinlichkeitsmaßes, d.h. eine Mengenfunktion $\mathbb{P} : \mathscr{A} \to [0, 1]$ ist ein Wahrscheinlichkeitsmaß wenn es (a) normiert, (b) additiv und (c) σ-stetig ist. □

§02.15 **Korollar** *(Ungleichungen von Boole)*. *Für jedes $\mathbb{P} \in \mathcal{W}(\mathscr{A})$ und jede Folge $(A_n)_{n \in \mathbb{N}}$ von Teilmengen aus $\mathscr{A}$ gilt:*

$$\mathbb{P}\Big(\bigcap_{n \in \mathbb{N}} A_n\Big) \leqslant \inf_{n \in \mathbb{N}} \mathbb{P}(A_n) \leqslant \sup_{n \in \mathbb{N}} \mathbb{P}(A_n) \leqslant \mathbb{P}\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) \leqslant \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$
$$\mathbb{P}(\liminf_{n \to \infty} A_n) \leqslant \liminf_{n \to \infty} \mathbb{P}(A_n) \leqslant \limsup_{n \to \infty} \mathbb{P}(A_n) \leqslant \mathbb{P}(\limsup_{n \to \infty} A_n).$$

§02.16 **Beweis** von Lemma §02.12. Übungsaufgabe. □

§02.17 **Definition**. Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathscr{B})$ wird $\mathbb{F} : \mathbb{R} \to [0, 1]$ mit $x \mapsto \mathbb{F}(x) := \mathbb{P}((-\infty, x])$ die zugehörige *Verteilungsfunktion* genannt. □

§02.18 **Lemma**. *Für jede Verteilungsfunktion $\mathbb{F}$ eines Wahrscheinlichkeitsmaß $\mathbb{P} \in \mathcal{W}(\mathscr{B})$ gilt:*

(i) *$\mathbb{F}$ ist monoton wachsend, d.h., für alle $x, y \in \mathbb{R}$ mit $x \leqslant y$ gilt $\mathbb{F}(x) \leqslant \mathbb{F}(y)$.*

(ii) *$\mathbb{F}$ ist rechtsstetig, d.h., für alle $x_n \downarrow x$ in $\mathbb{R}$ gilt $\mathbb{F}(x_n) \downarrow \mathbb{F}(x)$.*

(iii) *$\mathbb{F}$ besitzt einen linksseitigen Grenzwert, das heißt für alle $x_n \uparrow x$ in $\mathbb{R}$ gilt $\mathbb{F}(x_n) \uparrow \mathbb{F}(x-) := \mathbb{P}((-\infty, x))$.*

(iv) *Die Sprünge von $\mathbb{F}$ korrespondieren mit den Atomen von $\mathbb{P}$, d.h. $\mathbb{F}(x) - \mathbb{F}(x-) = \mathbb{P}(\{x\})$ für alle $x \in \mathbb{R}$.*

(v) *Es gilt $\lim_{x\to-\infty} \mathbb{F}(x) = 0$ und $\lim_{x\to\infty} \mathbb{F}(x) = 1$.*

§02.19 **Beweis** von Lemma §02.18. Übungsaufgabe. □

§02.20 **Bemerkung**. Aufgrund von Lemma §02.05 und Lemma §02.18 (iv) ist die Anzahl der Unstetigkeitsstellen einer Verteilungsfunktion $\mathbb{F}$ abzählbar. □

§03 Dynkin'scher π-λ-Satz

Seien $(\Omega, \mathscr{A})$ ein messbarer Raum und $\mathscr{E} \subseteq 2^\Omega$ ein Erzeuger von $\mathscr{A}$, also $\mathscr{A} = \sigma(\mathscr{E})$, so stellt sich die Frage: Ist ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathscr{A}$ schon eindeutig festgelegt durch seine Werte auf $\mathscr{E}$? Im Allgemeinen ist die Antwort „nein". Die Antwort ist aber „ja", wenn das Mengensystem schnittstabil ist. Dies ist eine Schlussfolgerung aus dem Dynkin'schen π-λ-Satz.

§03.01 **Definition**. Ein Teilmengensystem $\mathscr{E}$ aus 2^Ω heißt

(a) *$\cap$-stabil* (sprich: *schnittstabil*) oder ein *π-System*, falls für je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \cap B \in \mathscr{E}$;

(b) *σ-$\cap$-stabil* (sprich: *sigma-schnittstabil*), falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen aus $\mathscr{A}$ gilt, dass auch $\cap_{n\in\mathbb{N}} A_n \in \mathscr{E}$;

(c) *$\cup$-stabil* (sprich: *vereinigungsstabil*), falls für je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \cup B \in \mathscr{E}$;

(d) *σ-$\cup$-stabil* (sprich: *sigma-vereinigungsstabil*), falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen aus $\mathscr{A}$ gilt, dass auch $\cup_{n\in\mathbb{N}} A_n \in \mathscr{E}$;

(e) *$\setminus$-stabil* (sprich: *differenzmengenstabil*), falls je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \setminus B \in \mathscr{E}$;

(f) *komplementstabil*, falls mit jeder Menge $A \in \mathscr{E}$ auch $A^c \in \mathscr{E}$ gilt. □

§03.02 **Bemerkung**.

(i) Ein Teilmengensystem $\mathscr{E}$ aus 2^Ω mit $\Omega \in \mathscr{E}$, das komplementstabil und σ-$\cup$-stabil ist, ist somit eine σ-Algebra.

(ii) Für ein komplementstabiles Teilmengensystem $\mathscr{E}$ aus 2^Ω folgen aus den de Morgan'schen Regeln die Äquivalenzen von $\cup$-stabil und $\cap$-stabil als auch von σ-$\cup$-stabil und σ-$\cap$-stabil. □

§03.03 **Definition**. Ein Teilmengensystem $\mathscr{D}$ aus 2^Ω heißt Dynkin-System (oder λ-System), wenn es folgenden Bedingungen genügt:

(λ_1) $\Omega \in \mathscr{D}$;

(λ_2) $\mathscr{D}$ ist komplementstabil;

(λ_3) für je abzählbar viele, paarweise disjunkte Menge $A_1, A_2, \dots$ in $\mathscr{D}$ gilt $\biguplus_{n=1}^{\infty} A_n \in \mathscr{D}$. □

§03.04 **Bemerkung**. Die Bedingung (λ_2), das $\mathscr{D}$ komplementstabil ist; kann äquivalent ersetzt werden durch die scheinbar stärkere Bedingung

(λ_2') für alle $A, B \in \mathscr{D}$ mit $A \subseteq B$ gilt $B \setminus A \in \mathscr{D}$.

Da jedes Dynkin-System auch (λ_2') erfüllt. Denn für $A, B \in \mathscr{D}$ mit $A \subseteq B$ sind A und B^c disjunkt und es gilt $B \setminus A = (A \uplus B^c)^c \in \mathscr{D}$. □

§03.5 **Anmerkung**. Jede σ-Algebra ist ein Dynkin-System. Die Umkehrung gilt nicht, da (λ_3) nur für Folgen paarweise disjunkter Ereignisse gefordert ist. Zum Beispiel für $\Omega = \{1, 2, 3, 4\}$ ist $\mathscr{D} = \{\emptyset, \{1, 2\}, \{1, 4\}, \{2, 3\}, \{3, 4\}, \Omega\}$ ein Dynkin-System aber keine σ-Algebra. Der Unterschied ist allerdings nicht sehr groß. □

§03.06 **Lemma**. *Ein Dynkin-System $\mathscr{D} \subseteq 2^{\Omega}$ ist genau dann $\cap$-stabil, wenn es eine σ-Algebra ist.*

§03.07 **Beweis** von Lemma §03.06. In der Vorlesung. □

§03.08 **Korollar**. *Für ein Teilmengensystem $\mathscr{E}$ aus 2^{Ω} ist*

$$\boxed{\delta(\mathscr{E})} := \bigcap \{\mathscr{D} \mid \mathscr{D} \text{ ist Dynkin-System auf } \Omega \text{ und } \mathscr{E} \subseteq \mathscr{D}\}$$

das kleinstes Dynkin-System auf Ω, das $\mathscr{E}$ enthält. $\mathscr{E}$ heißt Erzeuger von $\delta(\mathscr{E})$ und $\delta(\mathscr{E})$ das von $\mathscr{E}$ erzeugte Dynkin-System auf Ω.

§03.09 **Beweis** von Korollar §03.08. Der Beweis §01.07 lässt sich direkt auf Dynkin-Systeme übertragen. □

§03.10 **Bemerkung**. Da jede σ-Algebra ein Dynkin-System ist, gilt stets $\delta(\mathscr{E}) \subseteq \sigma(\mathscr{E})$. □

§03.11 **π-λ-Satz**. *Für jedes $\cap$-stabile $\mathscr{E}$ ist das erzeugte Dynkin-System $\delta(\mathscr{E})$ auch $\cap$-stabil.*

§03.12 **Beweis** von Satz §03.11. In der Vorlesung. □

§03.13 **Korollar** *(Dynkin'scher π-λ-Satz)*. *Für jedes $\cap$-stabile $\mathscr{E}$ gilt $\sigma(\mathscr{E}) = \delta(\mathscr{E})$.*

§03.14 **Beweis** von Korollar §03.13. „$\supseteq$" Dies ist klar nach Bemerkung §03.10.
„$\subseteq$" Nach Satz §03.11 ist das Dynkin-System $\delta(\mathscr{E})$ $\cap$-stabil, und somit nach Lemma §03.06 auch eine σ-Algebra, die definitionsgemäß $\mathscr{E}$ enthält. □

§03.15 **Korollar**. *Seien $\mathscr{D}$ ein Dynkin-System und $\mathscr{E} \subseteq \mathscr{D}$ $\cap$-stabil. Dann gilt $\sigma(\mathscr{E}) \subseteq \mathscr{D}$.*

§03.16 **Beweis** von Korollar §03.15. Für das erzeugte Dynkin-System $\delta(\mathscr{E})$ gilt definitionsgemäß $\delta(\mathscr{E}) \subseteq \mathscr{D}$. Da $\mathscr{E}$ $\cap$-stabil ist, folgt $\sigma(\mathscr{E}) = \delta(\mathscr{E})$ aus Korollar §03.13, also auch $\sigma(\mathscr{E}) \subseteq \mathscr{D}$. □

§03.17 **Beweisstrategie.**. Möchten wir zeigen, dass alle Ereignisse einer σ-Algebra eine bestimmte Eigenschaft, sagen wir (R) besitzen, so ist eine häufig angewendete Beweisstrategie:

(Schritt 1) Zeige, dass $\{A \in \mathscr{A} : A \text{ erfüllt } (R)\}$ ein Dynkin-System ist;

(Schritt 2) Finde einen $\cap$-stabilen Erzeuger $\mathscr{E}$ von $\mathscr{A}$, d.h. ein π-System $\mathscr{E}$ mit $\mathscr{A} = \sigma(\mathscr{E})$;

(Schritt 3) Zeige, dass alle Elemente aus $\mathscr{E}$ die Eigenschaft (R) besitzen.

