Fortbildung Stochastik

Universität Freiburg, 1. März 2018

Norbert Henze, Institut für Stochastik, Email: Norbert.Henze@kit.edu

Karlsruher Institut für Technologie

Am KIT gibt es seit 2013 eine eigens für Lehramtsstudierende konzipierte Vorlesung:

Youtube-Videos: Kanal „Stochastikclips“,
Einführung in die Stochastik (SS 2015)

Folien zum Download: http://www.math.kit.edu/stoch/ henze/

Im Folgenden wird aus nachstehenden Büchern zitiert:

- F. Barth, R.Haller: Stochastik Leistungskurs, Ehenwirth Verlag 1992 [BH]
- Mathematik neue Wege: Arbeitsbuch für Gymnasien 6, Verlag Schroedel, 2009 [MNW6]
- Lambacher Schweizer: Mathematik für Gymnasien 3, Verlag Klett, 2005 [LS3]
- Lambacher Schweizer: Mathematik für Gymnasien 10, Verlag Klett, 2016 [LS10]
- Lambacher Schweizer: Mathematik für Gymnasien, Kursstufe, Verlag Klett, 2013 [LS-K]
- Lambacher Schweizer: Stochastik Grundkurs, Verlag Klett, 1996 [LS-G]
- A. Eichler, M. Vogel: Leitfaden Stochastik: Für Studierende und Ausübende des Lehramts, Verlag Vieweg+Teubner, 2011 [LFS]
- Sigma Grundkurs Stochastik, Verlag Klett 1990 [SGK].

0 Verständnisfragen

Aus [LS10]: Eine Bernoulli-Kette besteht aus mehreren voneinander unabhängigen Durchführungen eines Bernoulli-Experimentes.

Was bedeutet voneinander unabhängige Durchführungen?

Ist diese Situation bei Basketballfreiwürfen gegeben?

Aus [LS10]: Eine $B_{n;p}$-verteilte Zufallsgröße hat den Erwartungswert $\mathbb{E}(X) = \mu = n \cdot p$.

Warum?

Aus [LS10]: Falls μ ganzzahlig ist, hat diese Trefferanzahl die höchste Wahrscheinlichkeit, und das entsprechende Rechteck im Histogramm ist das höchste. Falls μ nicht ganzzahlig ist, hat eine der benachbarten Trefferanzahlen die maximale Wahrscheinlichkeit.

Warum?

Aus [LS10]: Eine $B_{n;p}$-verteilte Zufallsgröße hat die Standardabweichung $\sigma = \sqrt{n \cdot p \cdot (1-p)}$.

Warum?

Warum führt man überhaupt die Standardabweichung ein?

Anmerkung: $\sigma = \sqrt{(x_1-\mu)^2 \cdot \mathbb{P}(X = x_1) + \ldots + (x_n-\mu)^2 \cdot \mathbb{P}(X = x_n)}$

Aus [LS10]: Man kann zeigen: Wenn bei einer Binomialverteilung die Parameter n und p genügend groß sind, dann beträgt die Wahrscheinlichkeit, dass die Trefferanzahl um höchstens σ vom Erwartungswert μ abweicht, ca. 68%.

Anmerkung: „p genügend groß“ ist falsch!

Was heißt „n genügend groß“?

Warum gerade 68% ?

Aus [LS10]: Warum stimmen die Zahlen im Pascalschen Dreieck mit den Binomialkoeffizienten überein?

Antwort: Es gilt die (durch Rechnung nachgeprüfte) Formel

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

Kann ich *begreifen, warum* diese Formel gilt?

1 Grundräume, Ereignisse, Zufallsvariablen

1.1 Definition (Grundraum)

Sei $\Omega := \{\omega_1, \omega_2, \ldots\}$ eine abzählbare Menge.

Ω heißt (elementarer) Grundraum (Ergebnismenge, Ausgangsmenge).

Ω steht für die Menge der Ergebnisse eines stochastischen Vorgangs.

Sprechweise: $\omega \in \Omega$ heißt Ergebnis.

1.2 Beispiel (n-facher Würfelwurf)

$$\begin{aligned} \Omega &:= \{(a_1, \ldots, a_n) : a_j \in \{1, \ldots, 6\} \text{ für } j = 1, \ldots, n\} \\ &= \{1, 2, 3, 4, 5, 6\}^n \end{aligned}$$

a_j beschreibt das Ergebnis des j-ten Wurfs.

1.3 Beispiel (n nicht unterscheidbare Würfel gleichzeitig werfen)

$$\Omega := \{(b_1, \ldots, b_n) : 1 \leq b_1 \leq b_2 \leq \ldots \leq b_n \leq 6\}$$

b_j steht für die j-kleinste Augenzahl.

Aus [LS-G]: Ein Zufallsexperiment ist erst dann ausreichend beschrieben, wenn eine Menge S von möglichen Ergebnissen $e_1, \ldots, e_k$ so festgelegt ist, dass bei jeder Durchführung des Zufallsexperimentes genau eines dieser Ergebnisse eintreten muss. Diese Menge heißt Ergebnismenge (Ausgangsmenge).

Es wurden damals auch unendliche Ergebnismengen thematisiert:

Welches ist die Ergebnismenge für das Zufallsexperiment: Werfen einer Münze und Feststellen der Anzahl der Würfe, bis zum ersten Mal Wappen fällt?

1.4 Definition (Ereignis)

Jede Teilmenge A von Ω heißt Ereignis.

$\{\omega\}$ heißt Elementarereignis, $\omega \in \Omega$.

Sprechweise: „A tritt ein“ $\iff \omega \in A$

1.5 Beispiel (n-facher Würfelwurf)

Ereignis *verbal*: „Es tritt mindestens eine Sechs auf“.

Ereignis als Teilmenge von $\Omega = \{1, \ldots, 6\}^n$:

$$A = \left\{ (a_1, \ldots, a_n) \in \Omega : \max_{j=1\ldots,n} a_j = 6 \right\}$$

Aus [LS-G]: Ein Zufallsexper. habe die Ergebnismenge $S = \{e_1, e_2, \ldots, e_k\}$. Dann nennt man jede Teilmenge A von S ein zu diesem Zufallsexperiment gehörendes Ereignis. Endet die Durchführung des Zufallsexperimentes mit einem Ergebnis aus A, so ist das Ereignis A eingetreten.

1.6 Mengentheoretische Verknüpfungen von Ereignissen

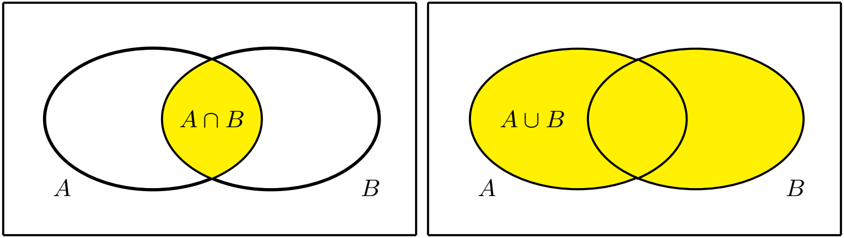
Seien $A, B \subseteq \Omega$ Ereignisse.

$A \cap B \quad = \quad \{\omega \in \Omega : \omega \in A$ *und* $\omega \in B\}$ (Durchschnitt von A und B)

$\qquad\qquad verbal$: A *und* B treten beide ein

$A \cup B \quad = \quad \{\omega \in \Omega : \omega \in A$ *oder* $\omega \in B\}$ (Vereinigung von A und B)

$\qquad\qquad verbal$: A *oder* B tritt ein (evtl. beide!)



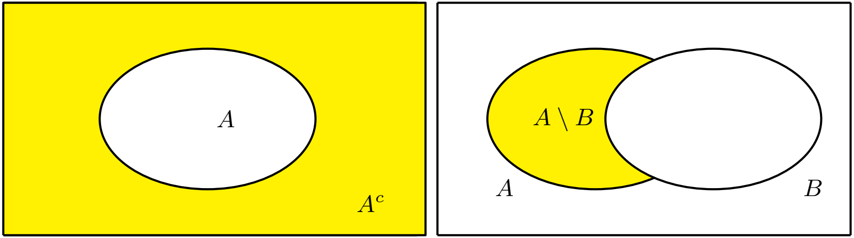
$$A^c \quad = \quad \{\omega \in \Omega : \omega \notin A\} \quad (\text{Komplement von } A)$$

in Worten : A tritt *nicht* ein

$$A \backslash B \quad = \quad A \cap B^c \quad ((\text{Mengen})\text{-Differenz von } A \text{ und } B)$$

in Worten : A tritt ein *und* (*aber*) B *nicht*

(„A ohne B“)

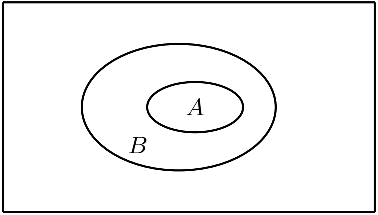


Im Fall $A \subseteq B$ Sprechweisen:

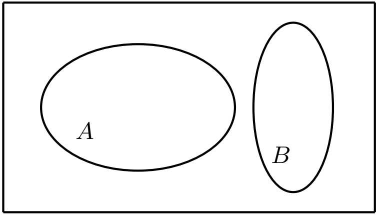
- „Aus A folgt B".
- „Wenn A eintritt, so auch B".
- „Das Eintreten von A zieht das Eintreten von B nach sich".

Im Fall $A \cap B = \emptyset$ Sprechweisen:

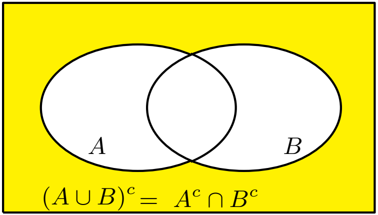
- „A und B disjunkt",
- „A und B unvereinbar",
- „A und B schließen sich gegenseitig aus".



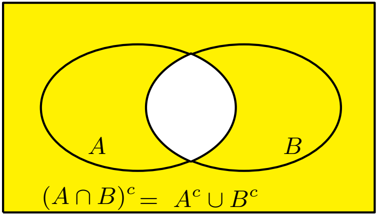
$A \subseteq B$



$A \cap B = \emptyset$



Verbal: Es tritt genau dann nicht mindestens eines der Ereignisse A und B ein, wenn keines dieser Ereignisse eintritt.



Verbal: Es treten genau dann nicht beide der Ereignisse A und B ein, wenn mindestens eines dieser Ereignisse nicht eintritt.

1.7 Definition (Zufallsvariable)

Jede Abbildung

$$X : \Omega \rightarrow \mathbb{R}$$

heißt (reelle) Zufallsvariable.

Für $\omega \in \Omega$ heißt $X(\omega)$ Realisierung von X (zum Ausgang ω).

Zufallsvariablen beschreiben reellwertigen Aspekt eines stochastischen Vorgangs, z.B. bei $\Omega = \{1, 2, 3, 4, 5, 6\}^n$:

- $X(a_1, \ldots, a_n) := a_1 + \ldots + a_n$ (Augensumme beim n-fachen Würfelwurf),
- $X(a_1, \ldots, a_n) := \max_{1 \leq j \leq n} a_j$ (größte Augenzahl beim n-fachen WW),
- $X(a_1, \ldots, a_n) := a_2$ (zweite Augenzahl beim n-fachen WW).

[MNW6]: Spielst Du z.B. Monopoly, dann kommt es auf die Augensumme an, bei anderen Spielen auf den Unterschied der beiden Augenzahlen.Wir nennen die Größe, auf die es uns ankommt, Zufallsvariable.

[SGK]: Eine Abbildung $X : \Omega \to \mathbb{R}$ heißt eine Zufallsgröße auf Ω.

Wichtiger Zweck von Zufallsvariablen: Sie beschreiben Ereignisse!

1.8 Beispiel (Zweifacher Würfelwurf, $X :=$ Augensumme)

(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)

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

1.9 Definition (Indikator(funktion), Indikatorvariable)

Sei $A \subseteq \Omega$ ein Ereignis. Die durch

$$\mathbf{1}_A(\omega) := \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \notin A, \end{cases}$$

definierte Zufallsvariable $\mathbf{1}_A : \Omega \to \mathbb{R}$ heißt Indikatorfunktion von A oder Indikator von A.

1.10 Satz (Rechenregeln für Indikatorfunktionen)

a) $\mathbf{1}_\emptyset \equiv 0, \quad \mathbf{1}_\Omega \equiv 1,$

b) $\mathbf{1}_A^2 = \mathbf{1}_A,$

c) $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A,$

d) $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B,$

e) $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B},$

f) $A \subseteq B \iff \mathbf{1}_A \leq \mathbf{1}_B.$

Memo: $\mathbf{1}_A(\omega) = 1$, falls $\omega \in A$; $\mathbf{1}_A(\omega) = 0$, falls $\omega \notin A$

Beachte: Seien $A_1, \ldots, A_n \subseteq \Omega$ Ereignisse. Dann modelliert die Indikatorsumme (Zählvariable)

$$X := \mathbf{1}_{A_1} + \mathbf{1}_{A_2} + \ldots + \mathbf{1}_{A_n}$$

die Anzahl der eintretenden Ereignisse unter $A_1, \ldots, A_n$.

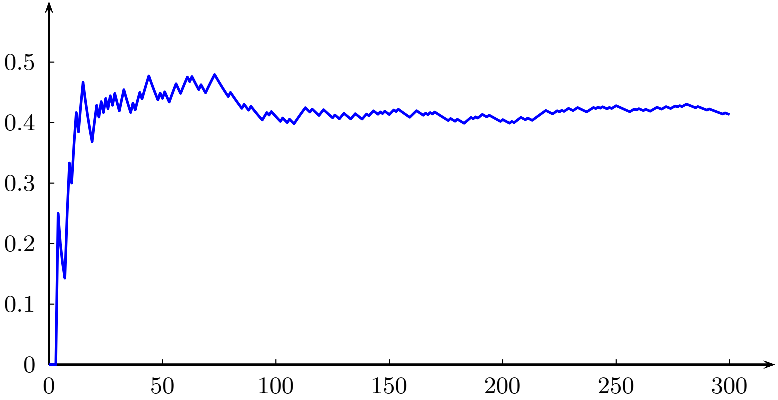
Beispiele: Anzahl der

- Treffer in einer Bernoulli-Kette der Länge n,
- gezogenen roten Kugeln beim Ziehen ohne Zurücklegen,
- Fixpunkte einer zufälligen Vertauschung von Zahlen,
- freien Fächer beim Verteilen von Teilchen auf Fächer,
- Mehrfachgeburtstage in einer Gruppe von Personen usw.

Möglicher Grundraum:

$$\Omega := \{0,1\}^n = \{(a_1, \ldots, a_n) : a_j \in \{0,1\} \text{ für } j = 1, \ldots, n\}.$$
$$A_j := \{(a_1, \ldots, a_n) \in \Omega : a_j = 1\}$$

2 Diskrete Wahrscheinlichkeitsräume



Fortlaufend notierte relative Häufigkeiten für „Spitze nach oben“ beim Werfen einer Reißzwecke

Ideales Zufallsexperiment:

Experiment mit *zufälligem Ausgang*, das *beliebig oft* unter *gleichen, sich gegenseitig nicht beeinflussenden Bedingungen* wiederholt werden kann, z.B.

- Münz- oder Würfelwurf,
- Ziehen aus einer Urne mit Zurücklegen,
- Drehen eines Glücksrades,
- Kartenverteilungen,
- Ausspielungen beim Lotto.

Empirisches Gesetz über die Stabilisierung relativer Häufigkeiten:

Wächst bei einem *idealen Zufallsexperiment* die Anzahl der Wiederholungen, so stabilisieren sich die relativen Häufigkeiten des Eintretens eines Ereignisses erfahrungsgemäß um einen gewissen (unbekannten) Wert.

Ideales Zufallsexperiment, Ergebnise modelliert durch Grundraum Ω.

n mal „in unabhängiger Folge“ wiederholen

Ergebnisse $\in \Omega^n \;=\; \{(a_1,\ldots,a_n) : a_j \in \Omega$ für $j = 1,\ldots,n\}$

Sei $A \subseteq \Omega$, $(a_1,\ldots,a_n) \in \Omega^n$ fest

$$h_n(A) \;:=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_A(a_j) \quad \text{(relative Häufigkeit von } A \text{ zu } (a_1,\ldots,a_n))$$

Für die relative Häufigkeitsfunktion $h_n : \mathcal{P}(\Omega) \to \mathbb{R}$ gelten:

- $0 \le h_n(A) \;\le\; 1, \qquad A \subseteq \Omega,$
- $h_n(\Omega) \;=\; 1,$
- $h_n(A \cup B) \;=\; h_n(A) + h_n(B), \quad \text{falls } A \cap B \;=\; \emptyset,$
- $h_n(A) \;\rightsquigarrow\; ?$ bei $n \to \infty$.

[LS3]: Bei einem Zufallsversuch gibt die Wahrscheinlichkeit eines Ergebnisses an, welchen Anteil man für das Ergebnis bei Versuchswiederholungen erwartet.

2.1 Definition (diskreter Wahrscheinlichkeitsraum)

$(\Omega, \mathbb{P})$ heißt diskreter Wahrscheinlichkeitsraum $:\Longleftrightarrow$

- $\Omega \neq \emptyset$ elementarer Grundraum
- $\mathbb{P} : \mathcal{P}(\Omega) \to \mathbb{R}$ mit folgenden Eigenschaften:

(P1) $\mathbb{P}(A) \geq 0, \qquad A \subseteq \Omega,$

(P2) $\mathbb{P}(\Omega) = 1,$

(P3) $A_1, A_2, \ldots \subseteq \Omega$ paarweise disjunkt $\Longrightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

(sog. σ-Additivität von $\mathbb{P}$)

$\mathbb{P}$ heißt Wahrscheinlichkeitsverteilung oder Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$.

$\mathbb{P}(A)$ heißt Wahrscheinlichkeit von A.

Beachte: Keine inhaltliche Definition von Wahrscheinlichkeit!

