

Stochastik Wintersemester 2009

Leibniz Universität Hannover

Vorlesungsmitschrift

Dozent: [Prof. Dr. L. Baringhaus](#)

Mitschrift von
[Ronald Becher](#)

19. November 2009

Zusammenfassung

Diese Mitschrift wird erstellt im Zuge der Vorlesung [Stochastik A](#) im Wintersemester 2009 an der [Leibniz Universität Hannover](#). Obwohl mit großer Sorgfalt geschrieben, werden sich sicherlich Fehler einschleichen. Diese bitte ich zu melden, damit sie korrigiert werden können. Ihr könnt euch dazu an jeden der genannten Autoren (mit Ausnahme des Dozenten) wenden.

Dieses Skript wird primär über [Github](#) verteilt und „gepflegt“. Dort kann man es auch forken, verbessern und dann (idealerweise) ein „Pull Request“ losschicken. Siehe auch [Github Guides](#). Auch wenn ihr gute Grafiken zur Verdeutlichung beitragen könnt, dürft ihr diese gerne schicken oder selbst einfügen (am besten als \LaTeX -geeignete Datei (Tikz, SVG, PNG, ...)).

Hinweis: Es wird o.B.d.A. davon abgeraten, die Vorlesung zu versäumen, nur weil eine Mitschrift angefertigt wird! Selbiges gilt für den Übungsbetrieb! Stochastik besteht ihr nur, wenn ihr auch zu der Veranstaltung geht!

Inhaltsverzeichnis

I	Einführung / Organisatorisches	5
1	Literatur	5
II	Grundlegende Wahrscheinlichkeitstheorie	5
2	Wahrscheinlichkeitsräume	5
2.1	Eigenschaften von Wahrscheinlichkeits-Maßen	7
2.2	Beispiel	8
3	Grundformeln der Kombinatorik	9
3.1	Zur Erklärung	9
3.2	Anschauliche Beispiele	9
3.3	Rechenbeispiele	10
3.4	Multinomialscher Lehrsatz	12
3.5	Beispiel	12
3.6	Die Siebformel	13
3.7	Beispiel	13
3.8	Weitere Beispiele für W-Maße	14
3.9	Die Borelsche σ -Algebra	15
3.10	Verteilungsfunktionen	15
3.11	Verteilungsdichten	16
3.11.1	Spezialfälle	16
4	Weitere Eigenschaften von Wahrscheinlichkeitsmaßen	18
5	Stetigkeitseigenschaften von W-Maßen	19
5.1	Beweis zu (1)	19
5.2	Beweis zu (2)	20
5.3	Beispiel	20
III	Zufallsvariablen und ihre Verteilung	21
6	Spezialfälle	21
7	Beispiel: Würfelwurf von 2 echten Würfeln	22

8	Beispiele für Zufallsvariablen und ihre Verteilungen	24
8.1	Die Rechteck- oder Gleichverteilung	24
8.2	Die Normalverteilung	24
8.2.1	Beispiele für Normalverteilungen	25
8.3	Die Exponential-Verteilung	25
IV	Bedingte Wahrscheinlichkeiten	26
9	Bedingte Wahrscheinlichkeiten	26
9.1	Definition	26
9.2	Satz	27
9.2.1	Beweis	27
9.3	Beispiel: Lostrommel	27
10	Stochastische Unabhängigkeit von Ereignissen und Zufallsvariablen	28
10.1	Definition	28
10.2	Einsicht / Beweis	29
10.3	Definition	29
10.3.1	Bemerkung	29
10.4	Definition	29
10.5	Spezialfälle	30
10.5.1	Diskrete Verteilung	30
10.5.2	Stetige Verteilung	30
10.6	Beispiel	30
11	Binomialverteilung	32
11.1	Definition	32
11.2	Anwendung	32
11.3	Beispiel Urnenmodell	32
11.3.1	Mit Zurücklegen	32
11.3.2	Ohne Zurücklegen	32
12	Hypergeometrische Verteilung	33
12.1	Definition	33
V	Erwartungswerte	34
13	Erwartungswerte von Zufallsvariablen	34
14	Eigenschaften von Erwartungswerten	34
15	Rechnen mit Erwartungswerten	36

16 Spezialfälle	36
16.1 Stetige Verteilung	36
16.2 Diskrete Verteilung	36
17 Anwendungen	37
17.1 Erwartungswert der Rechteck-Verteilung	37
17.2 Erwartungswert der Normalverteilung	37
17.3 Beispiel	38
17.3.1 Binomialverteilt	38
17.3.2 Hypergeometrisch verteilt	38
18 Die Varianz	38
18.1 Definition	39
18.2 Bemerkung	39
18.3 Spezialfälle	39
18.3.1 Zufallsvariablen mit diskreter Verteilung	39
18.3.2 Zufallsvariablen mit stetiger Verteilung	39
18.4 Beispiele	39
18.4.1 Rechteckverteilt	39
18.4.2 Normalverteilt	40
18.4.3 Exponentialverteilt	40
19 Die Kovarianz	40
19.1 Definition	40
19.2 Bemerkung	41
19.3 Spezialfall	41
19.4 Beispiele	41
19.4.1 Binomialverteilt	41
19.4.2 Hypergeometrisch verteilt	42

Teil I

Einführung / Organisatorisches

Die Klausur wird am **01.02.2010 von 9:15 - 10:45** stattfinden, es wird wöchentliche Haus- und Stundenübungen geben.

Die abgegebenen Hausübungen werden korrigiert und in den Übungsgruppen zurückgegeben werden, es gibt allerdings keine Bonuspunktregelung. Übungsgruppen wird es folgende geben:

- Mo, 12:15 - 13:00 Uhr (F309)
- Mo, 14:15 - 15:00 Uhr (G123) [max. 20 Teilnehmer]
- Di, 10:15 - 11:00 Uhr (F428)
- Di, 12:15 - 13:00 (F442)
- Mi, 12:15 - 13:00 (B305)

1 Literatur

Der Dozent empfiehlt das Buch “Stochastik für Einsteiger” von Herrn N. Henze (Vieweg Verlag)

Teil II

Grundlegende Wahrscheinlichkeitstheorie

„Was ist das richtige Modell, um ein Experiment zu beschreiben?“ ist die wichtigste Frage, die die mathematische Statistik zu beantworten hat, und diese Modelle an die Wirklichkeit anzupassen, ist die hauptsächliche Aufgabe der Stochastik. Der letzte Teil wird vor allem in der Vorlesung **Stochastik B** behandelt.

2 Wahrscheinlichkeitsräume

Zufallsexperimente werden beschrieben durch

- die Menge der möglichen Ergebnisse Ω
Im **Würfelwurf** sei beispielsweise $\Omega := \{1, 2, 3, 4, 5, 6\}$

- die Menge $\mathfrak{A} \subset \mathfrak{P}(\Omega)^1 := \{A; A \subset \Omega\}$ als Menge der interessierenden Ergebnisse.
 \mathfrak{A} ist eine so genannte **σ -Algebra** auf Ω , d.h.

1.

$$\Omega \in \mathfrak{A} \tag{1}$$

(Die Grundmenge Ω ist in \mathfrak{A} enthalten)

2.

$$A \in \mathfrak{A} \iff A^c \in \mathfrak{A} \tag{2}$$

(Wenn \mathfrak{A} eine Teilmenge $A \in \Omega$ als mögliches Ergebnis enthält, dann auch deren Komplement $A^c = \Omega \setminus A$)

3.

$$A_1, A_2, \dots \in \mathfrak{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A} \tag{3}$$

(Wenn für jede natürliche Zahl n die Menge $A_n \in \mathfrak{A}$ ist, so ist auch die abzählbare Vereinigung aller $A_n \in \mathfrak{A}$)Im **Würfelwurf** seien beispielsweise gültige Ereignisse:

- $A := \{2, 4, 6\}$ das Ereignis „Es fällt eine gerade Zahl“, und
- $A := \{5, 6\}$ das Ereignis „Es fällt eine Zahl größer gleich 5“.
- durch ein **Wahrscheinlichkeitsmaß** P (lat: „Probabilitas“=Glaubwürdigkeit) mit $P : \mathfrak{A} \rightarrow [0, 1]$, also eine Mengenfunktion auf \mathfrak{A} mit den Eigenschaften

1.

$$P(\Omega) = 1 \tag{4}$$

(Die so genannte „Normiertheit“).

2. Für jede Folge von **paarweise disjunkten** $A_1, A_2, \dots \in \mathfrak{A}$ gilt:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \tag{5}$$

 σ -Additivität von P Für $A \in \mathfrak{A}$ ist $P(A)$ die Wahrscheinlichkeitsfunktion (für das Eintreten) von A Das Tripel $(\Omega, \mathfrak{A}, P)$ heißt **Wahrscheinlichkeitsraum**.

¹Potenzmenge

$\omega \in \Omega$ nennen wir das Ergebnis des Zufallsexperimentes. Das Ereignis $A \in \mathfrak{A}$ tritt genau dann ein, wenn $\omega \in A$.

$$\omega \in \bigcup_{n=1}^{\infty} A_n \iff \exists n \in \mathbb{N} \wedge \omega \in A_n \quad (6)$$

Das durch $\bigcup_{n=1}^{\infty} A_n$ beschriebene Ereignis tritt genau dann ein, wenn mindestens eines der Ereignisse A_n , $n \in \mathbb{N}$ eintritt.

Auf die Frage, ob auch \emptyset eine gültige Menge ist, sei gesagt, dass

$$\emptyset = \Omega^c \in \mathfrak{A}$$

das „unmögliche Ereignis“ genannt wird (Die Antwort ist also „ja“).