Nach Korollar §03.15 besitzen dann alle Ereignisse aus $\mathscr{A} = \sigma(\mathscr{E})$ die Eigenschaft (R). □

§03.18 **Satz** *(Eindeutigkeit eines Wahrscheinlichkeitsmaßes). Seien $(\Omega, \mathscr{A})$ ein messbarer Raum, $\mathscr{E}$ ein $\cap$-stabiler Erzeuger von $\mathscr{A}$ und $\mathbb{P}$, $\widetilde{\mathbb{P}}$ Wahrscheinlichkeitsmaße auf $\mathscr{A}$. Falls die Einschränkungen $\mathbb{P}\big|_{\mathscr{E}}$ und $\widetilde{\mathbb{P}}\big|_{\mathscr{E}}$ von $\mathbb{P}$ bzw. $\widetilde{\mathbb{P}}$ auf $\mathscr{E}$ übereinstimmen, d.h. $\mathbb{P}(E) = \widetilde{\mathbb{P}}(E)$ für alle $E \in \mathscr{E}$ gilt, dann stimmen auch $\mathbb{P}$ und $\widetilde{\mathbb{P}}$ überein, $\mathbb{P} = \widetilde{\mathbb{P}}$.*

§03.19 **Beweis** von Satz §03.18. Übungsaufgabe unter Verwendung der Beweisstrategie §03.17 mit $(R) =$„$\mathbb{P}(A) = \widetilde{\mathbb{P}}(A)$“. □

§03.20 **Korollar** *(Eindeutigkeit der Verteilungsfunktion). Ist $\mathbb{F} : \mathbb{R} \to [0, 1]$ monoton wachsend, rechtsstetig mit $\lim_{x\to-\infty} \mathbb{F}(x) = 0$ und $\lim_{x\to\infty} \mathbb{F}(x) = 1$, so existiert genau ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathscr{B})$ mit $\mathbb{P}((x, y]) = \mathbb{F}(y) - \mathbb{F}(x)$ für alle $x, y \in \mathbb{R}$ mit $x < y$. Insbesondere ist $\mathbb{F}$ die Verteilungsfunktion von $\mathbb{P}$.*

§03.21 **Beweis** von Korollar §03.20. Die *Existenz* wird in der Vorlesung Wahrscheinlichkeitstheorie 1 gezeigt, die *Eindeutigkeit* folgt direkt aus Satz §03.18 mit $\cap$-stabilen $\mathscr{E} = \{(-\infty, x] \,|\, x \in \mathbb{R}\}$. □

§04 Diskreter Wahrscheinlichkeitsraum

§04.01 **Definition**. Ist Ω eine abzählbare Menge, so wird ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathscr{A} = 2^\Omega$ *diskret* genannt und $(\Omega, \mathscr{A}, \mathbb{P})$ heißt *diskreter Wahrscheinlichkeitsraum*. Die Abbildung $\boxed{\mathbb{p}} : \Omega \to [0, 1]$ mit $\omega \mapsto \mathbb{p}(\omega) := \mathbb{P}(\{\omega\})$ wird *Zähldichte* des Wahrscheinlichkeitsmaßes $\mathbb{P}$ genannt. □

§04.02 **Lemma**.

(i) *Ist $(\Omega, \mathscr{A}, \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum, so ist $\mathbb{P}$ eindeutig durch seine Zähldichte $\mathbb{p}$ festgelegt.*

(ii) *Ist andererseits Ω eine abzählbare Menge und besitzt $\mathbb{p} : \Omega \to [0, 1]$ die Eigenschaft $\sum_{\omega\in\Omega} \mathbb{p}(\omega) = 1$, so wird durch $A \mapsto \mathbb{P}(A) := \sum_{\omega\in A} \mathbb{p}(\omega)$ ein diskretes Wahrscheinlichkeitsmaß $\mathbb{P} : 2^\Omega \to [0, 1]$ definiert, dessen Zähldichte gerade $\mathbb{p}$ ist.*

§04.03 **Beweis** von Lemma §04.02. Übungsaufgabe. □

§04.04 **Lemma**. *In einer Urne liegen N Kugeln mit den Aufschriften $1, 2, \ldots, N$. Es werden n Kugeln gezogen. Dann gilt für die Grundmenge und die Anzahl verschiedenen Ergebnisse:*

Ziehen mit Zurücklegen, mit Betrachtung der Reihenfolge:
$\Omega_1 = \{(k_i)_{i\in[\![n]\!]} \,|\, k_i \in [\![N]\!], \forall\, i \in [\![n]\!]\} = [\![N]\!]^n$, $|\Omega_1| = N^n$;

Ziehen ohne Zurücklegen, mit Betrachtung der Reihenfolge für $n \leqslant N$:
$\Omega_2 = \{(k_i)_{i\in[\![n]\!]} \in [\![N]\!]^n \,|\, k_1, \ldots, k_n \text{ paarweise verschieden}\}$, $|\Omega_2| = \frac{N!}{(N-n)!}$;

Ziehen ohne Zurücklegen, ohne Betrachtung der Reihenfolge für $n \leqslant N$:
$\Omega_3 = \{A \subseteq [\![N]\!] \,|\, |A| = n\}$, $|\Omega_3| = \binom{N}{n}$;

Ziehen mit Zurücklegen, ohne Betrachtung der Reihenfolge:
$\Omega_4 = \{(k_i)_{i\in[\![n]\!]} \in [\![N]\!]^n \,|\, k_1 \leqslant k_2 \leqslant \cdots \leqslant k_n\}$, $|\Omega_4| = \binom{N+n-1}{n}$.

§04.05 **Beweis** von Lemma §04.04. In der Vorlesung. □

§04.06 **Beispiel.**

(a) Die Menge aller Ergebnisse im Lotto „6 aus 49“ ist Ω_3 mit $N = 49$ und $n = 6$, so dass es $\binom{49}{6} = 13983816$ verschiedene Ergebnisse gibt.

(b) Wie groß ist die Wahrscheinlichkeit, dass in dem Hörsaal keine zwei Personen am selben Tag Geburtstag haben? Nehmen wir an, dass Jahr hat 365 Tage, so ist die Menge aller Geburtstagskombinationen für $n \in \mathbb{N}$ Personen gerade $\Omega_1 = [\![N]\!]^n$ mit $N = 365$. Das Ereignis „keine zwei Personen haben am selben Tag Geburtstag“ entspricht $\Omega_2 = \{(k_i)_{i\in[\![n]\!]} \in [\![N]\!]^n \mid k_1, \dots, k_n \text{ paarweise verschieden}\}$. Unter der Annahme einer Gleichverteilung folgt $\frac{|\Omega_2|}{|\Omega_3|} = \frac{N!}{(N-n)!N^n} = 1 \cdot (1 - \frac{1}{N}) \cdots (1 - \frac{n-1}{N})$. Für $n = 25$, $n = 50$ und $n = 100$ ergibt sich approximativ 0.431, 0.03 und $3.1 \cdot 10^{-7}$. □

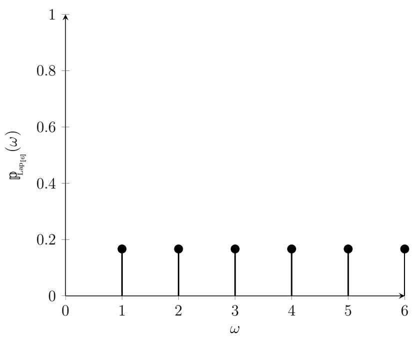
§04.7 **Anmerkung.** Sei $\emptyset \neq \Omega \subseteq \mathbb{R}$ abzählbar und $\mathbb{p}$ eine Zähldichte auf Ω. Das Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathscr{B})$ mit $\mathbb{P}(B) := \sum_{\omega\in\Omega} \mathbb{p}(\omega)\delta_\omega$ für $B \in \mathscr{B}$ und die zugehörige Verteilungsfunktion $\mathbb{F} : \mathbb{R} \to [0,1]$ mit $\mathbb{F}(x) = \mathbb{P}\big((-\infty, x]\big) = \sum_{\omega\in\Omega} \mathbb{p}(\omega)\delta_\omega\big((-\infty, x]\big) = \sum_{\omega\in\Omega} \mathbb{1}_{(-\infty,x]}(\omega)\mathbb{p}(\omega) =: \sum_{\omega\leqslant x} \mathbb{p}(\omega)$ für $x \in \mathbb{R}$ werden *diskret* genannt. □

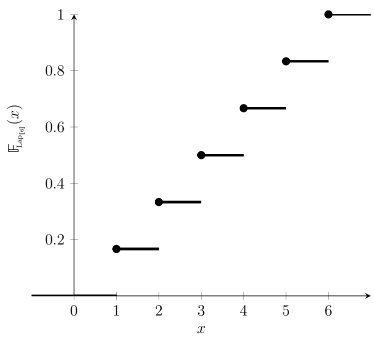
§04.08 **Beispiele.** Folgende Zähldichten beschreiben häufig auftretende diskrete Verteilungen:

(a) *Laplace-/Gleich-Verteilung*, kurz Lap_Ω, mit $|\Omega| < \infty$:

$$\mathbb{p}_{\mathrm{Lap}_\Omega}(\omega) = \frac{1}{|\Omega|} \text{ für } \omega \in \Omega.$$