Im Folgenden Schreibweise:

$$A+B := A \cup B, \text{ falls } A \cap B = \emptyset, \quad \sum_{j=1}^{n} A_j := \bigcup_{j=1}^{n} A_j, \text{ falls } A_i \cap A_j = \emptyset \, \forall i \neq j.$$

Memo: (P1) $\mathbb{P}(A) \geq 0$, (P2) $\mathbb{P}(\Omega) = 1$, (P3) $\mathbb{P}(\sum_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$

2.2 Folgerungen Für $A, B, A_1, A_2, \ldots \subseteq \Omega$ gelten:

a) $\mathbb{P}(\emptyset) = 0$,

b) $\mathbb{P}\left(\sum_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mathbb{P}(A_j)$ (endliche Additivität),

c) $0 \leq \mathbb{P}(A) \leq 1$,

d) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ (Regel von der komplementären W'),

e) $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$ (Monotonie),

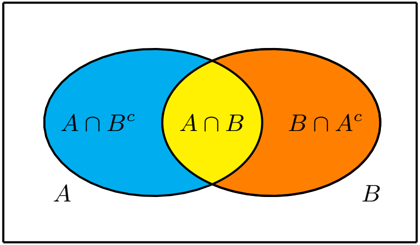
f) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ (Additionssatz),

g) $$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_k})$$

(Formel des Ein- und Ausschließens)

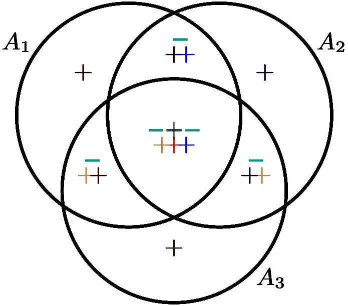
Memo: 2.2.b): $\mathbb{P}\left(\sum_{j=1}^{n} A_j\right) \; = \; \sum_{j=1}^{n} \mathbb{P}(A_j)$

f) $\mathbb{P}(A \cup B) \; = \; \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$



Memo: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$$\begin{aligned}
\mathbb{P}(A_1 \cup A_2 \cup A_3) &= \mathbb{P}(A_1 \cup A_2) + \mathbb{P}(A_3) - \mathbb{P}((A_1 \cup A_2) \cap A_3) \\
&= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) \\
&\quad -\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) \\
&\quad + \mathbb{P}(A_1 \cap A_2 \cap A_3).
\end{aligned}$$

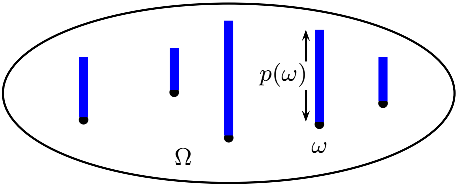


2.3 Definition (Wahrscheinlichkeitsfunktion)

Sei $(\Omega, \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum und

$$p(\omega) \;:=\; \mathbb{P}(\{\omega\}), \qquad \omega \in \Omega.$$

Die Funktion $p : \Omega \to [0,1]$ heißt Wahrscheinlichkeitsfunktion von $\mathbb{P}$.



Beachte: $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega), \qquad A \subseteq \Omega.$

2.4 Definition (Endlicher W-Raum, Laplacescher W-Raum)

Es sei $(\Omega, \mathbb{P})$ ein diskreter W-Raum. Ist Ω endlich, so heißt $(\Omega, \mathbb{P})$ endlicher W-Raum.

Gilt für jedes $A \subseteq \Omega$

$$\mathbb{P}(A) \; := \; \frac{|A|}{|\Omega|} \; = \; \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl aller möglichen Fälle}},$$

so heißt $(\Omega, \mathbb{P})$ Laplacescher W-Raum (der Ordnung $|\Omega|$).

$\mathbb{P}$ heißt (diskrete) Gleichverteilung auf Ω.

$(\Omega, \mathbb{P})$ heißt auch Laplace-Modell.

Es gilt dann insbesondere

$$\mathbb{P}(\{\omega\}) \; = \; \frac{1}{|\Omega|}, \quad \omega \in \Omega.$$

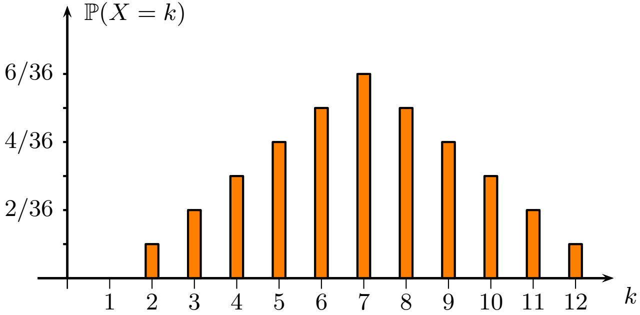
Sprechweisen: Echte Münze, fairer (homogener) Würfel, rein zufälliges Ziehen.

2.5 Beispiel (Zweifacher Würfelwurf: Augensumme)

$$\Omega \; := \; \{(i,j) : 1 \le i,j \le 6\}, \quad \mathbb{P}(\{(i,j)\}) := \frac{1}{36} =: p(i,j), \qquad 1 \le i,j \le 6.$$

X = Augensumme, d.h. $X(i,j) := i + j, \; \omega \; = \; (i,j) \in \Omega$.

$$\mathbb{P}(X=5) = p(1,4) + p(2,3) + p(3,2) + p(4,1) = 4 \cdot \frac{1}{36} = \frac{1}{9}.$$



Stabdiagramm der *Verteilung (?) von* X

Seien $(\Omega, \mathbb{P})$ ein W-Raum und $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann ist

$$\mathbb{P}^X : \begin{cases} \mathcal{P}(\mathbb{R}) \to \mathbb{R}_{\geq 0}, \\ M \mapsto \mathbb{P}^X(M) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \in M\}) \end{cases}$$

ein W–Maß auf den Teilmengen von $\mathbb{R}$, d.h. es gelten $\mathbb{P}^X(M) \geq 0$, $\mathbb{P}^X(\mathbb{R}) = 1$ und

$$\mathbb{P}^X\left(\sum_{j=1}^{\infty} M_j\right) = \sum_{j=1}^{\infty} \mathbb{P}^X(M_j), \quad \text{falls } M_1, M_2, \ldots \subseteq \mathbb{R} \text{ paarweise disjunkt.}$$

2.6 Definition (Verteilung einer Zufallsvariablen)

Das Wahrscheinlichkeitsmaß $\mathbb{P}^X$ auf $\mathcal{P}(\mathbb{R})$ heißt Verteilung von X.

Da X höchstens abzählbar viele Werte $x_1, x_2, \ldots$ annehmen kann, gilt

$$\mathbb{P}^X(M) = \sum_{j : x_j \in M} \mathbb{P}(X = x_j), \qquad M \subseteq \mathbb{R}$$

d.h. $\mathbb{P}^X$ ist durch Werte $\mathbb{P}(X = x_j)$, $j \geq 1$, festgelegt ($\to$ Stabdiagramm!)

Sei X eine (auf irgendeinem W-Raum $(\Omega, \mathbb{P})$ definierte) Zufallsvariable.

Besitzt X die Verteilung Q, so schreiben wir hierfür

$$X \sim Q \quad (:\Longleftrightarrow \mathbb{P}^X = Q).$$

Manche Verteilungen sind so wichtig, dass sie eigene Namen und Bezeichungen erhalten, z.B.

$$X \sim \text{Bin}(n,p) \;:\Longleftrightarrow\; \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0,1,\ldots,n\}$$

(sogenannte Binomialverteilung mit Parametern n und p, $n \in \mathbb{N}, 0 \leq p \leq 1$).

3 Kombinatorik

3.1 Satz (Erstes Fundamentalprinzip des Zählens)

Es seien M und N endliche Mengen. Dann gilt:

$$|M| = |N| \iff \exists\, f : M \to N,\ f \text{ bijektiv.}$$

3.2 Satz (Zweites Fundamentalprinzip des Zählens, Multiplikationsregel)

Seien M eine endliche Menge und $m_1, \ldots, m_k \in \{1, \ldots, |M|\}$.

Sukzessive werden Elemente $a_1, a_2, \ldots, a_k \in M$ so ausgewählt, dass es für

das 1. Element a_1 m_1 Möglichkeiten,

das 2. Element a_2 bei festgelegtem a_1 stets m_2 Möglichkeiten,

das 3. Element a_3 bei festgelegten a_1, a_2 stets m_3 Möglichkeiten,

$\vdots$

das k. Element a_k bei festgelegten $a_1, \ldots, a_{k-1}$ stets m_k Möglichkeiten gibt.

Die Anzahl verschiedener k-Tupel $(a_1, \ldots, a_k)$ ist dann das Produkt

$$\boxed{m_1 \cdot m_2 \cdot \ldots \cdot m_k.}$$

3.3 Definition (k-Permutationen)

Sei M eine n-elementige Menge.

a) Ein k-Tupel $(a_1, \ldots, a_k)$ mit $a_j \in M \quad \forall j = 1, \ldots, k$ heißt k-Permutation aus M mit Wiederholung (m.W.).

$$\mathrm{Per}_k^M(mW) := M^k$$

sei die Menge aller k-Permutationen m.W. aus M.

b) Gilt zusätzlich $a_i \neq a_j$ für alle $i \neq j$, so heißt $(a_1, \ldots, a_k)$ eine k-Permutation aus M ohne Wiederholung (o.W.).
Hierfür muss $k \leq n$ gelten.

$$\mathrm{Per}_k^M(oW) := \{(a_1, \ldots, a_k) \in M^k : a_i \neq a_j \ \forall\ i \neq j\}$$

sei die Menge aller k-Permutationen o.W. aus M.

Die n-Permutationen aus M heißen kurz Permutationen von M.

3.4 Definition (k-Kombinationen)

Sei M durch eine Relation $\leq$ geordnet.

a) Ein k-Tupel $(a_1, \ldots, a_k)$ aus M^k mit $a_1 \leq a_2 \leq \ldots \leq a_k$ heißt k-Kombination aus M mit Wiederholung.

$$\text{Kom}_k^M(mW) := \{(a_1, \ldots, a_k) \in M^k : a_1 \leq a_2 \leq \ldots \leq a_k\}$$

sei die Menge aller k-Kombinationen m.W. aus M.

b) Ein k-Tupel $(a_1, \ldots, a_k)$ aus M^k mit $a_1 < a_2 < \ldots < a_k$ heißt k-Kombination aus M ohne Wiederholung.

$$\text{Kom}_k^M(oW) := \{(a_1, \ldots, a_k) \in M^k : a_1 < a_2 < \ldots < a_k\}$$

bezeichne die Menge aller k-Kombinationen aus M ohne Wiederholung.

Ist speziell $M = \{1, \ldots, n\}$, so

$$\mathrm{Per}_k^n := \mathrm{Per}_k^M, \qquad \mathrm{Kom}_k^n := \mathrm{Kom}_k^M.$$

Memo: $\mathrm{Per}_k^n(oW) = \{(a_1, \ldots, a_k) \in \{1, \ldots, n\}^k : a_i \neq a_j \ \forall i \neq j\}$

Memo: $\mathrm{Per}_k^n(mW) = \{1, \ldots, n\}^k$

Memo: $\mathrm{Kom}_k^n(oW) = \{(a_1, \ldots, a_k) \in \{1, \ldots, n\}^k : 1 \leq a_1 < a_2 < \ldots < a_k\}$

Memo: $\mathrm{Kom}_k^n(mW) = \{(a_1, \ldots, a_k) \in \{1, \ldots, n\}^k : a_1 \leq a_2 \leq \ldots \leq a_k\}$

3.5 Satz (Grundformeln der Kombinatorik)

Die Anzahl der k-Permutationen mit/ohne Wiederholung und der k-Kombinationen mit/ohne Wiederholung aus M ist

	m.W.	o.W. ($k \leq n$)
k-Permutationen	n^k	$n^{\underline{k}}$
k-Kombinationen	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Hierbei ist $$n^{\underline{k}} = n\,(n-1)\,\ldots\,(n-(k-1)) = \frac{n!}{(n-k)!}$$

$$\binom{m}{\ell} := \frac{m!}{\ell!\,(m-\ell)!} \qquad m! = 1 \cdot 2 \cdot \ldots \cdot m, \quad 0! = 1$$

Memo: $|\mathrm{Kom}_s^r(oW)| = \binom{r}{s}$ Behauptung: $|\mathrm{Kom}_k^n(mW)| = \binom{n+k-1}{k}$

Beweis: Sei $a = (a_1, \ldots, a_k) \in \mathrm{Kom}_k^n(mW)$, also $1 \le a_1 \le a_2 \le \ldots \le a_k \le n$.

Idee: Ziehe Komponenten $a_1, \ldots, a_k$ auseinander

$b_1 := a_1,\ b_2 := a_2 + 1,\ b_3 := a_3 + 2,\ \ldots,\ b_k := a_k + k - 1$

$$\Longrightarrow\ b\ :=\ (b_1, \ldots, b_k)\ =:\ f(a)\ \in \mathrm{Kom}_k^{n+k-1}(oW)$$

$f : \mathrm{Kom}_k^n(mW) \to \mathrm{Kom}_k^{n+k-1}(oW)$ ist injektiv!

f ist surjektiv! (da $b_j = a_j + j - 1 \Longleftrightarrow a_j = b_j - j + 1$)

$$\Longrightarrow |\mathrm{Kom}_k^n(mW)|\ =\ |\mathrm{Kom}_k^{n+k-1}(oW)|\ =\ \binom{n+k-1}{k}$$

Mit dieser Beweismethode bestimmt man z.B. die W' für einen Zwilling beim Lotto.

4 Urnen- und Fächer-Modelle

4.1 Urnenmodelle

In einer Urne liegen gleichartige, von 1 bis n nummerierte Kugeln. Wir betrachten vier verschiedene Arten, k Kugeln aus dieser Urne zu ziehen.

(1) Ziehen unter Beachtung der Reihenfolge mit Zurücklegen

Nach jedem Zug Kugel-Nummer notieren und Kugel zurücklegen.
a_j sei die Nummer der beim j-ten Zug erhaltenen Kugel.
Geeigneter Ergebnisraum:

$$\mathrm{Per}_k^n(mW) = \{(a_1, \ldots, a_k) : 1 \leq a_j \leq n \text{ für } j = 1, \ldots, k\}$$

(k-Permutationen aus $1, 2, \ldots, n$ mit Wiederholung)

(2) Ziehen unter Beachtung der Reihenfolge ohne Zurücklegen

Sei $k \leq n$. Ziehen wie oben, aber ohne Zurücklegen.
Geeigneter Ergebnisraum:

$$\mathrm{Per}_k^n(oW) = \{(a_1, \ldots, a_k) \in \{1, 2, \ldots, n\}^k : a_i \neq a_j \text{ für } 1 \leq i \neq j \leq k\}$$

(k-Permutationen aus $1, \ldots, n$ ohne Wiederholung)

(3) Ziehen ohne Beachtung der Reihenfolge mit Zurücklegen

Ziehen mit Zurücklegen; am Ende nur Info, *wie oft* jede der n Kugeln gezogen wurde. Geeigneter Ergebnisraum:

$$\mathrm{Kom}_k^n(mW) \quad = \quad \{(a_1, \ldots, a_k) \in \{1, 2, \ldots, n\}^k : a_1 \leq \ldots \leq a_k\}$$

(k-Kombinationen aus $1, \ldots, n$ mit Wiederholung).

a_j ist j-kleinste Nummer der gezogenen Kugeln.

(4) Ziehen ohne Beachtung der Reihenfolge ohne Zurücklegen

Ziehen wie in (3), aber ohne Zurücklegen (vgl. Lotto), $k \leq n$
Geeigneter Ergebnisraum:

$$\mathrm{Kom}_k^n(oW) \quad = \quad \{(a_1, \ldots, a_k) \in \{1, 2, \ldots, n\}^k : a_1 < \ldots < a_k\}$$

(k-Kombinationen aus $1, \ldots, n$ ohne Wiederholung)

4.2 Fächer–Modelle

Es werden k Teilchen auf n von 1 bis n nummerierte Fächer verteilt.

(1) Unterscheidbare Teilchen, Mehrfachbesetzungen zugelassen

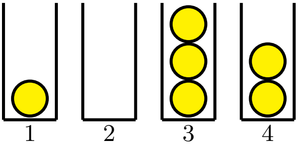
Geeigneter Grundraum $= \mathrm{Per}_k^n(mW)$.
$a_j =$ Nummer des Fachs, in dem das j-te Teilchen liegt.

(2) Unterscheidbare Teilchen, keine Mehrfachbesetzungen

Geeigneter Merkmalraum $= \mathrm{Per}_k^n(oW)$.

(3) Nichtunterscheidbare Teilchen, Mehrfachbesetzungen zugelassen

Geeigneter Merkmalraum $= \mathrm{Kom}_k^n(mW)$.



$(1, 3, 3, 3, 4, 4) \in Kom_6^4(mW)$

(4) Nichtunterscheidbare Teilchen, keine Mehrfachbesetzungen

Geeigneter Merkmalraum $= \mathrm{Kom}_k^n(oW)$

Beachte: Urnen- und Fächer-Modelle sind begrifflich gleichwertig!

Teilchen in Fach Nr. j legen	$\Longleftrightarrow$	Kugel Nr. j ziehen.
Teilchen unterscheidbar	$\Longleftrightarrow$	Reihenfolge beachten.
Mit Zurücklegen	$\Longleftrightarrow$	Mehrfachbesetzungen zugelassen.

- Kollisionsprobleme (Geburtstagsproblem, Lotto-Gewinnreihenwiederholung)
- Zwei-Drittel-Gesetz beim Roulette
- Vollständige Serien, Sammelbilder-Probleme

4.3 Beispiel (Das Paradoxon der ersten Kollision)

Erstmals im Lotto dieselbe Zahlenreihe

Stuttgart *(dpa/lsw). Die Staatliche Toto-Lotto GmbH in Stuttgart hat eine Lottosensation gemeldet: Zum ersten Mal in der 40jährigen Geschichte des deutschen Zahlenlottos wurden zwei identische Gewinnreihen festgestellt. Am 21. Juni dieses Jahres kam im Lotto am Mittwoch in der Ziehung A die Gewinnreihe 15–25–27–30–42–48 heraus. Genau dieselben Zahlen wurden bei der 1628. Ausspielung im Samstaglotto schon einmal gezogen, nämlich am 20. Dezember 1986. Welch ein Lottozufall: Unter den 49 Zahlen sind fast 14 Millionen verschiedene Sechserreihen möglich.*

In der 3016. Ausspielung war zum ersten Mal eine Gewinnreihenwiederholung aufgetreten!

Es gibt

$$n \;:=\; \binom{49}{6} \;=\; 13\ 983\ 816$$

mögliche Gewinnreihen. Gedankliche Durchnummerierung:

Nr. 1:	1	-	2	-	3	-	4	-	5	-	**6**
Nr. 2:	1	-	2	-	3	-	4	-	5	-	**7**
Nr. 3:	1	-	2	-	3	-	4	-	5	-	**8**
$\vdots$					$\vdots$						$\vdots$
Nr. 44:	1	-	2	-	3	-	4	-	5	-	**49**
Nr. 45:	1	-	2	-	3	-	4	-	**6**	-	**7**
Nr. 46:	1	-	2	-	3	-	4	-	6	-	**8**
$\vdots$					$\vdots$						$\vdots$
Nr. n:	44	-	45	-	46	-	47	-	48	-	**49**

Gewinnreihenermittlung ist rein zufälliges Besetzen eines von insgesamt n verschiedenen Fächern.

Modellierung: Sei

$$X_n \quad := \quad \text{Zeitpunkt der ersten Kollision beim sukzessiven} \\ \text{rein zufälligen Besetzen von } n \text{ Fächern.}$$

Welche Werte nimmt X_n an? Antwort: $2, 3, \ldots, n+1$.

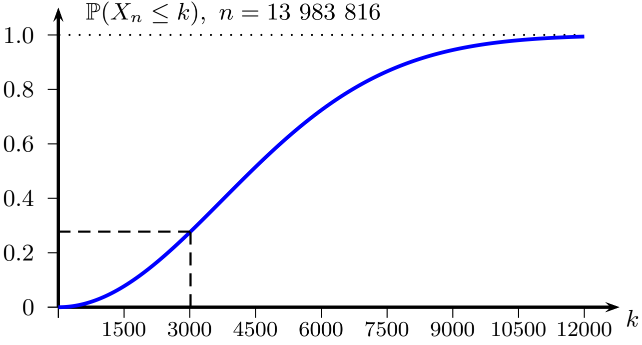
$$X_n \geq k+1 \iff \text{die ersten } k \text{ Teilchen fallen in verschiedene Fächer}$$

$$\begin{aligned} \mathbb{P}(X_n \geq k+1) &= \frac{n \cdot (n-1) \cdot (n-2) \cdot \ \ldots \ \cdot (n-k+1)}{n^k} \\ &= \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) \end{aligned}$$

$k = 1, 2, \ldots, n+1$. Komplementbildung liefert

$$\mathbb{P}(X_n \le k) \; = \; 1 - \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right), \qquad k = 1, 2, 3, \ldots, n+1.$$

Für $n = 13\ 983\ 816$ gilt $\mathbb{P}(X_n \le 3016) \; = \; 0.2775\ldots \approx 10/36$.