Seien $A_1, A_2, \dots \in \mathfrak{A}$. Dann

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathfrak{A} \quad (7)$$

Anmerkung: $\bigcap_{n=1}^{\infty} A_n$ tritt genau dann ein, wenn alle A_n , $n \in \mathbb{N}$ eintreten.

Wenn nun $A, B \in \mathfrak{A}$, dann gilt

•

$$A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \dots \in \mathfrak{A}$$

(Die Vereinigung von A, B ist auch ein interessierendes Ereignis und steht für „das Ereignis A oder B treten ein“)

•

$$A \cap B = A \cap B \cap \Omega \cap \Omega \cap \dots \in \mathfrak{A}$$

(Die Schnittmenge von A, B ist auch ein interessierendes Ereignis und steht für „das Ereignis A **und** B treten ein“)

•

$$A \Delta B = (A \cap B^c) \cup (B \cap A^c)$$

(Die „symmetrische Differenz“ beschreibt das Ereignis „Entweder A oder B “)

- $A \subseteq B$ heißt: „Das Ereignis A hat das Ereignis B zur Folge“.
(Kurz erklärt: Wenn A eintritt, folgt $\omega \in A \wedge A \subseteq B \Rightarrow \omega \in B$)

2.1 Eigenschaften von Wahrscheinlichkeits-Maßen

- Aus $P(\emptyset) = P(\emptyset \cup \emptyset \cup \dots) = \sum_{n=1}^{\infty} P(\emptyset)$ folgt $P(\emptyset) = 0$.

- Für $A, B \in \mathfrak{A}$ mit $A \subseteq B$ ist

$$P(B) = P(A \cup B \cap A^c) = P(A \cup B \cap A^c \cup \emptyset \cup \dots) = P(A) + P(B \cap A^c) \geq P(A) \quad (8)$$

Also gelten folgende Eigenschaften / Gesetzmäßigkeiten:

$$\text{---} \quad P(A) \leq P(B) \quad (9)$$

(Isotonie)

$$\text{---} \quad P(B \cap A^c) = P(B) - P(A) \quad (10)$$

(Subtraktivität)

2.2 Beispiel

$\emptyset \neq \Omega$ sei endlich, $|\Omega|$ Anzahl der Elemente von Ω und $\mathfrak{A} = \mathfrak{P}(\Omega)$.

Sei P definiert durch

$$P(A) = \frac{|A|}{|\Omega|}, \quad A \subset \Omega \quad (11)$$

P heißt „diskretes Laplace’sches W-Maß“, auf Ω .

Es ist

$$P(\{\omega\}) = \frac{1}{|\Omega|}, \quad \omega \in \Omega \quad (12)$$

also folgert man: „Alle möglichen Ergebnisse sind **gleich wahrscheinlich**.“

3 Grundformeln der Kombinatorik

$n, r \in \mathbb{N}$, M_n ist eine n -elementige Menge, o.B.d.A. $M_n = \{1, \dots, n\}$.

•

$$P_n^r = \{(x_1, \dots, x_r); x_i \in M_n, 1 \leq i \leq r\} \quad (13)$$

das sind die r -Permutationen aus M_n mit Wiederholung.

•

$$P_n^{(r)} = \{(x_1, \dots, x_r); x_i \in M_n, 1 \leq i \leq r, \text{ paarweise verschieden} \} \quad (r \leq n) \quad (14)$$

das sind die r -Permutationen aus M_n ohne Wiederholung.

•

$$K_n^{(r)} = \{(x_1, \dots, x_r); x_i \in M_n, 1 \leq i \leq r, x_1 < x_2 < \dots < x_n\} \quad (r \leq n) \quad (15)$$

das sind die r -Kombinationen aus M_n ohne Wiederholung.

•

$$K_n^r = \{(x_1, \dots, x_r); x_i \in M_n, 1 \leq i \leq r, x_1 \leq x_2 \leq \dots \leq x_n\} \quad (r \leq n) \quad (16)$$

das sind die r -Kombinationen aus M_n mit Wiederholung.

3.1 Zur Erklärung

1. Tupel in Permutationen sind nicht angeordnet, hier ist die Reihenfolge wichtig!
2. Tupel in Kombinationen sind o.B.d.A. angeordnet, wir können sie also “vergleichen”.
3. Bei Varianten **mit Wiederholung**, dürfen also zwei Elemente “nebeneinander” auch gleich sein, bei Varianten **ohne Wiederholung** ist das nicht der Fall, hier müssen die Elemente also echt unterschiedlich sein.
4. Bei einer r -Kombination haben wir eine r -elementige Teilmenge (die also kleiner oder gleich groß der “großen” Menge M_n ist).
5. “mit Wiederholung” nennt man auch “mit Zurücklegen”

3.2 Anschauliche Beispiele

- Bei Kombinationen interessiert uns ausschließlich, welche Elemente wir aus M_n bekommen, ein gutes Beispiel für eine Kombination **ohne Wiederholung** ist also das übliche Lotto-Spiel “6 aus 49”. Hier ist es uns egal, ob die Reihenfolge $(1, 2, 3, 4, 5, 6)$ ist oder $(6, 5, 4, 3, 2, 1)$.

mit Wiederholung ist die Frage, ob wir mindestens zwei Mal in fünf Versuchen eine Zahl ≥ 5 würfeln

- Bei Permutationen interessiert uns neben der Auswahl auch die Anordnung (man sagt: die Elemente sind “unterscheidbar”). Eine Permutation

ohne Wiederholung ist z.B. ein Kinobesuch. Ob Tim zwischen Tina und Claudia sitzt, ist für ihn nicht gleichwertig mit der Variante, zwischen Klaus und Jochen zu sitzen.

Anders ausgedrückt gilt in Permutationen

$$(Claudia, Tim, Tina, Klaus, Jochen) \neq (Claudia, Tina, Klaus, Tim, Jochen)$$

mit Wiederholung ist die Frage, ob ich beim Roulette einen aufsteigenden Run habe, also erst eine 0 fällt, dann eine 1, dann eine 2, usw. Das Standardbeispiel hierbei ist eine Urne, wo ich Kugeln oder Zettel heraushole und wieder hereinlege.

Eine Permutation mit Wiederholung ist also eine “Auswahl mit Wiederholungen unter Beachtung der Reihenfolge”.

3.3 Rechenbeispiele

Es gilt

•

$$|P_n^r| = n^r \quad (17)$$

(Erklärung: Man greift r Mal in die Urne, schaut sich einen der n Zettel an und legt ihn wieder herein, es gibt also danach genauso viele Möglichkeiten zu ziehen (also n))

•

$$|P_n^{(r)}| = \begin{cases} r \leq n & n \cdot (n-1) \cdot \dots \cdot (n-r+1) \\ r = n & r! \end{cases} \quad (18)$$

(Erklärung: Lege ich meine Zettel nicht wieder in die Urne zurück, habe ich beim nächsten Versuch natürlich eine Möglichkeit weniger, welche Zettel ich ziehen könnte)

•

$$r! \cdot |K_n^{(r)}| = |P_n^{(r)}| \quad (19)$$

\Downarrow

$$|K_n^{(r)}| = \frac{n(n-1) \dots (n-r+1)}{r!} = \binom{n}{r} \quad (20)$$

\Downarrow

$$|\{A; A \subset M_n, |A| = r\}| = \binom{n}{r} \quad (21)$$

\Downarrow

$$|\mathfrak{P}(M_n)| = \sum_{r=0}^r \binom{n}{r} 1^r 1^{n-r} = (1+1)^n = 2^n \quad (22)$$

(21) Gilt auch noch für $r = 0$.

Erklärungen:

(19) und äquivalent (20) kann man sich so vorstellen, dass man sich seine r -elementige Teilmenge (ohne Zurücklegen) zusammensuchen kann, aber die verschiedenen Möglichkeiten, sie zu permutieren ("die Kinobesucher umzusetzen"), interessieren den Betrachter hier nicht, also müssen wir die aus der Rechnung rausnehmen.

(21) ist die Anzahl aller r -elementigen Teilmengen von M_n hoch r , und wegen (20) ebenfalls $\binom{n}{r}$.

(22) ist die Länge der Potenzmenge, also aller möglichen Teilmengen von Elementen aus M_n

$$\bullet \quad K_n^r \rightarrow (x_1, \dots, x_r) \xrightarrow[\text{ist bijektiv}]{} (x_1, x_1 + 1, \dots, x_r + r - 1) \in K_{n+r-1}^{(r)}$$

$$\Rightarrow |K_n^r| = |K_{n+r-1}^{(r)}| = \binom{n+r-1}{r}$$

Identifiziere $(x_1, \dots, x_r) \in K_n^r$ mit dem Besetzungszahlvektor (k_1, \dots, k_n) , wobei $k_i = |\{j \in \{1, \dots, r\}; x_j = i\}|$, $1 \leq i \leq n$. Dann ist $k_1 + \dots + k_n = r$.

$$\text{Dann ist } |\{(k_1, \dots, k_n) \in \mathbb{N}_0^n; k_1 + \dots + k_n = r\}| = \binom{n+r-1}{r}$$

$$\text{Folgerung: } |\{(k_1, \dots, k_n) \in \mathbb{N}^n; k_1 + \dots + k_n = r\}| = \binom{n+(r-n)-1}{r-n} = \binom{r-1}{r-n} = \binom{r-1}{n-1}, r \geq n$$

Aus der Menge der Permutationen betrachten wir eine spezifische Teilmenge.