Betrachten wir den *Wurf eines Würfels*, so ist $\Omega = [\![6]\!]$. Ist der Würfel fair, so erhalten wir die Zähldichte $\mathbb{p}_{\mathrm{Lap}_{[\![6]\!]}}(\omega) = 1/6$, $\omega \in [\![1,6]\!]$, mit der zugehörigen Verteilungsfunktion $\mathbb{F}_{\mathrm{Lap}_{[\![6]\!]}}(x) = \frac{1}{6}\mathbb{1}_{[1,2)}(x) + \frac{2}{6}\mathbb{1}_{[2,3)}(x) + \frac{3}{6}\mathbb{1}_{[3,4)}(x) + \frac{4}{6}\mathbb{1}_{[4,5)}(x) + \frac{5}{6}\mathbb{1}_{[5,6)}(x) + \mathbb{1}_{\mathbb{R}^+}(x-6)$, $x \in \mathbb{R}$:

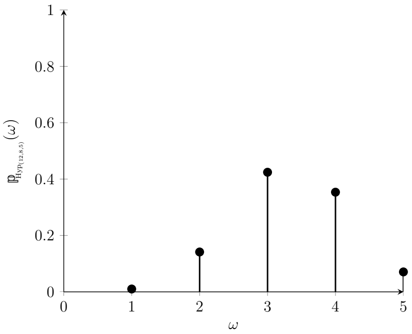


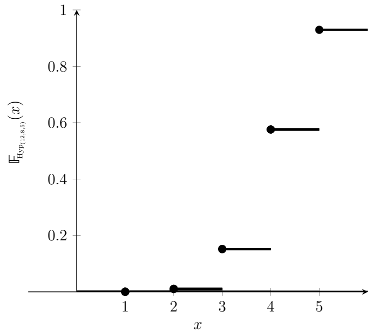


(b) *hypergeometrische Verteilung*, kurz $\mathrm{Hyp}_{(N,K,n)}$, mit Parametern $N \in \mathbb{N}$, $n \in [\![0,N]\!]$, $K \in [\![0,N]\!]$ und $\Omega = [\![0 \vee (n+K-N), n \wedge K]\!]$:

$$\mathbb{p}_{\mathrm{Hyp}_{(N,K,n)}}(\omega) = \frac{\binom{N-K}{n-\omega}\binom{K}{\omega}}{\binom{N}{n}} \text{ für } \omega \in [\![0 \vee (n+K-N), n \wedge K]\!].$$

Ziehen wir n Kugeln aus einer Urne mit K schwarzen und $N-K$ weißen Kugeln, so ist die Wahrscheinlichkeit, dass davon ω schwarz sind, gerade durch $\mathbb{p}_{\mathrm{Hyp}_{(N,K,n)}}(\omega)$ gegeben. Für $N = 12$, $K = 8$ und $n = 5$ erhalten wir die Zähldichte und die zugehörigen Verteilungsfunktion:

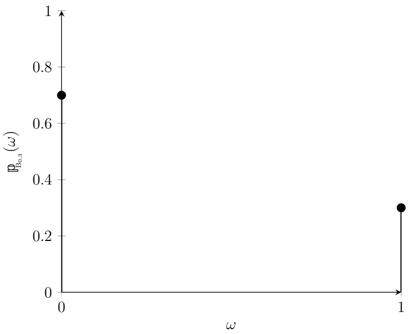


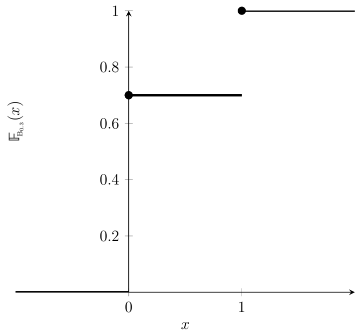


(c) *Bernoulli-Schema*, kurz B_p^n, mit Erfolgswahrscheinlichkeit $p \in [0,1]$, Länge $n \in \mathbb{N}$ und $\Omega = \{0,1\}^n$:

$$\mathbb{p}_{\mathrm{B}_p^n}(\omega) = p^{\sum_{i=1}^n \omega_i}(1-p)^{n-\sum_{i=1}^n \omega_i} \text{ für } \omega = (\omega_i)_{i\in[\![n]\!]} \in \{0,1\}^n.$$

Im Fall $n = 1$, beschreibt eine *Bernoulli-Verteilung*, kurz B_p, also $\mathbb{p}_{\mathrm{B}_p}(\omega) = p^\omega(1-p)^{1-\omega}$ für $\omega \in \{0,1\}$, gerade den Wurf einer Münze, wobei wir die Ergebnisse 1=„Kopf“ als Erfolg und 0=„Zahl“ als Misserfolg auffassen. Für $\mathrm{B}_{0.3}$ (die Münze ist nicht fair, da keine Laplace-Verteilung vorliegt), erhalten wir die Zähldichte und die zugehörigen Verteilungsfunktion:

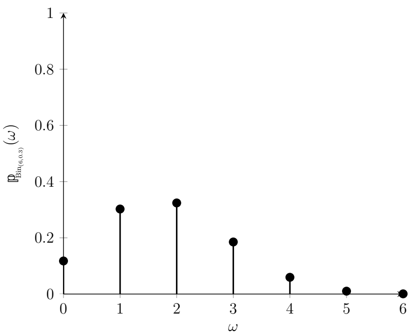


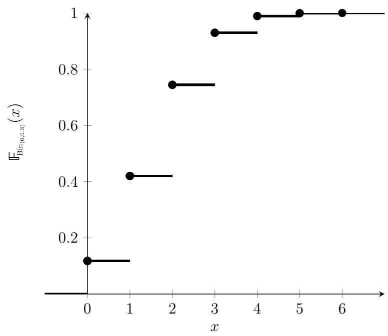


(d) *Binomial-Verteilung*, kurz $\mathrm{Bin}_{(n,p)}$, mit Erfolgswahrscheinlichkeit $p \in [0,1]$, Länge $n \in \mathbb{N}$ und $\Omega := [\![0,n]\!]$:

$$\mathbb{p}_{\mathrm{Bin}_{(n,p)}}(\omega) = \binom{n}{\omega} p^\omega (1-p)^{n-\omega} \text{ für } \omega \in [\![0,n]\!].$$

Die Anzahl „Zahl“ bei 6 Würfen einer Münze (wie in (c)) kann mit einer $\mathrm{Bin}_{(6,0.3)}$ beschrieben werden, für die wir folgende Zähldichte und zugehörige Verteilungsfunktion erhalten:





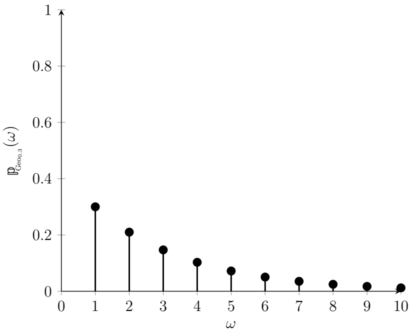
Ziehen wir wie in (b) Kugeln aus einer Urne aber mit Zurücklegen, so erhalten wir für die

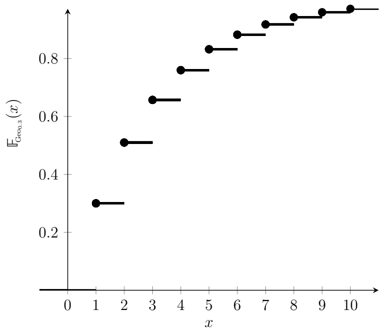
Anzahl der gezogenen schwarzen Kugeln eine $\mathrm{Bin}_{(n,K/N)}$-Verteilung. Die Binomialapproximation einer Hypergeometrischen Verteilung ist eine Übungsaufgabe.

(e) *geometrische Verteilung*, kurz Geo_p, mit Erfolgswahrscheinlichkeit $p \in [0,1]$ und $\Omega = \mathbb{N}$:

$$\mathbb{p}_{\mathrm{Geo}_p}(\omega) = (1-p)^{\omega-1} p \text{ für } \omega \in \mathbb{N}.$$

Werfen wir eine Münze wie in (c) solange bis das erste Mal „Kopf“ fällt, so kann die Anzahl der benötigten Würfe durch eine $\mathrm{Geo}_{0.3}$-Verteilung beschrieben werden, für die wir folgende Zähldichte und zugehörige Verteilungsfunktion erhalten:

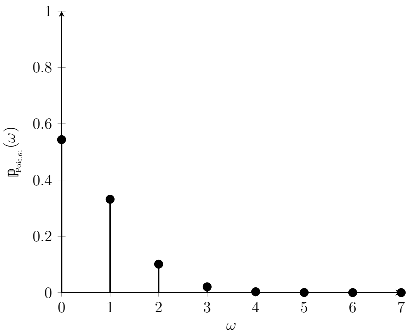


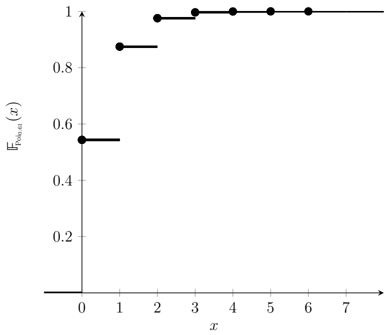


(f) *Poissonverteilung*, kurz Poi_λ, mit Parameter $\lambda \in \mathbb{R}^+_{\setminus 0}$ und $\Omega = \mathbb{N}_0$:

$$\mathbb{p}_{\mathrm{Poi}_\lambda}(\omega) = \exp(-\lambda)\frac{\lambda^\omega}{\omega!} \text{ für } \omega \in \mathbb{N}_0.$$

Ladislau von Bortkewitsch [1898] beschreibt die Anzahl an Todesfällen in der preußischen Kavallerie durch den Schlag eines Pferdes (vgl. Beispiel §04.12) durch eine $\mathrm{Poi}_{0.61}$-Verteilung, für die wir folgende Zähldichte und zugehörigen Verteilungsfunktion erhalten:





□

§04.09 **Poissonscher Grenzwertsatz**. *Es sei eine Folge von reellen Zahlen $(p_n)_{n\in\mathbb{N}}$ aus dem Intervall $[0,1]$ mit $\lambda := \lim_{n\to\infty} np_n > 0$ gegeben. Dann gilt für alle $\omega \in \mathbb{N}_0$:*

$$\lim_{n\to\infty} \mathbb{p}_{\mathrm{Bin}_{(n,p_n)}}(\omega) = \mathbb{p}_{\mathrm{Poi}_\lambda}(\omega).$$