Wahrscheinlichkeit für die erste Gewinnreihenwiederholung im Lotto nach höchstens k Ziehungen

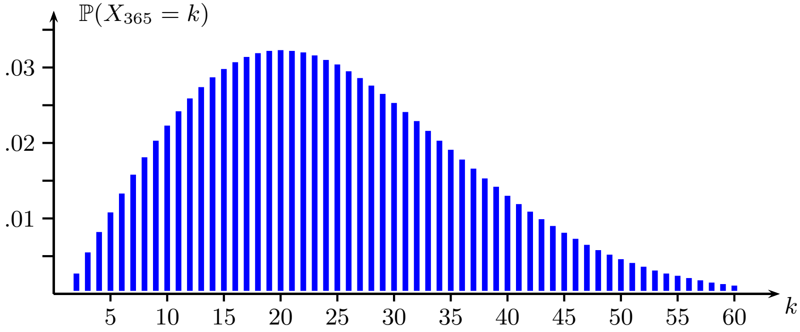
Memo: $\mathbb{P}\left(X_n \le k\right) \;=\; 1 - \prod_{j=1}^{k-1}\left(1 - \frac{j}{n}\right), \qquad k = 1, 2, 3, \ldots, n+1.$

Beachte: $\{X_n \le k-1\} + \{X_n = k\} = \{X_n \le k\}$. Also:

$$\begin{aligned}\mathbb{P}(X_n = k) &= \mathbb{P}(X_n \le k) - \mathbb{P}(X_n \le k-1)\\ &= \;\ldots\; = \frac{k-1}{n} \cdot \prod_{j=1}^{k-2}\left(1 - \frac{j}{n}\right),\end{aligned}$$

$k = 2, 3, \ldots, n+1.$

$n = 365$ (Tage des Jahres, ohne 29. Februar)



$\mathbb{P}(X_{365} \le 23) \approx 0.503$ (Geburtstags-Paradoxon)

Es gilt

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_n}{\sqrt{n}} \le t\right) = 1 - \exp\left(-\frac{t^2}{2}\right), \quad t > 0.$$

5 Der Erwartungswert

Motivation: Stochastischer Vorgang (z.B. Glücksspiel) mit Ergebnismenge $\Omega = \{\omega_1, \ldots, \omega_s\}$.

Sei $X(\omega_j)$ der Gewinn bei Ausgang ω_j.

Vorgang n mal unter gleichen, sich gegenseitig nicht beeinflussenden Bedingungen durchführen.

H_j mal trete der Ausgang ω_j auf, $j = 1, \ldots, s$.

Gesamtgewinn: $X(\omega_1) \cdot H_1 + \ldots + X(\omega_s) \cdot H_s$.

Durchschnittlicher Gewinn pro Vorgang:

$$X(\omega_1) \cdot \underbrace{\frac{H_1}{n}}_{\rightsquigarrow \mathbb{P}(\{\omega_1\})} + \ldots + X(\omega_s) \cdot \underbrace{\frac{H_s}{n}}_{\rightsquigarrow \mathbb{P}(\{\omega_s\})}.$$

(empirisches Gesetz über die Stabilisierung relativer Häufigkeiten)

5.1 Definition (Erwartungswert)

Seien $(\Omega, \mathbb{P})$ ein endlicher W-Raum, wobei $\Omega := \{\omega_1, \ldots, \omega_s\}$, und $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann heißt

$$\mathbb{E}(X) \;:=\; \sum_{j=1}^{s} X(\omega_j) \cdot \mathbb{P}(\{\omega_j\})$$

der Erwartungswert von X.

Beachte: Ist $\Omega = \{\omega_j : j \geq 1\}$ abzählbar-unendlich, so fordert man

$$\sum_{j=1}^{\infty} |X(\omega_j)|\, \mathbb{P}(\{\omega_j\}) < \infty$$

und definiert unter dieser Bedingung

$$\mathbb{E}(X) \;:=\; \sum_{j=1}^{\infty} X(\omega_j)\, \mathbb{P}(\{\omega_j\}).$$

Memo: $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}(\{\omega\})$

5.2 Satz (Strukturelle Eigenschaften der Erwartungswertbildung)

Es seien X und Y Zufallsvariablen auf Ω. Dann gelten:

a) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,

b) $\mathbb{E}(a \cdot X) = a \cdot \mathbb{E}(X), \qquad a \in \mathbb{R}$,

c) $X \leq Y \implies \mathbb{E}(X) \leq \mathbb{E}(Y)$,

d) $\mathbb{E}(\mathbf{1}_A) = \mathbb{P}(A), \quad A \subseteq \Omega$.

Botschaft: Die Zuordnungsvorschrift $X \mapsto \mathbb{E}(X)$ ist additiv, homogen und monoton.

Memo: $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y), \quad \mathbb{E}(\mathbf{1}_A) = \mathbb{P}(A)$

5.3 Folgerung (Erwartungswert einer Zählvariablen)

Seien $A_1, \ldots, A_n \subseteq \Omega$ Ereignisse und

$$X \; = \; \sum_{j=1}^{n} \mathbf{1}_{A_j}$$

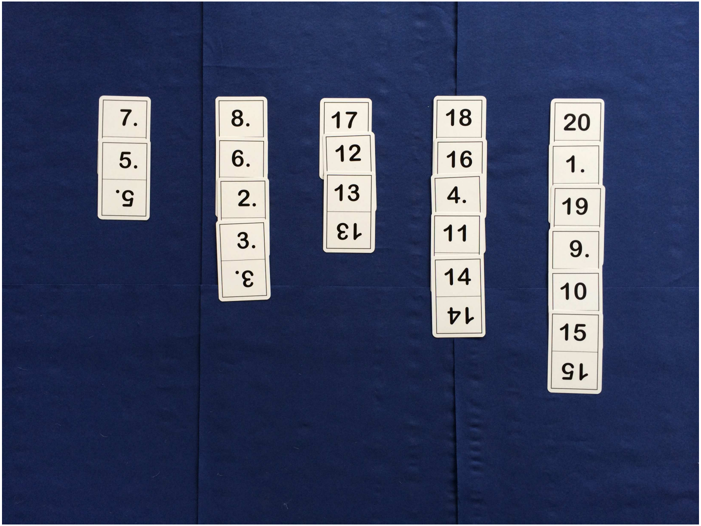
die Anzahl der eintretenden A_j. Dann gilt $\mathbb{E}(X) \; = \; \sum_{j=1}^{n} \mathbb{P}(A_j)$.

Gilt speziell $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \ldots = \mathbb{P}(A_n) =: p$, so folgt $\mathbb{E}(X) = n \cdot p$.

5.4 Folgerung Sei X die Anzahl der gezogenen roten Kugeln beim n-maligen rein zufälligen Ziehen (mit oder ohne Zurücklegen!) aus einer Urne mit r roten und s schwarzen Kugeln. Dann gilt

$$\mathbb{E}(X) = n \cdot \frac{r}{r+s} \qquad \text{(warum ?)}$$

5.5 Beispiel (Anzahl der Rekorde in rein zufälliger Permutation)



Sei $\Omega := \mathrm{Per}_n^n(oW)$ die Menge der Permutationen von $1, 2, \ldots, n$.

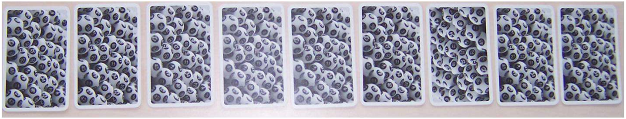
$\mathbb{P}$ sei die Gleichverteilung auf Ω.

Für $j = 1, \ldots, n$ sei

$$A_j := \{(a_1, \ldots, a_n) \in \Omega : a_j = \max(a_1, a_2, \ldots, a_j)\}.$$

Sei $X_n := \mathbf{1}_{A_1} + \ldots + \mathbf{1}_{A_n}$ die Anzahl der Rekorde.

$$\mathbb{P}(A_j) \; = \; ?\frac{1}{j} \quad (!)$$



$$\Longrightarrow \; \mathbb{E}(X_n) \; = \; 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$$

$\mathbb{E}(X_{32}) \approx 4.06, \qquad \mathbb{E}(X_{1000000000}) \approx 20.3.$

5.6 Satz (Darstellungsformel für $\mathbb{E}(X)$ mithilfe der Verteilung)

Nimmt die Zufallsvariable X die verschiedenen Werte $x_1, \ldots, x_k$ an, so gilt:

$$\mathbb{E}(X) = \sum_{j=1}^{k} x_j \cdot \mathbb{P}(X = x_j) \qquad \text{(ist Definition von } \mathbb{E}(X) \text{ in [LS-K])}$$

Beweis: Sei $A_j := \{\omega \in \Omega : X(\omega) = x_j\}$, $j = 1, \ldots, k$. Es folgt

$$\begin{aligned} \mathbb{E}(X) &= \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}(\{\omega\}) = \sum_{j=1}^{k} \sum_{\omega \in A_j} X(\omega) \cdot \mathbb{P}(\{\omega\}) \\ &= \sum_{j=1}^{k} x_j \cdot \sum_{\omega \in A_j} \mathbb{P}(\{\omega_j\}) = \sum_{j=1}^{k} x_j \cdot \mathbb{P}(X = x_j). \end{aligned}$$

Allgemeiner gilt: Ist $g : \mathbb{R} \to \mathbb{R}$ eine Funktion, so folgt:

$$\mathbb{E}(g(X)) = \sum_{j=1}^{k} g(x_j) \cdot \mathbb{P}(X = x_j).$$

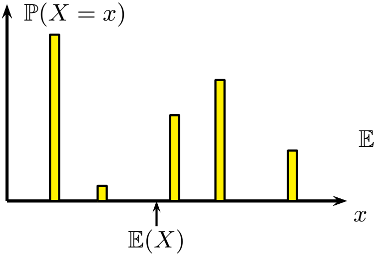
5.7 Beispiel

Die Zufallsvariable X besitze eine Gleichverteilung auf $\{1, 2, \ldots, k\}$, d.h. es gilt

$$\mathbb{P}(X = j) \;=\; \frac{1}{k}, \quad j = 1, \ldots, k. \qquad (\text{Beachte: } \Omega \text{ wird nicht spezifiziert!})$$

$$\Longrightarrow \mathbb{E}(X) \;=\; \sum_{j=1}^{k} j\,\mathbb{P}(X = j) \;=\; \frac{1}{k}\sum_{j=1}^{k} j \;=\; \frac{1}{k}\cdot\frac{k(k+1)}{2} \;=\; \frac{k+1}{2}.$$

Der Erwartungswert von X muss also keine Realisierung von X sein!



$\mathbb{E}(X)$ ist physikalischer Schwerpunkt!

Memo: $\mathbb{E}\left(\sum_{j=1}^n \mathbf{1}_{A_j}\right) = \sum_{j=1}^n \mathbb{P}(A_j)$.

5.8 Beispiel (Anzahl der Fixpunkte in einer rein zufälligen Permutation)

$$\Omega = \begin{Bmatrix} 1234 & 2134 & 3124 & 4123 \\ 1243 & 2143 & 3142 & 4132 \\ 1324 & 2314 & 3214 & 4213 \\ 1342 & 2341 & 3241 & 4231 \\ 1423 & 2413 & 3412 & 4312 \\ 1432 & 2431 & 3421 & 4321 \end{Bmatrix}$$

Sei

$$A_j = \{a_1a_2a_3a_4 \in \Omega : a_j = j\}, \quad j = 1, 2, 3, 4.$$

Es gilt

$$\mathbb{P}(A_1) = \frac{1}{4} = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \mathbb{P}(A_4).$$

Es folgt

$$\mathbb{E}\left(\sum_{j=1}^4 \mathbf{1}_{A_j}\right) = 4 \cdot \frac{1}{4} = 1.$$

Allgemein bei einer rein zufälligen Vertauschung der Zahlen $1, 2, \ldots, n$?

Das Schnur-Orakel

n Schnüre (im Bild $n = 4$) werden in der Mitte festgehalten, so dass $2n$ Enden frei sind. Diese Enden werden rein zufällig verknotet.

Welchen Erwartungswert (welche Verteilung) besitzt die Anzahl R_n der dabei entstehenden (geschlossenen) Ringe?

$$\text{Sei } A_j \quad := \quad \{j\text{-te Verknotung führt zu einem Ring}\}, \quad j = 1, \ldots, n$$

$$\Longrightarrow \; R_n \quad = \quad \sum_{j=1}^{n} \mathbf{1}_{A_j}$$

Bei $n = 4$ Schnüren: $\mathbb{P}(A_1) = \frac{1}{7}$, $\mathbb{P}(A_2) = \frac{1}{5}$, $\mathbb{P}(A_3) = \frac{1}{3}$, $\mathbb{P}(A_4) = 1$

$$\text{Allg.: } \mathbb{P}(A_1) = \frac{1}{2n-1}, \; \mathbb{P}(A_2) = \frac{1}{2n-3}, \ldots, \mathbb{P}(A_{n-1}) = \frac{1}{3}, \; \mathbb{P}(A_n) = 1.$$

$$\mathbb{E}(R_n) \; = \; \sum_{j=1}^{n} \mathbb{P}(A_j) = 1 + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{2n-1} = \sum_{k=1}^{2n} \frac{1}{k} - \sum_{j=1}^{n} \frac{1}{2j}$$

$$\Longrightarrow \mathbb{E}(R_n) = \ldots \approx \frac{\log n}{2} + 0.98175 \quad (!!!).$$

n	3	10	100	1000	10^6	10^9
$\mathbb{E}(R_n)$	1.53	2.18	3.28	4.44	7.89	11.34

6 Binomialverteilung und hypergeometrische Verteilung

Situation: Urne mit r roten und s schwarzen Kugeln.

Die Kugeln seien gedanklich nummeriert: $1, 2, \ldots, r, r+1, \ldots, r+s$.
Es wird n mal rein zufällig mit Zurücklegen gezogen.

$$\Omega := \mathrm{Per}_n^{r+s}(mW), \quad \mathbb{P} := \text{Gleichverteilung auf } \Omega$$

$$A_j := \{(a_1, \ldots, a_n) \in \Omega : a_j \leq r\} \quad (j\text{-te Kugel rot})$$

$$X := \sum_{j=1}^{n} \mathbf{1}_{A_j} \qquad (= \text{Anzahl der gezogenen roten Kugeln})$$

$$|\Omega| = (r+s)^n$$

$$|A_j| = r \cdot (r+s)^{n-1}$$

$$\mathbb{P}(A_j) = \frac{|A_j|}{|\Omega|} = \frac{r}{r+s} =: p, \qquad j = 1, 2, \ldots, n$$

Memo: $X = \sum_{j=1}^{n} \mathbf{1}_{A_j}$, $\mathbb{P}(A_j) = p$, $A_j = \{(a_1, \ldots, a_n) \in \mathrm{Per}_n^{r+s}(mW) : a_j \leq r\}$

6.1 Satz Es gelten:

a) $\mathbb{E}(X) \;=\; n\,p,$

b) $\mathbb{P}(X = k) \;=\; \binom{n}{k} p^k \, (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$

Beweis von b)

$$
\begin{aligned}
|\{X = k\}| &= |\{(a_1, \ldots, a_n) : \text{ genau } k \text{ der } a_j \text{ sind } \leq r\}| \\
&= \binom{n}{k} \cdot |\{(a_1, \ldots, a_n) : a_1 \leq r, \ldots, a_k \leq r, a_{k+1} > r, \ldots, a_n > r\}| \\
&= \binom{n}{k} \cdot r^k \cdot s^{n-k}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}(X = k) &= \frac{|\{X = k\}|}{(r+s)^n} = \binom{n}{k} \cdot \left(\frac{r}{r+s}\right)^k \cdot \left(\frac{s}{r+s}\right)^{n-k} \\
&= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}
\end{aligned}
$$

Jetzt: Ziehen ohne Zurücklegen, wobei $n \leq r+s$.

$$\Omega := \mathrm{Per}_n^{r+s}(oW), \qquad \mathbb{P} := \text{Gleichverteilung auf } \Omega,$$

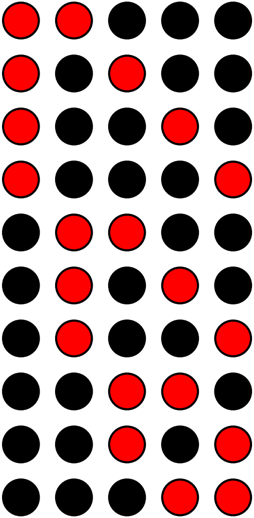
$$A_j := \{(a_1, \ldots, a_n) \in \Omega : a_j \leq r\} \quad (j\text{-te Kugel rot})$$

$$X := \sum_{j=1}^{n} \mathbf{1}_{A_j} \qquad (= \text{Anzahl der gezogenen roten Kugeln})$$

$$|\Omega| = (r+s)\cdot(r+s-1)\cdot\ldots\cdot(r+s-(n-1)) = (r+s)^{\underline{n}}$$

$$|A_j| = r\cdot(r+s-1)\cdot\ldots\cdot(r+s-(n-1)) = r\cdot(r+s-1)^{\underline{n-1}}$$

$$\mathbb{P}(A_j) = \frac{|A_j|}{|\Omega|} = \frac{r}{r+s}, \qquad j = 1, \ldots, n. \quad (!)$$



Memo: $X = \sum_{j=1}^{n} \mathbf{1}_{A_j}, \; A_j = \{(a_1, \ldots, a_n) \in \mathrm{Per}_n^{r+s}(oW) : a_j \leq r\}$

6.2 Satz Es gelten:

a) $\mathbb{E}(X) \; = \; n \cdot \dfrac{r}{r+s},$

b) $\mathbb{P}(X = k) \; = \; \dfrac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}}, \qquad k = 0, 1, \ldots, n, \; \binom{m}{\ell} := 0,$ falls $m < \ell$.

$$
\begin{aligned}
\text{b: } |\{X = k\}| &= |\{(a_1, \ldots, a_n) : \text{ genau } k \text{ der } a_j \text{ sind } \leq r\}| \\
&= \binom{n}{k} \cdot |\{(a_1, \ldots, a_n) : a_1 \leq r, \ldots, a_k \leq r, a_{k+1} > r, \ldots, a_n > r\}| \\
&= \binom{n}{k} \cdot r(r-1) \ldots (r-(k-1)) \cdot s(s-1) \ldots (s-(n-k-1))
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}(X = k) &= \frac{|\{X = k\}|}{(r+s)^{\underline{n}}} = \binom{n}{k} \cdot \frac{r!}{(r-k)!} \cdot \frac{s!}{(s-(n-k))!} \cdot \frac{(r+s-n)!}{(r+s)!} \\
&= \frac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}} \qquad \text{(sog. hypergeometrische Verteilung } \mathrm{Hyp}(n, r, s))
\end{aligned}
$$

Selbsttest:

Können Sie für jeden der beiden Fälle zeigen, dass

$$\sum_{k=0}^{n} k \cdot \mathbb{P}(X = k) \; = \; n \cdot p, \qquad p = \frac{r}{r+s},$$

gilt?

