Einführung in die Wahrscheinlichkeitstheorie und Statistik

Prof. Dr. Niko Naumann

2. Juni 2015

Inhaltsverzeichnis

1 Erinnerungen an die Maßtheorie

1.1 Defintion

Sei Ω eine Menge, $\mathcal{P}(\Omega) := \{A \mid A \subseteq \Omega\}$ die Potenzmenge von Ω auf $\mathcal{A} \in \mathcal{P}(\Omega)$.

a) $\mathcal{A}$ heißt **σ-Algebra** (auf Ω), falls gelten:

   i) $\emptyset \in \mathcal{A}$

   ii) $\forall A \in \mathcal{A}$ gilt $A^C := \{\omega \in \Omega \mid \omega \notin A\} \in \mathcal{A}$

   iii) $\forall A_n \in \mathcal{A}(n \in \mathbb{N})$ gilt $\bigcup_n A_n \in \mathcal{A}$

   Ein **Messraum** ist ein Paar $(\Omega, \mathcal{A})^{n\in\mathbb{N}}$ bestehend aus einer Menge Ω und einer $\mathcal{A} \in \mathcal{P}(\Omega)$.

b) Falls in a) statt iii) nur gilt:

   iii*) Für $A, B \in \mathcal{A}$ gilt $A \cup B \in \mathcal{A}$

   so heißt $\mathcal{A}$ eine **Algebra** (auf Ω)

c) Falls in a) statt iii) nur gilt:

   iii**) Falls $A_n \in \mathcal{A}(n \in \mathbb{N})$ paarweise disjunkt sind (d.h. $A_m \cap A_n \neq \emptyset$ für $n \neq m$), dann ist $\dot\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$

so heißt $\mathcal{A}$ ein **Dynkin System**.

1.2 Bemerkung & Definition

a) Die größte σ-Algebra auf Ω ist $\mathcal{P}(\Omega)$, die kleinste ist $\{\emptyset, \Omega\}$. Falls $I \neq \emptyset$ eine Menge und für jedes $i \in I$ $\mathcal{A}_i \subseteq \mathcal{P}(\Omega)$ eine σ-Algebra ist, so auch $\bigcap_{i\in\mathbb{N}} \mathcal{A}_i \subseteq \mathcal{P}(\Omega)$.
Sei $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ beliebig, dann ist

$$\sigma(\mathcal{E}) := \bigcap\{\mathcal{A} \mid \mathcal{A} \subseteq \mathcal{P}(\Omega) \text{ ist eine } \sigma \text{ -Algebra mit } \mathcal{E} \subseteq \mathcal{A}\} \subseteq \mathcal{P}\Omega$$

die kleinste σ-Algebra, die $\mathcal{E}$ enthält (**die von $\mathcal{E}$ erzeugte σ-Algebra**). Analog existieren ein kleinstes Dynkin-System $\mathcal{D}(\mathcal{E})$ und eine kleinste Algebra, die $\mathcal{E}$ die enthalten.
Die von $\mathcal{E}$ erzeugte Algebra kann man angeben, nämlich $\{\bigcup_{i=1}^{n} \bigcap_{j=1}^{n} A_{ij} \mid A_{ij} \in \mathcal{E} \cup \mathcal{E}^C, n \geq I\}$, wobei $\mathcal{E}^C := \{A^C \mid A \in \mathcal{E}$. Für σ -Algebren gilt nichts analoges.

b) Ein **Ring** (auf Ω) ist eine Teilmenge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ mit i) iii*) und ii*) Für alle $A, B \in \mathcal{A}$ ist $A \backslash B \in \mathcal{A}$.

1.3 Lemma

a) Ein $\cap$-stabiles Dynkin-System ist eine σ-Algebra.

b) **(Dynkin-Lemma)** Sei $\mathcal{E} \in \mathcal{P}(\Omega)$ $\cap$-stabil, dann ist $\mathcal{D}(\mathcal{E}) = \sigma(\mathcal{E})$

Beweis: Kommt später!

1.4 Definition

a) Sei $(\Omega, \mathcal{A})$ ein Messraum. Dann heißt eine Abbildung $\mu : \mathcal{A} \to [0, \infty] := [0, \infty) \cup \{\infty\}$ ein **Maß** (auf $(\Omega, \mathcal{A})$, falls gelten:

   i) $\mu(\emptyset) = 0$

   ii) σ-Additivität: Für alle $(\mathcal{A})_{n\geq 1} \subseteq \mathcal{A}$ mit $A_n \cap A_m = \emptyset$ für $n \neq m$ gilt $\mu(\dot\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} \mu(A_n)$

b) Ein Maß μ heißt

i) **endlich** $:\Leftrightarrow \mu(\Omega) \leq \infty$

ii) **σ-finit** $:\Leftrightarrow$ Es exisitiert eine Folge $(A_n)_{n\geq 1} \subseteq \mathcal{A}$ mit $\mu(A_n) \leq \infty$ $\forall n \leq 1$ und $\bigcup_{n\geq 1} A_n = \Omega$

iii) **Wahrscheinlichkeitsmaß** (W-Maß), falls gilt $\mu(\Omega) = 1$.

c) Sei $\mathcal{A}$ ein Ring auf Ω, dann heißt eine Abbildung $\mu : \mathcal{A} \to [0, \infty]$ ein **Inhalt**, falls gelten:

i) $\mu(\emptyset) = 0$

ii) $\mu(A \cup B) = \mu(A) + \mu(B)$ $\forall A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ **(endliche Additivität)**.

1.5 Beispiel

a) **Dirac-Maß**: Zu $x \in \Omega$ definiere für alle $A \in \mathcal{P}(\Omega) : \delta_x(A) := \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$. Dann ist ein δ_x ein Maß auf $\mathcal{P}(\Omega)$ (und damit auf jeder σ-Algebra auf Ω).

b) **Elementargeometrischer Inhalt**: Betrachte $\mathcal{E} := \{(a, b] \mid \ \leq a \leq b \leq +\infty\} \subseteq \mathcal{P}(\Omega := \mathbb{R})$ und $\mathcal{R} := \{\dot\bigcup_{i=1}^{n} A_i \mid A_i \in \mathcal{E}, A_i \cap A_j \neq \emptyset$ für $i \neq j, n \geq 1\} \subseteq \mathcal{P}(\mathbb{R})$. Dann ist $\mathcal{R}$ ein Ring. Setzte $\lambda(\bigcup_{i=1}^{n} A_i := \sum_{i=1}^{n}(b_i - a_i)$ für $A_i = (a_i, b_i] \in \mathcal{E}$ paaarweise disjunkt. Dann ist $\lambda : \mathbb{R} \to [0, \infty)$ ein Inhalt.

c) Sei $\Omega = \mathbb{N}, \mathcal{A} := \{A \subseteq \mathbb{N} \mid A$ oder A^C ist endlich $\}$ und

$$\mu(\mathcal{A}) := \begin{cases} 0, & A \text{ endlich} \\ 1, & A^C \text{ endlich} \end{cases} \quad \forall A \in \mathcal{A}$$

. Dann ist $\mathcal{A}$ eine Algebra auf $\mathbb{N}$. Und μ ist ein Inhalt auf $\mathcal{A}$, aber $\mathcal{A}$ ist keine σ-Algebra (und μ ist nicht σ-additiv).

d) **Zählmaß**: Für $A \subseteq \Omega$ definiere $\mu(A) := \begin{cases} |A|, & \text{falls } A \leq \infty \\ \infty, & \text{sonst} \end{cases}$. Dann ist μ ein Maß auf $\mathcal{P}(\Omega)$, welches genau dann σ-finit ist, wenn Ω abzählbat ist.