Sei $k_1, \dots, k_n \in \mathbb{N}_0$ mit $k_1 + \dots + k_n = r$. Einen so genannter "Besetzungszahlvektor" drückt man beispielsweise so aus:

$$\begin{aligned} & \{(x_1, \dots, x_n) \in \mathbb{P}_n^r; |\{j; j \in \{1, \dots, r\}, x_j = i\}| = k_i, 1 \leq i \leq n\} \\ = & \{(x_1, \dots, x_n) \in \mathbb{P}_n^r; \text{Genau } k_i \text{ der } r \text{ Komponenten } (x_1, \dots, x_r) \text{ sind gleich } i, 1 \leq i \leq n\} \end{aligned} \quad (23)$$

Die Länge oder Mächtigkeit dieser Menge ist wie folgt:

$$\begin{aligned} |23| &= \binom{r}{k_1} \binom{r-k_1}{k_2} \binom{r-k_1-k_2}{k_3} \dots \binom{r-k_1-\dots-k_{n-1}}{k_n} \\ &= \frac{r!}{k_1! \cancel{(r-k_1)!}} \cdot \frac{\cancel{(r-k_1)!}}{k_2! \cancel{(r-k_1-k_2)!}} \dots \frac{\cancel{(r-k_1-\dots-k_{n-1})!}}{k_n!} \\ &= \frac{r!}{k_1! \dots k_n!} \end{aligned}$$

3.4 Multinomialer Lehrsatz

Es folgt aus dem (bekannten) binomischen Lehrsatz durch Verallgemeinerung der multinomialen Lehrsatz.

$$n^r = \sum_{\substack{(k_1, \dots, k_n) \in \mathbb{N}_0^n \\ k_1 + \dots + k_n = r}} \frac{r!}{k_1! \dots k_n!} \quad (24)$$

$$\text{Anzahl der Summanden aus 24} = \binom{n+r-1}{r} \quad (25)$$

3.5 Beispiel

O.B.d.A. sind r Personen im Raum, $n = 365$ seien die Anzahl der Tage im Jahr und $r \leq n$.

Gesucht: Wahrscheinlichkeit, dass mindestens 2 der r Personen am gleichen Tag Geburtstag haben.

- $\Omega = \{(x_1, \dots, x_r); x_i \in \{1, \dots, n\}\}$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|}, A \subseteq \Omega$
- Alle möglichen r -Tupel von Geburtstagen sind gleich wahrscheinlich, also
- $A = \{(x_1, \dots, x_r) \in \Omega; i, j \in \{1, \dots, r\}, i \neq j, x_i = x_j\}$

- $P(A) = 1 - P(A^c)$
In vielen Fällen ist es ein valider “Trick”, statt der Ergebnismenge das Gegenereignis zu berechnen
- $A^c = \{(x_1, \dots, x_r) \in \Omega; x_1, \dots, x_r \text{ paarweise verschieden}\}$
- $|A^c| = n(n-1) \cdots (n-r+1)$
- $P(A) = 1 - \frac{n(n-1) \cdots (n-r+1)}{n^r}$

$$P(A) = \begin{cases} > 1/2 & \text{wenn } r > 23, \\ 0,706 & \text{wenn } r = 30, \\ 0,994 & \text{wenn } r = 60. \end{cases}$$

3.6 Die Siebformel

Seien A_1, A_n Ereignisse in W-Raum $(\Omega, \mathfrak{A}, P)$

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1 \cup (A_2 \cap (A_1 \cap A_2)^c)) \\ &= P(A_1) + P(A_2 \cap (A_1 \cap A_2)^c) \\ &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \end{aligned} \quad (26)$$

Dies führt uns zu folgendem: Seien A_1, \dots, A_n Ereignisse. Dann

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq j_1 < \dots < j_k \leq n} P(A_{j_1} \cap \dots \cap A_{j_k}) \quad (27)$$

Wie leicht überprüfbar ist, ist (26) eine Anwendung von (27).

3.7 Beispiel

Sei

- $\Omega = \{\pi; \pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \text{ bijektiv}\}$
- $|\Omega| = n!$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- P das diskrete Laplace'sche W-Maß auf Ω
- $A_i = \{\pi \in \Omega; \pi(i) = i\}, 1 \leq i \leq n$
- $\bigcup_{i=1}^n A_i = \{\pi \in \Omega; \text{Es gibt ein } i \in \{1, \dots, n\} \text{ mit } \pi(i) = i\}$

Dann

$$\begin{aligned}
 P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq j_1 < \dots < j_k \leq n} \frac{|A_{j_1} \cap \dots \cap A_{j_k}|}{|\Omega|} \\
 &= \sum_{k=1}^n (-1)^{k-1} \sum_{\dots} \frac{(n-k)!}{n!} \\
 &= \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \cdot \frac{(n-k)!}{n!} \\
 &= \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!}
 \end{aligned} \tag{28}$$

und wegen

$$\begin{aligned}
 \bigcap_{i=1}^n A_i^c &= \left(\bigcup_{i=1}^n A_i\right)^c \\
 &= \{\pi \in \Omega; \forall i \in \{1, \dots, n\} \text{ ist } \pi(i) \neq i\}
 \end{aligned} \tag{29}$$

ist

$$\begin{aligned}
 P\left(\bigcap_{i=1}^n A_i^c\right) &= 1 - \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!} \\
 &= \sum_{k=0}^n (-1)^k \frac{1}{k!} \\
 &\xrightarrow{n \rightarrow \infty} \sum_{k=0}^{\infty} (-1)^k \frac{1}{k!} \\
 &= e^{-1} \\
 &\approx 0.37
 \end{aligned} \tag{30}$$

3.8 Weitere Beispiele für W-Maße

Sei $\emptyset \neq \Omega$, \mathfrak{A} σ -Algebra mit $\{\omega\} \in \mathfrak{A}$ für jedes $\omega \in \Omega$.

Sei P W-Maß auf Ω mit der Eigenschaft, dass eine abzählbare Menge $\Omega_o \subseteq \Omega$ existiert mit $P(\Omega_o) = 1$

Dann

$$P(\Omega_o^c) = 1 - P(\Omega_o) = 0$$

Für $A \in \mathfrak{A}$ gilt

$$\begin{aligned}
 P(A) &= P((A \cap \Omega_0) + (A \cap \Omega_0^c)) \\
 &= P(A \cap \Omega_0) + P(A \cap \Omega_0^c) \\
 &= P\left(\sum_{\omega \in A \cap \Omega_0} \{\omega\}\right) \\
 &= \sum_{\omega \in A \cap \Omega_0} P(\{\omega\}) \\
 &= \sum_{\omega \in \Omega} P(\{\omega\}) \delta_\omega(A)
 \end{aligned} \tag{31}$$

Für $\omega \in \Omega$ heißt $\delta_\omega : \mathfrak{A} \rightarrow [0, 1]$, definiert durch

$$\delta_\omega(A) = \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A. \end{cases}$$

das Einpunktmaß oder **Dirac-Maß** in ω .

Es gilt also

$$P = \sum_{\omega \in \Omega_0} P(\{\omega\}) \cdot \delta_\omega.$$

Solche W-Maße heißen **diskrete W-Maße**.

3.9 Die Borelsche σ -Algebra

Betrachte den Fall $\Omega = \mathbb{R}$

$$\mathfrak{B} = \mathfrak{A} = \bigcap_{\substack{\mathfrak{A}^* \text{ } \sigma\text{-Algebra auf } \mathbb{R} \\ \mathfrak{A}^* \supseteq \{(-\infty, x]; x \in \mathbb{R}\}}} \mathfrak{A}^*$$

ist die Borelsche σ -Algebra auf \mathbb{R} .

Beachte

$$\mathfrak{B} \neq \mathfrak{P}(\mathbb{R})$$

\mathfrak{B} enthält alle offenen, abgeschlossenen und ohne Intervalle Teilmengen von \mathbb{R} .

3.10 Verteilungsfunktionen

Sei $F : \mathbb{R} \rightarrow \mathbb{R}$ Funktion mit folgenden Eigenschaften:

1. F ist monoton wachsend

2. F ist rechtsseitig stetig

3.

$$\lim_{x \rightarrow -\infty} F(x) = 0$$

4.

$$\lim_{x \rightarrow +\infty} F(x) = 1$$

Es existiert genau ein W-Maß P auf \mathfrak{B} (auf \mathbb{R}) mit der Eigenschaft, dass

$$P((-\infty, x]) = F(x), \quad \forall x \in \mathbb{R} \quad (32)$$

F heißt die zu P gehörige **Verteilungsfunktion**.

Für $(a, b]$ mit $-\infty < a < b < +\infty$ ist

$$P((a, b]) = P((-\infty, b]) - P((-\infty, a]) = F(b) - F(a) \quad (33)$$

3.11 Verteilungsdichten

Sei $f : \mathbb{R} \rightarrow \mathbb{R}_+$ (uneigentlich) **Riemann-integrierbar** mit

$$\int_{-\infty}^{+\infty} f(t) dt = 1 \quad (34)$$

Setze

$$F(x) = \int_{-\infty}^{+\infty} f(t) dt, \quad -\infty < x < +\infty \quad (35)$$

F ist Verteilungsfunktion eines W-Maßes P auf \mathbb{R} .

f heißt **Wahrscheinlichkeitsdichte** (oder einfach „Dichte“) von P .