§04.10 **Beweis** von Satz §04.09. In der Vorlesung. □

§04.11 **Bemerkung**. Damit kann man für hinreichend großes n und np „mittelgroß“ die $\mathrm{Bin}_{(n,p)}$- Verteilung durch eine Poi_λ-Verteilung approximieren (zur Güte der Approximation vgl. Satz 5.9 in Krengel [2005]). Das heißt insbesondere, dass die Erfolgswahrscheinlichkeit p klein sein muss. Die Abhängigkeit von $p = p_n$ von n wird nur für die mathematische Beschreibung verwendet. Sie ist nicht so gemeint, dass p wirklich von n abhängt. □

§04.12 **Beispiel**. Ladislaus von Bortkewitsch [1898] gibt für 20 Jahre (1875-1894) und 10 Armeekorps der preußischen Kavallerie, folgende Anzahlen an Todesfällen durch den Schlag eines Pferdes pro $20 \times 10 = 200$ Korpsjahr an:

Todesfälle ω	0	1	2	3	4	$\geqslant 5$
Anzahl Korpsjahre mit ω Todesfällen	109	65	22	3	1	0
$200\mathbb{p}_{\mathrm{Poi}_{0.61}}(\omega)$	109	66	20	4	1	0

Der Parameter λ der Poissonverteilung Poi_λ wurde als mittlere Anzahl an Todesfällen pro Korpsjahr gewählt $\lambda = (1 \cdot 65 + 2 \cdot 22 + 3 \cdot 3 + 1 \cdot 4)/200 = 0.61$. □

§04.13 **Definition**. Sind $\mathbb{p}_1, \dots, \mathbb{p}_n$ Zähldichten auf Ω, so heißt

$$\mathbb{p}(\omega) := \prod_{i \in [\![n]\!]} \mathbb{p}_i(\omega_i) \quad \text{für } \omega = (\omega_i)_{i \in [\![n]\!]} \in \Omega^n$$

die *Produktzähldichte* der $(\mathbb{p}_i)_{i \in [\![n]\!]}$ auf Ω^n. □

§04.14 **Beispiel**. Die Zähldichte eines Bernoulli-Schema wie Beispiele §04.08 (d) ist die Produktzähldichte von n Zähldichten zu identischen Bernoulli-Verteilungen B_p. □

§04.15 **Satz** (Vitali (1905)). *Sei $\Omega = \{0,1\}^{\mathbb{N}}$ der Ergebnisraum eines unendlich oft wiederholten Münzwurfes. Für $n \in \mathbb{N}$ sei $T_n : \Omega \to \Omega$ mit*

$$\omega = (\omega_n)_{n \in \mathbb{N}} \mapsto T_n(\omega) := (\omega_1, \dots, \omega_{n-1}, 1 - \omega_n, \omega_{n+1}, \dots)$$

die Abbildung, welche das Ergebnis des n-ten Münzwurfes umdreht. Für $A \in 2^\Omega$ bezeichne $T_n A := \{T_n(\omega) : \omega \in A\}$ das Bild von A unter T_n. Dann gibt es kein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf der Potenzmenge 2^Ω, das folgender Invarianzeigenschaft genügt: Für alle $A \in 2^\Omega$ und $n \in \mathbb{N}$ gilt $\mathbb{P}(T_n A) = \mathbb{P}(A)$. (Dies drückt die Fairness der Münze und die Unabhängigkeit der Würfe aus.)

§04.16 **Beweis** von Satz §04.15. In der Vorlesung. □

§05 Stetiger Wahrscheinlichkeitsraum

§05.01 **Definition**. Ist $\mathbb{f} : \mathbb{R}^n \to \mathbb{R}^+$ eine Lebesgue-integrierbare Funktion mit $\int_{\mathbb{R}^n} \mathbb{f}(x)dx = 1$, so heißt $\mathbb{f}$ *Wahrscheinlichkeitsdichte* (kurz *Dichte*) auf $\mathbb{R}^n$. □

§05.02 **Satz**. *Jede Wahrscheinlichkeitsdichte $\mathbb{f}$ auf $\mathbb{R}^n$ erzeugt mittels*

$$\mathbb{P}((a,b]) = \int_a^b \mathbb{f}(x)dx = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \mathbb{f}(x_1, \dots, x_n)dx_n \cdots dx_1 \text{ für } a, b \in \mathbb{R}^n,\ a < b,$$

ein eindeutiges Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^n, \mathscr{B}^n)$ und es gilt $\mathbb{P}(B) = \int_B \mathbb{f}(x)dx$ für alle $B \in \mathscr{B}^n$.

§05.03 **Beweis** von Satz §05.02. In der Vorlesung Wahrscheinlichkeitstheorie 1. □

§05.04 **Definition**. Ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^n, \mathscr{B}^n)$ heißt *stetig*, wenn eine Wahrscheinlichkeitsdichte $\mathbb{f}$ auf $\mathbb{R}^n$ mit $\mathbb{P}(B) = \int_B \mathbb{f}(x)dx$ für alle $B \in \mathscr{B}^n$ existiert. $(\mathbb{R}^n, \mathscr{B}^n, \mathbb{P})$ wird dann *stetiger Wahrscheinlichkeitsraum* genannt. □

§05.05 **Lemma**.

(i) *Ist $\mathbb{f}$ eine Dichte eines Wahrscheinlichkeitsmaßes $\mathbb{P}$ auf $\mathscr{B}$ mit Verteilungsfunktion $\mathbb{F}$, so gilt $\mathbb{F}(x) = \int_{-\infty}^{x} \mathbb{f}(y)dy$ für alle $x \in \mathbb{R}$.*

(ii) *Ist die Verteilungsfunktion $\mathbb{F}$ eines Wahrscheinlichkeitsmaßes $\mathbb{P}$ auf $\mathscr{B}$ (schwach) differenzierbar, so ist $\mathbb{f} := \mathbb{F}'$ die zugehörige Wahrscheinlichkeitsdichte.*

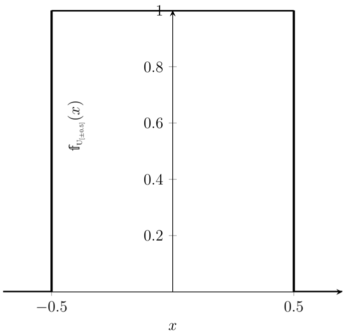
§05.06 **Beweis** von Lemma §05.05. In der Vorlesung Analysis 3. □

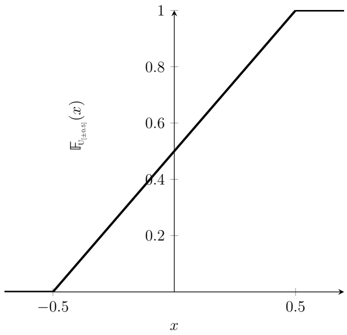
§05.07 **Beispiele**. Folgende Wahrscheinlichkeitsdichten beschreiben häufig auftretende stetige Verteilungen auf $(\mathbb{R}, \mathscr{B})$:

(a) *Gleich-/Uniformverteilung*, kurz U_G, auf $G \in \mathscr{B}$ mit Lebesgue-Maß $\lambda(G) \in \mathbb{R}^+_{\setminus 0}$:

$$\mathbb{f}_{\mathrm{U}_G}(x) = \frac{1}{\lambda(G)} \mathbb{1}_G(x) \text{ für } x \in \mathbb{R}.$$

Runden wir Messungen reelle Größen auf die jeweils nächstgelegene ganze Zahl hin auf bzw. ab, so kann der Abrundungsfehler durch eine Gleichverteilung $\mathrm{U}_{[\pm 0.5]}$ auf dem Intervall $[\pm 0.5] := [-0.5, 0.5]$ mit Dichte $\mathbb{f}_{\mathrm{U}_{[\pm 0.5]}}(x) = \mathbb{1}_{[\pm 0.5]}(x)$ und Verteilungsfunktion $\mathbb{F}_{\mathrm{U}_{[\pm 0.5]}}(x) = (x + 0.5)\mathbb{1}_{[\pm 0.5]}(x) + \mathbb{1}_{\mathbb{R}^+_{\setminus 0}}(x - 0.5)$ für $x \in \mathbb{R}$ beschrieben werden:

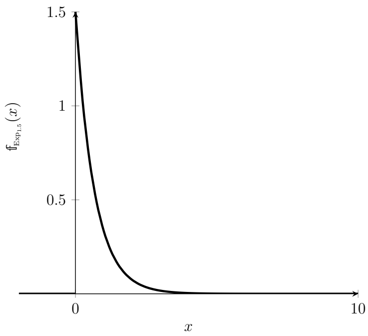


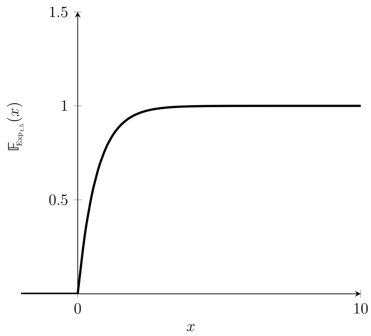


(b) *Exponentialverteilung*, kurz Exp_λ, mit Parametern $\lambda \in \mathbb{R}^+_{\setminus 0}$:

$$\mathbb{f}_{\mathrm{Exp}_\lambda}(x) = \lambda \exp(-\lambda x)\mathbb{1}_{\mathbb{R}^+}(x) \text{ für } x \in \mathbb{R}.$$

Häufig wird in Telefonzentralen die Dauer eines Telefongesprächs durch eine Exponentialverteilung beschrieben. Nimmt eine Servicehotline zum Beispiel für die Gesprächsdauer eine $\mathrm{Exp}_{1.5}$-Verteilung an, so besitzt diese die Dichte $\mathbb{f}_{\mathrm{Exp}_{1.5}}(x) = 1.5\exp(-1.5x)\mathbb{1}_{\mathbb{R}^+}(x)$ und die zugehörige Verteilungsfunktion $\mathbb{F}_{\mathrm{Exp}_{1.5}}(x) = (1 - \exp(-1.5x))\mathbb{1}_{\mathbb{R}^+}(x)$ für $x \in \mathbb{R}$:





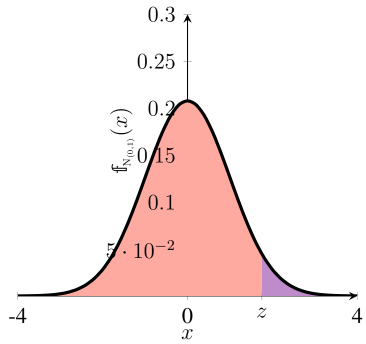
(c) *Normalverteilung*, kurz $\boxed{\mathrm{N}_{(\mu,\sigma^2)}}$, mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 \in \mathbb{R}^+_{\setminus 0}$:

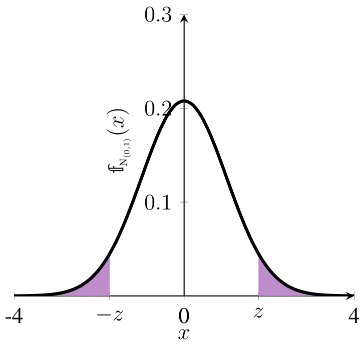
$$\mathbb{f}_{\mathrm{N}_{(\mu,\sigma^2)}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\tfrac{1}{2\sigma^2}(x-\mu)^2) \text{ für } x \in \mathbb{R}.$$

Die Verteilung zufälliger Größen (z. Bsp. der jährliche Wasserverbrauch eines Haushaltes) kann häufig durch eine Normalverteilung gut approximiert werden (vgl. **??**). Eine $\mathrm{N}_{(0,1)}$-Verteilung heißt *Standardnormalverteilung*. Wir bezeichnen mit $\boxed{\Phi}$ die Verteilungsfunktion einer Standardnormalverteilung, d.h., für alle $z \in \mathbb{R}$ gilt

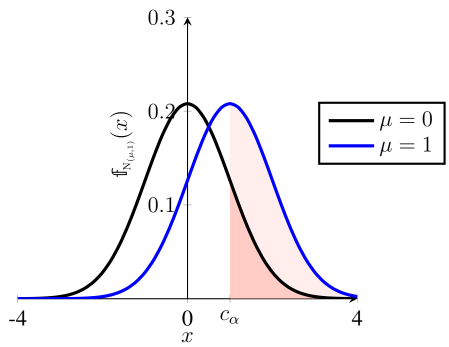
$$\Phi(z) = [\mathrm{N}_{(0,1)}]\big((-\infty, z]\big) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}x^2)dx.$$

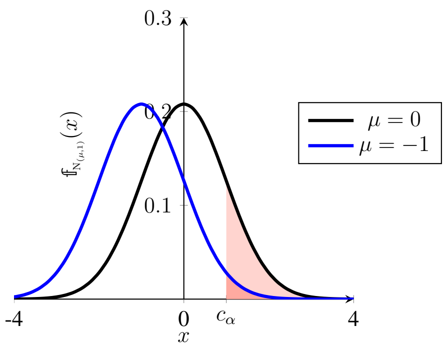
Für alle $z \in \mathbb{R}^+$ gilt dann $\Phi(-z) = 1 - \Phi(z)$:





Weiterhin heißt für $\alpha \in (0,1)$ der eindeutig bestimmte Wert $\boxed{z_\alpha}$ mit $\alpha = \Phi(z_\alpha)$ das α-Quantil einer Standardnormalverteilung. Typische Werte für Quantile der Standardnormalverteilung sind tabellarisiert, siehe Seite 25 im Anhang. Wir halten weiterhin fest, dass für $\alpha \in (0,1)$ und $c_\alpha \in \mathbb{R}$ mit $[\mathrm{N}_{(\mu,\sigma^2)}]\big([c_\alpha,\infty)\big) = \alpha$ gilt $[\mathrm{N}_{(\mu+c,\sigma^2)}]\big([c_\alpha,\infty)\big) > \alpha$ für $c > 0$ sowie $[\mathrm{N}_{(\mu+c,\sigma^2)}]\big([c_\alpha,\infty)\big) < \alpha$ für $c < 0$:





□

§05.08 **Definition**. Sind $\mathbb{f}_1, \dots, \mathbb{f}_n$ Wahrscheinlichkeitsdichten auf $\mathbb{R}$, so heißt

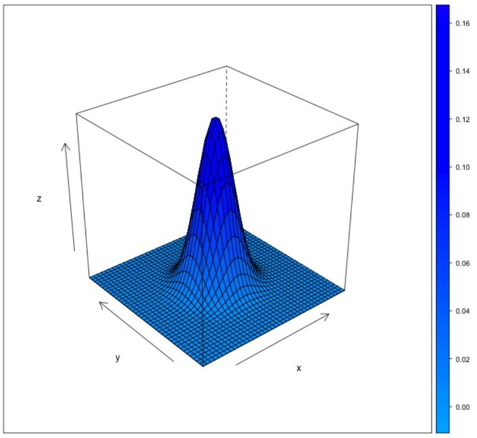
$$\mathbb{f}(x) := \prod_{i \in [\![n]\!]} \mathbb{f}_i(x_i) \text{für } x = (x_i)_{i \in [\![n]\!]} \in \mathbb{R}^n$$

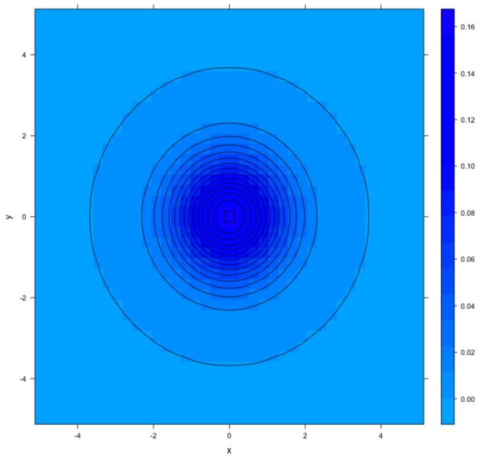
die Produktdichte der $(\mathbb{f}_i)_{i \in [\![n]\!]}$ auf $\mathbb{R}^n$. □

§05.09 **Beispiel**. Die *n-dimensionale Standard-Normalverteilung* $\boxed{\mathrm{N}_{(0,E_n)}}$ im $\mathbb{R}^n$, wobei E_n die n-dimensionale Einheitsmatrix und 0 den n-dimensionalen Nullvektor bezeichnet, ist definiert durch die Dichte

$$\mathfrak{f}_{\mathrm{N}_{(0,E_n)}}(x) := \prod_{i\in[\![n]\!]} \mathfrak{f}_{\mathrm{N}_{(0,1)}}(x_i) = (2\pi)^{-n/2}\exp(-\tfrac{1}{2}\sum_{i\in[\![n]\!]} x_i^2) \text{ für } x\in\mathbb{R}^n;$$

Für $n=2$ erhalten wir die Dichte $\mathfrak{f}_{\mathrm{N}_{(0,E_2)}}(x_1,x_2) = (2\pi)^{-1}\exp(-\frac{1}{2}(x_1^2+x_2^2))$ für $x_1,x_2\in\mathbb{R}$:





□

§06 Statistisches Modell

§06.01 **Erinnerung**. $\mathcal{W}(\mathscr{F})$ bezeichnet die Menge aller Wahrscheinlichkeitsmaße auf einem messbaren Raum $(\mathcal{X},\mathscr{F})$. Für eine nicht-leere Indexmenge Θ wird eine Familie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ von Wahrscheinlichkeitsmaßen auf $\mathscr{F}$ formal durch die Abbildung $\Theta\to\mathcal{W}(\mathscr{F})$ mit $\theta\mapsto\mathbb{P}_\theta$ definiert. □

§06.02 **Definition**. Sei $\mathbb{P}_\Theta := (\mathbb{P}_\theta)_{\theta\in\Theta}$ eine Familie von Wahrscheinlichkeitsmaßen auf einem messbaren Raum (*Stichprobenraum*) $(\mathcal{X},\mathscr{F})$. Die nicht-leere Indexmenge Θ wird *Parametermenge* genannt. Wir bezeichnen das Tripel $\boxed{(\mathcal{X},\mathscr{F},\mathbb{P}_\Theta)}$ als *statistisches Experiment* oder *statistisches Modell*. Ein statistisches Experiment $(\mathcal{X},\mathscr{F},\mathbb{P}_\Theta)$ heißt *adäquat* für ein zufälliges Experiment, wenn dieses durch den Wahrscheinlichkeitsraum $(\mathcal{X},\mathscr{F},\mathbb{P}_\theta)$ für ein $\theta\in\Theta$ beschrieben wird. In diesem Fall wird der Parameter $\theta\in\Theta$ auch *wahrer Parameter* genannt. □

§06.03 **Sprechweise**. Wir nennen ein statistisches Experiment $(\mathcal{X},\mathscr{F},\mathbb{P}_\Theta)$

diskret, wenn der Stichprobenraum $\mathcal{X}$ abzählbar und $\mathscr{F} = 2^{\mathcal{X}}$ ist. In diesem Fall bezeichnet $\boxed{\mathbb{p}_\Theta} := (\mathbb{p}_\theta)_{\theta\in\Theta}$ die zur Familie $\mathbb{P}_\Theta$ von diskreten Wahrscheinlichkeitsmaßen gehörige Familie von Zähldichten;

stetig, wenn $(\mathbb{R}^n,\mathscr{B}^n)$ der Stichprobenraum ist und für jeden Parameter $\theta\in\Theta$ ist $\mathbb{P}_\theta$ ein stetiges Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n,\mathscr{B}^n)$. In diesem Fall bezeichnet $\boxed{\mathfrak{f}_\Theta} := (\mathfrak{f}_\theta)_{\theta\in\Theta}$ die zur Familie $\mathbb{P}_\Theta$ von stetigen Wahrscheinlichkeitsmaßen gehörige Familie von Wahrscheinlichkeitsdichten. □

§06.04 **Beispiel**. Im Folgenden geben wir statistische Modelle für jeweils ein Ergebnis der im Kapitel 1 Prolog vorgestellten Beispiele an.