1.6 Bemerkung

Sei μ ein Inhalt auf einem Ring $\mathcal{A}$. Dann gelten:

a) μ ist **monoton**, d.h. für $A, B \in \mathcal{A}$ mit $A \subseteq B$ gilt $\mu(A) \leq \mu(B)$

b) μ ist **subtraktiv**, d.h. für $A, B \in \mathcal{A}$ mit $A \in B$ und $\mu(A) \leq \infty$ gilt $\mu(B \backslash A) = \mu(B) - \mu(A)$

c) μ ist **subadditiv**, d.h. für $A_1, ..., A_n \in \mathcal{A}$ gilt $\mu(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mu(A_i)$

1.7 Satz

Sei μ ein Inhalt auf einem Ring $\mathcal{A}$. Betrachte folgende Aussagen:

a) μ ist σ-additiv

b) Für alle $\{A_1, A_2, ...\} \subseteq \mathcal{A}$ mit $A_1 \subseteq A_2 \subseteq ...$ und $A := \bigcup_{n \geq 1} A_n \in \mathcal{A}$ gilt $\mu(A) = \lim_{n \to \infty} \mu(A_n)$ (d.h. μ ist **stetig von unten**).

c) Für alle $\{A_1, A_2, ...\} \subseteq \mathcal{A}$ mit $A_1 \supseteq A_2 \supseteq ..., \bigcap_{n \geq 1} A_n = \emptyset$ und $\mu(A_1) \leq \infty$ gilt $\lim_{n \to \infty} \mu(A_n) = 0$ (d.h.) μ ist **stetig von oben**. Dann gilt $a) \Leftrightarrow b) \Rightarrow c)$. Falls μ endlich ist, so gilt außerdem $c) \Leftrightarrow b)$

Beweis: Kommt später!

1.8 Definition

Seien Ω eine Menge, μ ein σ-additiver Inhalt auf einem Ring $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ und $B \subseteq \Omega$

a) Es heißt $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$, definiert durch

$$\mu^*(B) := \begin{cases} \inf\{\sum_{n \geq 1} \mu(A_n) \mid A_n \in \mathcal{A}, B \subseteq \bigcup_{n \geq 1} A_n\} & , \text{ falls } \{...\} \neq \emptyset \\ \infty & , \text{ sonst} \end{cases}$$

das dazugehörige Maß.

b) Eine Menge $A \subseteq \Omega$ heißt **maßlos**, falls gilt $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^C)$ für alle $B \in \mathcal{P}(\Omega)$

1.9 Satz

(Caratheodory) Seien Ω eine Menge und μ ein σ-additiver, σ-finiter Inhalt auf einem Ring $\mathcal{A}$ auf Ω. Dann ist $\sigma(\mathcal{A}) \subseteq \overline{\sigma}(\mathcal{A}) := \{B \subseteq \Omega \mid B \text{ messbar } \}$ eine σ-Algebra und $\mu^*|_{\overline{\sigma}(\mathcal{A})}$ ist ein Maß. Das Maß $\mu^*|_{\overline{\sigma}(\mathcal{A})}$ ist die einzige Maßfortsezung von μ. (ohne Beweis) Die σ-Algebra $\overline{\sigma}$ hat eine Bedeutung weiter unten. (siehe Satz 1.11)

1.10 Definition

Ein Maß μ auf einer σ-Algebra $\mathcal{A}$ heoßt **vollständig**, falls gilt: Aus $A \subseteq B \subseteq \Omega, B \subseteq \mathcal{A}$ und $\mu(B) = 0$ folgt $A \subseteq \mathcal{A}$ (und natürlich $\mu(A) = 0$). Ein vollständiges Maß $\mu : A \to [0, \infty]$ heißt **Vervollständigung des Maßes $\boldsymbol{\mu_0 : \mathcal{A} \to [0; \infty]}$**, falls $\mathcal{A}_0 \subseteq \mathcal{A},\ \mu|_{\mathcal{A}_0} = \mu_0$ und folgende Eigenschaft gilt: Sei $\mu^* : \mathcal{A}^* \to [0, \infty]$ eine vollständige Fortsetzung von μ_0. Dann gelten $\mathcal{A} \subseteq \mathcal{A}^*$ und $\mu^*|_{\mathcal{A}} = \mu$.
Definiton: In der Situation von 1.8) setze $d_{\mu^*}(A, B) := \mu^*(A\Delta B)$ für alle $A, B \subseteq \Omega$, wobei $A\Delta B := (A\backslash B) \cup (B\backslash A)$ die **symmetrische Differenz von A und B** ist.

1.11 Satz

Seien Ω eine Menge, $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ ein Ring und μ ein σ-finiter und σ-additiver Inhalt auf $\mathcal{A}$. Dann gelten:

a) $\mu^*|_{\overline{\delta}(\mathcal{A})}$ ist die Vervollständigung von $\mu^*|_{\delta(\mathcal{A})}$

b) $\overline{\delta}(\mathcal{A}) := \{B \in \mathcal{P}(\Omega) \mid \forall \epsilon \geq 0\ \exists A \in \mathcal{A} \text{ mit } d_{\mu^*}(A, B) \leq \epsilon\}$ (ohne Beweis)

1.12 Bemerkung

a) Die obigen Sätze besagen, dass ein σ-finiter und σ -additiver Inhalt auf einem Ring eindeutig ein Maß auf der erzeugten σ-Algebra definiert. Das zum elementargeometrischen Inhalt (Beispiel 1.5.c) gehörige Maß heißt **Lebesgue-Maß** λ, die σ-Algebra ist die **Borel-Algebra** $\mathcal{B}(\mathbb{R})$. Die zur Vervollständigung gehörende σ-Algebra heißt **das System der Lebesgue-messbaren Teilmengen auf** $\mathbb{R}$.

b) Es gilt zB.: $\overline{\sigma}(\mathcal{A}) = \{A\Delta N \mid A \in \sigma(\mathcal{A}), N \subseteq \tilde{N} \in \sigma(\mathcal{A}) \text{ mit } \mu(\tilde{N}) = 0\}$. Für das Lebesgue-Maß gilt: Die Kardinalität von $\mathcal{B}(\mathbb{R})$ ist die von $\mathbb{R}$, aber die Kardinalität der Lebesgue-messbaren Mengen ist (mindestens) $\lambda^{|\mathbb{R}|}$ und damit echt größer, denn:

es genügt, eine Lebesgue-messbare Teilmenge $C \subseteq \mathbb{R}$ zu finden, die überabzählbar ist (also $|\text{C}| = |Real|$) und für die $lambda = 0$ gilt. (Dann folgt $\mathcal{P}(C) \subseteq$ Lebesgue-messbar). Die Cantormenge C hat diese Eigenschaft.