3.11.1 Spezialfälle

•

$$f(t) = \begin{cases} 0 & , \quad x < a, \\ \frac{1}{b-a} & , \quad a \leq x \leq b, \\ 0 & , \quad x > b. \end{cases}$$

$$F(x) = \begin{cases} 0 & , \quad x < a, \\ \frac{x-a}{b-a} & , \quad a \leq x \leq b, \\ 1 & , \quad x > b. \end{cases}$$

•

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(t-a)^2}{\sigma^2}\right), \quad -\infty < t < +\infty, \quad (a \in \mathbb{R}, \sigma^2 > 0)$$

(Anmerkung: Diese Verteilungsdichte wird auch Standardnormalverteilung genannt. Die Verteilungsfunktion dafür produziert die so genannte Gauß'sche Glockenkurve)

Dann ist

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

und somit

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(t-a)^2}{\sigma^2}\right) dt \\ &= \phi\left(\frac{x-a}{\sigma}\right) \\ \phi(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2}\right) dt, x \in \mathbb{R} \end{aligned}$$

Sei $(\Omega, \mathfrak{B}, P)$ der W-Raum unserer Wahl. P habe Dichte f , also eine Teilfunktion

$$f : \mathbb{R} \rightarrow \mathbb{R}^+ \text{ mit } \int_{-\infty}^{+\infty} f(t) dt = 1$$

$$F(x) = P((-\infty, x]) = \int_{-\infty}^{+\infty} f(t) dt, \quad x \in \mathbb{R}$$

Allgemein gilt $P(E) = \int_E f(t) dt$, $E \in \mathfrak{B}(\text{geeignet})$.

Beispiel:

$$f(t) = \begin{cases} 0 & t < 0 \\ \lambda \cdot \exp(-\lambda t) & t \geq 0 \end{cases}$$

Aus

$$\int_0^x \lambda \cdot \exp(-\lambda t) dt = [-\exp(-\lambda t)]_0^x = 1 - \exp(-\lambda x)$$

folgt

$$F(x) = \int_{-\infty}^{+\infty} f(t) dt = \begin{cases} 0 & x < 0 \\ 1 - \exp(-\lambda x) & x \geq 0 \end{cases}$$

4 Weitere Eigenschaften von Wahrscheinlichkeitsmaßen

Sei (R, \mathfrak{A}, P) W-Raum. Seien $A_1, A_2, \dots \in \mathfrak{A}$. Dann gilt

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n) \quad (36)$$

Diese Eigenschaft wird als **Sub- σ -Additivität** bezeichnet.

$$\bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty} (A_n \cap A_1^c \cap \dots \cap A_{n-1}^c) \quad (37)$$

Wir können die Vereinigung auch als Summe von disjunkten Ereignissen auffassen und gelangen zu dieser Schreibweise in (36). Deshalb gilt auch (37)

$$\begin{aligned}
& P\left(\bigcup_{n=1}^{\infty} A_n\right) \\
&= \sum P(A_1) + \sum_{n=2}^{\infty} P(A_n \cap A_1^c \cap \dots \cap A_{n-1}^c) \\
&\leq \sum_{n=1}^{\infty} P(A_n)
\end{aligned} \tag{38}$$

Speziell folgt hieraus die so genannte “Sub-Additivität von P”:

$$P\left(\bigcup_{j=1}^n A_j\right) \leq \sum_{j=1}^n P(A_j) \tag{39}$$

Seien $A_n \in \mathfrak{A}, n = 1, 2, \dots, A \in \mathfrak{A}$. Wir schreiben

- $A_n \uparrow A$ für $A_1 \subset A_2 \subset \dots, A = P\left(\bigcup_{j=1}^{\infty} A_j\right)$
Man sagt “ A_n konvergiert **isoton** gegen A”)
- $A_n \downarrow A$ für $A_1 \supset A_2 \supset \dots, A = P\left(\bigcap_{j=1}^{\infty} A_j\right)$
Man sagt “ A_n konvergiert **antiton** gegen A”)

5 Stetigkeitseigenschaften von W-Maßen

$$1. A_n \uparrow A \implies \lim_{n \rightarrow \infty} P(A_n) = P(A)$$

$$2. A_n \downarrow A \implies \lim_{n \rightarrow \infty} P(A_n) = P(A)$$

5.1 Beweis zu (1)

$$A = \bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty} (A_n \cap A_{n-1}^c) \tag{40}$$

Also

$$\begin{aligned}
P(A) &= P(A) + \sum_{n=2}^{\infty} P(A_n \cap A_{n-1}^c) \\
&= P(A) + \sum_{n=2}^{\infty} P(A_n - A_{n-1}^c) \\
&= P(A) + \lim_{m \rightarrow \infty} \sum_{n=2}^m P(A_n - A_{n-1}) \\
&= P(A_1) + \lim_{m \rightarrow \infty} (P(A_m - A_1)) \\
&= \lim_{m \rightarrow \infty} P(A_m)
\end{aligned} \tag{41}$$

5.2 Beweis zu (2)

Mit Gegenmenge ...

5.3 Beispiel

$\Omega = R, x \in R, (x_n) \subset R$ Folge, die monoton fallend gegen x konvergiert.

$$(-\infty, x] = \bigcap_{n=1}^{\infty} (-\infty, x_n]$$

P ist W-Maß auf R

$$F(x) = P((-\infty, x]) = \lim_{n \rightarrow \infty} P((-\infty, x_n]) = \lim_{n \rightarrow \infty} F(x_n)$$

Hieraus folgt allgemein:

Ist $(x_n) \subset R$ eine Folge, die von rechts gegen x konvergiert (d.h. $x \leq x_n$ für jedes $n \leftarrow N$ und $\lim_{n \rightarrow \infty} x_n = x$), so gilt

$$\lim_{n \rightarrow \infty} F(x_n) = F(x)$$

(“ F ist rechtsseitig stetig”)

Teil III

Zufallsvariablen und ihre Verteilung

Wir gehen von einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ aus, R sei unsere gesuchte Bild-Menge, \mathfrak{S} σ -Algebra auf R .

$$X : \Omega \rightarrow R$$

sei Wahrscheinlichkeitsfunktion mit der Eigenschaft, dass für jedes $B \in \mathfrak{S}$ gilt

$$X^{-1}(B) = \{\omega \in \Omega; X(\omega) \in B\} \in \mathfrak{A} \quad (42)$$

heißt Zufallsvariable.

•

$$\{X \in B\} = \{\omega \in \Omega; X(\omega) \in B\} = X^{-1}(B) \quad (43)$$

•

$$P(X \in B) \text{ für } P(\{X \in B\}) = P(X^{-1}(B)) \quad (44)$$

• Durch

$$P^X : \mathfrak{S} \rightarrow [0, 1]$$

definiert durch

$$P(B) = P(X^{-1}(B)) = P(X \in B), \quad B \in \mathfrak{S}$$

wird ein W-Maß P^X auf \mathfrak{S} definiert.

•

$$X^{-1}\left(\sum_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} X^{-1}(B_n)$$

P^X heißt die **Verteilung von X** .

6 Spezialfälle

1. $R = \mathbb{R}, \mathfrak{S} = \mathfrak{B}$. Eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ heißt **reelle Zufallsvariable**.

$F(x) = P^X((-\infty, x]) = P(X \in ((-\infty, x])) = P(X \leq x), x \in \mathbb{R}$ ist die Verteilungsfunktion von P^X , kurz: Die Verteilungsfunktion von X .

Hat P^X die Dichte f , so ist

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt, x \in \mathbb{R}$$

Sprechweise: X hat die Dichte f .

2. $R \neq \emptyset$, \mathfrak{G} enthalte die Mengen $\{x\}$ für $x \in R$.

Es gibt eine abzählbare Menge $R_0 \subset R$ mit $P(X \in R_0) = 1$.

Dann ist $P(X \notin R_0) = 1 - P(X \in R_0) = 0$

Es gilt daher

$$P(X \in R_0^c \cap B) = 0, \forall B \in \mathfrak{G}$$

und

$$\begin{aligned} P(X \in B) &= P(X \in R_0 \cap B) + \cancel{P(X \in R_0^c \cap B)} \\ &= \sum_{x \in R_0 \cap B} P(X = x) \\ &= \sum_{x \in R_0} P(X = x) \delta_x(B) \end{aligned}$$

Also

$$P^X = \sum_{x \in R_0} P(X = x) \delta_x$$

X heißt dann diskret verteilt.

P^X ist festgelegt durch die Werte $P(X = x), x \in R_0$

7 Beispiel: Würfelwurf von 2 echten Würfeln

- $\Omega = \{(\omega_1, \omega_2); \omega_1, \omega_2 \in \{1, \dots, 6\}\}$
- $|\Omega| = 6^2 = 36$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|}, A \subset \Omega$
- $X : \Omega \rightarrow \mathbb{R}, X(\omega_1, \omega_2) = \omega_1 + \omega_2, (\omega_1, \omega_2) \in \Omega$
- $P(X = k)$

Also

$$\begin{aligned} P(X = k) &= P(\{(\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 = k\}) \\ &= \frac{|\{(\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 = k\}|}{36} \\ &= \begin{cases} \frac{1}{36} & , k = 2 \\ \frac{2}{36} & , k = 3 \\ \frac{3}{36} & , k = 4 \\ \vdots & \\ \frac{1}{36} & , k = 12 \end{cases} \end{aligned} \tag{45}$$

Natürlich kann man das auch allgemeiner schreiben:

Sei $R = \mathbb{R}^d = \mathfrak{B}^d$. X heißt dann d -dimensionaler Zufallsvektor $X : \Omega \rightarrow \mathbb{R}^d$.

$X = (X_1, \dots, X_d)$ mit reellen Zufallsvariablen X_1, \dots, X_n die Komponenten von X . X heißt nun **Zufallsvektor**.

Sei $f : \mathbb{R}^d \rightarrow \mathbb{R}_+$ mit

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \dots, t_d) dt_1 \cdots dt_d = 1$$

f heißt Dichte des Zufallsvektors X , wenn gilt

$$\begin{aligned} F(x_1, \dots, x_d) &= P(X_1 \leq x_1, \dots, X_d \leq x_d) \\ &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \dots, t_d) dt_1 \cdots dt_d, \quad (x_1, \dots, x_d) \in \mathbb{R}^d \end{aligned}$$

F heißt Verteilungsfunktion des Zufallsvektors.