Beispiel I: Setzen wir Eins für das weibliche Geschlecht und Null für das männliche Geschlecht eines Konsumierenden, so beschreiben wir das zufällige Geschlecht eines Konsumierenden durch ein *Bernoulliverteilungsmodell* $\big(\{0,1\}, 2^{\{0,1\}}, (\mathrm{B}_p)_{p\in[0,1]}\big)$, vgl. Beispiele §04.08 (c).

Beispiel II: Die zufällige Anzahl beschädigter Schrauben in einem Beutel mit 50 Schrauben beschreiben wir durch ein *Binomialverteilungsmodell* $\big([\![0,50]\!], 2^{[\![0,50]\!]}, (\mathrm{Bin}_{(50,p)})_{p\in[0,1]}\big)$, vgl. Beispiele §04.08 (d).

Beispiel III Die zufällige Anzahl der eingegangen Anrufe innerhalb einer Viertelstunde beschreiben wir durch ein *Poissonverteilungsmodell* $\big(\mathbb{N}_0, 2^{\mathbb{N}_0}, (\mathrm{Poi}_\lambda)_{\lambda\in\mathbb{R}^+_{\setminus 0}}\big)$, vgl. Beispiele §04.08 (f).

Beispiel IV Die zufällige Wartezeit an der U-Bahn Haltestelle an einem Tag (in Minuten) beschreiben wir durch ein *Uniformverteilungsmodell* $\big(\mathbb{R}, \mathscr{B}, \big(\mathrm{U}_{[0,\theta]}\big)_{\theta\in\mathbb{R}^+_{\setminus 0}}\big)$, vgl. Beispiele §05.07 (a).

Beispiel V Die zufällige Lebensdauer einer Glühlampe (in Stunden) beschreiben wir durch ein *Exponentialverteilungsmodell* $\big(\mathbb{R}, \mathscr{B}, (\mathrm{Exp}_\lambda)_{\lambda\in\mathbb{R}^+_{\setminus 0}}\big)$, vgl. Beispiele §05.07 (b).

Beispiel VI Die zufälligen Fehler in gemessenen Werten beschreiben wir durch ein *Normalverteilungsmodell* $\big(\mathbb{R}, \mathscr{B}, (\mathrm{N}_{(\mu,\sigma^2)})_{(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}^+_{\setminus 0}}\big)$, vgl. Beispiele §05.07 (c). □.

§06.01 Testen einfacher Hypothesen

§06.05 **Beispiel** *(Binomialverteilungsmodell)*. Im Beispiel I im Kapitel 1 Prolog zählte der Verbraucherservice 699 Konsumentinnen unter den 1000 befragten Personen. Dieses zufällige Experiment lässt sich durch ein *Binomialverteilungsmodell* $\big([\![0,1000]\!], 2^{[\![0,1000]\!]}, (\mathrm{Bin}_{(1000,p)})_{p\in[0,1]}\big)$ beschreiben (formale Begründung in **??**, **??**). Zur (unrealistischen) Vereinfachung geht der Verbraucherservice zunächst davon aus, dass $\{0.5, 0.7\}$ die Parametermenge ist, das heißt, das statistische Modell $\big([\![0,1000]\!], 2^{[\![0,1000]\!]}, (\mathrm{Bin}_{(1000,0.5)}, \mathrm{Bin}_{(1000,0.7)})\big)$ beschreibt adäquat das zufällige Experiment. In anderen Worten, entweder ist $\mathbb{P}_0 = \mathrm{Bin}_{(1000,0.5)}$ oder $\mathbb{P}_1 = \mathrm{Bin}_{(1000,0.7)}$ die wahre Verteilung. □

§06.06 **Definition**. Sei $\big(\mathcal{X}, \mathscr{F}, (\mathbb{P}_0, \mathbb{P}_1)\big)$ ein (binäres) statistisches Experiment. Die Entscheidung, ob die *Nullhypothese* $H_0 : \{\mathbb{P}_0\}$ oder die *Alternativhypothese* $H_1 : \{\mathbb{P}_1\}$ vorliegt, wird *(statistisches) Testproblem mit einfachen Hypothesen* genannt. Eine Entscheidungsfunktion $\varphi : \mathcal{X} \to \{0,1\}$, also Entscheidungen nur anhand einer Stichprobe x aus dem Stichprobenraum $\mathcal{X}$, mit dem Ereignis $\varphi^{-1}(\{1\}) \in \mathscr{F}$ aller Stichproben, die zu einer Entscheidung gegen die Nullhypothese H_0, also für die Alternativhypothese H_1 führen, sowie dem Ereignis $\varphi^{-1}(\{0\}) \in \mathscr{F}$ aller Stichproben, die zu einer Entscheidung für die Nullhypothese H_0 führen, heißt *(statistischer) Hypothesentest*, kurz *Test*. □

§06.07 **Sprechweise**. Die Sprechweise Testproblem mit *einfachen* Hypothesen bedeutet in diesem Zusammenhang, dass sowohl die Nullhypothese als auch die Alternativhypothese jeweils einelementig ist. Üblicherweise bezeichnen wir eine Entscheidung für die Alternativhypothese H_1 als *Ablehnen der Nullhypothese*, wogegen eine Entscheidung für die Nullhypothese H_0 *nicht Ablehnen der Nullhypothese* genannt wird. Für einen Test φ bezeichnen wir das Ereignis $\mathcal{A} := \varphi^{-1}(\{1\}) \in \mathscr{F}$ aller Stichproben, die zum Ablehnen der Nullhypothese H_0 führen, als *Ablehnbereich* des Tests φ. Da wir hier nur die zwei Möglichkeiten Ablehnen oder nicht Ablehnen für einen Test zulassen, ist $\varphi = \mathbb{1}_{\mathcal{A}}$ eine Indikatorfunktion und $\mathcal{A}^c = \varphi^{-1}(0) \in \mathscr{F}$

das Ereignis aller Stichproben, die nicht zu einem Ablehnen der Nullhypothese führen, genannt *Annahmebereich*, wobei definitionsgemäß $\mathcal{A} \biguplus \mathcal{A}^c = \mathcal{X}$ gilt. □

§06.08 **Beispiel** *(Binomialverteilungsmodell §06.05 fortgesetzt)*. Wir sind an einer Entscheidung unter Vorliegen der Stichprobe $x = 677$ aus dem Stichprobenraum $\mathcal{X} = [\![0, 1000]\!]$ interessiert. Wir legen *vor* der Erhebung der Stichprobe einen Test $\varphi : [\![0, 1000]\!] \to \{0, 1\}$ fest. Ergibt ein Auswerten des Testes $\varphi(677) = 1$, so lehnen wir die Nullhypothese $H_0 : \left\{\mathrm{Bin}_{(1000,0.5)}\right\}$ gegen die Alternativhypothese $H_1 : \left\{\mathrm{Bin}_{(1000,0.7)}\right\}$ ab. Nimmt der Test dagegen den Wert $\varphi(677) = 0$ an, so lehnen wir die Nullhypothese H_0 nicht ab. □

§06.09 **Bemerkung**. Durch Angabe des Ablehnbereiches $\mathcal{A}$ ist ein Test $\varphi = \mathbb{1}_{\mathcal{A}}$ eindeutig festgelegt. Offensichtlich können in einem statistisches Testproblem mit einfachen Hypothesen nur zwei Fehlentscheidungen auftreten, die Nullhypothese $H_0 : \{\mathbb{P}_0\}$ wird abgelehnt, also der Ablehnbereich $\mathcal{A}$ tritt ein, obwohl $\mathbb{P}_0$ vorliegt, oder die Nullhypothese wird nicht gegen die Alternativhypothese $H_1 : \{\mathbb{P}_1\}$ abgelehnt, also der Annahmebereich $\mathcal{A}^c$ tritt ein, obwohl $\mathbb{P}_1$ vorliegt. Man beachte, dass wir für einen Test gefordert haben, dass der Ablehnbereich $\mathcal{A}$ und somit auch der Annahmebereich $\mathcal{A}^c$ messbar ist, also $\mathcal{A} \in \mathscr{F}$ gilt. Damit können wir den Ereignissen $\mathcal{A}$ und $\mathcal{A}^c$ Wahrscheinlichkeiten zuordnen. □

§06.10 **Definition**. Sei $\left(\mathcal{X}, \mathscr{F}, (\mathbb{P}_0, \mathbb{P}_1)\right)$ ein (binäres) statistisches Experiment und $\mathcal{A} \in \mathscr{F}$ der Ablehnbereich eines Tests $\varphi = \mathbb{1}_{\mathcal{A}}$ der einfachen Nullhypothese $H_0 : \{\mathbb{P}_0\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathbb{P}_1\}$. Dann bezeichnet

Fehler 1. Art: die Wahrscheinlichkeit $\mathbb{P}_0(\mathcal{A})$ die Nullhypothese abzulehnen, sich also für $\mathbb{P}_1$ zu entscheiden, obwohl $\mathbb{P}_0$ vorliegt;

Fehler 2. Art: die Wahrscheinlichkeit $\mathbb{P}_1(\mathcal{A}^c)$ die Nullhypothese nicht abzulehnen, sich also für $\mathbb{P}_0$ zu entscheiden, obwohl $\mathbb{P}_1$ vorliegt. □

§06.11 **Anmerkung**. Möchten wir zwei Tests miteinander vergleichen, so erscheint es sinnvoll, die Fehler 1. Art und 2. Art zu vergleichen. Offensichtlich würden wir gern einen Test finden der sowohl den Fehler 1. Art als auch den Fehler 2. Art minimiert. Betrachten wir den konstanten Test $\varphi = \mathbb{1}_{\mathcal{X}}$ mit Ablehnbereich $\mathcal{A} = \varphi^{-1}(\{1\}) = \mathcal{X}$, das heißt, wir lehnen immer die Nullhypothese ab. Dann ist $\mathbb{P}_0(\mathcal{A}) = 1$ aber auch $\mathbb{P}_1(\mathcal{A}^c) = \mathbb{P}_1(\emptyset) = 0$. Damit können wir im Allgemeinen nicht beide Fehler gleichzeitig minimieren. Es gibt verschiedene Möglichkeiten, dieses Problem zu umgehen. Zum Beispiel könnte eine gewichtete Summe der beiden Fehler minimiert werden. Da die beiden Fehler häufig unterschiedliche Konsequenzen haben, betrachten wir im Folgenden nur Testfunktionen, die eine Obergrenze für den Fehler 1. Art einhalten. Innerhalb dieser Klasse von Tests suchen wir dann denjenigen, der den Fehler 2. Art minimiert. □