1.13 Definition

Seien $(\Omega, \mathcal{A})$ und $(S, \mathcal{S})$ Maßräume. Für eine Abbildung $X : \Omega \to S$ definiere für $B \subseteq S$:

$$X^{-1}(B) := \{X \in B\} := \{\omega \in \Omega \mid X(\omega) \in B\}$$

Dann heißt X $\mathcal{A} - \mathcal{S}$-messbar, falls für alle $B \in \mathcal{S}$ gilt $X^{-1}(B) \in \mathcal{A}$, d.h. es gilt $X^{-1}(\mathcal{S}) \in \mathcal{A}$.

1.14 Bemerkung

a) Jede konstante Abbildung is messbar bzgl. jeder σ-Algebra $\mathcal{A}$ (denn $X^{-1}(B) \in \{\emptyset, \Omega\} \in \mathcal{A}$).

b) Sind $X : (\Omega, \mathcal{A}) \to (S_1, \mathcal{S}_1)$ und $Y : (S_1, \mathcal{S}_1) \to (S_2, \mathcal{S}_2)$ messbar, so auch $Y \circ X(\Omega, \mathcal{A}) \to (S_2, \mathcal{S}_2)$, denn $(Y \circ X)^{-1}(\mathcal{S}_2) = X^{-1}(Y^{-1}(\mathcal{S}_2)) \overset{\text{Y messbar}}{\overset{\downarrow}{\subseteq}} X^{-1}(\mathcal{S}_1) \overset{\text{X messbar}}{\overset{\downarrow}{\subseteq}} \mathcal{A}$

1.15 Lemma

(Prüfen der Messbarkeit auf einem Erzeugendensystem). Seien $(\Omega, \mathcal{A})$ und $(S, \mathcal{S})$ Maßräume auf $\mathcal{S} = \sigma(\mathcal{E})$ für ein $\mathcal{E} \subseteq \mathcal{P}(\mathcal{S})$ Dann ist eine Abbildung $X : \Omega \to S$ genau dann $\mathcal{A} - \mathcal{S}-$messbar, wenn $X^{-1}(\mathcal{E}) \subseteq \mathcal{A}$ gilt.

Beweis: Kommt später!

Beispiel: Falls in 1.15 $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ oder $(S, \mathcal{S}) = (\overline{\mathbb{R}}, \overline{\mathcal{B}}(\mathbb{R}))$ gilt, so folgt: Eine Abbildung $X : (\Omega, \mathcal{A}) \to (S, \mathcal{S})$ ist genau dann messbar wenn für alle $\alpha \in \mathbb{R}$ gilt $\{X \leq \alpha\} \in \mathcal{A} = X^{-1}((-\infty, \alpha])$. (Beachte $\mathcal{B}(\mathbb{R}) = \sigma(\{(-\infty, \alpha] \mid \alpha \in \mathbb{R}\})$)

2 Zufallsexperiment und Wahrscheinlichkeit

2.1 Definition

Ein **Zufallsexperiment** ist ein W-Raum $(\Omega, \mathcal{A}, P)$, d.h. ein Maßraum mit $P(\Omega) = 1$. Ḿan sagt:

a) Ein $\omega \in \Omega$ heißt **Ergebnis** des Experiments.

b) Ein $A \in \mathcal{A}$ heißt **Ereignis** (die Wahl von A ergibt sich aus dem Experiment).

c) Für jedes $A \in \mathcal{A}$ heißt $P(A)$ **die Wahrscheinlichkeit für das Eintreten des Ereignisses** $A \in \mathcal{A}$. Das Maß $P : \mathcal{A} \to [0, 1]$ heißt die zum Experiment gehörige **Wahrscheinlichkeitsverteilung (W-Verteilung)**.

2.2 Bemerkung

a) Die Forderung, dass $\mathcal{A}$ ein Algebra ist, ist naheliegend. Dabei entspricht $\emptyset$ dem unmmöglichen Ereignis, $A \cap B$ dem gleichzeitigen Eintreten von A und B und A^C der logischen Negation von A. Die Forderung, dass P ein Inhalt ist, ist naheliegend wegen der Häufigkeitsinterpretation von Wahtscheinlichkeit. Die Bedingung, dass $\mathcal{A}$ ein σ-Algebra und P ein Maß sind, ist eine mathematische Idealisierung (vgl. Satz 1.7, der dies als eine Stetigkeitseigenschaft charakterisiert).

b) Warum betrachtet man überhaupt σ-Algebren $\mathcal{A} \subsetneqq \mathcal{P}(\Omega)$? Ein Satz von Ulan besagt, dass es kein W-Maß P auf $\mathcal{P}([0, 1])$ mit $P\{x\} = 0$ für alle $x \in [0, 1]$ gibt. Falls Ω abzählbar ist (insbesondere zB. endlich) ist, wird man immer $\mathcal{A} = \mathcal{P}(\Omega)$ wählen.

2.3 Beispiel

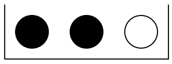
a) **Laplace-Experiment:** Hier ist $|\Omega| < \infty$, d.h. wir betrachten ein Experiment mit nur endlich vielen möglichen Ergebnissen, $\mathcal{A} = \mathcal{P}(\Omega\cdot)$ und $P(A) = \frac{|A|}{|\Omega|}$ für alle $A \in \mathcal{P}(\Omega)$. (Äquivalent: $\forall x \in \Omega : P(\{x\}) = \frac{1}{|\Omega|}$.)

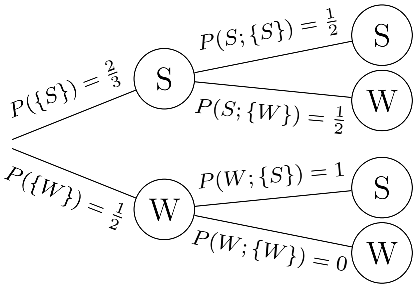
b) **Gleichverteilung auf dem Intervall $[a, b]$ $(-\infty < a < b < \infty)$:** Hier ist $\Omega = [a, b]$, $\mathcal{A} = \mathcal{B}(\Omega)$ und $P(A) = \frac{\lambda(A)}{\lambda(\Omega)}$ für alle $A \in \mathcal{A}$.

c) **Mehrstufige Experimente:** Hier wird ein Zufallsexperiment $(\Omega_1, \mathcal{A}_1, P_1)$ n-fach wiederholt. Man erhält $\Omega = \Omega_1 \times ... \times \Omega_1 = (\Omega_1)^n$, $\mathcal{A} = \mathcal{A}_1 \otimes ... \otimes \mathcal{A}_1$ und für das zugehörige Maß P gilt, falls $|\Omega_1| < \infty$, die Darstellung $P(\{x_1, ..., x_n\}) =$

$P_1(\{x_1\}) \cdot P_1(x_1;\{x_2\}) \cdot \ldots \cdot P_1(x_1,...,x_{n-1};\{x_n\})$ wobei $P(x_1,...,x_k;-)$ die Übergangswahrscheinlichkeit für die k-te Wiederholung ist (Abhängig von $x_1,...,x_k$)).

Beispiel: Ziehen **ohne** Zurücklegen von zwei Kugeln aus einer Urne mit zwei schwarzen und einer weißen Kugel. Hier ist $n = 2$, $\Omega_1 = \{S, W\}$ und $\mathcal{A}_1 = \mathcal{P}(\Omega_1)$. Die Übergangswahrscheinlichkeiten werden durch ein Baumdiagramm beschrieben.