7 Modellierung mehrstufiger Experimente

Viele stochastische Vorgänge bestehen aus Teilexperimenten (Stufen).

Ergebnisse eines n-stufigen Experiments sind n-Tupel $\omega = (a_1, a_2, \ldots, a_n)$.

Dabei sei a_j der Ausgang des j-ten Teilexperiments.

Sei Ω_j die Ergebnismenge des j-ten Teilexperiments.

$$\Omega \quad := \quad \Omega_1 \times \cdots \times \Omega_n \; = \; \{\omega = (a_1, \ldots, a_n) : a_j \in \Omega_j \;\text{ für }\; j = 1, \ldots, n\}$$

ist natürlicher Grundraum für das Gesamt–Experiment.

7.1 Beispiel (Pólya-Urnenschema)

Eine Urne enthalte eine rote und drei schwarze Kugeln.

Rein zufällig Kugel ziehen und Farbe notieren.

Diese sowie eine weitere Kugel derselben Farbe in die Urne zurücklegen.

Nach Mischen wieder Kugel ziehen. Mit welcher W' ist diese rot?

Ziehen einer roten (bzw. schwarzen) Kugel: „r“ (bzw. „s“)

$\Omega := \Omega_1 \times \Omega_2$ mit $\Omega_1 = \Omega_2 = \{r, s\}$

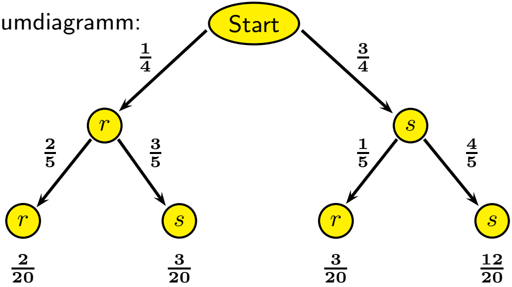
$B := \{(r,r),(s,r)\}$ (die beim zweiten Mal gezogene Kugel ist rot)

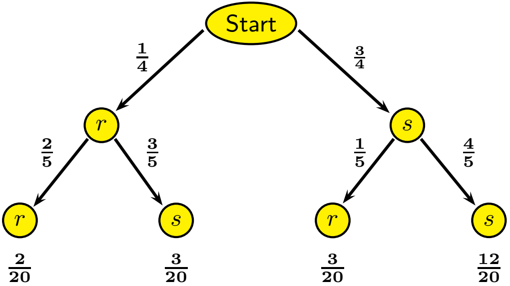
Festlegung der Wahrscheinlichkeiten $p(\omega) := \mathbb{P}(\{\omega\})$:

$$p(r,r) := \frac{1}{4}\cdot\frac{2}{5}, \qquad p(r,s) := \frac{1}{4}\cdot\frac{3}{5}, \qquad \text{(„Erste Pfadregel“)}$$

$$p(s,r) := \frac{3}{4}\cdot\frac{1}{5}, \qquad p(s,s) := \frac{3}{4}\cdot\frac{4}{5}. \qquad \text{(motiviert durch relat. Häufigkeiten!)}$$

Baumdiagramm:





$$\mathbb{P}(B) \;=\; p(r,r) + p(s,r) \;=\; \frac{2}{20} + \frac{3}{20} \;=\; \frac{1}{4} \quad (!)$$

(„Zweite Pfadregel“)

Allgemein erhält man ein begründetes W-Maß auf (den Teilmengen von) Ω mithilfe einer Start-Verteilung und Übergangswahrscheinlichkeiten.

7.2 Satz und Definition

Ω_1 und Ω_2 seien elementare Grundräume, $\mathbb{P}_1$ ein W-Maß auf Ω_1, (sog. Startverteilung), mit W-Funktion $p_1(\omega_1) := \mathbb{P}_1(\{\omega_1\})$, $\omega_1 \in \Omega_1$.

$p_2 : \Omega_1 \times \Omega_2 \to \mathbb{R}$ sei eine Übergangs-W-Funktion von Ω_1 nach Ω_2,

d.h. eine Funktion mit den Eigenschaften

$$\begin{aligned} p_2(\omega_1, \omega_2) &\geq 0, \quad (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2, \\ \sum_{\omega_2 \in \Omega_2} p_2(\omega_1, \omega_2) &= 1, \quad \omega_1 \in \Omega_1. \end{aligned}$$

Dann definiert

$$p(\omega_1, \omega_2) := p_1(\omega_1) \cdot p_2(\omega_1, \omega_2) \qquad \text{(sog. Erste Pfadregel)}$$

eine W-Funktion auf $\Omega := \Omega_1 \times \Omega_2$. Das durch

$$\mathbb{P}(A) := \sum_{\omega \in A} p(\omega), \quad A \subseteq \Omega,$$

definierte W-Maß $\mathbb{P}$ zu p heißt Kopplung von p_1 und p_2.

7.3 Satz (Zweite Pfadregel)

$\mathbb{P}$ sei die Kopplung von p_1 und p_2. Sei $A_2 \subseteq \Omega_2$ und $A := \Omega_1 \times A_2$.

(A bezieht sich nur auf das zweite Teilexperiment!)

Dann gilt:

$$\mathbb{P}(A) \;=\; \sum_{\omega_2 \in A_2} \left(\sum_{\omega_1 \in \Omega_1} p_1(\omega_1) \cdot p_2(\omega_1, \omega_2) \right)$$

(Summation der W'en aller Pfade, die zu einem Endknoten in A_2 führen).

Interpretation der Übergangswahrscheinlichkeit $p_2(\omega_1, \omega_2)$:

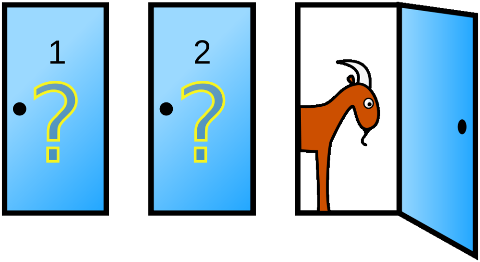
W', dass das zweite Teilexperiment den Ausgang ω_2 hat, wenn das erste Teilexperiment den Ausgang ω_1 hat.

Die Modellierung n-stufiger Experimente geschieht induktiv.

Falls für jedes ω_2 die ÜW $p_2(\omega_1, \omega_2)$ nicht von ω_1 abhängt, so sind die Experimente „unabhängig voneinander“ ($\rightarrow$ stochastische Unabhängigkeit).

8 Bedingte Wahrscheinlichkeiten

Das Ziegen-Problem (Drei-Türen-Problem, Monty-Hall-Problem)



Hinter einer von drei Türen befindet sich ein Auto, hinter den beiden anderen jeweils eine Ziege. Der Kandidat zeigt auf Tür 1; diese bleibt zunächst verschlossen. Der Moderator weiß, hinter welcher Tür sich das Auto befindet. Er darf die Auto-Tür nicht öffnen, muss aber eine Ziege zu erkennen geben. Der Moderator öffnet Tür 3 und bietet an, von Tür 1 zu Tür 2 zu wechseln.

Soll man das tun?

Über einen stochastischen Vorgang sei bekannt, dass ein Ereignis B eingetreten ist. Wie beeinflusst diese Bedingung (Information) die Aussicht auf das Eintreten eines Ereignisses A?

Motivierung bedingter Wahrscheinlichkeiten mithilfe relativer Häufigkeiten

Experiment mit Grundraum Ω n mal „in unabhängiger Folge“ wiederholen.

Seien $A, B \subseteq \Omega$.

Sei $H_n(B)$ die Anzahl der Male, bei denen das Ereignis B eintritt,
$H_n(A \cap B)$ die Anzahl der Male, bei denen sowohl A als auch B eintreten.

$$\frac{H_n(A \cap B)}{H_n(B)} = \frac{\frac{1}{n} H_n(A \cap B)}{\frac{1}{n} H_n(B)} \quad \begin{array}{l} \rightsquigarrow \mathbb{P}(A \cap B) \\ \rightsquigarrow \mathbb{P}(B) \end{array}$$

ist relativer Anteil derjenigen Fälle unter allen Fällen, in denen B eintritt, in denen auch noch A eintritt.

Memo: $\dfrac{H_n(A \cap B)}{H_n(B)} = \dfrac{\frac{1}{n} H_n(A \cap B)}{\frac{1}{n} H_n(B)} \quad \begin{array}{l} \rightsquigarrow \mathbb{P}(A \cap B) \\ \rightsquigarrow \mathbb{P}(B) \end{array}$

8.1 Definition (bedingte Wahrscheinlichkeit)

Seien $(\Omega, \mathbb{P})$ ein diskreter W-Raum, $A, B \subseteq \Omega$ mit $\mathbb{P}(B) > 0$. Dann heißt

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

die bedingte Wahrscheinlichkeit von A unter der Bedingung B.

Achtung! Es gibt kein Ereignis $A|B$, vgl. [LFS], S. 116.

8.2 Zusammenhang mit Übergangswahrscheinlichkeiten

Betrachte zweistufiges Experiment:

$$\Omega = \Omega_1 \times \Omega_2, \quad \omega = (a_1, a_2), \quad p(\omega) := p_1(a_1) p_2(a_1, a_2)$$

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega), \qquad A \subseteq \Omega.$$

Seien $a_1 \in \Omega_1$, $a_2 \in \Omega_2$.

$$B := \{a_1\} \times \Omega_2 \qquad \text{(beim 1. Teilexperiment tritt Ergebnis } a_1 \text{ auf)},$$
$$A := \Omega_1 \times \{a_2\} \qquad \text{(beim 2. Teilexperiment tritt Ergebnis } a_2 \text{ auf)}.$$

Es gilt $A \cap B = \{(a_1, a_2)\} \qquad \mathbb{P}(A \cap B) = p_1(a_1) \cdot p_2(a_1, a_2)$,

$$\mathbb{P}(B) = \sum_{\widetilde{a}_2 \in \Omega_2} p_1(a_1) \cdot p_2(a_1, \widetilde{a}_2) = p_1(a_1) \cdot \underbrace{\sum_{\widetilde{a}_2 \in \Omega_2} p_2(a_1, \widetilde{a}_2)}_{=1} = p_1(a_1)$$

$$\Longrightarrow \mathbb{P}(A|B) = p_2(a_1, a_2).$$

Übergangswahrscheinlichkeiten sind bedingte Wahrscheinlichkeiten!

Beachte: Es gilt

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

Hierbei meist $\mathbb{P}(A)$ und $\mathbb{P}(B|A)$ als Modellbausteine gegeben!

8.3 Satz (Multiplikationsformel)

Es seien $A_1, \ldots, A_n$ Ereignisse mit $\mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$. Dann gilt

$$\begin{aligned} & \mathbb{P}(A_1 \cap \ldots \cap A_n) \\ = \; & \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1}). \end{aligned}$$

Die erste Pfadregel ist ein Spezialfall der Multiplikationsformel!

8.4 Satz (Formel von der totalen Wahrscheinlichkeit, Bayes-Formel)

Es seien $A_1, \ldots, A_n$ paarweise disjunkte Ereignisse mit $\mathbb{P}(A_j) > 0$ für jedes j und $\sum_{j=1}^n A_j = \Omega$. Weiter sei B ein Ereignis. Dann gelten:

a) $\mathbb{P}(B) \; = \; \sum\limits_{j=1}^n \mathbb{P}(A_j) \cdot \mathbb{P}(B|A_j)$ (Formel von der totalen Wahrscheinlichkeit)

b) Falls $\mathbb{P}(B) > 0$, so gilt für jedes k

$$\mathbb{P}(A_k|B) \quad = \quad \frac{\mathbb{P}(A_k) \cdot \mathbb{P}(B|A_k)}{\sum\limits_{j=1}^n \mathbb{P}(A_j) \cdot \mathbb{P}(B|A_j)} \qquad \text{(Bayes-Formel)}$$

Beweis:

$$\mathbb{P}(B) \; = \; \sum_{j=1}^n \mathbb{P}(B \cap A_j) \; = \; \sum_{j=1}^n \mathbb{P}(A_j)\mathbb{P}(B|A_j).$$

Bemerkung: Im Zusammenhang mit der Bayes-Formel nennt man

$\mathbb{P}(A_1), \ldots, \mathbb{P}(A_n) \ldots$ a-priori-Wahrscheinlichkeiten

$\mathbb{P}(A_1|B), \ldots, \mathbb{P}(A_n|B) \ldots$ a-posteriori-Wahrscheinlichkeiten.

8.5 Beispiel (Test auf eine seltene Krankheit)

$\Omega \ := \ \{(0,\ominus),(0,\oplus),(1,\ominus),(1,\oplus)\}$,

1 (0) in erster Komponente: krank bzw. gesund

$\oplus$ $(\ominus)$ in zweiter Komponente: positives (negatives) Testergebnis

$K := \{(1,\ominus),(1,\oplus)\}$ „krank“, $\quad G := \{(0,\ominus),(0,\oplus)\}$ „gesund“ ,

$\ominus = \{(1,\ominus),(0,\ominus)\}$ „Test negativ“ , $\quad \oplus = \{(1,\oplus),(0,\oplus)\}$ „Test positiv“

Modellannahmen: $\mathbb{P}(K) \ = \ q$,

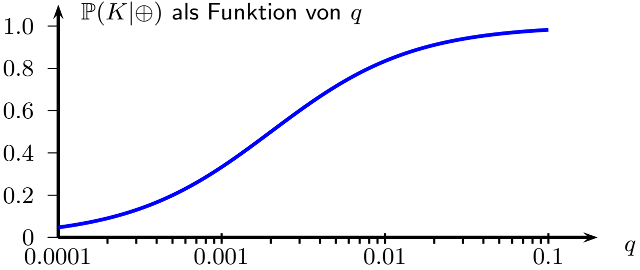
$$\mathbb{P}(\oplus|K) \ = \ p_{se}, \quad \text{„Sensitivität“}, \quad \mathbb{P}(\ominus|G) \ = \ p_{sp} \quad \text{„Spezifität“}.$$

Bayes-Formel $\Longrightarrow$ $\mathbb{P}(K|\oplus) \ = \ \dfrac{\mathbb{P}(K)\cdot\mathbb{P}(\oplus|K)}{\mathbb{P}(K)\cdot\mathbb{P}(\oplus|K) \ + \ \mathbb{P}(G)\cdot\mathbb{P}(\oplus|G)}$.

$\mathbb{P}(G) = 1-q, \quad \mathbb{P}(\oplus|G) = 1-p_{sp} \Longrightarrow$

$$\mathbb{P}(K|\oplus) \ = \ \frac{q\cdot p_{se}}{q\cdot p_{se} \ + \ (1-q)\cdot(1-p_{sp})} \ .$$

$p_{se} = p_{sp} = 0.998$



1 000 000 Menschen Bluttest unterziehen, davon 1000 krank.

Von den 1000 Kranken erhalten ca. 998 ein positives Ergebnis.

Von den 999 000 Gesunden erhalten ca. 1 998 ein positives Ergebnis.

Insgesamt gibt es ca. 2 996 Personen mit positivem Testergebnis.
Davon ist nur ca. ein Drittel krank.

Sinn und Unsinn von Reihenuntersuchungen (!)

8.6 Beispiel (Eine männerfeindliche Universität?)

	Frauen		Männer	
	Bewerberinnen	zugelassen	Bewerber	zugelassen
Fach 1	900	720	200	**180**
Fach 2	100	20	800	**240**
Summe	**1000**	**740**	**1000**	**420**

$$\mathbf{0.74 \;=\; 0.9 \cdot 0.8 \;+\; 0.1 \cdot 0.2}, \qquad \mathbf{0.42 \;=\; 0.2 \cdot 0.9 \;+\; 0.8 \cdot 0.3}$$

8.7 Beispiel (Das Simpson-Paradoxon)

Für Ereignisse $A, B \subseteq \Omega$ sowie $\Omega = K_1 + \ldots + K_n$ kann Folgendes gelten:

$$\mathbb{P}(B|A \cap K_j) > \mathbb{P}(B|A^c \cap K_j) \text{ für jedes } j = 1, \ldots, n$$

und (!) $\mathbb{P}(B|A) < \mathbb{P}(B|A^c)$. (Simpson-Paradoxon)

Im obigen Beispiel: $n = 2$, $K_j = \{\text{Bewerbung in Fach } j\}$,

$$B = \{\text{aus allen 2000 Bewerbern zufällig gewählte Person wird zugelassen}\},$$
$$A = \{\text{aus allen 2000 Bewerbern zufällig gewählte Person ist männlich}\}.$$

Beachte: Nach der Formel von der totalen W' für $\mathbb{P}_A$ und $\mathbb{P}_{A^c}$ gilt

$$\mathbb{P}_A(B) = \sum_{j=1}^{n} \mathbb{P}_A(K_j) \cdot \mathbb{P}_A(B|K_j) \qquad (0.42 = 0.2 \cdot 0.9 + 0.8 \cdot 0.3)$$

$$\mathbb{P}_{A^c}(B) = \sum_{j=1}^{n} \mathbb{P}_{A^c}(K_j) \cdot \mathbb{P}_{A^c}(B|K_j) \qquad (0.74 = 0.9 \cdot 0.8 + 0.1 \cdot 0.2)$$

9 Stochastische Unabhängigkeit

Ergebnisse von 25 Würfelwürfen:

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

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

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

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

Begriffliche Schwierigkeiten!

- Stein-Schere-Papier (kann ich unabhängig wechseln?)
- Roulette (werden Zahlen bei längerem Ausbleiben wahrscheinlicher?)
- Lotto (s.o.)

Die bedingte Wahrscheinlichkeit $\mathbb{P}(A|B)$ von A unter der Bedingung B ist

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

Meist gilt

$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \ \neq \ \mathbb{P}(A).$$

Beispiel: Ziehen ohne Zurücklegen aus „r/s-Urne".

$$A \ = \ \{\text{„zweite Kugel rot"}\}, \qquad B \ = \ \{\text{„erste Kugel rot"}\}$$

$$\mathbb{P}(A) \ = \ \frac{r}{r+s}, \qquad \mathbb{P}(A|B) \ = \ \frac{r-1}{r+s-1}.$$

Falls $\mathbb{P}(A|B) = \mathbb{P}(A)$, so hat Eintreten von B wahrscheinlichkeitstheoretisch keinen Einfluss auf das Eintreten von A.

In gleicher Weise bedeutet

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

Die Wahrscheinlichkeit des Eintretens von B ist „unabhängig“ von der Information „A geschieht“.

Beachte:

$$\begin{aligned}\mathbb{P}(A|B) &= \mathbb{P}(A) \iff \mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B),\\ \mathbb{P}(B|A) &= \mathbb{P}(B) \iff \mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B).\end{aligned}$$

In diesem Fall heißen A und B (stochastisch) unabhängig.

Dabei ist auch $\mathbb{P}(A) = 0$ oder $\mathbb{P}(B) = 0$ zugelassen.