§06.12 **Definition**. Sei $\left(\mathcal{X}, \mathscr{F}, (\mathbb{P}_0, \mathbb{P}_1)\right)$ ein (binäres) statistisches Experiment. Ein Test $\varphi = \mathbb{1}_{\mathcal{A}}$ mit Ablehnbereich $\mathcal{A} \in \mathscr{F}$ der einfachen Nullhypothese $H_0 : \{\mathbb{P}_0\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathbb{P}_1\}$ hält das (*Signifikanz-*) *Niveau* $\alpha \in [0, 1]$ ein (oder kurz ist ein *α-Test*), wenn der Fehler 1. Art $\mathbb{P}_0(\mathcal{A}) \leqslant \alpha$ erfüllt. Ein Test $\varphi = \mathbb{1}_{\mathcal{A}}$ mit Ablehnbereich $\mathcal{A} \in \mathscr{F}$ heißt *bester Test zum Niveau $\alpha \in [0, 1]$*, falls er das Niveau $\alpha \in [0, 1]$ einhält und der Fehler 2. Art $\mathbb{P}_1(\widetilde{\mathcal{A}}^c)$ eines jeden anderen α-Tests $\varphi = \mathbb{1}_{\widetilde{\mathcal{A}}}$ mit Ablehnbereich $\widetilde{\mathcal{A}} \in \mathscr{F}$ nicht kleiner ist, dass heißt, $\mathbb{P}_1(\mathcal{A}^c) \leqslant \mathbb{P}_1(\widetilde{\mathcal{A}}^c)$ gilt. □

§06.13 **Bemerkung**. In der Definition eines besten α-Tests werden die Fehler 1. und 2. Art nicht symmetrisch einbezogen. Dies ist eine gewollte Eigenschaft, da so sicher gestellt wird, dass ein Fehler 1. Art klein ist. Andererseits sollte somit die Festlegung der Nullhypothese und Alternativhypothese dies widerspiegeln. Vereinfacht gesprochen, das Ziel ist es, die Nullhypothese abzulehnen, da nur dann die Wahrscheinlichkeit sich zu irren, immer klein ist. □

§06.14 **Definition**. Sei $\big(\mathcal{X}, \mathscr{F}, (\mathbb{P}_0, \mathbb{P}_1)\big)$ ein (binäres) diskretes statistisches Experiment mit entsprechenden Zähldichten $\mathbb{p}_0$ und $\mathbb{p}_1$. Jeder Test $\varphi = \mathbb{1}_{\mathcal{A}_k}$ mit Ablehnbereich der Form

$$\mathcal{A}_k := \{x \in \mathcal{X} \mid \mathbb{p}_1(x) \geqslant k\mathbb{p}_0(x)\}$$

für einen *kritischen Wert* $k \in \mathbb{R}^+$ heißt *Neyman-Pearson-Test*. □

§06.15 **Beispiel** *(Binomialverteilungsmodell §06.08 fortgesetzt)*. Wir betrachten das Testproblem der einfachen Nullhypothese $H_0 : \{\mathrm{Bin}_{(1000,0.5)}\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathrm{Bin}_{(1000,0.7)}\}$ im binären diskreten statistischen Experiment $\big([\![0, 1000]\!], 2^{[\![0,1000]\!]}, (\mathrm{Bin}_{(1000,p)})_{p\in\{0.5,0.7\}}\big)$. Setzen wir $p = 0.5$ und $q = 0.7$, so ist für einen kritischem Wert $k \in \mathbb{R}^+$ ein Neyman-Pearson-Test $\varphi = \mathbb{1}_{\mathcal{A}_k}$ gegeben durch den Ablehnbereich

$$\begin{aligned}\mathcal{A}_k &= \left\{x \in [\![0, 1000]\!] \;\middle|\; \binom{1000}{x} q^x (1-q)^{1000-x} \geqslant k \binom{1000}{x} p^x (1-p)^{1000-x}\right\} \\ &= \left\{x \in [\![0, 1000]\!] \;\middle|\; \left(\tfrac{q/(1-q)}{p/(1-p)}\right)^x \left(\tfrac{1-q}{1-p}\right)^{1000} \geqslant k\right\}\end{aligned}$$

Da $q > p$ gilt, folgt $\frac{q/(1-q)}{p/(1-p)} > 1$ und die Funktion $L_{q,p}(x) := \left(\frac{q/(1-q)}{p/(1-p)}\right)^x \left(\frac{1-q}{1-p}\right)^{1000}$, $x \in \mathbb{R}^+$ ist somit streng monoton wachsend. Damit gibt es zu jedem kritischen Wert $k \in \mathbb{R}^+$ einen Wert $x^* \in [\![0, 1000]\!]$, derart dass $\mathcal{A}_k = [\![x^*, 1000]\!]$ gilt, so dass jeder Neyman-Pearson-Test durch einen Ablehnbereich dieser Form gegeben ist. □

§06.16 **Neyman-Pearson Lemma**. *Sei $\big(\mathcal{X}, \mathscr{F}, (\mathbb{P}_0, \mathbb{P}_1)\big)$ ein (binäres) diskretes statistisches Experiment. Für das Testproblem der einfachen Nullhypothese $H_0 : \{\mathbb{P}_0\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathbb{P}_1\}$ ist jeder Neyman-Pearson-Test $\varphi = \mathbb{1}_{\mathcal{A}_k}$ mit kritischem Wert $k \in \mathbb{R}^+$ und Ablehnbereich $\mathcal{A}_k \in \mathscr{F}$ wie in Definition §06.14 ein bester Test zum Niveau $\mathbb{P}_0(\mathcal{A}_k) \in [0, 1]$.*

§06.17 **Beweis** von Satz §06.16. In der Vorlesung. □

§06.18 **Bemerkung**. Variieren wir den kritischen Wert k im Neyman-Pearson-Test $\varphi = \mathbb{1}_{\mathcal{A}_k}$, so ändert sich auch der entsprechende Fehler 1. Art $\mathbb{P}_0(\mathcal{A}_k)$. Insbesondere gilt $\mathbb{P}_0(\mathcal{A}_k) \downarrow 0$ für $k \to \infty$ und $\mathbb{P}_0(\mathcal{A}_k) \uparrow 1$ für $k \to 0$. Diese Änderung ist aber im Allgemeinen nicht stetig, so dass zu einem vorgegebenen Niveau $\alpha \in (0, 1)$ nicht immer ein kritischer Wert k_α und entsprechender Neyman-Pearson-Test $\varphi = \mathbb{1}_{\mathcal{A}_{k_\alpha}}$ mit $\mathbb{P}_0(\mathcal{A}_{k_\alpha}) = \alpha$ gewählt werden kann. In diesem Fall wählen wir zu einem vorgegebenen Niveau $\alpha \in (0, 1)$ den kritischen Wert

$$k_\alpha := \arg\max\left\{k \in \mathbb{R}^+ \;\middle|\; \alpha \geqslant \mathbb{P}_0(\mathcal{A}_k)\right\}.$$

Offensichtlich gilt dann auch $\mathbb{P}_0(\mathcal{A}_{k_\alpha}) \leqslant \alpha$ und der entsprechende Neyman-Pearson-Test $\varphi = \mathbb{1}_{\mathcal{A}_{k_\alpha}}$ mit Ablehnbereich $\mathcal{A}_{k_\alpha}$ ist ein α-Test, der aber im Allgemeinen nicht mehr ein bester α-Test ist, wie wir in der Vorlesung Statistik I sehen werden. □

§06.19 **Beispiel** *(Binomialverteilungsmodell §06.15 fortgesetzt).* Für jedes $x^* \in [\![0, 1000]\!]$ ist ein Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x^*,1000]\!]}$ mit Ablehnbereich $[\![x^*, 1000]\!]$ ein bester Test zum Niveau

$$\mathrm{Bin}_{(1000,0.5)}\big([\![x^*, 1000]\!]\big) = \sum_{x\in[\![x^*,1000]\!]} \binom{1000}{x} 0.5^x (1-0.5)^{1000-x} = 0.5^{1000} \sum_{x\in[\![x^*,1000]\!]} \binom{1000}{x}$$

wobei der entsprechende Fehler 2. Art

$$\mathrm{Bin}_{(1000,0.7)}\big([\![0, x^*-1]\!]\big) = \sum_{x\in[\![0,x^*-1]\!]} \binom{1000}{x} 0.7^x (1-0.7)^{1000-x}$$

minimal ist. Zum Beispiel für $x^* = 538$ erhalten wir $\mathrm{Bin}_{(1000,0.5)}\big([\![538, 1000]\!]\big) \approx 0.0088$ und $\mathrm{Bin}_{(1000,0.7)}\big([\![0, 537]\!]\big) \approx 0$ sowie für $x^* = 537$ gilt $\mathrm{Bin}_{(1000,0.5)}\big([\![537, 1000]\!]\big) \approx 0.0105$ und $\mathrm{Bin}_{(1000,0.7)}\big([\![0, 536]\!]\big) \approx 0$. Sind wir an einem Test zu einem vorgegebenen Niveau $\alpha \in (0, 1)$ interessiert, so wählen wir

$$x_\alpha := \arg\min \left\{ x^* \in [\![0, 1000]\!] \;\middle|\; \alpha \geqslant \mathrm{Bin}_{(1000,0.5)}\big([\![x^*, 1000]\!]\big) \right\}.$$

Dann ist der entsprechende Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x_\alpha,1000]\!]}$ mit Ablehnbereich $[\![x_\alpha, 1000]\!]$ ein α-Test. Ein typischer Wert für das Niveau ist $\alpha = 0.01$, so dass mit der obigen Rechnung $x_{0.01} = 538$ und der entsprechende Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![538,1000]\!]}$ den Ablehnbereich $[\![538, 1000]\!]$ besitzt. Im Beispiel I im Prolog §1 zählte der firmeneigene Verbraucherservice 699 Konsumentinnen unter den 1000 befragten Personen, so dass die Nullhypothese, der Anteil der Konsumentinnen beträgt 50%, gegen die Alternativhypothese, der Anteil der Konsumentinnen beträgt 70%, zum Niveau $\alpha = 0.01$ abgelehnt werden kann. Offensichtlich hängt ein Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x^*,1000]\!]}$ mit Ablehnbereich $[\![x^*, 1000]\!]$ nicht von dem Parameter $q \in (p, 1]$ der Alternativhypothese ab. Damit ist der Neyman-Pearson-Test mit Ablehnbereich $[\![538, 1000]\!]$ auch bester Test zum Niveau $\alpha = \mathrm{Bin}_{(1000,0.5)}\big([\![538, 1000]\!]\big)$ der einfachen Nullhypothese $H_0 : \big\{\mathrm{Bin}_{(1000,0.5)}\big\}$ gegen die einfache Alternativhypothese $H_1 : \big\{\mathrm{Bin}_{(1000,0.55)}\big\}$ im binären diskreten statistischen Experiment $\big([\![0, 1000]\!], 2^{[\![0,1000]\!]}, (\mathrm{Bin}_{(1000,p)})_{p\in\{0.5,0.55\}}\big)$. Man beachte, dass der minimale Fehler 2. Art dann $\mathrm{Bin}_{(1000,0.55)}\big([\![0, 537]\!]\big) \approx 0.213$ ist, und dieser hängt natürlich von der Alternativhypothese ab.