2.4 Definition

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum. Dann heißt P

a) **auf $\boldsymbol{A \in \mathcal{A}}$ konzentriert**, falls $P(A) = 1$ gilt.

b) **diskret**, falls P auf einer abzählbaren Menge konzentriert ist.

c) diskrete W-Räume gestatten folgende Charakterisierungen:

2.5 Satz

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum mit $\{x\} \in \mathcal{A}$ für alle $x \in \Omega$, dann ist die Teilmenge $\Omega \supseteq \Omega_0 := \{\omega \in \Omega \mid P(\{\omega\} > 0\}$ äquivalent:

a) P diskret

b) $P(\Omega_0) = 1$

c) $\int_\Omega f \, dP = \sum_{x\in\Omega_0} f(x) \cdot P(\{x\})$ für alle beschränkten und $\mathcal{A}$-messbaren Abbildungen $f : \Omega \to \mathbb{R}$

d) Es gilt $P = \sum_{x\in\Omega_0} P(\{x\}) \cdot \delta_x$
(Dabei ist a) als Gleichheit von Abbildungen $\mathcal{A} \to [0,1]$ gemeint, d.h.: $\forall A \in \mathcal{A}$:

$P(A) = \sum_{x \in \Omega_0} P(\{x\}) \cdot \delta_x(A) = \sum_{x \in A \cap \Omega_0} P(\{x\})$).

Beweis: Kommt Später!

Beispiel: Für das Lebesgue-Maß λ auf einem Intervall $[a, b]$ $(-\infty < a < b < \infty)$ gilt $\lambda(\{x\}) = 0$ für alle $x \in \Omega = [a, b]$. Es folgt $\Omega_0 = \emptyset$ und $\lambda(\Omega_0) \neq 1$, also ist $\frac{1}{|b-a|} \cdot \lambda$ ein nicht diskretes W-Ma0 auf $[a, b]$

3 Endliche Wahrscheinlichkeitsräume

In diesem Abschnitt betrachten wir W-Räume $(\Omega, \mathcal{A}, P)$ mit $|\Omega| < \infty$ und $\mathcal{A} = \mathcal{P}(\Omega)$. Dann ist P diskret, aber wir nehmen **nicht** an, dass $P\{x\} = \frac{1}{|\Omega|} \forall x \in \Omega$ (vgl. 2.3.a).

3.1 Definition

a) Für $k, n \in \mathbb{N}$ heißt $\binom{n}{k} := \begin{cases} \frac{n!}{k!(n-k)!} & \text{, falls } k \leq n \\ 0 & \text{, sonst} \end{cases}$ der **Binomialkoeffizient „n über k “**.

b) Für $r > 1$; $k_1, ..., k_r \geq 0$ und $n := k_1 + ... + k_r$ heißt $\binom{n}{k_1, ..., k_r} := \frac{n_1}{(k_1!) \cdot (k_2!) \cdot ... \cdot (k_r!)}$ **Multinomialkoeffizient**.

Beispiel: Für $n, k \geq 0$ gilt $\binom{n}{k} = \binom{n}{k, n-k}$

3.2 Satz

Sei M eine Menge mit $0 \leq |M| =: n \leq \infty$

a) Für $k \geq 0$ besitzt M genau $\binom{n}{k}$ Teilmengen mit k Elementen, d.h. es gilt

$$| \{X \mid X \subseteq M \text{ mit } |X| = k\} | = \binom{|M|}{k} = \binom{n}{k}$$

b) Mit den Bezeichnungen aus 3.1.b ist $(k_1, ..., k_r$ die Anzahl an Möglichkeiten, n paarweise verschiedene Objekte auf r Schachteln der Größen $k_1, ..., k_r$ aufzuteilen.

Beispiel: Es ist $\binom{n}{k}$ nach 3.2.a. die Anzahl der k-elementigen Teilmengen von $\{1, ..., n\}$, wegen $\binom{n}{k} = \binom{n}{k, n-k}$ nach 3.2.b aber auch die Anzahl der Möglichkeiten, die Objekte $\{1, ..., n\}$ auf 2 Schachteln der Größe k und $(n-k)$ zu verteilen.

Beweis: Kommt später!

3.3 Satz

(Urnenmodell) Aus einer Urne mit n paarweise verschiedenen Kugeln werden k Kugeln nacheinander gezogen. Dann gibt es folgende Anzahl möglicher Ereignisse:

	mit zurücklegen	ohne zurücklegen
unter Beachtung der Reihenfolge	n^k	$n \cdot (n-1) \cdot ... \cdot (n-k+1)$
ohne Beachtung der Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Beweis: Kommt später!

3.4 Satz

Eine urne beinhalte n (paarweise verschiedene) Kugeln, von denen m weiß und $n-m$ schwarz seien. Man zieht l Kugeln, und $X \in \{0, ..., l\}$ sei die Anzahl der gezogenen weißen Kugeln. Bei Annahme eines Laplace Experimentes erhält man folgende Wahrscheinlichkeiten für die Werte von X.

a) **Hypergeomerische Verteilung:** Beim Ziehen ohne Zurücklegen ist die Wahrscheinlichkeit für $X = k$ gegeben durch $H(l, m, n)\{k\} := \frac{\binom{n}{k}\binom{n-m}{l-k}}{\binom{n}{l}}$ für $0 \leq k \leq \min\{m, l\}$

b) **Bionomialverteilung:** Bei Ziehen mit Zurücklegen ist die Wahrscheinlichkeit für $X = k$ gegeben durch $B(l, p)\{k\} := \binom{l}{k} p^k (1-P)^{l-k}$ für $0 \leq k \leq l$ und $p := \frac{m}{n}$

Beweis: Kommt später!

3.5 Definiton

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum (nicht notwendig endlich). Für $A, B \in \mathcal{A}$ mit $P(B) \geq 0$ heißt $P(A|B) = \frac{P(A \cap B)}{P(B)}$ die **bedingte Wahrscheinlichkeit von A unter der Bedingung B**.

3.6 Bemerkung

a) In der Situation von 3.5 ist $(B, \mathcal{A} \cap B, P(|B|))$ ein W-Raum.

b) Sei $\Omega = \sum_{i \in \mathbb{N},\ P(B_i)>0} P(B_i)\ P(A|B_i)$ (Berechnung einer Wahrscheinlichkeit $P(A)$ durch die Fallunterscheidund $\Omega = \dot\bigcup_{i \in \mathbb{N}} B_i$). Fals $A \in \mathcal{A}$ mit $P(A) \geq 0$ gegeben ist gilt:

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{j \in \mathbb{N},\ P(B_j) \geq 0} P(B_j) \cdot P(A|B_j)} \qquad \textbf{(Formel von Bayes)}$$

Beweis: Kommt später!