9.1 Definition (stochastische Unabhängigkeit)

Ereignisse $A_1, \ldots, A_n \subseteq \Omega$ heißen (stochastisch) unabhängig, falls gilt:

$$\mathbb{P}\left(\bigcap_{j \in T} A_j\right) = \prod_{j \in T} \mathbb{P}(A_j)$$

für *jede* Menge $T \subseteq \{1, 2, \ldots, n\}$ mit $|T| \geq 2$. ($2^n - n - 1$ Gleichungen)

Folgerung: Drei Ereignisse A, B und C sind unabhängig $\Longleftrightarrow$

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(A) \cdot \mathbb{P}(B), \\
\mathbb{P}(A \cap C) &= \mathbb{P}(A) \cdot \mathbb{P}(C), \\
\mathbb{P}(B \cap C) &= \mathbb{P}(B) \cdot \mathbb{P}(C), \\
\mathbb{P}(A \cap B \cap C) &= \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C).
\end{aligned}$$

Die ersten 3 Gleichungen bedeuten die paarweise Unabhängigkeit von A, B, C.

9.2 Wichtige Punkte im Zusammenhang mit Unabhängigkeit

a) A, B, C paarweise unabhängig $\not\Longrightarrow$ A, B, C unabhängig

Gegenbeispiel: Echte Münze zweimal in unabhängiger Folge werfen. Mögliche Ergebnisse: $(0,0), (0,1), (1,0), (1,1)$, je mit W' $1/4$. Das Paar (a_1, a_2) deterministisch durch $a_3 := a_1 + a_2 \pmod 2$ zu Tripel (a_1, a_2, a_3) ergänzen.

$$\Omega := \left\{ \begin{matrix} (0,0,0) \\ (0,1,1) \\ (1,0,1) \\ (1,1,0) \end{matrix} \right\}$$

Sei $A_j := \{(a_1, a_2, a_3) \in \Omega : a_j = 1\}$, $j = 1, 2, 3$. Dann gelten:

$$\begin{aligned} \mathbb{P}(A_1) &= \frac{1}{2} = \mathbb{P}(A_2) = \mathbb{P}(A_3), \\ \mathbb{P}(A_1 \cap A_2) &= \frac{1}{4} = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3), \\ \mathbb{P}(A_1 \cap A_2 \cap A_3) &= 0. \end{aligned}$$

Folg.: A_1, A_2, A_3 paarweise unabhängig, aber nicht unabhängig.

b) $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) \;\not\Longrightarrow\; A, B, C$ unabhängig

Gegenbeispiel: $\Omega = \{1,2,3,4,5,6,7,8\}$, Laplace-Modell. Sei

$$A := B := \{1,2,3,4\}, \quad C := \{1,5,6,7\}.$$

Es gilt $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$, $\mathbb{P}(A \cap B \cap C) = 1/8$. A und B sind nicht unabhängig!

c) Unabhängigkeit hat nichts mit Disjunktheit zu tun!

d) Unabhängigkeit ist von realer Beeinflussung zu unterscheiden!

Beispiel: Urne mit 2 roten und einer schwarzen Kugel; Ziehen ohne Z.

$A := \{$„erste Kugel rot“$\}$, $B := \{$„zweite Kugel rot“$\}$

$\mathbb{P}(B) = 2/3, \quad \mathbb{P}(B|A) = 1/2 \;\Longrightarrow\; A$ und B *nicht* unabhängig

B ist real beeinflusst von A, aber nicht A von B !

Mit A und B sind auch A und B^c unabhängig, da

$$\begin{aligned}\mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A) \cdot \mathbb{P}(B) \\ &= \mathbb{P}(A) \cdot (1 - \mathbb{P}(B)) \\ &= \mathbb{P}(A) \cdot \mathbb{P}(B^c).\end{aligned}$$

Damit sind auch A^c und B^c unabhängig!

Induktiv gilt:

9.3 Satz (Unabhängigkeit und Komplementbildung)

Es seien $(\Omega, \mathbb{P})$ ein W-Raum und $A_1, \ldots, A_n \subseteq \Omega$ Ereignisse, $n \geq 2$.

Dann sind folgende Aussagen äquivalent:

a) $A_1, \ldots, A_n$ sind stochastisch unabhängig.

b) Es gilt $\mathbb{P}\left(\bigcap_{i \in I} A_i \cap \bigcap_{j \in J} A_j^c\right) = \prod_{i \in I} \mathbb{P}(A_i) \cdot \prod_{j \in J} \mathbb{P}\left(A_j^c\right)$ für jede Wahl disjunkter Teilmengen I und J aus $\{1, 2, \ldots, n\}$.

9.4 Definition (Unabhängigkeit von Zufallsvariablen)

Zufallsvariablen $X_1, \ldots, X_n$ heißen (stochastisch) unabhängig, falls gilt:

$$\mathbb{P}\left(X_1 = k_1, \ldots, X_n = k_n\right) = \prod_{j=1}^{n} \mathbb{P}(X_j = k_j)$$

für alle $k_1, \ldots, k_n$ mit $\mathbb{P}(X_j = k_j) > 0$, $j = 1, \ldots, n$.

10 Bernoulli-Kette, Binomialverteilung

10.1 Satz (Erzeugungsweise der Binomialverteilung)

Es seien $A_1, \ldots, A_n$ unabhängige Ereignisse in einem W-Raum $(\Omega, \mathbb{P})$ mit $\mathbb{P}(A_j) =: p$, $j = 1, \ldots, n$. Dann gilt:

$$X := \sum_{j=1}^{n} \mathbf{1}_{A_j} \sim \text{Bin}(n,p), \text{d.h. } \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ k = 0, \ldots, n.$$

BEWEIS: Sei $N := \{1, \ldots, n\}$, $\{N\}_k := \{T \subseteq N : |T| = k\} \implies$

$$\{X = k\} = \sum_{T \in \{N\}_k} \left(\bigcap_{i \in T} A_i \cap \bigcap_{j \in N \setminus T} A_j^c \right)$$

$$\mathbb{P}\left(\bigcap_{i \in T} A_i \cap \bigcap_{j \in N \setminus T} A_j^c \right) = \prod_{i \in T} \mathbb{P}(A_i) \cdot \prod_{j \in N \setminus T} \mathbb{P}(A_j^c) = p^k \cdot (1-p)^{n-k}$$

$$|\{N\}_k| = \binom{n}{k} \qquad \checkmark$$

10.2 Definition (Bernoulli-Kette)

Die Situation von Satz 10.1 ist gegeben mit

- $\Omega \;:=\; \Omega_1 \times \ldots \times \Omega_n$,
- $\Omega_j \;:=\; \{0, 1\}$ für $j = 1, \ldots, n$,
- $\mathbb{P}_j(\{1\}) \;:=\; p \;=\; 1 - \mathbb{P}_j(\{0\})$ für $j = 1, \ldots, n$,
- $\mathbb{P}(\{(a_1, \ldots, a_n)\}) = p^{a_1+\ldots+a_n}(1-p)^{n-a_1-\ldots-a_n}, \quad \omega = (a_1, \ldots, a_n) \in \Omega$
- $A_j \;:=\; \{(a_1, \ldots, a_n) \in \Omega : a_j = 1\}$

Der W-Raum $(\Omega, \mathbb{P})$ (und das damit einhergehende Experiment) heißt
Bernoulli-Kette der Länge n mit Trefferwahrscheinlichkeit p.

Dabei steht 1 bzw. 0 für Treffer bzw. Niete.

Es gilt: Die Ereignisse $A_1, \ldots, A_n$ sind stochastisch unabhängig, wobei $\mathbb{P}(A_j) = p$, $1 \le j \le n$. (wurde in [BH] bewiesen).

Die Bernoulli-Kette wurde so in [BH], S. 221, eingeführt.

10.3 Beispiel (Bernoulli-Kette der Länge 4)

$$\Omega = \left\{ \begin{matrix} (0,0,0,0), & (0,0,0,1), & (0,0,1,0), & (0,0,1,1), \\ (0,1,0,0), & (0,1,0,1), & (0,1,1,0), & (0,1,1,1), \\ (1,0,0,0), & (1,0,0,1), & (1,0,1,0), & (1,0,1,1), \\ (1,1,0,0), & (1,1,0,1), & (1,1,1,0), & (1,1,1,1) \end{matrix} \right\}.$$

Sei $q := 1 - p$. Es gilt

$$\begin{aligned} \mathbb{P}(A_1) &= p \cdot \left(q^3 + 3\,pq^2 + 3\,p^2 q + p^3\right) \\ &= p \cdot (p+q)^3 = p = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \mathbb{P}(A_4), \\ \mathbb{P}(A_1 \cap A_2) &= p^2 \cdot \left(q^2 + 2pq + p^2\right) = p^2 \cdot (p+q)^2 = p^2 = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \\ &= \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_1 \cap A_4) \text{ usw.} \end{aligned}$$

$$\begin{aligned} \mathbb{P}(A_1 \cap A_2 \cap A_3) &= p^3 \cdot (q+p) = p^3 = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) \\ &= \mathbb{P}(A_1 \cap A_2 \cap A_4) = \mathbb{P}(A_1 \cap A_3 \cap A_4) \text{ usw.} \\ \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) &= p^4 = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) \cdot \mathbb{P}(A_4). \end{aligned}$$

Man erkennt allgemeinen Sachverhalt in Bernoulli-Kette der Länge n: (!)

$$\text{Falls } 1 \le i_1 < \ldots < i_k \le n, \text{ so } \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right) = p^k \cdot (p+q)^{n-k}$$

[LFS], S.145: Die n-malige paarweise stochastisch unabhängige Wiederholung eines Bernoulli-Experimentes heißt Bernoulli-Kette der Länge n.

Hiermit kann die Binomialverteilung $\mathrm{Bin}(n,p)$ nicht als Verteilung der Anzahl der Treffer hergeleitet werden. Ein Gegenbeispiel ist für $n = 3$:

$$\Omega := \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\}$$

mit der Gleichverteilung $\mathbb{P}$ auf Ω sowie

$$A_1 := \{(1,0,1),(1,1,0)\},\ A_2 := \{(0,1,1),(1,1,0)\},\ A_3 := \{(0,1,1),(1,0,1)\},$$

vgl. 9.2.a). Die Ereignisse A_1, A_2, A_3 sind paarweise stochastisch unabhängig, und A_j bezieht sich auf ein j-tes Bernoulli-Experiment.

Die Zufallsvariable

$$X := \sum_{j=1}^{3} \mathbf{1}_{A_j}$$

ist nicht $\mathrm{Bin}(3,1/2)$-binomialverteilt, da $\mathbb{P}(X = 1) = 0$ gilt.

10.4 Satz (Binomische Formel) Es gilt

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k}. \qquad \text{(Übungsaufgabe in [BH]!)}$$

Begrifflicher Beweis: Es gilt

$$(x+y)^n \quad = \quad \underbrace{(x+y)\cdot(x+y)\cdot(x+y)\cdot\ldots\cdot(x+y)}_{n \text{ Klammern}}$$

Nach den Regeln der Klammerrechnung muss man aus jeder Klammer entweder x oder y auswählen und das Produkt der ausgewählten Zahlen bilden.

Hat man k mal x (und damit $n-k$ mal y) gewählt, so entsteht als Produkt $x^k \cdot y^{n-k}$.

Es gibt $\binom{n}{k}$ Möglichkeiten, aus den n Klammern (Objekten) k hierfür zu wählen.

Am Ende muss man alle $\binom{n}{k} \cdot x^k \cdot y^{n-k}$ über k von 0 bis n summieren.

Mit der binomischen Formel erhält man leicht den Erwartungswert der Binomialverteilung über das „Erwartungswertbildung-Rezept“:
Gilt $X \sim \mathrm{Bin}(n,p)$, so folgt

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{k=1}^{n} \underbrace{k \cdot \binom{n}{k}}_{= \frac{k \cdot n \cdot (n-1)!}{k \cdot (k-1)! \cdot (n-k)!}} \cdot p^k \cdot (1-p)^{n-k} \\
&= n \cdot p \cdot \sum_{k=1}^{n} \underbrace{\frac{(n-1)!}{(k-1)! \cdot (n-1-(k-1))!}}_{= \binom{n-1}{k-1}} \cdot p^{k-1} \cdot (1-p)^{n-1-(k-1)} \\
&= n \cdot p \cdot \sum_{j=0}^{n-1} \binom{n-1}{j} \cdot p^j \cdot (1-p)^{n-1-j} \\
&= n \cdot p \cdot (p + 1 - p)^{n-1} \; = \; n \cdot p.
\end{aligned}
$$

Für welchen Wert k wird $\mathbb{P}(X = k)$ maximal?

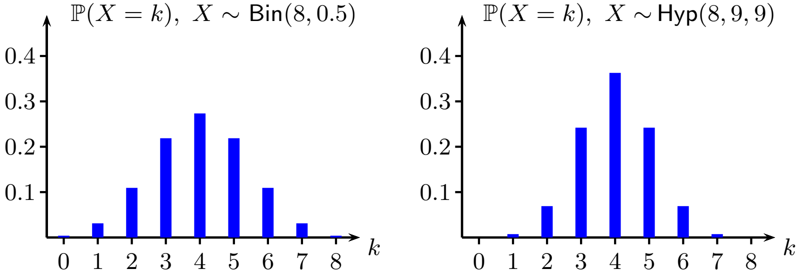
[LS10], S. 140: Ist $n \cdot p$ ganzzahlig, so hat diese Trefferanzahl die höchste Wahrscheinlichkeit. Andernfalls hat eine der beiden benachbarten Trefferanzahlen die maximale Wahrscheinlichkeit (ohne Beweis).

Für $k = 1, 2, \ldots, n$ gilt

$$\begin{aligned} \frac{\mathbb{P}(X = k)}{\mathbb{P}(X = k-1)} &= \frac{\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}{\binom{n}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-(k-1)}} \\ &= \frac{p}{1-p} \cdot \frac{n-(k-1)}{k} \\ \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} & \; 1 \iff k \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} np + p. \end{aligned}$$

Hieraus folgt die Behauptung.

11 Varianz



Zwei Verteilungen mit gleichem Erwartungswert, aber unterschiedlicher „Streuung“

Aufgabe aus [LS-G]: Aus den natürlichen Zahlen von 1 bis n wird eine zufällig ausgewählt. X sei die Zufallsvariable für die ausgewählte Zahl. Berechne $\mathbb{E}(X)$ und $\mathbb{V}(X)$.

Motivation der Varianz über den *durchschnittlichen Wert pro Vorgang der quadratischen Abweichung von X um den Erwartungswert* $\mu = \mathbb{E}(X)$, also

$$(X(\omega_1) - \mu)^2 \cdot \frac{H_1}{n} + \ldots + (X(\omega_s) - \mu)^2 \cdot \frac{H_s}{n}$$

11.1 Definition (Varianz und Standardabweichung)

Für eine Zufallsvariable X heißen (mit $\mu := \mathbb{E}(X)$)

$$\mathbb{V}(X) \; := \; \mathbb{E}(X-\mu)^2 \quad \left(:= \mathbb{E}\left[(X-\mu)^2\right]\right)$$

die Varianz (der Verteilung) von X und

$$+\sqrt{\mathbb{V}(X)}$$

die Standardabweichung oder Streuung (der Verteilung) von X.

$\mathbb{V}(X)$ ist der Erwartungswert der Zufallsvariablen $g(X)$ mit $g(x) = (x-\mu)^2$.

Memo: $\mathbb{V}(X) = \mathbb{E}(X-\mu)^2, \quad \mathbb{E}\,g(X) = \sum_{j\geq 1} g(x_j)\,\mathbb{P}(X=x_j)$

11.2 Satz (Darstellungsformeln für die Varianz)

a) Es gilt $\mathbb{V}(X) = \mathbb{E}X^2 - \mu^2$.

b) Falls X die Werte $x_1,\ldots,x_n$ annimmt, so gilt

$$\begin{aligned}\mathbb{V}(X) &= \sum_{j=1}^{n} (x_j-\mu)^2\,\mathbb{P}(X=x_j)\\ &= \sum_{j=1}^{n} x_j^2\,\mathbb{P}(X=x_j) - \mu^2.\end{aligned}$$

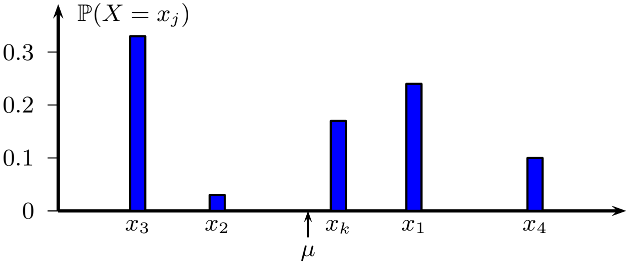
BEWEIS von a): $(X-\mu)^2 = X^2 - 2\cdot\mu\cdot X + \mu^2 \Longrightarrow$

$$\mathbb{V}(X) = \mathbb{E}(X^2) - 2\cdot\mu\cdot\mu + \mu^2 = \mathbb{E}\left(X^2\right) - \mu^2.$$

11.3 Beispiel (Gleichverteilung auf $\{1,2,\ldots,k\}$)

Sei $\mathbb{P}(X=j) = 1/k$ für $j=1,\ldots,k$. Dann gilt

$$\mathbb{V}(X) = \sum_{j=1}^{k} j^2\,\frac{1}{k} - \left(\frac{k+1}{2}\right)^2 = \ldots = \frac{k^2-1}{12}.$$



Stabdiagramm als Masseverteilung mit Schwerpunkt $\mu = \mathbb{E}(X)$.
Mit Winkelgeschwindigkeit v um μ rotieren lassen

$$v_j := |x_j - \mu|\, v \qquad \text{Rotationsgeschwindigkeit von } x_j,$$

$$E_j := \frac{1}{2}\,\mathbb{P}(X = x_j)\, v_j^2 \qquad \text{Rotationsenergie von } x_j$$

$$\sum_{j=1}^{k} E_j = \frac{v^2}{2} \underbrace{\sum_{j=1}^{k} (x_j - \mu)^2\, \mathbb{P}(X = x_j)}_{\text{Trägheitsmoment}} \qquad \text{gesamte Rotationsenergie}$$

Memo: $\mathbb{V}(X) \;= \mathbb{E}(X-\mu)^2 \;=\; \sum_{\omega\in\Omega} (X(\omega)-\mu)^2 \cdot \mathbb{P}(\{\omega\})$

Memo: $\mathbb{V}(X) \;= \mathbb{E}X^2 - \mu^2$

11.4 Satz (Eigenschaften der Varianz)

a) $\mathbb{V}(X) \geq 0, \;\; \surd$

$\mathbb{V}(X) = 0 \quad \Longleftrightarrow \quad \mathbb{P}(X = \mu = 1, \;\; \surd$

b) $\mathbb{V}(a\,X+b) \;=\; a^2\,\mathbb{V}(X), \qquad a,b\in\mathbb{R},$

d) $\mathbb{V}(\mathbf{1}_A) \;=\; \mathbb{P}(A)(1-\mathbb{P}(A)), \qquad A\subseteq\Omega. \;\; \surd$

$$\begin{aligned}\text{b)}\;\; \mathbb{V}(aX+b) \;&=\; \mathbb{E}\left(aX+b-\mathbb{E}(aX+b)\right)^2 = \mathbb{E}\left(aX+b-a\cdot\mu-b\right)^2\\ &=\; \mathbb{E}\left(a^2(X-\mu)^2\right) \;=\; a^2\mathbb{V}(X)\;\surd\end{aligned}$$

11.5 Satz (Varianz einer Indikatorsumme)

Es seien $A_1, \ldots, A_n \subseteq \Omega$ und $X := \mathbf{1}_{A_1} + \ldots + \mathbf{1}_{A_n}$. Dann gilt:

$$\mathbb{V}(X) = \sum_{i=1}^{n} \mathbb{P}(A_i)\,(1 - \mathbb{P}(A_i)) + 2 \sum_{1 \le i < j \le n} (\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j))$$

Speziell: $\mathbb{P}(A_j) = \mathbb{P}(A_1) \;\; \forall j,\;\; \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_1 \cap A_2) \;\; \forall i \neq j \implies$

$$\mathbb{V}(X) \;=\; n \cdot \left\{\mathbb{P}(A_1)(1 - \mathbb{P}(A_1)) + (n-1)\left(\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)^2\right)\right\}.$$

Beweis: Wir verwenden $\mathbb{V}(X) = \mathbb{E}(X^2) - \mu^2$.

$$\begin{aligned} X^2 &= \left(\sum_{i=1}^{n} \mathbf{1}_{A_i}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{1}_{A_i} \cdot \mathbf{1}_{A_j} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{1}_{A_i \cap A_j} \\ &= \sum_{i=1}^{n} \mathbf{1}_{A_i} + 2 \sum_{1 \le i < j \le n} \mathbf{1}_{A_i \cap A_j} \implies \\ \mathbb{E}\left(X^2\right) &= \sum_{i=1}^{n} \mathbb{P}(A_i) + 2 \sum_{1 \le i < j \le n} \mathbb{P}(A_i \cap A_j) \text{ usw.} \end{aligned}$$

Memo: Falls $X = \sum_{j=1}^{n} \mathbf{1}_{A_j}$, so:

$$\mathbb{V}(X) = \sum_{i=1}^{n} \mathbb{P}(A_i)\,(1 - \mathbb{P}(A_i)) + 2 \sum_{1 \le i < j \le n} (\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j))$$

11.6 Beispiel

a) Binomialverteilung

Falls $X \sim \text{Bin}(n,p)$, so gilt $\mathbb{V}(X) \;=\; np(1-p)$.

b) Hypergeometrische Verteilung

Falls $X \sim \text{Hyp}(n,r,s)$, so gilt wegen $\mathbb{P}(A_i \cap A_j) = \dfrac{r}{r+s} \cdot \dfrac{r-1}{r+s-1}$:

$$\mathbb{V}(X) \;=\; np(1-p)\left(1 - \frac{n-1}{r+s-1}\right).$$

Mit Darstellungsformel?