8 Beispiele für Zufallsvariablen und ihre Verteilungen

Sei $(\Omega, \mathfrak{A}, P)$ W-Raum und $X : \Omega \rightarrow \mathbb{R}$ reelle Zufallsvariable. X hat dann eine Dichtefunktion (oder kurz „Dichte“) $f(\geq 0)$, wenn gilt

$$F(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f(t)dt, \quad -\infty < x < +\infty \quad (46)$$

....

8.1 Die Rechteck- oder Gleichverteilung

X heißt gleichverteilt oder Rechteck-verteilt auf $[a, b]$. Die Schreibweise hierfür sieht aus wie folgt

$$X \sim \mathfrak{R}(a, b)$$

Soviel zu Rechteck- und Gleich-verteilung.

8.2 Die Normalverteilung

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(t-\mu)^2}{\sigma^2}\right), \quad -\infty < t < +\infty (\mu \in \mathbb{R}, \sigma^2 > 0) \quad (47)$$

Die Dichte sieht aus wie folgt

$$F(x) = P(X \leq x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(t-\mu)^2}{\sigma^2}\right)dt \quad (48)$$

Für

- $s := \frac{t-\mu}{\sigma}$
- $t = \mu + \sigma s$
- $dt = \sigma ds$

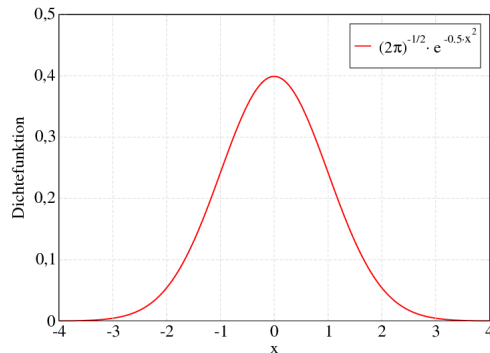
ist (oben)

$$\begin{aligned} () &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot s^2\right)dt \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \\ \Phi(x) &= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} \cdot s^2\right)dt, \quad -\infty < x < +\infty \end{aligned} \quad (49)$$

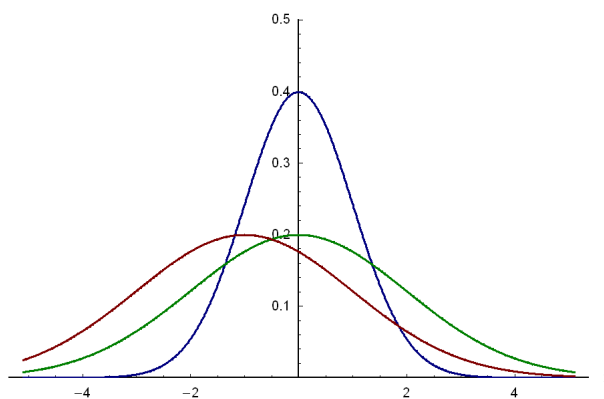
X heißt Normal-verteilt mit dem Parameter $\mu \in \mathbb{R}$ und $\sigma^2 > 0$. Schreibweise:

$$X \sim N(\mu, \sigma^2) \quad (50)$$

8.2.1 Beispiele für Normalverteilungen



- Das ² ist die so genannte Standard-Normal-Verteilung, auch bekannt als Gauß'sche Glockenkurve



- Zum Vergleich Dichtefunktionen der Normalverteilungen $N(0, 1)$ (blau), $N(0, 2)$ (grün) und $N(-1, 2)$ (rot)

8.3 Die Exponential-Verteilung

$$f(t) = \begin{cases} 0 & , t < 0 \\ \lambda \cdot \exp(-\lambda t) & , t \geq 0 \end{cases} \quad (51)$$

Hierbei sei $\lambda > 0$.

Dann

$$F(x) = P(X \leq x) = \begin{cases} 0 & , x < 0 \\ 1 - \exp(-\lambda x) & , x \geq 0 \end{cases} \quad (52)$$

X heißt Exponential-verteilt mit dem Parameter $\lambda > 0$. Schreibweise:

$$X \sim \text{Exp}(\lambda) \quad (53)$$

²Quelle [Wikipedia](#)

$X : \Omega \rightarrow \mathbb{R}^2$ sei ein 2-dimensionaler Zufallsvektor, $X = (X_1, X_2)$. X habe die Dichte $f(\geq 0)$, d.h.

$$\begin{aligned} P(X_1 \leq x_1, X_2 \leq x_2) &= \int_{-\infty}^{x_2} \left(\int_{-\infty}^{x_1} f(t_1, t_2) dt_2 \right) dt_1 \\ &= \int_{-\infty}^{x_1} \left(\int_{-\infty}^{x_2} f(t_1, t_2) dt_2 \right) dt_1 \end{aligned} \quad (54)$$

Dann ist

$$\begin{aligned} P(X_1 \leq x_1) &= \int_{-\infty}^{x_1} \left(\int_{-\infty}^{+\infty} f(t_1, t_2) dt_2 \right) dt_1 \\ &= \int_{-\infty}^{x_1} f_{X_1}(t_1) dt_1 \end{aligned} \quad (55)$$

Also ist

$$f_{X_1}(t_1) = \int_{-\infty}^{+\infty} f(t_1, t_2) dt_2, \quad -\infty < t_1 < +\infty \quad (56)$$

Dichte von X_1

Analog errechnet sich die Dichte von X_2

Teil IV

Bedingte Wahrscheinlichkeiten

9 Bedingte Wahrscheinlichkeiten

$(\Omega, \mathfrak{A}, P)$ ist W-Raum. Zunächst $|\Omega| < \infty$, $\mathfrak{A} = \mathfrak{P}(\Omega)$, $P(A) = \frac{|A|}{|\Omega|}$, $A \subset \Omega$, $A \subset \Omega$, $B \subset \Omega$, $|B| \geq 1$. Dann

$$\frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)} \quad (57)$$

„Wahrscheinlichkeit für das Eintreten von A unter Bedingung, dass B eintritt“, lese als „Wahrscheinlichkeit von A unter B“

9.1 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum (bel), $A, B \in \mathfrak{A}$, $P(B) > 0$. Dann heißt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (58)$$

die bedingte Wahrscheinlichkeit von A unter B („A gegeben B“).

9.2 Satz

$(\Omega, \mathfrak{A}, P)$ W-Raum, $\emptyset \neq I$ abzählbare Indexmenge. Wir unterstellen $A_i \in \mathfrak{A}, i \in I$ seien paarweise disjunkt, $P(A_i) > 0$ für jede $i \in I$. Dann halten wir fest

1. Für $A \in \mathfrak{A}$ mit der Eigenschaft $A \in \sum_{i \in I} A_i$ ist $P(A) = \sum_{i \in I} P(A|A_i)P(A_i)$ (bekannt als „Satz von der totalen Wahrscheinlichkeit“)
2. Sei $A \in \mathfrak{A}$ mit $P(A) > 0, i_0 \in I$. Dann

$$P(A_{i_0}|A) = \frac{P(A|A_{i_0})P(A_{i_0})}{\sum_{i \in I} P(A|A_i)P(A_i)} \quad (59)$$

(bekannt als „Formel von Bayes“)

3. Sei $I = \{0, 1, \dots, n\}$, ($n \in \mathbb{N}$ fest). Dann gilt

$$P(A_0 \cap A_1 \cap \dots \cap A_n) = P(A_0)P(A_1|A_0)P(A_2|A_0 \cap A_1) \cdots P(A_n|A_0 \cap \dots \cap A_{n-1}) \quad (60)$$

(bekannt als „Multiplikationsformel“)

9.2.1 Beweis

1. Aus

$$A = \sum_{i \in I} (A \cap A_i)$$

folge sofort

$$P(A) \sum_{i \in I} P(A \cap A_i) = \sum_{i \in I} P(A|A_i)P(A_i)$$

2. folgt sofort aus 1.

- 3.

$$\begin{aligned} & P(A_0)P(A_1|A_0)P(A_2|A_0 \cap A_1) \cdots P(A_n|A_0 \cap \dots \cap A_{n-1}) \\ &= \cancel{P(A_0)} \cdot \frac{P(A_1 \cap \cancel{A_0})}{\cancel{P(A_0)}} \cdot \frac{P(A_0 \cap A_1 \cap A_2)}{\cancel{P(A_0 \cap A_1)}} \cdots \\ &= P(A_0 \cap A_1 \cap \dots \cap A_n) \end{aligned} \quad (61)$$

9.3 Beispiel: Lostrommel

Eine Lostrommel enthält a Gewinnlose und b Nieten, $A \geq 1, b \geq 2$. Es gibt 2 Spielmöglichkeiten:

1. Ein Los wird gezogen. Das gezogene Los ist ein Gewinnlos oder eine Niete. Das Spiel ist beendet.
2. Ein Los wird gezogen und unbesehen weggeworfen. Daraufhin entfernt der Losverkäufer eine Niete aus der Trommel. Ein weiteres Los wird gezogen. Dieses Los ist ein Gewinnlos oder eine Niete.

Sei A das Ereignis, dass bei der ersten Ziehung ein Gewinnlos gezogen wird.
 Sei B das Ereignis, dass bei der zweiten Ziehung ein Gewinnlos gezogen wird.