3.7 Proposition

(Einschluss-Ausschluss-Prinzip) Es seien $(\Omega, \mathcal{A}, P)$ ein W-Raum (nicht unbedingt endlich), $n \geq 1$ und $A_1, ..., A_n \in \mathcal{A}$. Für eine Teilmenge $J \subseteq I := \{1, ..., n\}$ setze $B_J := \bigcap_{j \in J} A_j \cap (\bigcap_{j \in I \setminus J} A_j^C) \in \mathcal{A}$. Dann gelten:

a) Für jede Teilmenge $J \in I$ gilt $P(\bigcap_{k \in K} A_k) = \sum_{K \subseteq J \subseteq K} P(B_J)$

b) Für jede Teilmenge $J \in I$ gilt $P(B_J) = \sum_{J \subseteq K \subseteq J} (-1)^{|K \setminus J|}\ P(\bigcap_{k \in K} A_k)$.
Inbesondere gilt:

c) $P(\bigcup_{i \in J} A_i) = 1 - P(B_\emptyset) = \sum_{\emptyset \neq k \subseteq J} (-1)^{|K|+1}\ P(\bigcap_{k \in K} A_k)$ also zB.
$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) \qquad (n = 2)$
$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cup A_2) - P(A_1 \cup A_3) - P(A_2 \cup A_3) + P(A_1 \cup A_2 \cup A_3) \qquad (n = 3)$

Beweis: Kommt später!

3.8 Beispiel

(Fixpunktfreie Permutationen) Zwei identische Kartenspiele von je 52 Karten werden gemischt und nebeneinander gelegt. Dann werden 52 mal die beiden jeweils obersten karten aufgedeckt. Wie hoch ist die Wahrscheinlichkeit wenigstens einmal dieselbe Karte aufzudecken?
Lösung: Seien $N := 52,\ I := \{1, ..., N\},\ \Omega := \sum(I) := \{\varphi : I \to I \mid \varphi \text{ bijektiv}\}$ also $|\Omega| = N!$ und $(\Omega, \mathcal{P}(\Omega), P(\omega) = \frac{1}{|\Omega|} \forall \omega \in \Omega)$ der zugehörige Laplace W-Raum (vgl. 23.a). Numeriert man die Karten des ersten Stapels als $1, ..., N$ durch, so ist die Reihenfolge des zweiten Stapels durch genau eine Permutation $\omega \in \Omega$ festgelegt. Da das Mischen fair sein soll, nehmen wir $P(\omega) = \frac{1}{|\Omega|} \forall \omega \in \Omega$ an. Das fragliche Ereignis ist nun

$$A := \{\omega \in \Omega \mid \exists 1 \leq i \leq N : \omega(i) = i\} \subseteq \mathcal{P}(\Omega)$$

Offenbar gilt

a) $A = \bigcup_{i \in I} A_i$

b) $P(\bigcap_{k \in K} A_k) = \frac{(N-|K|)!}{N!}$ für jede Teilmenge $K \subseteq I$

Nun folgt

$$P(A) \overset{(a)}{=} \sum_{\emptyset \neq K \subseteq I} (-1)^{|K|+1} \cdot P(\bigcap_{k \in K} A_k) \overset{(b)}{=} \sum_{\emptyset \neq K \subseteq I} (-1)^{|K|+1} \cdot \frac{(N-|K|)!}{N!}$$

(Summanden hängen von K nur über $|K|$ ab.)

$$\begin{aligned}
&\sum_{k=1}^{N} (-1)^{|K|+1} \cdot \frac{(N-|K|)!}{N!} \cdot |\ \{k \mid \emptyset \neq K \subseteq I : |K| = k\}\ | = \\
=&\sum_{k=1}^{N} (-1)^{k+1} \cdot \frac{(N-k)}{N!} \cdot \binom{N}{k} = \\
=&\sum_{k=1}^{N} (-1)^{k+1} \cdot \frac{1}{k!} = \\
=&1 - e^{-1} - \underbrace{\sum_{k \geq N+1} (-1)^{k+1} \cdot \frac{1}{k!}}_{R(N)}
\end{aligned}$$

Für $N = 52$ ist $|R(N)| \leq \frac{1}{N!} \leq 10^{-60}$, also $P(A) \approx 10^{-60}$, also $P(A) \approx 1 - e^{-1} \approx 0.63$

Tabelle für kleines N:

N	1	2	3	4
$P(A)$	1	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{8}$

Fixpunkte für $\omega \in \Omega = \sum(\{1,2,3\})$:

	1	1	2	2	3	3
	2	3	1	4	1	1
	3	2	3	1	2	1
Fixpunkte	3	1	1	0	0	1

$$\rightsquigarrow \qquad \frac{4}{6} = \frac{2}{3} \approx 0.66$$

4 Zufallsvariablen, Erwartungswert und Varianz

4.1 Definition

a) Seien Ω eine Menge und $(X_i)_{i \in I}$ eine Familie von Abbildungen $X_i : \Omega \to S_i$, wobei $(S_i, \mathcal{S}_i)$ ein Messraum ist. Dann heißt $\sigma((X_i)_{i\in I} := \sigma(\bigcup_{i \in I} X_i^{-1}(\mathcal{S}_i)) \subseteq \mathcal{P}(\Omega)$ **die von $(X_i)_{i\in I}$ erzeugte σ-Algebra.**

b) Seien $(\Omega, \mathcal{A})$ und $(S, \mathcal{S})$ W-Räume. Dann heißt eine Abbildung $X : \Omega \to S$ eine **Zufallsfunktion**, falls $X : \mathcal{A} \to \mathcal{S}$ messbar ist. D.h. wenn für alle $S \in \mathcal{S}$ gilt: $X^{-1}(\mathcal{S}) \in \mathcal{A}$. Das Bildmaß $P \circ X^{-1} : \mathcal{S} \to [0,1], B \to P(X^{-1}(B)) =: P(\{X \in B\})$ heißt die **Verteilung oder W-Verteilung von X.**

c) Nun seien speziell $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ messbar. Dann heißt X **Zufallsvariable (ZV)** und die Abblildung $F_X : \mathbb{R} \to [0,1], F_X(t) := P(X^{-1}(-\infty, t]) = P\{(X \leq t\}$ heißt die **Verteilungsfunktion von X.**

Die ZV X heißt **diskret verteilt**, falls $P \circ X^{-1}$ diskret ist und **stetig verteilt**, falls $P \circ X^{-1}$ absolutstetig bzgl. des Lebesgue-Maßes ist. D.h. eine Dichte f_X exisitiert mit $P \circ X^{-1} = f_X(t)\, d\lambda$ (Gleichheit von Maßen auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$).

4.2 Bemerkung

a) F_X ist monoton wachsend, rechtsstetig (r.c.), d.h. es gilt $\lim_{t\to-\infty} F_X(t+h) = F_X(t)\; \forall\, t \in \mathbb{R}$, und es gelten $\lim_{t\to-\infty} F_X(t) = 0$ und $\lim_{t\to+\infty} F_X(t) = 1$. Außerdem gilt für alle $t \in \mathbb{R}$:

$$F_X(t) - \lim_{h>0, h\to 0} F_X(t-h) = (P \circ X^{-1}(\{t\})$$

, also sind die Sprungstellen von F_X genau diejenigen $t \in \mathbb{R}$ mit $P(X = t) \neq 0$.

b) X ist genau dann stetig verteilt, wenn F_X absolut stetig ist, d.h. wenn $F_X(t) = \int_{-\infty}^{t} F_X(s)\,ds$ mit einer messbaren Funktion $f_X : \mathbb{R} \to \mathbb{R}_f$ und $\int_{\mathbb{R}} f_X(s)ds = 1$ gilt(zB. wenn F_X stetig differenzierbar ist).