Memo: $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b, \quad \mathbb{V}(aX+b) = a^2\mathbb{V}(X)$

11.7 Satz und Definition (Standardisierung)

Sei X eine Zufallsvariable mit positiver Varianz. Dann heißt

$$\widetilde{X} := \frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}}$$

die zu X standardisierte Zufallsvariable oder Standardisierung von X.

Der Übergang von X zu $\dfrac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}}$ heißt Standardisierung.

Es gilt $\mathbb{E}(\widetilde{X}) = 0, \quad \mathbb{V}(\widetilde{X}) = 1.$

Beim Standardisieren wird durch die Standardabweichung geteilt! Beispiel:

$$X \sim \text{Bin}(n,p) \implies \widetilde{X} = \frac{X - np}{\sqrt{np(1-p)}}.$$

Memo: $U \leq V \Longrightarrow \mathbb{E}(U) \leq \mathbb{E}(V), \quad \mathbb{E}(\mathbf{1}_A) = \mathbb{P}(A)$

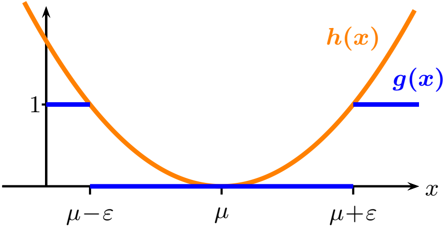
11.8 Satz (Tschebyschow-Ungleichung)

Für jedes (noch so große) $\varepsilon > 0$ gilt (mit $\mu := \mathbb{E}(X)$)

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

$$\underbrace{\mathbf{1}\{|X - \mu| \geq \varepsilon\}}_{= g(X)} \leq \underbrace{\left(\frac{X - \mu}{\varepsilon}\right)^2}_{= h(X)}$$

$$\mathbb{E}\, g(X) \leq \mathbb{E}\, h(X) \quad \checkmark$$



Beachte: Ist $\widetilde{X}$ standardisiert, so gilt $\mathbb{P}(|\widetilde{X}| \geq \varepsilon) \leq \dfrac{1}{\varepsilon^2}$.

12 Wartezeitverteilungen

Gegeben sei eine Bernoulli-Kette mit Trefferwahrscheinlichkeit p, $0 < p < 1$.

Sei X die Anzahl der Nieten (0) vor dem ersten Treffer (1).

Welche Verteilung besitzt X?

$$X = k \iff \underbrace{0\,0\,0\ldots0}_{k}\,1$$

Die Wahrscheinlichkeit hierfür ist $(1-p)^k\,p$ (1. Pfadregel)

Möglicher Grundraum: $\Omega = \{1, 01, 001, 0001, 00001, \ldots\}$

12.1 Definition und Satz (geometrische Verteilung)

Die Zufallsvariable X hat eine geometrische Verteilung mit Parameter p, $0 < p < 1$, kurz: $X \sim \mathsf{G}(p)$, falls gilt:

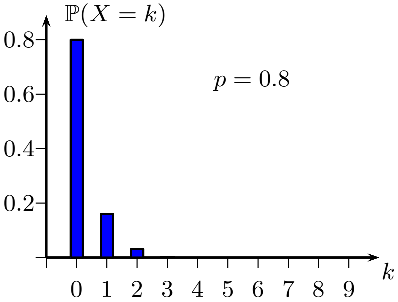
$$\mathbb{P}(X = k) \;=\; (1-p)^k\, p, \qquad k \in \mathbb{N}_0.$$

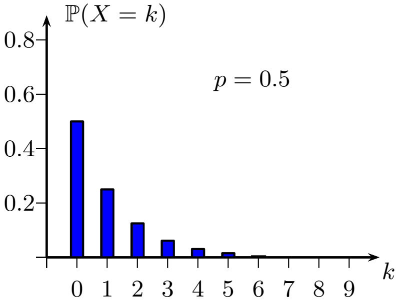
Falls $X \sim \mathsf{G}(p)$, so gelten (warum?):

$$\mathbb{E}(X) \;=\; \frac{1-p}{p} \;=\; \frac{1}{p} \;-\; 1\,, \qquad \mathbb{V}(X) \;=\; \frac{1-p}{p^2}.$$

Achtung! $X + 1$ modelliert die Zahl der Versuche bis zum ersten Treffer (einschließlich des letzten Versuchs, der den ersten Treffer ergibt).

Also: $p = 1/6 \implies \mathbb{E}(X+1) = 6.$





Stabdiagramme geometrischer Verteilungen

12.2 Beispiel

Ein Lottospieler gibt regelmäßig 10 verschiedene Tippreihen ab.

Pro Woche finden zwei Ausspielungen statt.

Wie groß ist der Erwartungswert (in Jahren) bis zum ersten „Sechser“?

Modellierung der Folge der Ausspielungen als Bernoulli-Kette

mit Trefferwahrscheinlichkeit $p = 10/\binom{49}{6}$.

Sei X die Anzahl der Ausspielungen bis zum ersten Sechser.

Es gilt $X - 1 \sim \mathsf{G}(p)$.

$$\mathbb{E}\,X \;=\; \frac{1}{p} \;=\; 1\;398\;381,6 \quad \text{(Ausspielungen)}$$

Ca. 104 Ausspielungen pro Jahr $\Longrightarrow$

$$\mathbb{E}\left(\frac{X}{104}\right) \;=\; \frac{1\;398\;381,6}{104} \;\approx\; 13\;446 \quad \text{(Jahre)}$$

12.3 Satz (Gedächtnislosigkeit der geometrischen Verteilung)

Sei $X \sim \mathrm{G}(p)$. Dann gilt:

$$\mathbb{P}(X = m + k | X \geq k) \;=\; \mathbb{P}(X = m) \quad \text{für jede Wahl von } k, m \in \mathbb{N}_0.$$

Beweis: Es gilt

$$\begin{aligned}
\mathbb{P}(X = m + k | X \geq k) &= \frac{\mathbb{P}(X = m + k)}{\mathbb{P}(X \geq k)} \\
&= \frac{(1-p)^{m+k} \cdot p}{p^k} \\
&= (1-p)^m \cdot p \\
&= \mathbb{P}(X = m).
\end{aligned}$$

Bernoulli-Kette: Verallgemeinerung der Fragestellung:

Sei $r \geq 1$ fest.

Welche Verteilung besitzt die Anzahl X der Nieten vor dem r-ten Treffer?

$$X = k \iff \underbrace{\star \star \star \star \star \star \star \star \star \star \star \star \star \star \star \cdots \star}_{\text{Hier } k \text{ Nullen und } r-1 \text{ Einsen}} 1$$

W' für eine konkrete Folge aus k Nullen, $r-1$ Einsen und einer Eins am Ende:

$$(1-p)^k \cdot p^{r-1} \cdot p.$$

Anzahl der Auswahlen von k Nullen aus $k+r-1$ Plätzen: $\binom{k+r-1}{k}$

12.4 Definition und Satz (negative Binomialverteilung)

Die Zufallsvariable X besitzt eine negative Binomialverteilung mit Parametern r und p $(r \in \mathbb{N},\ 0 < p < 1)$, kurz: $X \sim \mathrm{Nb}(r,p)$, falls gilt:

$$\mathbb{P}(X = k) \;=\; \binom{k+r-1}{k} p^r \, (1-p)^k, \qquad k \in \mathbb{N}_0.$$

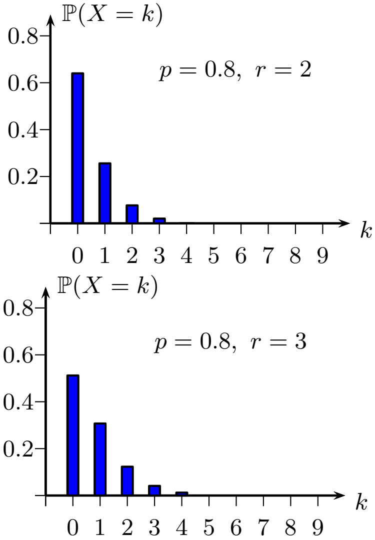
Falls $X \sim \mathrm{Nb}(r,p)$, so gelten:

$$\mathbb{E}(X) \;=\; r \cdot \frac{1-p}{p}\,, \qquad \mathbb{V}(X) \;=\; r \cdot \frac{1-p}{p^2}.$$

Beachte: (Warum negative Binomialverteilung?)

$$\binom{k+r-1}{k} = \frac{(k+r-1)!}{k!(r-1)!} = \frac{(k+r-1)(k+r-2)\ldots(r+1)r}{k!}$$

$$= \frac{(-1)^k}{k!} \cdot (-r)(-r-1)\ldots(-r-(k-2))(-r-(k-1)) = (-1)^k \binom{-r}{k}$$



Stabdiagramme von negativen Binomialverteilungen

Offenbar gilt $\mathsf{G}(p) = \mathsf{Nb}(1,p)$.

Besteht ein Zusammenhang zwischen $\mathsf{Nb}(r,p)$ und $\mathsf{G}(p)$ für $r \geq 2$?

$$\underbrace{0\ \ 0}_{X_1}\ 1\ \underbrace{0\ \ 0\ \ 0\ \ 0}_{X_2}\ 1\ \underbrace{0\ \ 0}_{X_3}\ 1\ \underbrace{0\ \ 0\ \ 0\ \ 0\ \ 0}_{X_4}\ 1\ \underbrace{0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0}_{X_5}\ 1$$

Vermutung: Ist $X \sim \mathsf{Nb}(r,p)$, so sollte $X \sim X_1 + \ldots + X_r$ gelten.

Dabei: $X_1, \ldots, X_r$ stochastisch unabhängig und je $\mathsf{G}(p)$-verteilt.

12.5 Satz (Struktur- und Additionsgesetz für Nb(r,p))

Sind $X_1, \ldots, X_r$ unabhängig und $X_j \sim \mathsf{G}(p)$ $\forall j$, so gilt $\sum_{j=1}^{r} X_j \sim \mathsf{Nb}(r,p)$.

Für einen formalen Beweis hilfreich ist hier die Binomialreihe

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad |x| < 1,\ \alpha \in \mathbb{R}.$$

13 Die Poisson-Verteilung

- eine grundlegende diskrete Verteilung,
- entsteht u.a. als Approximation der Binomialvertelung $\mathrm{Bin}(n, p)$ bei großem n und kleinem p
- tritt approximativ bei Zählvorgängen mit seltenen Ereignissen auf (Unfälle, starke Erdbeben, Gewitter, Selbstmorde ...)

Betrachte Folge von Binomialverteilungen $(\mathrm{Bin}(n, p_n))_{n \geq 1}$ mit

$$n \cdot p_n \;=\; \lambda, \quad n \geq 1, \quad \text{wobei } \lambda \in (0, \infty).$$

Die erwartete Trefferzahl np_n bleibt also konstant.

Sei $k \in \mathbb{N}_0$ fest. Für $n \in \mathbb{N}$ mit $n \geq k$ gilt

$$\begin{aligned}
\binom{n}{k} \cdot p_n^k \cdot (1-p_n)^{n-k} &= \frac{(n \cdot p_n)^k}{k!} \cdot \frac{n^{\underline{k}}}{n^k} \cdot \left(1 - \frac{n \cdot p_n}{n}\right)^{-k} \cdot \left(1 - \frac{n \cdot p_n}{n}\right)^n \\
&= \frac{\lambda^k}{k!} \cdot \underbrace{\frac{n^{\underline{k}}}{n^k}}_{\substack{\downarrow \\ 1}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\substack{\downarrow \\ 1}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\substack{\downarrow \\ \mathrm{e}^{-\lambda}}} \\
&\to \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \quad \text{für} \quad n \to \infty.
\end{aligned}$$

13.1 Satz (Gesetz seltener Ereignisse)

Sei (p_n) eine Folge in $[0,1]$ mit $\lim_{n\to\infty} np_n = \lambda$, wobei $0 < \lambda < \infty$. Dann gilt

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

13.2 Definition (Poisson-Verteilung)

Die Zufallsvariable X besitzt eine Poisson-Verteilung *mit Parameter* λ, $\lambda > 0$, kurz: $X \sim \mathrm{Po}(\lambda)$, falls gilt:

$$\mathbb{P}(X = k) \; = \; \mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!}\,, \qquad k = 0, 1, 2, \ldots$$

13.3 Satz (Eigenschaften der Poisson-Verteilung)

a) Falls $X \sim \mathrm{Po}(\lambda)$, so gilt $\mathbb{E}(X) \; = \; \mathbb{V}(X) \; = \; \lambda$.

b) Sind X, Y unabhängig, $X \sim \mathrm{Po}(\lambda)$, $Y \sim \mathrm{Po}(\mu)$, so gilt

$$X + Y \; \sim \; \mathrm{Po}(\lambda + \mu) \qquad \text{(Additionsgesetz)}$$

Beweis: a)

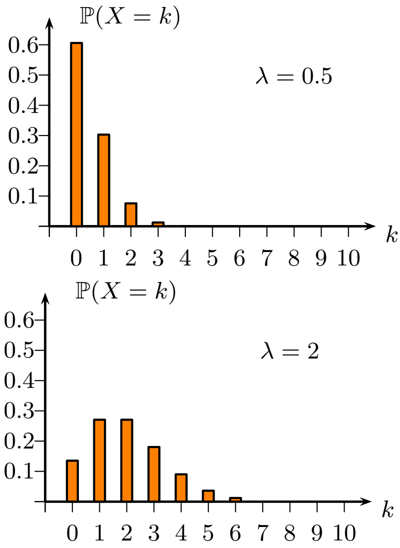
$$\mathbb{E}(X) \; = \; \sum_{k=0}^{\infty} k\,\mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!} \; = \; \mathrm{e}^{-\lambda}\,\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \; = \; \lambda\,\mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda} \; = \; \lambda$$

$$\mathbb{E}[X(X-1)] \; = \; \sum_{k=0}^{\infty} k(k-1)\mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!} \; = \; \lambda^2 \mathrm{e}^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \; = \; \lambda^2 \; \checkmark$$

b) z.z.: X, Y unabhängig, $X \sim \mathrm{Po}(\lambda), Y \sim \mathrm{Po}(\mu) \Longrightarrow X + Y \sim \mathrm{Po}(\lambda + \mu)$

Beweis: Sei $k \in \mathbb{N}_0$ beliebig. Es gilt

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{j=0}^{k} \mathbb{P}(X=j, Y=k-j) \\
&= \sum_{j=0}^{k} \mathbb{P}(X=j) \cdot \mathbb{P}(Y=k-j) \\
&= \sum_{j=0}^{k} \mathrm{e}^{-\lambda} \frac{\lambda^j}{j!} \mathrm{e}^{-\mu} \frac{\mu^{k-j}}{(k-j)!} \\
&= \mathrm{e}^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!} \sum_{j=0}^{k} \binom{k}{j} \left(\frac{\lambda}{\lambda+\mu}\right)^j \left(\frac{\mu}{\lambda+\mu}\right)^{k-j} \\
&= \mathrm{e}^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!} \cdot 1 \;\checkmark
\end{aligned}$$



Stabdiagramme von Poisson-Verteilungen

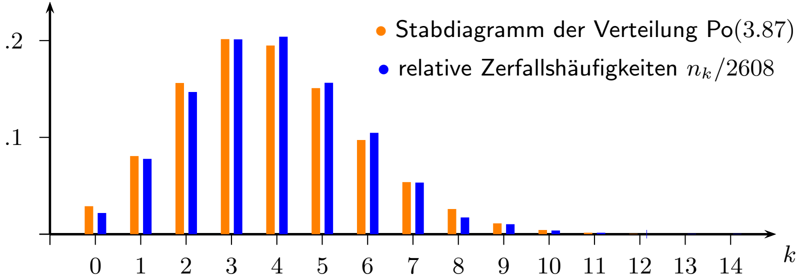
13.4 Beispiel (Das Rutherford-Geiger-Experiment)

Experiment: Radioaktives Präparat über 2608 Zeitintervalle von je 7.5 Sekunden Länge untersuchen. Insgesamt 10097 Zerfälle gezählt, also im Durchschnitt 3.87 Zerfälle innerhalb von 7.5 Sekunden.

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
n_k	57	203	383	525	532	408	273	139	45	27	10	4	0	1	1

Werte zum Rutherford-Geiger-Versuch

Dabei: n_k = Anzahl der Zeitintervalle, in denen k Zerfälle beobachtet wurden.



14 Grenzwertsätze

In diesem Kapitel behandeln wir

- das Schwache Gesetz großer Zahlen

sowie

- Zentrale Grenzwertsätze

im Zusammenhang mit diskreten Verteilungen.