Dann ergibt sich bei diesen Fällen

1.
$$P(A) = \frac{a}{a+b} \quad (62)$$

2.
$$\begin{aligned} P(B) &= P(B|A) \cdot P(A) + P(B|A^c)P(A^c) \\ &= \frac{a-1}{a+b-2} \cdot \frac{a}{a+b} + \frac{a}{a+b-2} \cdot \frac{b}{a+b} \\ &= \frac{a}{a+b} \cdot \underbrace{\left(\frac{a+b-1}{a+b-2} \right)}_{\geq 1} > \frac{a}{a+b} > P(A) \end{aligned} \quad (63)$$

Sei nun $a = 1$, $b = 2$. Dann folgt sofort

$$\begin{aligned} P(A) &= \frac{1}{3} \\ P(B) &= \frac{1}{3} \cdot 2 = \frac{2}{3} \end{aligned}$$

Sei $A, B \in \mathfrak{A}$, $P(B) > 0$. Dann gilt offenbar

$$P(A|B) = P(A) \iff P(A \cap B) = P(A)P(B) \quad (64)$$

10 Stochastische Unabhängigkeit von Ereignissen und Zufallsvariablen

10.1 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum, $A, B \in \mathfrak{A}$. A, B heißen (stochastisch) unabhängig, wenn

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

gilt.

10.2 Einsicht / Beweis

$$\begin{aligned}
 P(A \cap B^c) &= P(A) - P(A \cap B) \\
 &= P(A) - P(A)P(B), \text{ falls } A, B \text{ unabhängig} \\
 &= (1 - P(B))P(A) \\
 &= P(A)P(B^c)
 \end{aligned} \tag{65}$$

10.3 Definition

- $\emptyset \neq I$ endliche Indexmenge, $A_i \in \mathfrak{A}, i \in I$ Ereignisse.
 A_i heißen (stochastisch) unabhängig, wenn für jede $\emptyset \neq J \subset I$ gilt

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j) \tag{66}$$

- $\emptyset \neq I$ beliebige Indexmenge, $A_i \in \mathfrak{A}, i \in I$ Ereignisse.
 $\mathfrak{A}_i, i \in I$ heißen unabhängig, wenn für jede Teilmenge $\emptyset \neq J \subset I$, Jendlich die Ereignisse $A_j, j \in J$ unabhängig sind.

10.3.1 Bemerkung

$A_i, i \in I$ seien Ereignisse. Dann gilt

$$A_i, i \in I \text{ unabhängig} \iff B_i, i \in I \text{ unabhängig}, B_i \in \{A_i, A_i^c\} \text{ beliebig} \tag{67}$$

10.4 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum, $\emptyset \neq I$ Indexmenge, (R_i, \mathfrak{S}) Messräume, $i \in I$, $X_i : \Omega \rightarrow R_i$ Zufallsvariable, $i \in I$.

$X_i, i \in I$ heißen (stochastisch) unabhängig, wenn für jede Wahl von $B_i \in \mathfrak{S}_i, i \in I$ die Ereignisse $A_i = X_i^{-1}(B_i), i \in I$ unabhängig sind.

$(\Omega, \mathfrak{A}, P)$ sei W-Raum. Dann $X_i : \Omega \Rightarrow R_i, i = 1, \dots, n$ Zufallsvariable X_1, \dots, X_n sind unabhängig, wenn gilt

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$$

für jede Auswahl von Ereignissen $B_i \in \mathfrak{A}, i = 1, \dots, n$

...

10.5 Spezialfälle

10.5.1 Diskrete Verteilung

Die X_1, \dots, X_n haben je eine diskrete Verteilung. Dann sind

$$X_1, \dots, X_n \text{ unabhängig} \iff P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$$

für jede Auswahl von Elementen $x_i \in R_i, i = 1, \dots, n$.

(Beachte hier: $\{x\} \in R_i, \forall x \in R, \forall i \in I$)

10.5.2 Stetige Verteilung

Sei $X = (X_1, \dots, X_n)$ mit den reellen Zufallsvariablen X_1, \dots, X_n . Es habe X eine Dichte f , das heißt

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_1, \dots, dt_n, (x_1, \dots, x_n \in \mathbb{R}^n)$$

Dann hat X_i die Dichte

$$f_i(x_i) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \dots, t_{i-1}, x_i, t_{i+1}, \dots, t_n) dt_1, \dots, dt_n$$

Gilt

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i), (x_1, \dots, x_n) \in \mathbb{R}^n \quad (68)$$

so sind die Zufallsvariablen X_1, \dots, X_n unabhängig.

10.6 Beispiel

Ein Einzelexperiment, bei dem ein interessierendes Ereignis A mit Wahrscheinlichkeit $p \in [0, 1]$ eintritt, wird n -mal unter identischen Versuchsbedingungen ohne gegenseitige Beeinflussung (unabhängige Versuchswiederholungen) wiederholt.

Beschreibe das Einzelexperiment durch den W-Raum $(\Omega_0, \mathfrak{A}_0, P_0) = (\{0, 1\}, \mathfrak{P}(\{0, 1\}), P_0)$ mit

- $P_0(\{0\}) = 1 - p$

- $P_0(\{1\}) = p$
- $0 \cong A$ tritt nicht ein
- $1 \cong A$ tritt ein

Beschreibe das Gesamtexperiment durch

$$\begin{aligned}
 (\Omega, \mathfrak{A}, P) \text{ mit } \Omega = \{0, 1\}^n = \{(\omega_1, \dots, \omega_n); \omega_i \in \{0, 1\}, i = 1, \dots, n\}, \mathfrak{A} = \mathfrak{P}(\Omega) \\
 P(\{\omega_1, \dots, \omega_n\}) = P_0(\{\omega_1\}) \cdot P_0(\{\omega_n\}) \\
 = p^{\omega_1} \cdot (1-p)^{1-\omega_1} \dots p^{\omega_n} \cdot (1-p)^{1-\omega_n} \\
 = p^{\omega_1 + \dots + \omega_n} \cdot (1-p)^{n - (\omega_1 + \dots + \omega_n)}, (\omega_1, \dots, \omega_n) \in \Omega
 \end{aligned} \tag{69}$$

$X_i : \Omega \rightarrow \{0, 1\}$ sei definiert durch

$$X_i(\omega_1, \dots, \omega_n) = \omega_i, (\omega_1, \dots, \omega_n) \in \Omega, i = 1, \dots, n$$

Die X_1, \dots, X_n sind unabhängig, denn

$$\begin{aligned}
 &P(X_1 = \epsilon_1, \dots, X_n = \epsilon_n) \\
 &= P(\omega = \{1, \dots, n\} \in \Omega; X_1(\omega_1, \dots, \omega_n) = \epsilon_1, \dots, X_n(\omega_1, \dots, \omega_n) = \epsilon_n) \\
 &= P(\{(\epsilon_1, \dots, \epsilon_n)\}) \\
 &= p^{\epsilon_1 + \dots + \epsilon_n} \cdot (1-p)^{n - (\epsilon_1 + \dots + \epsilon_n)} \\
 &= \prod_{i=1}^n p_i^{\epsilon_i} \cdot (1-p)^{1-\epsilon_i}
 \end{aligned} \tag{70}$$

$$\begin{aligned}
 P(X_1 = \epsilon_1) &= P(\{(\omega_1, \dots, \omega_n) \in \Omega; X_1(\omega_1, \dots, \omega_n) = \epsilon_1\}) \\
 &= P(\{(\epsilon_1, \omega_2, \dots, \omega_n); \omega_i \in \{0, 1\}, 2 \leq i \leq n\}) \\
 &= \sum_{nnn} P \dots
 \end{aligned} \tag{71}$$

$$P(X = k) = \dots \tag{72}$$

11 Binomialverteilung

11.1 Definition

Eine reelle (oder $\{0, 1\}$ -wertige) Zufallsvariable X heißt Binomialverteilt mit den Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$, wenn gilt

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}, k \in \{0, 1, \dots, n\} \quad (73)$$

Schreibweise:

$$X \sim \mathfrak{B}(n, p)$$

11.2 Anwendung

Sei $(\Omega, \mathfrak{A}, P)$ W-Raum, $A \in \mathfrak{A}$ Ereignis. Dann heißt $I_A : \Omega \rightarrow \{0, 1\}$ **Indikator** (oder Indikatorvariable) von A , definiert durch

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A. \end{cases}$$

$I_A \sim \mathfrak{B}(1, p)$ mit $p = P(A)$.

Seien $A_1, \dots, A_n \in \mathfrak{A}$ unabhängige Ereignisse mit $p = P(A_i), i = 1, \dots, n$.

Dann gilt $\underbrace{X_1}_{=I_{A_1}}, \dots, \underbrace{X_n}_{=I_{A_n}}$ unabhängig und $X_i \sim \mathfrak{B}(1, p)$. $X = X_1 + \dots + X_n \sim \mathfrak{B}(n, p)$.

11.3 Beispiel Urnenmodell

Eine Urne enthält r rote und s schwarze Kugeln. Uns können verschiedene Modi interessieren.

11.3.1 Mit Zurücklegen

Es wird n -mal mit Zurücklegen je eine Kugel gezogen.

Sei X Zufallsvariable und bezeichne die Anzahl der gezogenen roten Kugeln in n Ziehungen.

Dann $X \sim \mathfrak{B}\left(n, \frac{r}{r+s}\right)$

11.3.2 Ohne Zurücklegen

Es wird n -mal ohne Zurücklegen je eine Kugel gezogen mit $n \leq r + s =: a$.

Sei X Zufallsvariable und bezeichne die Anzahl der gezogenen roten Kugeln in n Ziehungen.