4.3 Satz

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum.

a) Sind X und Y Zufallsvariablen auf Ω, so auch $\max\{X, y\}$, $\min\{X, y\}$, $|X|$ für $r > 0$ und X^r für $r \in \mathbb{N}$

b) Ist $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf Ω, so sind $\inf_n X_n$, $\sup_n X_n$, $\liminf_n X_n$ und $\limsup_n X_n$ Zufallsfunktionen von Ω nach $(\overline{\mathbb{N}}, \overline{\mathbb{R}}$.

Beweis: Kommt später!

4.4 Satz

(Fouriertransformation) Seien $\mathcal{M}$ die Menge aller W-Maße $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und $\mathcal{F}$ die Menge aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$, welche monoton wachsen, r.c. sind und gegen 0 (bzw. 1) konvergieren für $t \to -\infty$ (bzw. $+\infty$). Dann ist die Abbildung

$$\mathcal{M} \to \mathcal{F},\ \mu \to F_\mu \text{ mit } F_\mu(t) := \mu((-1, t])$$

eine Bijektion.

4.5 Definition

Zwei ZV $X, Y :\to (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ heißen **stochastisch äquivalent** (geschrieben $X \sim Y$), falls $P \circ X^{-1} = P \circ Y^{-1}$.

4.6 Beispiel

Für $X \sim Y$ folgt nicht $X = Y$, und noch nicht mal $\mathrm{P(X = Y)} = 1$, wie wie folgendes Beispiel zeigt: Es sei $(\Omega, \mathcal{A}, P) := (\{0,1\}, \mathcal{P}(\{0,1\}), \frac{1}{2}(\delta_0 + \delta_1))$. Setzt man $X(t) := t$ und $Y(t) := 1 - t$ für $t \in \Omega$, so ist $P \circ X^{-1} = P \circ Y^{-1} = 1/2(\delta_0 + \delta_1)$, also gilt $X(t) \neq Y(t) \forall t \in \Omega$ und damit $P\{X = Y\} = 0$.

4.7 Satz

(Erinnerungen an das Lebesgue-Integral) Seien X und Y Zufallsvariablen auf dem W-Raum $(\Omega, \mathcal{A}, P)$.

a) X ist genau dann integrierbar, wenn $\int_\Omega |X| dP < \infty$

b) (Linearität) Seien X, Y integrierbar auf $a, b \in \mathbb{R}$. Dann gilt $\int_\Omega (aX + bY) dP = a \cdot \int_\Omega X dP + b \cdot \int_\Omega Y dP$

c) Seien X integrierbar, $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkt und $A := \bigcup_{n \geq 1} A_n$. Dann ist $\int_A X dP = \sum_{n \geq 1} X dP$

d) (Positivität) Aus $X \geq 0$ **P-fast sicher** (P-f.s., d.h. $P\{\omega \in \Omega \mid X(\omega) \leq 0\}$) folgt $\int_\Omega X dP \geq 0$

e) (Monotonie) Seien X, Y integrierbar mit $X \leq Y P - f.s.$. Dann gilt $\int_\Omega X dP \leq \int_\Omega Y dP$

f) Seien X integrierbar und $a, b \in \mathbb{R}$ ist $a \leq X \leq b$ P-f.s. auf einer Menge $A \in \mathcal{A}$. Dann gilt $a \cdot P(A) \leq \int_A X dP \leq b \cdot P(A)$

g) (Dreiecksungleichung) Für X integrierbar gilt $|\int_\Omega X dP| \leq \int_\Omega |X| dP$

h) (Satz von majorisierter Konvergenz) Seien X_n, X und Y Zufallsvariablen mit

   i) $|X_n| \leq Y$ P-f.s. $\forall n \in \mathbb{N}$

   ii) $\int_\Omega Y dP < \infty$

   iii) $X_n \to X$ für $n \to \infty$ P-f.s.

   Dann gilt: $\lim_{n \to \infty} \int_\Omega X_n dP = \int_\Omega X dP$

i) (gliedweise Integration) Seien X_n Zufallsvariablen mit $\sum_n (\int_\Omega |X_n| dP) \leq \infty$. Dann gilt $\int_\Omega (\sum_n |X_n|) dp \leq \infty$ P-f.s. und es gilt $\int_\Omega (\sum_n X_n) dP = \sum_n (\int_\Omega X_n dp)$

j) (Satz von majorisierbarer Konvergenz) Seien $X_n \geq 0$ Zufallsvariablen und X eine Zufallsvariable mit $X_n \nearrow X$ für $n \to \infty$ P-f.s.. Dann gilt $\lim_{n \to \infty} \int_\Omega X_n dP = \int_\Omega X dP$ (wobei beide Seiten $+\infty$ sein können)

k) (Lemma von Fatou) Seien $X_n \geq 0$ Zufallsvariablen. Dann ist

$$\int_\Omega \liminf_{n\to\infty} X_n dP \leq \liminf_{n\to\infty} \int_\Omega X_n dP$$

4.8 Definition

Sei X eine Zufallsvariable auf dem W-Raum $(\Omega, \mathcal{A}, P)$

a) Falls $X \in L^1(P) := \{X| \int_\Omega |X|dP \leq \infty\}$ gilt, so heißt $\mathrm{E}X := \int_\Omega X dP$ der **Erwartungswert von X** (auch sinnvoll falls X mit möglichem Wert $+\infty$).

b) Falls $X \in L^2(P) := \{X| \int_\Omega |X^2|dP < \infty\}$ (und damit auch $X \in L^1(P)$) gilt, so heißt $\mathrm{Var}X := \mathrm{E}[(X - \mathrm{E}X)^2] = \int_\Omega (X - \mathrm{E}X)^2 dP \geq 0$ die **Varianz von X**. Die Zahl $\sigma X := +\sqrt{\mathrm{Var}(X)}$ heißt **die Streuung von X**.

c) Für $n \geq 1$ heißt $\mu_n := \mathrm{E}(X^n)$ **das n-te Moment von X** (falls es existiert).

Bemerkung:

a) Die Eigenschaften des Integrals übertragen sich auf den Erwartungswert, zB. existiert $\mathrm{E}X$ genau dann, wen $\mathrm{E}X \leq \infty$ gilt.

b) Es gilt $\mathrm{E}1 = 1$

c) Nach dem Transformationslemma gilt $X \in L^1(P)$ genau dann, wenn $id_\mathbb{R} \in L^1(P \circ X^{-1})$. Die Existenz und der Wert von $\mathrm{E}X$ und $\mathrm{Var}X$ hängen nur von $P \circ X^{-1}$ ab, nämlich:

$$\mathrm{E}X = \int_\Omega X dP = \int_\mathbb{R} (P \circ X^{-1}) dt \text{ für } X \in L^1(P)$$

d) Allgemein in c) gilt für $g : \mathbb{R} \to \mathbb{R}$ mit $g \circ X \in L^1(P)$ (zB. $g : \mathbb{R} \to \mathbb{C}$)

$$\int_\Omega (g \circ X) dP = \int_\mathbb{R} g \; d(P \circ X^{-1}) = \begin{cases} \sum\limits_{x\in\mathbb{R}} g(x) \cdot P\{X = x\} & \text{, falls } X \text{ diskret} \\ \int_\mathbb{R} g f_x d\lambda & \text{, falls } X \text{ stetig} \end{cases}$$

$$\mathrm{E}X = \begin{cases} \sum\limits_{x\in\mathbb{R}} x \cdot P\{X = x\} & \text{, falls } X \text{ diskret} \\ \int_\mathbb{R} f_x d\lambda & \text{, falls } X \text{ stetig verteilt ist mit Dichte } f_x \end{cases}$$

4.9 Satz

Sei X eine Zufallsvariable auf dem W-Raum $(\Omega, \mathcal{A}, P)$ und $f : \mathbb{R} \to \mathbb{C}$ **stetig** mit $f \circ X \subseteq L^1(P)$. Dann gilt $\int_\Omega (f \circ X) dP = \int_{-\infty}^{\infty} f(t)\; dF_X(t)$, wobei auf der rechten Seite das Riemann-Integral steht (und F_X die Dichtefunktion von X ist.)