Diese Resultate sind grundlegende Ergebnisse der klassischen Wahrscheinlichkeitstheorie mit vielfältigen Anwendungen, auch in der Statistik.

Beide Resultate beziehen sich auf Summen unabhängiger Zufallsvariablen.

In der Folge schreiben wir kurz $\overline{X}_n := \frac{1}{n}\sum_{j=1}^{n} X_j$.

14.1 Satz (Schwaches Gesetz der großen Zahlen)

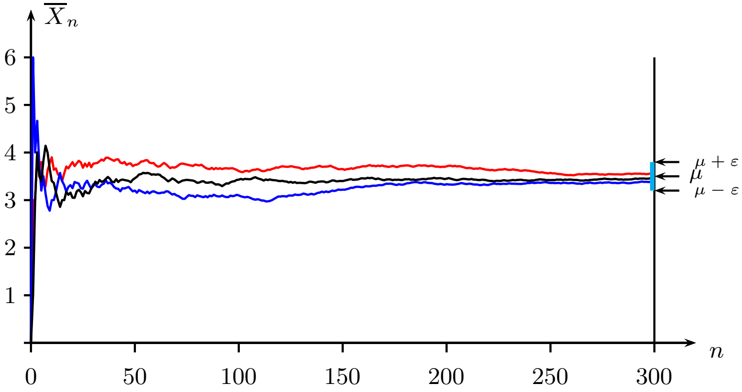
Seien $X_1, X_2, \ldots, X_n, \ldots$ unabhängige Zufallsvariablen mit gleichem Erwartungswert $\mu := \mathbb{E}X_1$ und gleicher Varianz $\sigma^2 := \mathbb{V}(X_1) < \infty$. Dann gilt:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) = 0 \quad \text{für jedes} \quad \varepsilon > 0.$$

Beweis: Sei $\varepsilon > 0$ beliebig. Es gilt $\mathbb{E}(\overline{X}_n) = \frac{1}{n}\sum_{j=1}^n \mathbb{E}(X_j) = \mu$.

$$\begin{aligned}
\mathbb{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) &\leq \frac{\mathbb{V}(\overline{X}_n)}{\varepsilon^2} && \text{(Tschebyschow-Ungleichung)}\\
&= \frac{1}{\varepsilon^2}\,\mathbb{V}\left(\frac{1}{n}\sum_{j=1}^n X_j\right)\\
&= \frac{1}{\varepsilon^2}\,\frac{1}{n^2}\,\mathbb{V}\left(\sum_{j=1}^n X_j\right) && (\mathbb{V}(aY) = a^2\mathbb{V}(Y))\\
&= \frac{1}{\varepsilon^2 n^2}\sum_{j=1}^n \mathbb{V}(X_j) && (X_1, \ldots, X_n \text{ unabh.})\\
&= \frac{1}{\varepsilon^2 n^2}\, n\,\sigma^2 \to 0 \text{ bei } n \to \infty.
\end{aligned}$$

Das schwache Gesetz großer Zahlen stellt einen Zusammenhang zwischen einem zufälligen arithmetischen Mittel und dem Erwartungswert her, vgl. das empirische Gesetz über die Stabilisierung relativer Häufigkeiten



Simulierte arithmetische Mittel der Augensumme beim Würfelwurf

14.2 Folgerung (Schwaches Gesetz großer Zahlen von Jakob Bernoulli)

Es seien $A_1, \ldots, A_n$ unabhängige Ereignisse mit gleicher Wahrscheinlichkeit p. Dann gilt:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{A_j} - p\right| \geq \varepsilon\right) \;=\; 0 \qquad \text{für jedes } \varepsilon > 0\ .$$

(Hauptergebnis der Ars conjectandi (1713) von Jakob Bernoulli)

$T_n \;:=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{A_j}$ ist zufällige relative Trefferhäufigkeit.

Zu jedem $\varepsilon > 0$ und zu jedem $\eta > 0$ existiert ein $n_0 = n_0(\varepsilon, \eta) \in \mathbb{N}$, so dass für jedes feste $n \geq n_0$ gilt:

$$\mathbb{P}(|T_n - p| < \varepsilon) \;\geq\; 1 - \eta.$$

In einer Bernoulli-Kette mit unbekannter Trefferwahrscheinlichkeit p ist die zufällige relative Trefferhäufgkeit ein sinnvoller Schätzer für p.

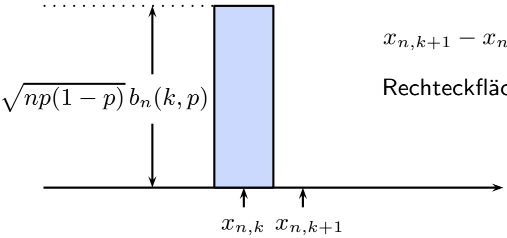
Sei $S_n \sim \mathsf{Bin}(n,p)$, $0 < p < 1$.

$$\widetilde{S}_n := \frac{S_n - np}{\sqrt{np(1-p)}}, \qquad x_{n,k} := \frac{k - np}{\sqrt{np(1-p)}}, \qquad k \in \mathbb{Z}.$$

$\widetilde{S}_n$ nimmt den Wert $x_{n,k}$ mit der Wahrscheinlichkeit

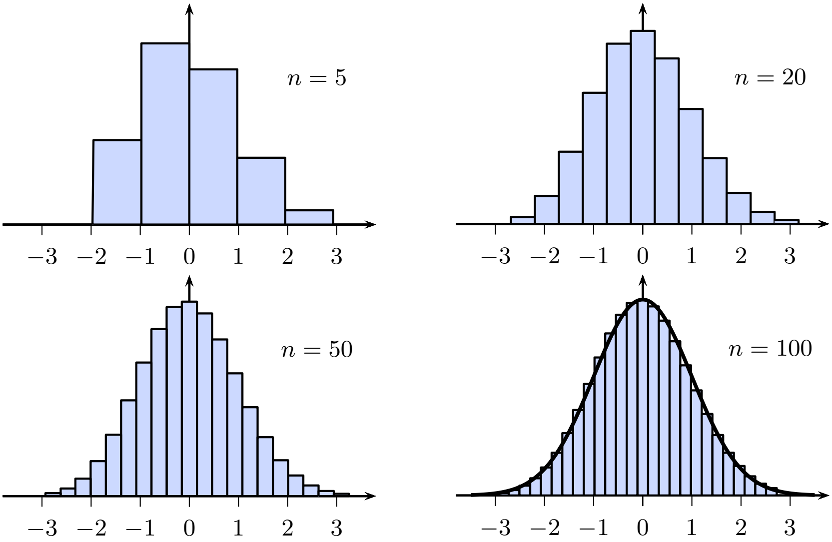
$$b_n(k,p) \ := \ \binom{n}{k} p^k \, (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

an.



$$x_{n,k+1} - x_{n,k} \ = \ \frac{1}{\sqrt{np(1-p)}}$$

Rechteckfläche $= b_n(k,p)$



Histogramme standardisierter Binomialverteilungen für $p = 0.3$

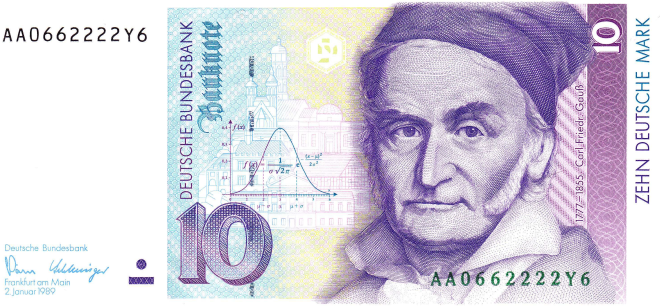
14.3 Definition (Dichte der Standard-Normalverteilung)

Die durch

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

definierte Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ heißt Dichte der Standard-Normalverteilung oder auch Gauß'sche Glockenkurve. Es gilt

$$\int_{-\infty}^{\infty} \varphi(x)\,\mathrm{d}x = 1.$$



14.4 Satz (Zentraler Grenzwertsatz von de Moivre–Laplace)

Sei $S_n \sim \mathrm{Bin}(n,p)$, $0 < p < 1$. Dann gelten für die standardisierte Zuf.variable

$$\widetilde{S}_n := \frac{S_n - np}{\sqrt{np(1-p)}} :$$

a) $\lim\limits_{n\to\infty} \mathbb{P}\left(a \le \widetilde{S}_n \le b\right) = \int_a^b \varphi(t)\,\mathrm{d}t, \qquad -\infty < a < b < \infty,$

b) $\lim\limits_{n\to\infty} \mathbb{P}\left(\widetilde{S}_n \le x\right) = \int_{-\infty}^x \varphi(t)\,\mathrm{d}t, \qquad x \in \mathbb{R}.$

Bemerkungen:

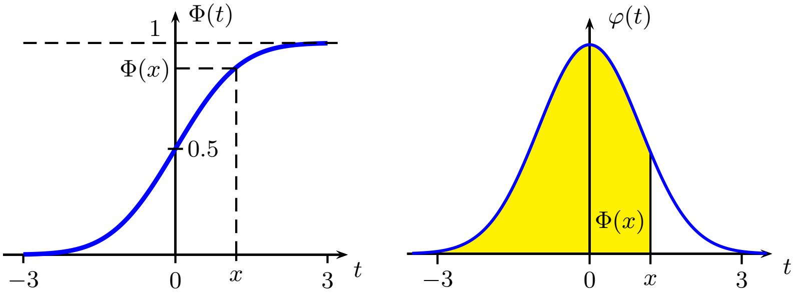
- in a) und b) kann jedes der $\le$-Zeichen durch das $<$-Zeichen ersetzt werden, ohne den jeweiligen Grenzwert zu ändern.
- Das in b) rechts stehende Integral erhält die folgende eigene Bezeichnung:

14.5 Definition (Verteilungsfunktion der Standard-Normalverteilung)

Die durch

$$\Phi(x) \ := \ \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t, \qquad x \in \mathbb{R},$$

definierte Funktion $\Phi : \mathbb{R} \to \mathbb{R}$ heißt Verteilungsfunktion der Standard-Normalverteilung.



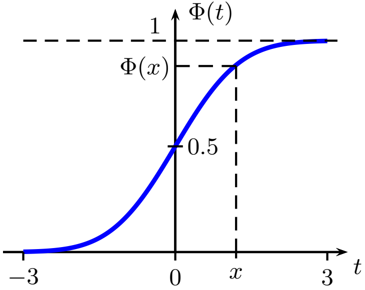
$\Phi(x)$ als Fläche unter der Gaußschen Glockenkurve

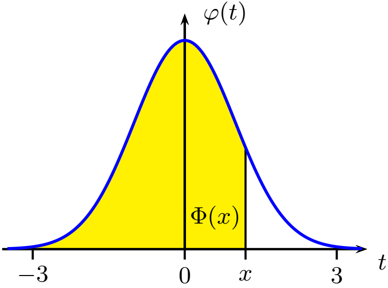
- Es gilt

$$\int_a^b \varphi(t)\,\mathrm{d}t \;=\; \Phi(b) - \Phi(a), \qquad -\infty < a < b < \infty.$$

- Wegen der Symmetrie von φ um 0 ist der Graph von Φ punktsymmetrisch zu $(0, 1/2)$, d.h. es gilt

$$\Phi(-x) \;=\; 1 - \Phi(x), \qquad x \in \mathbb{R}.$$





Allgemeiner gilt:

14.6 Satz (Zentraler Grenzwertsatz von Lindeberg-Lévy)

Es seien $Y_1, \ldots, Y_n, \ldots$ unabhängige, identisch verteilte (u.i.v.) Zufallsvariablen mit existierender, positiver Varianz. Dann gilt

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sum_{j=1}^n Y_j - n\,\mathbb{E}(Y_1)}{\sqrt{n\mathbb{V}(Y_1)}} \le x \right) = \Phi(x), \quad x \in \mathbb{R}.$$

14.7 Beispiel (Negative Binomialverteilung)

Seien $Y_1, \ldots, Y_n, \ldots$ stochastisch unabhängig und je G(p)-verteilt, $0 < p < 1$.

Additionsgesetz für die negative Binomialverteilung $\Longrightarrow$

$$S_n := Y_1 + \ldots + Y_n \ \sim \mathrm{Nb}(n,p). \text{ Mit } \mathbb{E}\,Y_1 = \frac{1-p}{p}, \ \ \mathbb{V}(Y_1) = \frac{1-p}{p^2} \text{ folgt}$$

$$\lim_{n\to\infty} \mathbb{P}\left(S_n - n\cdot\frac{1-p}{p} \le x \cdot \sqrt{n\,\frac{1-p}{p^2}} \right) = \Phi(x), \qquad x \in \mathbb{R}.$$

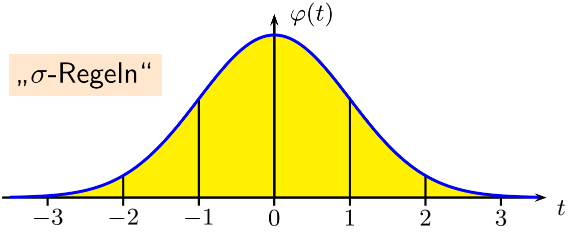
Memo: $Y_1, \ldots, Y_n, \ldots$ u.i.v., $\mu := \mathbb{E}\, Y_1$, $0 < \sigma^2 := \mathbb{V}(Y_1) < \infty \implies$

$$\lim_{n\to\infty} \mathbb{P}\left(a \le \frac{\sum_{j=1}^n Y_j - n\,\mu}{\sigma\sqrt{n}} \le b\right) = \int_a^b \varphi(x)\,\mathrm{d}x, \quad \sigma\sqrt{n} = \sqrt{\mathbb{V}\left(\sum_{j=1}^n Y_j\right)}$$

Speziell: $b = k \in \mathbb{N}$, $a = -b \implies$

$$\lim_{n\to\infty} \mathbb{P}\left(n\mu - k\,\sigma\sqrt{n} \le \sum_{j=1}^n Y_j \le n\mu + k\,\sigma\sqrt{n}\right) = \int_{-k}^k \varphi(x)\,\mathrm{d}x$$

„σ-Regeln“



$$\int_{-1}^1 \varphi(x)\,\mathrm{d}x \approx 0.6826, \quad \int_{-2}^2 \varphi(x)\,\mathrm{d}x \approx 0.9544, \quad \int_{-3}^3 \varphi(x)\,\mathrm{d}x \approx 0.9974$$

14.8 Beispiel (Bernoulli-Kette mit $p = 1/2$)

Seien $A_1, A_2, \ldots$ unabhängig, $\mathbb{P}(A_j) = 1/2$, $j \geq 1$.

$$Y_j := \mathbf{1}_{A_j}, \quad \mu = \mathbb{E}(Y_j) = \frac{1}{2}, \quad \sigma^2 = \mathbb{V}(Y_j) = \frac{1}{4}$$

$$S_n := \sum_{j=1}^{n} Y_j \sim \mathrm{Bin}\left(n, \frac{1}{2}\right).$$

Für großes n gilt

$$\mathbb{P}\left(\frac{n}{2} - 1 \cdot \frac{\sqrt{n}}{2} \leq S_n \leq \frac{n}{2} + 1 \cdot \frac{\sqrt{n}}{2}\right) \approx 0.6826,$$

$$\mathbb{P}\left(\frac{n}{2} - 2 \cdot \frac{\sqrt{n}}{2} \leq S_n \leq \frac{n}{2} + 2 \cdot \frac{\sqrt{n}}{2}\right) \approx 0.9544,$$

$$\mathbb{P}\left(\frac{n}{2} - 3 \cdot \frac{\sqrt{n}}{2} \leq S_n \leq \frac{n}{2} + 3 \cdot \frac{\sqrt{n}}{2}\right) \approx 0.9974.$$

Also für $n = 10000$:

$$\mathbb{P}(4950 \leq S_n \leq 5050) \approx 0.6826$$

$$\mathbb{P}(4900 \leq S_n \leq 5100) \approx 0.9544$$

$$\mathbb{P}(4850 \leq S_n \leq 5150) \approx 0.9974$$

15 Stetige Verteilungsmodelle

- Bislang: $(\Omega, \mathbb{P})$ diskreter W-Raum, d.h.:
- Ω abzählbar,
- $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ mit $\mathbb{P}(\Omega) = 1$ und

$$\mathbb{P}\left(\sum_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j) \quad \text{für paarweise disjunkte } A_1, A_2, \ldots$$

- $\mathbb{P}$ ist durch Punktmassen $\mathbb{P}(\{\omega_j\})$, $j \geq 1$, festgelegt
- Jede Zufallsvariable $X : \Omega \to \mathbb{R}$ nimmt höchstens abzählbar viele Werte mit positiver Wahrscheinlichkeit an.

Bei überabzählbarem Ω (z.B. $\Omega = \mathbb{R}$) können viele interessante Wahrscheinlichkeitsmaße nur noch auf gewissen Teilmengen von Ω definiert werden.

15.1 Definition (σ-Algebra)

Es sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt σ-Algebra (über Ω), falls gilt:

a) $\emptyset \in \mathcal{A}$,

b) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$,

c) $A_n \in \mathcal{A}\ (n \in \mathbb{N}) \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

15.2 Definition (Borelsche σ-Algebra)

Es bezeichne $\mathcal{O}^k$ das System der offenen Mengen des $\mathbb{R}^k$.

Die kleinste σ-Algebra über $\mathbb{R}^k$, die alle offenen Mengen enthält, heißt σ-Algebra der Borelmengen des $\mathbb{R}^k$.

15.3 Definition (Axiomensystem von Kolmogorow, 1933)

Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$.

Dabei sind $\Omega \neq \emptyset$ und $\mathcal{A}$ eine σ-Algebra über Ω.

Weiter ist $\mathbb{P} : \mathcal{A} \longrightarrow \mathbb{R}$ eine Funktion mit folgenden Eigenschaften:

a) $\mathbb{P}(A) \geq 0, \qquad A \in \mathcal{A}$, (Nichtnegativität)

b) $\mathbb{P}(\Omega) = 1$, (Normierung)

c) $\mathbb{P}(\sum_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$ (σ-Additivität)
für jede Folge $(A_j)_{j \geq 1}$ paarweise disjunkter Mengen aus $\mathcal{A}$.

$\mathbb{P}$ heißt Wahrscheinlichkeitsmaß auf $\mathcal{A}$ (kurz: W-Maß).

Jede Menge A aus $\mathcal{A}$ heißt Ereignis.

- Speziell: Diskreter W-Raum, ($\mathcal{A} = \mathcal{P}(\Omega)$, Ω abzählbar).

Bei Zufallsvariablen $X : \Omega \to \mathbb{R}$ muss jetzt die Bedingung

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

für jede Borelmenge B gefordert werden (sog. Messbarkeit).

Die Verteilung von X ist die Funktion

$$\mathbb{P}^X : \begin{cases} \mathcal{B}^1 \to [0,1], \\ B \mapsto \mathbb{P}^X(B) := \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) \end{cases}$$

Die Verteilung $\mathbb{P}^X$ ist eindeutig bestimmt durch die durch

$$F : \begin{cases} \mathbb{R} \to [0,1], \\ t \mapsto F(t) \ := \ \mathbb{P}(X \le t) \ = \ \mathbb{P}^X((-\infty, t]) \end{cases}$$

definierte Verteilungsfunktion von X ($\to$ Maßtheorie).