Geben wir uns das Niveau $\alpha \in (0, 1)$ vor, so können wir auch zu jedem $n \in \mathbb{N}$ im binären diskreten statistischen Experiment $\big([\![0, n]\!], 2^{[\![0,n]\!]}, (\mathrm{Bin}_{(n,p)})_{p\in\{0.5,0.55\}}\big)$ für das Testproblem der einfachen Nullhypothese $H_0 : \big\{\mathrm{Bin}_{(n,0.5)}\big\}$ gegen die einfache Alternativhypothese $H_1 : \big\{\mathrm{Bin}_{(n,0.55)}\big\}$ den Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x_{\alpha,n},n]\!]}$ mit Ablehnbereich $[\![x_{\alpha,n}, n]\!]$, derart dass

$$x_{\alpha,n} := \arg\min \big\{ x^* \in [\![0, n]\!] \;\big|\; \alpha \geqslant \mathrm{Bin}_{(n,0.5)}\big([\![x^*, n]\!]\big) \big\},$$

wählen. Der Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x_{\alpha,n},n]\!]}$ ist dann ein α-Test, sowie der beste Test zum Niveau $\mathrm{Bin}_{(n,0.5)}\big([\![x_{\alpha,n}, n]\!]\big) \leqslant \alpha$. Für $\alpha = 0.01$ erhalten wir zum Beispiel die kritischen Werte $x_{0.01,500} = 277$, $x_{0.01,1000} = 538$ und $x_{0.01,2000} = 1053$ mit entsprechendem Fehler 1. Art $\mathrm{Bin}_{(500,0.5)}\big([\![277, 500]\!]\big) \approx 0.0088$, $\mathrm{Bin}_{(1000,0.5)}\big([\![538, 1000]\!]\big) \approx 0.0088$ und
$\mathrm{Bin}_{(2000,0.5)}\big([\![1053, 2000]\!]\big) \approx 0.0094$, sowie Fehler 2. Art $\mathrm{Bin}_{(500,0.55)}\big([\![0, 276]\!]\big) \approx 0.553$, $\mathrm{Bin}_{(1000,0.55)}\big([\![0, 537]\!]\big) \approx 0.213$ und $\mathrm{Bin}_{(2000,0.55)}\big([\![0, 1052]\!]\big) \approx 0.016$, der offensichtlich von $n \in \mathbb{N}$ abhängt. Wir können uns also auch eine obere Schranke $\beta \in (0, 1)$ für den Fehler 2.

Art vorgegeben, und nach dem kleinsten Wert für n fragen, so dass der Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x_{\alpha,n},n]\!]}$ mit Ablehnbereich $[\![x_{\alpha,n}, n]\!]$ diese einhält, das heißt,

$$n_{\alpha,\beta} := \arg\min\big\{n \in \mathbb{N} \mid \beta \geqslant \mathrm{Bin}_{(n,0.55)}\big([\![0, x_{\alpha,n}[\![\big)\big\}.$$

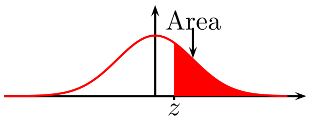
Für $\alpha = 0.01$ und $\beta = 0.01$ erhalten wir $n_{0.01,0.01} = 2170$, wobei $x_{0.01,2170} = 1140$, der Fehler 1. Art $\mathrm{Bin}_{(2170,0.5)}\big([\![1140, 2170]\!]\big) \approx 0.0096$ und der Fehler 2. Art $\mathrm{Bin}_{(2170,0.55)}\big([\![0, 1139]\!]\big) \approx 0.00997$ ist. Zusammenfassend, im Beispiel I im Prolog §1 müsste der firmeneigene Verbraucherservice mindestens 2170 Konsumierende befragen, um sicherzustellen, dass der Fehler 1. Art und der Fehler 2. Art des entsprechenden Neyman-Pearson-Tests nicht größer als 0.01 ist. □

§06.20 **Ausblick.** Betrachten wir für ein beliebiges $p \in (0,1)$, $q \in (p,1]$ und $n \in \mathbb{N}$ das Testproblem der einfachen Nullhypothese $H_0 : \{\mathrm{Bin}_{(n,p)}\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathrm{Bin}_{(n,q)}\}$ im diskreten statistischen Experiment $\big([\![0,n]\!], 2^{[\![0,n]\!]}, (\mathrm{Bin}_{(n,p)}, \mathrm{Bin}_{(n,q)})\big)$, so hängt der Ablehnbereich $[\![x^*, n]\!]$ mit $x^* \in [\![0,n]\!]$ eines Neyman-Pearson-Tests $\varphi = \mathbb{1}_{[\![x^*,n]\!]}$ nicht von dem Wert $q \in (p,1]$ der Alternativhypothese ab. Damit ist für jedes $q \in (p,1]$ ein Neyman-Pearson-Test $\varphi = \mathbb{1}_{[\![x^*,n]\!]}$ mit Ablehnbereich $[\![x^*, n]\!]$ bester Test zum Niveau $\mathrm{Bin}_{(n,p)}([\![x^*, n]\!])$ der *einfachen Nullhypothese* $H_0 : \{\mathrm{Bin}_{(n,p)}\}$ gegen die *einfache Alternativhypothese* $H_1 : \{\mathrm{Bin}_{(n,q)}\}$. Dies erlaubt uns, das diskrete statistische Experiment $\big([\![0,n]\!], 2^{[\![0,n]\!]}, (\mathrm{Bin}_{(n,q)})_{q\in[p,1]}\big)$ zu betrachten. Für jedes $x^* \in [\![0,n]\!]$ ist der Neyman-Pearson Test $\varphi = \mathbb{1}_{[\![x^*,n]\!]}$ mit Ablehnbereich $[\![x^*, n]\!]$ dann *gleichmäßig bester Test* zum Niveau $\alpha = \mathrm{Bin}_{(n,p)}([\![x^*, n]\!])$ der *einfachen Nullhypothese* $H_0 : \{\mathrm{Bin}_{(n,p)}\}$ gegen die *zusammengesetzte Alternativhypothese* $H_1 : \{\mathrm{Bin}_{(n,q)}, q \in (p,1]\}$, da der Fehler 2. Art für jeden anderen Test $\varphi = \mathbb{1}_{\mathcal{A}}$ zum Niveau α mit Ablehnbereich $\mathcal{A}$ gleichmäßig nicht kleiner ist, das heißt, $\mathrm{Bin}_{(n,q)}([\![x^*, n]\!]^c) \leqslant \mathrm{Bin}_{(n,q)}(\mathcal{A}^c)$ für alle $q \in (p,1]$ gilt. Wir halten fest, dass diese Schlussfolgerung möglich ist, da für jedes $p \in [0,1)$ und $q \in (p,1]$ die Funktion $L_{p,q}(x) := \big(\frac{q/(1-q)}{p/(1-p)}\big)^x \big(\frac{1-q}{1-p}\big)^n$, $x \in \mathbb{R}^+$, streng monoton wachsend ist. Bezeichnet $\mathbb{p}_p$ die Zähldichte der Verteilung $\mathrm{Bin}_{(n,p)}$ für $p \in [0,1]$, so gilt offensichtlich $L_{p,q}(x) = \frac{\mathbb{p}_q(x)}{\mathbb{p}_p(x)}$, $x \in [\![0,n]\!]$ und $L_{p,q}$ wird *Likelihood-Quotient* (oder *Dichtequotient*) genannt. Die Verteilungsfamilie $(\mathrm{Bin}_{(n,p)})_{p\in[0,1]}$ besitzt damit einen *monotonen Likelihood-Quotienten*. Im Beispiel I im Prolog §1 kann also der firmeneigene Verbraucherservice die einfache Nullhypothese, der Anteil der Konsumentinnen beträgt 50%, gegen die zusammengesetzte Alternativhypothese, der Anteil der Konsumentinnen beträgt mehr als 50%, zum Niveau $\alpha = 0.01$ ablehnen. □

Anhang

A.1 Normalverteilung

Figure 1: Normal Curve Areas. Standard normal probability in right-hand tail. For negative values of z, areas are found by symmetry.



z	0	.01	.02	.03	.04	.05	.06	.07	.08	.09
	Second decimal place of z									
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000 233									
4.0	.000 031 7									
4.5	.000 003 40									
5.0	.000 000 287									

Literaturverzeichnis

H. Bauer. *Maß- und Integrationstheorie*. Berlin etc.: Walter de Gruyter, 2., überarbeitete Auflage, 1992.

J. Elstrodt. *Maß- und Integrationstheorie.* Berlin: Springer, 7., überarbeitete und ergänzte Auflage, 2011.

H.-O. Georgii. *Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Berlin: De Gruyter, 5., überarbeitete und ergänzte Auflage, 2015.

A. Klenke. *Wahrscheinlichkeitstheorie*. Springer Spektrum, 3., überarbeitete und ergänzte Auflage, 2012.

U. Krengel. *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Braunschweig: Vieweg, 8., erweiterte Auflage, 2005.

L. von Bortkewitsch. Das Gesetz der kleinen Zahlen. Leipzig: B. G. Teubner., 1898.