Beweis: Kommt später!

4.10 Satz

(Diskrete Approximation des Erwartungswerts) Sei X eine Zufallsvariable auf dem W-Raum$(\Omega, \mathcal{A}, P)$. Dann gilt $\mathrm{E}|X| \leq 1 + \sum_{n \geq 1} P\{|X| \geq n\}$.

Beweis: Kommt später!

4.11 Lemma

Seien $(\Omega, \mathcal{A}, P)$ ein W-Raum und $X : \Omega \to \mathbb{R}$ eine ZV.

a) Für $X \in L^2(P)$ ist $\mathrm{Var}X = E(X^2) - (EX)^2$

b) Für $X \in L^2(P)$ und $\alpha, \beta \in \mathbb{R}$ gilt $\mathrm{Var}(\alpha X + \beta) = \alpha^2 \cdot \mathrm{Var}(X)$

Beweis: Kommt später!

4.12 Satz

(Ungleichung von Chebychev) Seien $(\Omega, \mathcal{A}, P)$ ein W-Raum, $X : \Omega \to \mathbb{R}$ eine ZV, $1 \leq p < \infty$ und es gelte $X \in L^p(P)$ (d.h. $||X||_{L^p}^p := \int_\Omega |X|^p dP < \infty$). Dann gilt für alle $c > 0$

$$P\{|X| \geq c\} \leq \frac{1}{c^p} \cdot ||X||_{L^p}^p$$

Insbesondere gilt für $X \in L^2(P)$

$$P\{|X - \mathrm{E}X| \geq c\} \leq \frac{1}{c^2} \cdot \mathrm{Var}(X)$$

(daher der Begriff Varianz).

Beweis: Kommt später!

4.13 Definition

Seien $(\Omega, \mathcal{A}, P)$ ein W-Raum und $X, Y : \Omega \to \mathbb{R}$ ZV. Es gelte $X, Y, XY \in L^1(P)$ (zB. $X, Y \in L^2(P)$, vgl. Hölderungleichung). Dann heißt

$$\mathrm{Cov}(X, Y) := \mathrm{E}[(X - \mathrm{E}X)(Y - \mathrm{E}Y)]$$

die Kovarianz von X und Y. Die ZV X, Y heißen **unkorreliert**, falls $\mathrm{Cov}(X, Y) = 0$ gilt. Gilt zusätzlich $\mathrm{Var}X, \mathrm{Var}Y > 0$, so heißt

$$\rho(X, Y) := \frac{\mathrm{Cov}(X, Y)}{(\mathrm{Var}X)^{\frac{1}{2}}(\mathrm{Var}Y)^{\frac{1}{2}}}$$

der **Korrelationskoeffizient von X and Y**.

4.14 Satz

Seien $(\Omega, \mathcal{A}, P)$ ein W-Raum und $X_1, ..., X_n : \Omega \to \mathbb{R}$ ZV mit $X_1, ..., X_n \in L^2(P)$. Dann gilt

$$\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}X_i + \sum_{i,j=1;\ i \neq j}^n \mathrm{Cov}(X_i, X_j)$$

Falls $X_1, ...X_n$ paarweise unkorreliert sind, so folgt

$$\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}X_i \qquad \text{(Formel von Bienayme)}$$

Beweis: Kommt später!

5 Beispiele wichtiger Wahrscheinlichkeitsverteilungen

Ist $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine ZV, so heißt das Maß $P \circ X^{-1}$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ die Verteilung von X. Viele Fragen hängen vonn X nur über $P \circ X^{-1}$ (und zB. nicht von $(\Omega, \mathcal{A}, P)$) ab. Deswegen wird im Folgenden nur das Maß $P \circ X^{-1}$ angegeben (obwohl weder P noch X definiert sind).

5.1 Beispiel

a) **Die Gleichverteilung auf $[\boldsymbol{a},\boldsymbol{b}]$** mit $-\infty < a < b < \infty$ ist

$$P \circ X^{-1} := \frac{1}{|b-a|} \cdot 1_{[a,b]} \cdot \lambda$$

b) **Die Binomialverteilung zu den Parametern $\boldsymbol{n} \in \mathbb{N}_{\geq 1}$ und $\boldsymbol{p} \in [0,1]$** ist definiert als

$$P \circ X^{-1} := B(n;p) := \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$$

Es gilt $\mathrm{E}(B(n;p)) = np$ und $\mathrm{Var}(B(n;p)) = np(1-p)$.

c) f **Die Poissonverteilung zum Parameter $\boldsymbol{\lambda} > 0$** ist definiert als

$$P \circ X := \Pi_\lambda|_{\mathcal{B}(\mathbb{R})}$$

wobei $\Pi_\lambda : \mathcal{P}(\mathbb{R}) \to [0,1]$ gegeben ist als

$$\Pi_\lambda := \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \cdot \delta_k$$

Es ist X diskret verteilt und auf $\mathbb{N}_0$ konzentriert. Es gilt $\mathrm{E}X = \mathrm{Var}X = \lambda$. Für $\mathrm{E}X$ sieht man das folgendermaßen:

$$\begin{aligned}\mathrm{E}X &= \int_{\mathbb{R}} id\ d(P \circ X^{-1}) = \sum_{k=0}^{\infty} k \cdot P\{X = k\} = \sum_{K=0}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \\ &= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k-1} = e^{-\lambda} \cdot \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda\end{aligned}$$

d) **Die Exponentialverteilung zum Parameter $\boldsymbol{\lambda} > 0$ ist definiert als**

$$P \circ X^{-1} = f_\lambda(t) := 1_{\mathbb{R}^+}(t) \cdot \lambda e^{-\lambda t}$$

(d.h. X ist stetig verteilt). Für die Verteilungsfunktion erhält man

$$F_X(t) := \int_{-a}^{t} f_x(s) ds = \begin{cases} 0, & t < 0 \\ 1 - e^{-\lambda t}, & t \geq 0 \end{cases}$$

Es gelten $\mathrm{E}X = \frac{1}{\lambda}$ und $\mathrm{Var}X = \frac{1}{\lambda^2}$.