Memo: Verteilungsfunktion von $X : F(t) = \mathbb{P}(X \leq t), \; t \in \mathbb{R}$.

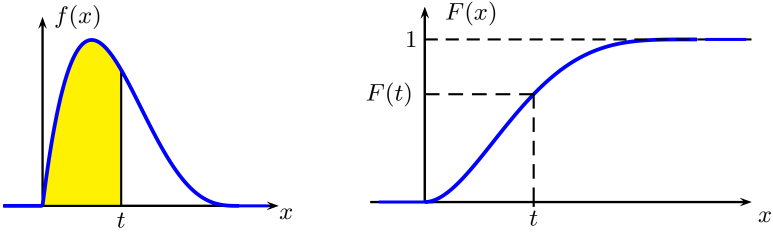
15.4 Definition (Stetige Zufallsvariable, Dichte)

Eine Zufallsvariable X heißt (absolut) stetig (verteilt)

$$:\Longleftrightarrow \quad \exists f : \mathbb{R} \to \mathbb{R}_{\geq 0} \text{ mit } \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1 \text{ und}$$

$$F(t) \;=\; \mathbb{P}(X \leq t) \;=\; \int_{-\infty}^{t} f(x)\,\mathrm{d}x, \qquad t \in \mathbb{R}.$$

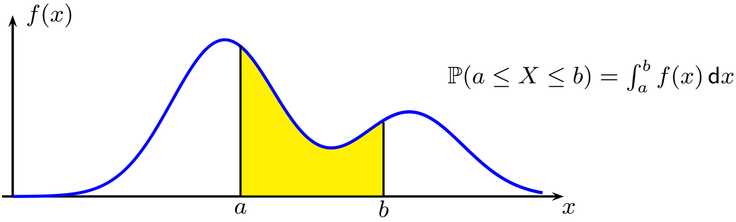
f heißt Dichte von X bzw. Dichte der Verteilung(sfunktion) von X.



Dichte und Verteilungsfunktion einer stetigen Zufallsvariablen

Sei X eine stetige Zufallsvariable mit Dichte f.

Dann gilt $\mathbb{P}(X = t) = 0$ für jedes $t \in \mathbb{R}$.



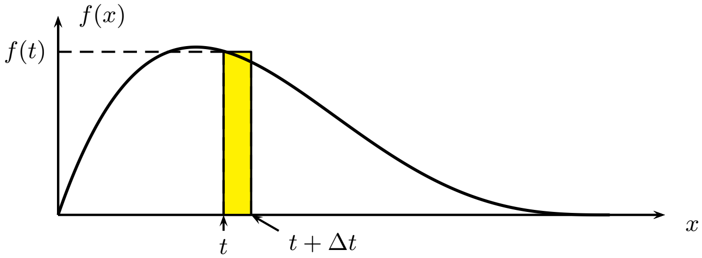
Eine Dichte kann an endlich vielen Stellen geändert werden, ohne die Verteilung von X zu ändern.

Sei X eine stetige Zufallsvariable mit Dichte f.

Sei t eine Stetigkeitsstelle von f. Für kleines $\Delta t > 0$ gilt

$$\mathbb{P}(t \le X \le t + \Delta t) \;=\; \int_t^{t+\Delta t} f(x)\,\mathsf{d}x \;\approx\; \Delta t\, f(t) \quad\Longrightarrow$$

$$f(t) \;\approx\; \frac{1}{\Delta t}\,\mathbb{P}(t \le X \le t + \Delta t)$$



Vgl. Massendichte (Grenzwert von Masse pro Volumeneinheit)

15.5 Definition (Gleichverteilung auf einem Intervall)

Seien $a, b \in \mathbb{R}$ mit $a < b$. Die Zufallsvariable X hat eine Gleichverteilung auf dem Intervall (a, b), falls X die Dichte

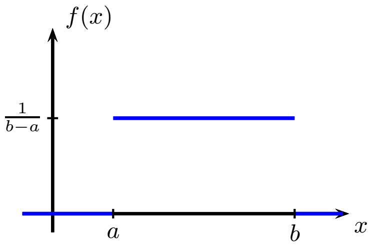
$$f(x) := \begin{cases} \dfrac{1}{b-a}, & \text{falls} \quad a < x < b, \\ 0, & \text{sonst} \end{cases}$$

besitzt. Wir schreiben hierfür kurz $X \sim \mathrm{U}(a, b)$.

Die Verteilungsfunktion von X hat die Darstellung

$$F(x) = \begin{cases} 0, & \text{falls} \quad x \leq a, \\ \dfrac{x-a}{b-a}, & \text{falls} \quad a < x < b, \\ 1, & \text{falls} \quad x \geq b. \end{cases}$$

Beachte: Pseudozufallszahlengeneratoren simulieren die Verteilung $\mathrm{U}(0, 1)$.



Dichte und Verteilungsfunktion der Gleichverteilung $\mathsf{U}(a,b)$

15.6 Satz Aus $X \sim \mathsf{U}(0,1)$ folgt $a + (b-a)X \sim \mathsf{U}(a,b)$.

15.7 Definition (Exponentialverteilung)

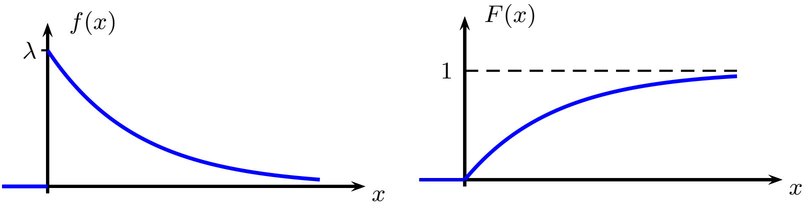
Die Zufallsvariable X hat eine Exponentialverteilung mit Parameter $\lambda > 0$, falls X die Dichte

$$f(x) \;:=\; \begin{cases} \lambda\,\mathrm{e}^{-\lambda x}, & \text{falls } x > 0, \\ 0, & \text{sonst}, \end{cases}$$

besitzt. Wir schreiben hierfür kurz $X \sim \mathrm{Exp}(\lambda)$.

Die Verteilungsfunktion von X ist

$$F(x) \;=\; \begin{cases} 1 - \mathrm{e}^{-\lambda x}, & \text{falls } x > 0, \\ 0, & \text{sonst}. \end{cases}$$



Dichte und Verteilungsfunktion der Exponentialverteilung

- Die Verteilung $\mathsf{Exp}(\lambda)$ ist das „gedächtnislose Analogon" der geometrischen Verteilung bei kontinuierlicher Zeitmessung.

  Für $t, h > 0$ gilt

$$\begin{aligned}
\mathbb{P}(X \geq t+h | X \geq t) &= \frac{\mathbb{P}(X \geq t+h,\ X \geq t)}{\mathbb{P}(X \geq t)} = \frac{\mathbb{P}(X \geq t+h)}{\mathbb{P}(X \geq t)} \\
&= \frac{1-F(t+h)}{1-F(t)} = \frac{\exp(-\lambda(t+h))}{\exp(-\lambda t)} = \mathbf{e}^{-\lambda h} \\
&= \mathbb{P}(X \geq h).
\end{aligned}$$

- Der Parameter λ bewirkt nur eine Skalenänderung: Es gilt

$$X \sim \mathsf{Exp}(1) \Longrightarrow \frac{1}{\lambda} X \sim \mathsf{Exp}(\lambda) \qquad (!)$$

- Es besteht ein direkter Zusammenhang mit der Gleichverteilung:

$$X \sim \mathsf{U}(0,1) \Longrightarrow -\frac{1}{\lambda} \log(1-X) \sim \mathsf{Exp}(\lambda) \qquad (!)$$

15.8 Definition (Normalverteilung)

Die Zufallsvariable X hat eine Normalverteilung mit Parametern μ und σ^2 ($\mu \in \mathbb{R}$, $\sigma > 0$), falls X die Dichte

$$f(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R},$$

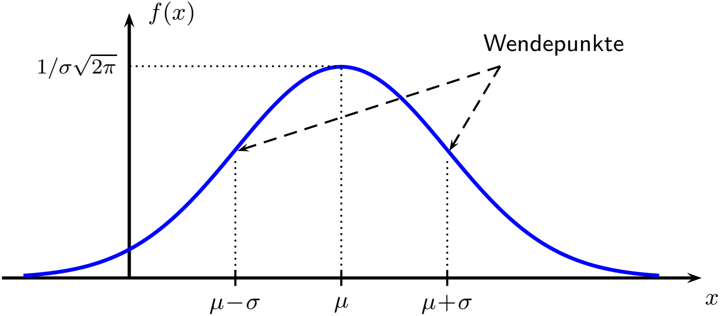
besitzt. Wir schreiben hierfür kurz $X \sim \mathrm{N}(\mu, \sigma^2)$.

Die Verteilung $\mathrm{N}(0,1)$ heißt Standard-Normalverteilung oder standardisierte Normalverteilung (vgl. Zentraler Grenzwertsatz).

Beachte: Mit $\varphi(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$ (Gaußsche Glockenkurve) gilt

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

Memo: $f(x) \;=\; \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right), \qquad x\in\mathbb{R}.$



Dichte der Normalverteilung $\mathrm{N}(\mu,\sigma^2)$

Verteilungsfunktion F der Normalverteilung $\mathrm{N}(\mu, \sigma^2)$?

Memo: $\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t, \quad \varphi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$

Memo: Dichte von $\mathrm{N}(\mu, \sigma^2)$: $f(t) = \frac{1}{\sigma}\,\varphi\left(\frac{t-\mu}{\sigma}\right)$

Also:

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\sigma}\,\varphi\left(\frac{t-\mu}{\sigma}\right)\mathrm{d}t \\
&= \int_{-\infty}^{(x-\mu)/\sigma} \varphi(z)\,\mathrm{d}z \quad \left(\text{mit } z := \frac{x-\mu}{\sigma} \right) \\
&= \Phi\left(\frac{x-\mu}{\sigma}\right).
\end{aligned}
$$

Fazit: Um Wahrscheinlichkeiten für eine beliebige Normalverteilung auszurechnen, reicht die Verteilungsfunktion Φ der Standard-Normalverteilung aus.

15.9 Satz

a) Aus $X \sim \mathsf{N}(0,1)$ folgt $\mu + \sigma X \sim \mathsf{N}(\mu, \sigma^2)$.

b) Aus $X \sim \mathsf{N}(\mu, \sigma^2)$ folgt $\dfrac{X-\mu}{\sigma} \sim \mathsf{N}(0,1)$.

- Die Normalverteilung ist eine der wichtigsten stetigen Verteilungen.
- Ihre Bedeutung beruht vor allem auf dem Zentralen Grenzwertsatz. Zur Erinnerung: $X_1, X_2, , \ldots$ unabhängig, identisch verteilt (u.i.v.), $\mu := \mathbb{E}(X_1)$, $0 < \sigma^2 := \mathbb{V}(X_1) < \infty$, $S_n := \sum_{j=1}^n X_j \Longrightarrow$

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\,\mu}{\sigma\sqrt{n}} \le x\right) \;=\; \Phi(x) \;=\; \int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t.$$

(Zentraler Grenzwertsatz von Lindeberg-Lévy).

15.10 Definition (Lognormalverteilung)

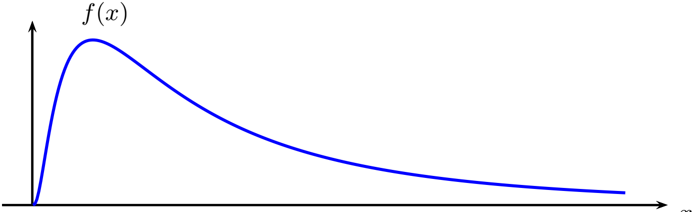
Die positive Zufallsvariable X besitzt eine Lognormalverteilung mit Parametern μ und σ^2 ($\mu \in \mathbb{R}, \sigma^2 > 0$), falls X die Dichte

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

für $x > 0$ und $f(x) = 0$, sonst, besitzt.

Wir schreiben hierfür kurz $X \sim \mathrm{LN}(\mu, \sigma^2)$.

Beachte: $X \sim \mathrm{LN}(\mu, \sigma^2) \Longrightarrow \log X \sim \mathrm{N}(\mu, \sigma^2)$.



Dichte der Lognormalverteilung

Auftreten der logarithmischen Normalverteilung:

- Verteilung von Aktienkursen im Black-Scholes-Modell der Finanzmathematik,
- Modell für Einkommensverteilungen,
- Modell für Schadenshöhen bei Versicherungen,
- Modell für die Größe von Lebewesen.

Sei $(\Omega, \mathbb{P})$ ein diskreter W-Raum, $X : \Omega \to \mathbb{R}$ Zufallsvariable. Dort:

$$\mathbb{E}(X) \ := \ \sum_{\omega \in \Omega} X(\omega)\, \mathbb{P}(\{\omega\}).$$

Erweiterung für einen allgemeinen W-Raum $(\Omega, \mathcal{A}, \mathbb{P})$:

$$\mathbb{E}(X) \ := \ \int_\Omega X(\omega)\, \mathbb{P}(\mathrm{d}\omega).$$

Dieses Integral definiert man für $X = \mathbf{1}_A$, $A \in \mathcal{A}$, zu $\mathbb{P}(A)$ und erweitert dann (durch „Linearisieren“, Approximation und Zerlegung in Positiv- und Negativteil).

Beachte:

- Man fordert absolute Konvergenz bzw. Integrierbarkeit!
- Rechenregeln für die Erwartungswertbildung gelten allgemein!

Ein allgemeiner Transformationssatz für Maß-Integrale liefert:

$$\mathbb{E}(X) = \int_\Omega X(\omega)\,\mathbb{P}(\mathrm{d}\omega) \;=\; \int_{-\infty}^{\infty} x\,\mathbb{P}^X(\mathrm{d}x).$$

Das rechts stehende Integral geht im Fall einer diskreten Zufallsvariablen mit Werten $x_1, x_2, \ldots$ in

$$\mathbb{E}(X) \;=\; \sum_{j\geq 1} x_j \cdot \mathbb{P}(X = x_j)$$

über. Ist X (absolut) stetig verteilt mit Dichte f, so:

$$\mathbb{E}(X) \;=\; \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x. \qquad \text{(Lebesgue-Integral)}$$

Allgemeine Darstellungsformel für $\mathbb{E}(g(X))$, wobei $g : \mathbb{R} \to \mathbb{R}$ (g messbar):

$$\mathbb{E}(g(X)) \;=\; \int_{-\infty}^{\infty} g(x) \cdot f(x)\,\mathrm{d}x.$$

Beachte: Absolute Konvergenz bzw. $\int_{-\infty}^{\infty} |g(x)| \cdot f(x)\,\mathrm{d}x < \infty$ vorausgesetzt!

15.11 Beispiel (Normalverteilung)

Es sei $X \sim \mathrm{N}(0,1) \implies f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$.

Aus Symmetriegründen gilt

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\,\varphi(x)\,\mathrm{d}x = 0.$$

Weiter gilt

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2\,\varphi(x)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot x\mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \left(-x\mathrm{e}^{-x^2/2}\Big]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x \right) \\ &= 1.\end{aligned}$$

Also gilt $\mathbb{V}(X) = 1$.

Sei jetzt $X \sim \mathrm{N}(\mu, \sigma^2)$. Dann gilt $X \sim \mu + \sigma X_0$, wobei $X_0 \sim \mathrm{N}(0,1)$.

Es folgt

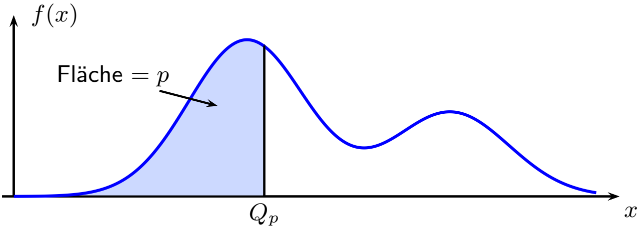
$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(\mu + \sigma X_0) = \mu + \sigma \mathbb{E}(X_0) = \mu, \\ \mathbb{V}(X) &= \mathbb{V}(\mu + \sigma X_0) = \sigma^2 \mathbb{V}(X_0) = \sigma^2.\end{aligned}$$

Sei X eine stetige Zufallsvariable mit Verteilungsfunktion F und Dichte f. Für $p \in (0,1)$ heißt der kleinste Wert x mit $F(x) = p$ das p-Quantil (der Verteilung) von X. Wir schreiben hierfür $Q_p = Q_p(X)$.

Speziell heißen

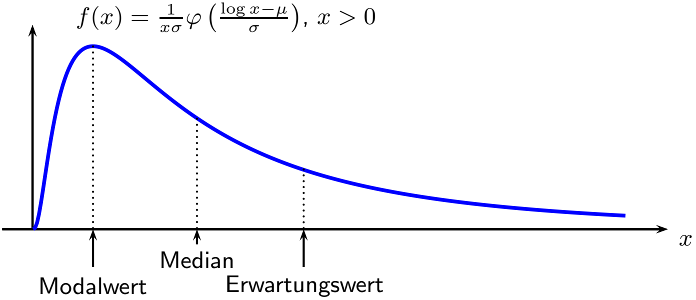
$Q_{1/2}$	der Median,
$Q_{1/4}$	das untere Quartil,
$Q_{3/4}$	das obere Quartil,
$Q_{3/4} - Q_{1/4}$	der Quartilsabstand

(der Verteilung) von X.



p-Quantil als „Flächen-Teiler“

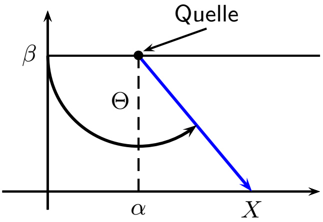
Memo: $Y \sim \mathrm{N}(\mu, \sigma^2) \implies X := \mathrm{e}^Y \sim \mathrm{LN}(\mu, \sigma^2)$

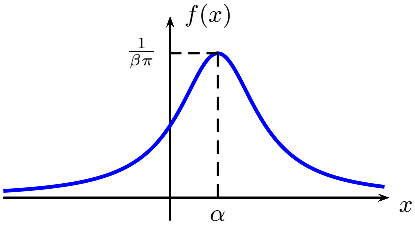


Die Lognormalverteilung $\mathrm{LN}(\mu, \sigma^2)$ hat eine rechtsschiefe Dichte.

$$
\begin{aligned}
\mathrm{Mod}(X) &= \exp(\mu - \sigma^2) && \text{(Modalwert = argmax } f\text{)} \\
\mathrm{Q}_{1/2}(X) &= \exp(\mu) && \text{(Median)} \\
\mathbb{E}(X) &= \exp(\mu + \sigma^2/2) && \text{(!)}
\end{aligned}
$$

15.12 Beispiel (Cauchy-Verteilung C(α, β))





- Im Punkt (α, β) ist eine Quelle angebracht.
- Quelle sendet unter zufälligem Winkel $\Theta \sim \mathrm{U}(0, \pi)$ Partikel in Richtung der x-Achse.
- Der zufällige Auftrittspunkt X hat die Dichte

$$f(x) \; = \frac{\beta}{\pi(\beta^2 + (x-\alpha)^2)}, \qquad x \in \mathbb{R}.$$

- Der Erwartungswert von X existiert nicht.
- $Q_{1/2}(X) = \alpha$, $Q_{3/4}(X) - Q_{1/4}(X) = 2\beta$ (!)