Nummeriere die Kugeln in folgender Weise:

$$\underbrace{1, \dots, r}_{\text{rot}}, \underbrace{r+1, \dots, \overbrace{r+s}^{=a}}_{\text{schwarz}}$$

Dann

- $\Omega = \{(\omega_1, \dots, \omega_n); \omega_i \in \{1, \dots, a\} \text{ paarweise verschieden}, 1 \leq i \leq n\}$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{a \cdot (a-1) \cdot \dots \cdot (a-n+1)}, A \subset \Omega$
-

$$\begin{aligned} P(X = k) &= \frac{|\{(\omega_1, \dots, \omega_n) \in \Omega; |j \in \{1, \dots, n\}, \omega_j \in \{1, \dots, r\}| = k\}|}{|\Omega|} \\ &= \frac{\binom{n}{k} \cdot r \cdot (r-1) \cdot \dots \cdot (r-k+1) \cdot s \cdot (s-1) \cdot \dots \cdot (s-(n-k)+1)}{a \cdot (a-1) \cdot \dots \cdot (a-n+1)} \quad (74) \\ &= \frac{\binom{r}{k} \cdot \binom{a-r}{n-k}}{\binom{a}{n}}, k \in \mathbb{N}, \max(0, a-r-n) \leq k \leq \min(n, r) \end{aligned}$$

12 Hypergeometrische Verteilung

12.1 Definition

Eine reelle Zufallsvariable mit der Verteilung

$$P(X = k) = \frac{\binom{r}{k} \cdot \binom{a-r}{n-k}}{\binom{a}{n}}$$

heißt hypergeometrisch verteilt mit den Parametern $a, r, n \in \mathbb{N}, r \leq a, n \leq a$ und der Schreibweise

$$X \sim \mathfrak{H}(a, r, n). \quad (75)$$

Das bekannteste Beispiel dürfte das Zahlenlotto “6 aus 49” sein.

Teil V**Erwartungswerte****13 Erwartungswerte von Zufallsvariablen**

Wir definieren Erwartungswerte von Zufallsvariablen schrittweise, beginnend mit Indikatorvariablen. Es sei X eine reelle Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$.

- (1) Für eine Indikatorvariable $X = I_A$ mit $A \in \mathfrak{A}$ sei $E(X) = E_p(X) = P(A)$
- (2) Für eine nicht negative einfache Zufallsvariable $X = \sum_{k=1}^n \alpha_k I_{A_k}$ mit nicht negativen reellen Zahlen $\alpha_1, \dots, \alpha_n$ und Ereignissen $A_1, \dots, A_n \in \mathfrak{A}$ sei

$$E(X) = E_p(X) = \sum \alpha_k \cdot P(A_k) \quad (76)$$

Der so definierte Wert $E(X)$ hängt nicht von der Darstellung von X ab, d.h. ist $X = \dots$

- (3) Für eine nicht negative Zufallsvariable $X \geq 0$ gibt es eine Folge von nicht negativen einfachen Zufallsvariablen X_n , $n \in \mathbb{N}$, mit $0 \leq X_1 \leq \dots$ mit $X = \sup_{n \in \mathbb{N}} X_n$. Es sei $E(X) = E - p(X) = \sup_{n \in \mathbb{N}} E(X)$
- (4) Jede reelle Zufallsvariable X lässt sich in der Form $X = X^+ - X^-$ mit den nicht negativen Zufallsvariablen $X^+ = \max(X, 0)$

14 Eigenschaften von Erwartungswerten

Es seien X, Y reelle Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$, α, β reelle Zahlen.

- (1) Sind $X, Y \geq 0$ und $\alpha, \beta \geq 0$, so ist

$$E(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot E(X) + \beta \cdot E(Y)$$

- (2) Sind X, Y integrierbar, so ist $\alpha \cdot X + \beta \cdot Y$ integrierbar und es ist

$$E(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot E(X) + \beta \cdot E(Y)$$

- (3) ist $0 \leq X \leq Y$, so ist $0 \leq E(X) \leq E(Y)$.
- (4) Existiert für die Zufallsvariable X der Erwartungswert, so ist $|E(X)| \leq E(|X|)$.
- (5) Sind X, Y integrierbar und ist $X \leq Y$, so ist $E(X) \leq E(Y)$.

- (6) X ist genau dann integrierbar, wenn $|X|$ integrierbar ist.
- (7) Ist $|X| \leq |Y|$ und Y integrierbar, so ist auch X integrierbar.
- (8) Sind $X, \dots, X_n, n \in \mathbb{N}$ nicht negative Zufallsvariablen mit $0 \leq X_1 \leq X_2 \leq \dots$ und $X = \sup_{n \in \mathbb{N}} X_n$, so ist $E(X) = \sup_{n \in \mathbb{N}} E(X_n)$
- (9) Sind $X_n, n \in \mathbb{N}$ nicht negative Zufallsvariablen, so ist

$$E(\liminf_{n \rightarrow \infty} X_n) \leq \liminf_{n \rightarrow \infty} E(X_n)$$

- (10) Sind $X, Y, X_n, n \in \mathbb{N}$ reelle Zufallsvariablen mit $X = \lim_{n \rightarrow \infty} X_n$ und $|X_n| \leq |Y|$ für jedes $n \in \mathbb{N}$, und es ist Y integrierbar, so sind auch X, X_n integrierbar und es gilt $\lim_{n \rightarrow \infty} E(|X_n - X|) = 0$. Insbesondere folgt hieraus auch $\lim_{n \rightarrow \infty} E(X) = E(X)$.
- (11) Ist $X \geq 0$, so ist $E(X) = 0$ genau dann, wenn $P(X > 0) = 0$ ist.

Aus der Eigenschaft (8) ergibt sich für eine Folge von nicht negativen, reellen Zufallsvariablen X_n mit der Eigenschaft, dass die Reihe $\sum_{k=1}^n X_k$ gegen die reelle Zufallsvariable X konvergiert, die Identität

$$E\left(\sum_{n=1}^{\infty} X_n\right) = \dots \quad (77)$$

... folgt hieraus

$$|X| = \sum_{n=1}^{\infty} |X| \cdot I(n-1 < X \leq n), \quad (78)$$

dass

$$\begin{aligned} E(|X|) &= \sum_{n=1}^{\infty} E(|X| \cdot I(n-1 < X \leq n)) \\ &\leq \sum_{n=1}^{\infty} P(n-1 < |X| \leq n) \\ &= \dots \\ &= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} P(n-1 < |X| \leq n) \\ &= \sum_{k=1}^{\infty} P(|X| > n-1) \\ &= \sum_{k=0}^{\infty} P(|X| > k) \end{aligned} \quad (79)$$

ist. Analog sieht man

$$E() \dots \quad (80)$$

...

15 Rechnen mit Erwartungswerten

() W-Raum, $\emptyset \neq R$ Menge, \mathfrak{L} - σ -Algebra auf R , $f : R \rightarrow \mathbb{R}$, $f^{-1}(B) \in \mathfrak{L}$ für jedes $B \in \mathfrak{B}$. Dann P^X Verteilung von X , P^X ist W-Maß auf R und (R, \mathfrak{L}, P^X) ist W-Raum für reelle Zufallsvariablen auf (R, \mathfrak{L}, P^X) .

Der Erwartungswert von $f \cdot X$ existiert genau da, wenn der Erwartungswert von f , aufgefasst als Erwartungswert auf (R, \mathfrak{L}, P^X) , existiert. Es gilt dann die **Transformationsformel**

$$E(f \cdot X) = E_P(f \cdot X) = E_{P^X}(f) \quad (81)$$

16 Spezialfälle

16.1 Stetige Verteilung

$(R, \mathfrak{L}) = (\mathbb{R}, \mathfrak{B})$, X habe die Dichte g .

$$(1) \quad f = I_{(-\infty, x]}$$

$$\begin{aligned} F(x) &= P(X \leq x) = \underbrace{E(I_{(-\infty, x]} \cdot X)}_{=E(f \cdot X)} \\ &= E_{P^X}(I_{(-\infty, x]}) \\ &= P^X((-\infty, x]) = \int_{-\infty}^x \\ &= \int_{-\infty}^x f(t)g(t)dt \end{aligned} \quad (82)$$

(2) Die Identität

$$E(f \cdot X) = \int_{-\infty}^x f(t)g(t)dt$$

gilt für "beliebige" $f : \mathbb{R} \rightarrow \mathbb{R}$ (mit $f^{-1}(B) \in \mathfrak{B}$ für jedes $B \in \mathfrak{B}$, mit existierendem $E(f \cdot X)$).

16.2 Diskrete Verteilung

X habe eine diskrete Verteilung, d.h. es gibt eine abzählbare Menge $A \subset R$ mit $P(X \in A) = 1$, $(\{x\} \in \mathfrak{L} \forall x \in R)$.

$$P^X = \sum_{a \in A} P(X = a) \cdot \delta_a$$

Für $f \geq 0$ ist

$$\begin{aligned} E(f \cdot X) &= E \left(\sum_{a \in A} f(a) \cdot I(X = a) + \underbrace{E(X \cdot I(X \neq A))}_{=0} \right) \\ &= \sum_{a \in A} f(a) \cdot E(I(X = a)) \\ &= \sum_{a \in A} f(a) \cdot P(X = a) \end{aligned} \tag{83}$$

Diese Gleichung ist besonders wichtig, da sie für beliebige f gilt mit existierendem $E(f \cdot X)$

17 Anwendungen

17.1 Erwartungswert der Rechteck-Verteilung

Mit $X \sim \mathfrak{R}(a, b)$ gilt

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

Dann ist der Erwartungswert

$$E(X) = \frac{1}{b-a} \cdot \int_a^b t \cdot dt = \frac{a+b}{2} \tag{84}$$

17.2 Erwartungswert der Normalverteilung

Sei $X \sim \mathfrak{N}(a, \sigma)$. Dann ist der Erwartungswert

...