5.2 Definition

5.3 Definition

5.4 Definition

5.5 Definition

5.6 Definition

5.7 Definition

6 Kapitel

6.1 Definition

6.2 Definition

6.3 Definition

6.4 Definition

6.5 Definition

6.6 Definition

7 Kapitel

7.1 Definition

7.2 Definition

7.3 Definition

7.4 Definition

7.5 Definition

7.6 Definition

7.7 Definition

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $X_n : \Omega \to \mathbb{R}$ ZV $(n \geq 1)$ mit $X_n \in L^1(P)$ $\forall n \geq 1$. Dann **genügt $(X_n)_n$ dem Gesetz der großen Zahlen** $:\Longleftrightarrow$ Für $S_n := X_1 + ... + X_n$ gilt $\frac{S_n - \mathrm{E}S_n}{n} \to 0$ stochastisch, d.h. wenn gilt

$$\forall \epsilon > 0, \delta > 0 \ \exists N = N(\epsilon, \delta) : \forall n \geq N : P\left\{\frac{|S_n - \mathrm{E}S_n|}{n} > \epsilon\right\} < \delta$$

Bemerkung: Seien $X_1, X_2, ...$ integrierbare ZV und $\forall i \geq 1$ gelte $\mathrm{E}X_i = \mathrm{E}X_1 =: \mathrm{E}$. Dann gilt

$$\frac{S_n - \mathrm{E}S_n}{n} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mathrm{E}$$

d.h. man betrachtet in der Def. 7.7 die Differenz zwischen dem Mittelwert („rel. Häufigkeit ") und dem Erwartungswert.

7.8 Beispiel

Sei $(X_n)_{n\geq 1}$ eine Folge unabhängiger ZV, die alle Bernoulliverteilt sind mit Parameter $p = \frac{1}{2}$ (zB. wiederholter Münzwurf einer fairen Münze). Wir sehen in 7.9, dass $(X_n)_{n\geq 1}$ dem schwachen Geesetz der großen Zahlen genügt, d.h. es gilt

$$\forall \epsilon > 0, \delta > 0 \ \exists N = N(\epsilon, \delta) : \forall n \geq N : P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{2}\right| > \epsilon\right\} < \delta$$

d.h. die relative Häufigkeit $\frac{1}{n}\sum_{i=1}^{n} X_i$ nähert sich dem Erwartungswert $\frac{1}{2}$ beliebig genau an.

Was ist aber mit den Ereignissen $A_n := \left\{\frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{2}\right\}$, d.h. das **exakte** Übereinstimmen von rel. Häufigkeit und Erwartungswert? Offenbar gilt $P(A_n) \overset{n\to\infty}{\longrightarrow} 0$. Wir betrachten dazu

Satz (Stirling-Formel)

$$\forall n \geq 0 \text{ gilt } n! = \sqrt{2\pi n} \cdot n^n e^{-n+\eta(n)} \text{ für } 0 < \eta(n) < \frac{1}{12n}$$

Insbesondere gilt:

$$\lim_{n\to\infty} \sqrt{2\pi n} \cdot n^n e^{-n} \cdot \frac{1}{n!} = \lim_{n\to\infty} e^{-\eta(n)} = 1 \qquad \text{(ohne Beweis)}$$

Damit berechnen wir für $n \geq 1$

$$P\left(\frac{1}{2n}\sum_{k=1}^{2n} X_k = \frac{1}{2}\right) = B(2n; \frac{1}{2})(\{n\}) = \binom{2n}{n}\left(\frac{1}{2}\right)^n\left(1-\frac{1}{2}\right)^{2n-n} = \frac{(2n)!}{(n!)^2} \cdot 2^{-2n}$$

und mit Stirling-Formel gilt

$$\lim_{n\to\infty} \frac{(2n)!}{(n!)^2} \cdot 2^{-2n} = \lim_{n\to\infty} \frac{\sqrt{2\pi 2n} \cdot (2n)^{2n} \cdot e^{-2n+\eta(2n)}}{2\pi n \cdot n^{2n} e^{-2n+2\eta(n)}} \cdot 2^{-2n} = \lim_{n\to\infty} \frac{1}{\sqrt{\pi n}} = 0$$

7.9 Satz

(Schwaches Gesetz der großen Zahlen) Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum, $X_n : \Omega \to \mathbb{R}$ $(n \geq 1)$ Zufallsvariablen, $(X_n)_{n\geq 1}$ paarweise unkorreliert (zB. paarweise unabhängig), $X_n \in L^2(P)$ $\forall n \geq 1$ und $v := \sup_{n\geq 1} \operatorname{Var}(X_n) < \infty$. Dann gilt

$$\forall \epsilon > 0 : P\left(\left|\frac{1}{2}\sum_{i=1}^n (X_i - \mathrm{E}X_i)\right| \geq \epsilon\right) \leq \frac{v}{n\epsilon^2} \qquad (\star)$$

Insbesondere gilt

$$\frac{1}{2}\sum_{i=1}^n (X_i - \mathrm{E}X_i) \to 0 \text{ stochastisch}$$

d.h. $(X_n)_{n\geq 1}$ genügt dem schwachen Gesetz der großen Zahlen. Gilt zusätzlich

$$\mathrm{E}X_n = \mathrm{E}X_1 \; \forall n \geq 1$$

so gilt

$$\frac{1}{n}\sum_{i=1}^n X_i \to \mathrm{E}X_1 \text{ stochastisch}$$

Beweis: Kommt später!

7.10 Beispiel

In einem Gefäß befinden sich $n = \frac{1}{4} \cdot 10^{23}$ Gasmoleküle, die sich jeweils unabhängig voneinander mit gleicher Wahrscheinlichkeit in der linken oder rechten Seite des Gefäßes befinden. Wie hoch ist die Wahrscheinlichkeit, dass sich in der rechten Kammer mindestens $\frac{1+10^{-8}}{2}$ aller Teilchen befinden. Seien $X_1, ..., X_n$ unabhängige, zum Parameter $p = \frac{1}{2}$ Binomialverteilte ZV ($X_i = 0$ heißt „i-tes Teilchen in der linken Hälfte“).

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq \frac{1+10^8}{2}\right) \overset{\text{Symmetrie}}{\overset{\downarrow}{=}} \frac{1}{2} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{2}\right)\right| \geq \frac{10^{-8}}{2}\right) \overset{(\star)\text{ in 7.9}}{\overset{\downarrow}{\leq}} \frac{1}{2} \cdot \frac{\mathrm{Var} X_1}{n(\frac{1}{2} \cdot 10^{-8})}$$

$$\overset{\mathrm{Var}X_1 = \frac{1}{4}}{\overset{\downarrow}{=}} \frac{1}{2} \cdot \frac{1}{4n \cdot 10^{-16}} \overset{n = \frac{1}{4} \cdot 10^{23}}{\overset{\downarrow}{=}} 2 \cdot 10^{-7}$$

(Man kann sogar zeigen, dass $P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq \frac{1+10^8}{2}\right) < 10^{-50000}$. Tschebychev sehr grobe Abschätzung!)

7.11 Defintion und Satz

Seien $n \geq 1, \mu_1, ..., \mu_n$ endliche Maße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (z.B. W-Maße) und $A_n : \mathbb{R}^n \to \mathbb{R}$ definiert durch $A_n((x_1, ..., x_n)) := x_1 + ... + x_n$. Dann heißt das Maß

$$\mu_1 * ... * \mu_n := (\mu_1 \otimes ... \otimes \mu_n) \circ A_n^{-1} \text{ auf } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

das Faltungsprodukt der Maße $\mu_1, ..., \mu_n$.
Gilt $\mu_i = P \circ X_i^{-1}$ für unabhängige Zufallsvariablen $X_1, ..., X_n$, so folgt

$$P \circ \left(\sum_{i=1}^{n} X_i\right)^{-1} = \mu_1 * ... * \mu_n = (P \circ X_1^{-1}) * ... * (P \circ X_n^{-1})$$

Beweis: Kommt später!