17.3 Beispiel

17.3.1 Binomialverteilt

$$X \sim \mathfrak{B}(n, p)$$

Ohne Einschränkung $X = \sum_{j=1}^n I_{A_j}$

$$p = P(A_j), \quad j = 1, \dots, n$$

$$E(X) = \sum_{j=1}^n \underbrace{E(I_{A_j})}_{=P(A)=p} = n \cdot p$$

17.3.2 Hypergeometrisch verteilt

$$X \sim \mathfrak{H}(a, r, n)$$

$$X = \sum_{j=1}^n I_{A_j}$$

A_j ist das Ereignis "Beim n -fachen Ziehen ohne Zurücklegen je einer Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln und bei der j -ten Ziehung wird eine rote Kugel gezogen".

$$E(X) = \sum_{j=1}^n P(A_j)$$

$$P(A_j) = \frac{r \cdot \cancel{(a-1)} \cdot \cancel{(a-2)} \cdots \cancel{(a-n+1)}}{a \cdot \cancel{(a-1)} \cdot \cancel{(a-2)} \cdots \cancel{(a-n+1)}} = \frac{r}{a} = \frac{r}{r+s}, \quad j = 1, \dots, n$$

Also: $E(X) = n \cdot \frac{r}{a}$

Beim n -fachen Ziehen mit Zurücklegen je eine Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln ist Anzahl X der gezogenen roten Kugeln $\mathfrak{B}(n, \frac{r}{a})$ -verteilt, also

$$E(X) = n \cdot \frac{r}{a}$$

18 Die Varianz

Sei $r \in \mathbb{N}$, X reelle Zufallsvariable mit $E(|X|^r) < \infty$.

Für $1 \leq s \leq r$, $s \in \mathbb{N}$ gilt dann wegen $|X|^s \leq |X|^r + 1$, dass $E(|X|^s) < \infty$ ist. Sei X reelle Zufallsvariable mit $E(X^2) < \infty$. Dann ist $E(|X|) < \infty$.

Für $a \in \mathbb{R}$ ist $|X - a|^2 \leq X^2 + 2 \cdot |a| \cdot |X| + a^2$, also $E(|X - a|^2) < \infty$. Es ist

$$\begin{aligned} E([X - a]^2) &= E\left(\left([X - E(X)] + [E(X) - a]\right)^2\right) \\ &= E\left([X - E(X)]^2\right) + [E(X) - a]^2 + \underbrace{E\left(2 \cdot [E(X) - a] \cdot [X - E(X)]\right)}_{=0} \\ &\geq E\left([X - E(X)]^2\right) \end{aligned}$$

18.1 Definition

Sei X reelle Zufallsvariable mit $E(X^2) < \infty$. Dann heißt

$$\text{Var}(X) = E([X - E(X)]^2) \quad (85)$$

die **Varianz** von X .

Es ist

$$\text{Var}(X) = \min_{a \in \mathbb{R}} E([X - a]^2) \quad (86)$$

18.2 Bemerkung

Es ist

$$\begin{aligned} \text{Var}(X) &= E([X - a]^2) = E(X^2 - 2 \cdot X \cdot E(X) + (E(X))^2) \\ &= E(X^2) - 2 \cdot (E(X))^2 + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \geq 0 \end{aligned} \quad (87)$$

18.3 Spezialfälle

18.3.1 Zufallsvariablen mit diskreter Verteilung

$$\begin{aligned} \text{Var}(X) &= \sum_{\substack{x \in \mathbb{R} \\ P(X=x) > 0}} (x - E(X))^2 \cdot P(X = x) \\ &= \sum_{\substack{x \in \mathbb{R} \\ P(X=x) > 0}} x^2 \cdot P(X = x) - \left(\sum_{\substack{x \in \mathbb{R} \\ P(X=x) > 0}} x \cdot P(X = x) \right)^2 \end{aligned} \quad (88)$$

18.3.2 Zufallsvariablen mit stetiger Verteilung

Hat X die Dichte f , so gilt

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{+\infty} (x - E(X))^2 \cdot f(x) \cdot dx \\ &= \int_{-\infty}^{+\infty} x^2 \cdot f(x) \cdot dx - \left(\int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx \right)^2 \end{aligned} \quad (89)$$

18.4 Beispiele

18.4.1 Rechteckverteilt

$$X \sim \mathfrak{R}(a, b)$$

18.4.2 Normalverteilt

$$X \sim \mathfrak{N}(a, \sigma^2)$$

18.4.3 Exponentialverteilt

$X \sim \text{Exp}(\lambda)$. X hat Dichte

$$f(x) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x) & , x \geq 0 \\ 0 & , \text{sonst} \end{cases}$$

Dann ist

$$E(X) = \frac{1}{\lambda} \quad (90)$$

$$\begin{aligned} \text{Var}(X) &= \lambda \cdot \int_0^{+\infty} [x^2 \cdot \exp(-\lambda \cdot x) \cdot dx] - \left(\frac{1}{\lambda}\right)^2 \\ &= \frac{1}{\lambda^2} \end{aligned} \quad (91)$$

19 Die Kovarianz

Seien X, Y reelle Zufallsvariablen mit $E(X^2) < \infty$ und $E(Y^2) < \infty$.

$$|X| \cdot |Y| \leq \frac{1}{2} \cdot (|X|^2 + |Y|^2) \iff 0 \leq (|X| - |Y|)^2 \Rightarrow E(|X \cdot Y|) < \infty$$

und es gilt

$$E((X + Y)^2) = E(X^2) + E(Y^2) + 2 \cdot E(X \cdot Y)$$

Also ist

$$\begin{aligned} \text{Var}(X + Y) &= E([X + Y - (E(X) + E(Y))])^2 \\ &= E([(X - E(X)) + (Y - E(Y))])^2 \\ &= \text{Var}(X) + \text{Var}(Y) + 2 \cdot E([X - E(X)] \cdot [Y - E(Y)]) \end{aligned} \quad (92)$$

19.1 Definition

X, Y sind reelle Zufallsvariablen mit $E(X^2) < \infty$, $E(Y^2) < \infty$. Dann heißt

$$\text{Cov}(X, Y) = E([X - E(X)] \cdot [Y - E(Y)]) \quad (93)$$

die Kovarianz von X und Y .

19.2 Bemerkung

Es ist

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

Also gilt

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y) \quad (94)$$

Mittels vollständiger Induktion folgt für reelle Zufallsvariablen X_1, \dots, X_n mit $E(X_j^2) < \infty, j = 1, \dots, n$.

$$\begin{aligned} \text{Var}(X_1 + \dots + X_n) &= \sum_{j=1}^n \text{Var}(X_j) + 2 \cdot \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) \\ &= \sum_{j=1}^n \text{Var}(X_j) + 2 \cdot \sum_{i,j=1, i \neq j}^n \text{Cov}(X_i, X_j) \end{aligned}$$

19.3 Spezialfall

$X = I_A, Y = I_B$. Dann

$$\text{Cov}(X, Y) = E(\underbrace{I_A \cdot I_B}_{I_{A \cap B}} - \underbrace{E(I_A)}_{P(A)} \cdot \underbrace{E(I_B)}_{P(B)})$$

$$\text{Cov}(I_A, I_B) = P(A \cap B) - P(A) \cdot P(B)$$

Also insbesondere $\text{Cov}(I_A, I_B) = 0$ falls A, B unabhängige Ereignisse. Ferner gilt für die Ereignisse A_1, \dots, A_n :

$$\begin{aligned} \text{Var}(I_{A_1} + \dots + I_{A_n}) &= \sum_{j=1}^n \underbrace{\text{Var}(I_{A_i})}_{P(A-i) \cdot (1-P(A_i))} + 2 \cdot \sum_{1 \leq i < j \leq n} [P(A_i \cap A_j) - P(A_i) \cdot P(A_j)] \\ &= \sum_{j=1}^n P(A-i) \cdot (1-P(A_i)) + 2 \cdot \sum_{1 \leq i < j \leq n} [P(A_i \cap A_j) - P(A_i) \cdot P(A_j)] \end{aligned}$$

19.4 Beispiele

19.4.1 Binomialverteilt

Sei $X \sim \mathfrak{B}(n, p)$. Dann wäre eigentlich

$$\text{Var}(X) = \sum_{k=0}^n (k - n \cdot p)^2 \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Wir vereinfachen das im Folgenden: O.B.d.A. sei $X = \sum_{j=1}^n I_{A_i}$ mit **unabhängigen** Ereignissen $A_i, i = 1, \dots, n$. Dann

$$\text{Var}(X) = \sum_{j=1}^n \text{Var}(I_{A_i}) = n \cdot p \cdot (1 - p) \quad (95)$$

19.4.2 Hypergeometrisch verteilt

Sei $X \sim \mathfrak{H}(a, r, n)$. Dann wäre eigentlich

$$\text{Var}(X) = \sum_{k=\max(0, n+r-a)}^{\min(r, n)} \left(k - n \cdot \frac{r}{a} \right)^2 \cdot \frac{\binom{r}{k} \cdot \binom{a-r}{n-k}}{\binom{a}{n}}$$

Wir vereinfachen das im Folgenden: O.B.d.A. fassen wir X auf als die Anzahl der gezogenen roten Kugeln in n Ziehungen bei n -fachem Ziehen ohne Zurücklegen je eine Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln.

Dann $X = \sum_{j=1}^n I_{A_i}$ und mit $P(A_i) = \frac{r}{a}$ sei

$$\begin{aligned} \text{Var}(X) &= n \cdot \frac{r}{a} \cdot \left(1 - \frac{r}{a} \right) + n \cdot (n-1) \cdot \left[\frac{r \cdot (r-1)}{a \cdot (a-1)} - \left(\frac{r}{a} \right)^2 \right] \\ &= \underbrace{n \cdot \frac{r}{a} \cdot \left(1 - \frac{r}{a} \right)}_{\text{Varianz beim Ziehen mit Zurücklegen}} \cdot \frac{a-n}{n-1} \end{aligned} \quad (96)$$

da

$$P(A_i \cap A_j) = \frac{r \cdot (r-1)(a-2) \cdots (a-n+1)}{a(a-1)(a-2) \cdots (a-n+1)} = \frac{r \cdot (r-1)}{a(a-1)}$$