Wahrscheinlichkeitstheorie
M. Schäl

§ 1 Einführung

Eines der Grundziele der Mathematik besteht darin, Erfahrungen des Menschen über Vorgänge der Natur in Modelle, d. h. Axiome, umzusetzen, diese mathematisch zu analysieren, also Eigenschaften und Gesetzmäßigkeiten (Theoreme) aus ihnen abzuleiten und durch Rückinterpretation in die Erfahrungswelt nutzbar zu machen. Dabei zeigt es sich häufig, daß Modifizierungen am mathematischen Modell nötig werden, die wiederum Anlaß zu erneuter Analyse und Rückinterpretation sind, u. s. w. . Das folgende **4 – Phasen – Schema** verdeutlicht diesen Kreislauf.

Erfahrungswelt | **Mathematik**

$$\begin{tikzcd}
\text{Erfahrung} \arrow[r, "\text{Modellierung}", "\text{(1. Phase)}"'] & \text{Axiomensystem} \arrow[d, "\text{Analyse}", "\text{(2. Phase)}"'] \\
\text{Interpretierte Theoreme} \arrow[u, "\text{Anwendung}"', "\text{(4. Phase)}"] & \text{Theoreme} \arrow[l, "\text{Rückinterpretation}"', "\text{(3. Phase)}"]
\end{tikzcd}$$

Ein wichtiges Ziel dieser Vorlesung ist die Entwicklung eines Basismodells zur Beschreibung und mathematischen Analyse des " Zufalls ". Der Begriff "zufällig" wird auch durch die Vokabel "**stochastisch**" ausgedrückt werden. Die Bezeichnung "**Stochastik**" wird als Sammelbegriff für die Gebiete Wahrscheinlichkeitstheorie und Mathematische Statistik verwendet.

Die Wahrscheinlichkeitstheorie (kurz **W–Theorie**) befaßt sich mit der Analyse (Beschreibung und Gesetzmäßigkeiten) von stochastischen Vorgängen. Anstelle des Begriffes " zufälliger Vorgang " führen wir den Terminus technicus " **Zufallsexperiment** " (kurz **Experiment**) ein. Der Begriff Zufallsexperiment steht für eine Situation, die ein vom Zufall beeinflußtes, also nicht vorhersagbares Ergebnis hervorbringt.

Der Übergang von der Wirklichkeit zum Modell ist dabei nie rein logisch begründbar. Er setzt in starkem Maß Erfahrung über die Natur des Experiments voraus. Das ist keine Besonderheit der Modelle für Zufallsexperimente. So genügt z. B. das ebene Modell der Erdoberfläche vollauf, wenn man eine Landkarte des Rhein–Sieg–Kreises herstellen will. Für feine geophysikalische Betrachtungen ist selbst das Modell der Erdkugel zu grob und man betrachtet im feineren Modell Abplattungen. Wir sehen daran gleich, daß die Wahl des Modells von der Zielsetzung mitbestimmt wird.

Wir halten fest, daß es keinen **prinzipiellen** Unterschied zwischen Modellen in der Geometrie und der W–Theorie gibt. **Praktisch** scheint es aber Unterschiede zu geben, weil das passende Modell für Zufallsexperimente oft weniger offensichtlich ist. Häufig läßt sich die Richtigkeit eines Modells nur empirisch prüfen, und das ist stets mit Unsicherheiten behaftet, die mit dem zufälligen Ausgang der Experimente zusammenhängen. So braucht man zur Überprüfung, ob ein Würfel unverfälscht ist, eine sehr große Zahl von Würfen und mit **absoluter** Sicherheit läßt sich auch dann noch nicht die Unverfälschtheit belegen. Man täuscht sich leicht darüber, ob ein Versuch wirklich hinreichend viele Symmetrien enthält, um die Annahme zu rechtfertigen, alle Versuchsausgänge seien gleichwahrscheinlich.

Zeitweise haben Mathematiker sogar geglaubt, es läge im Wesen der zufälligen Erscheinungen, daß sie sich nicht mathematisieren ließen. Jedenfalls hat es – im Gegensatz z. B. zur Geometrie – bis in dieses Jahrhundert hinein gedauert, bis man eine gesicherte axiomatische Grundlegung gegeben hat. Andererseits macht gerade diese Tatsache, daß man über Zufallsereignisse mathematisch rigorose Resultate beweisen kann, einen Reiz dieses Gebiets aus.

Hilbert stellte 1900 eine Liste der 23 wichtigsten damals ungelösten mathematischen Probleme auf. Das 6. Problem fragte gerade nach einer Begründung der Wahrscheinlichkeitstheorie, die den neuen, strengen Ansprüchen der Mathematik gerecht wird.

Der Beginn der Wahrscheinlichkeitsrechnung liegt im 17. Jahrhundert. (Dabei wird der Begriff W–Rechnung anstelle von W–Theorie verwendet, wenn die Theorie ohne Maßtheorie entwickelt wird.)

Die heute allgemein akzeptierte axiomatische Einführung der Wahrscheinlichkeit als normiertes Maß auf einem meßbaren Raum wurde im Jahre 1933 durch A. N. Kolmogoroffs "Grundbegriffe der Wahrscheinlichkeitsrechnung " (Ergebnisse der Mathematik Bd. 2, Heft 3) gegeben. Damit wurde die Wahrscheinlichkeitstheorie formal ein Teilgebiet einer anderen, weit entwickelten mathematischen Disziplin, nämlich der Maßtheorie. In der Wahrscheinlichkeitstheorie spielen jedoch viele in der allgemeinen Maßtheorie unmotivierte Begriffe eine bedeutende Rolle.

Die W–Theorie ist besonders wertvoll zur Beschreibung von zufälligen Experimenten, die oft und unter gleichen Bedingungen und im Idealfall ohne gegenseitige Beeinflussung wiederholt werden. Dazu zählen etwa die Notierung von Lebenszeiten oder Schadensfällen in einem Versicherungs–unternehemen, die Beobachtung des Heilungserfolges nach Verabreichung eines medizinischen Medikaments oder die Notierung der Nachfrage nach einer bestimmten Ware an verschiedenen Tage in einem Wirtschaftssunternehmen. Ein ideales Beispiel ist das folgende Experiment. Wir drehen ein Glücksrad.

1
|
/ Zeiger

3/4 – – 1/4

|
1/2

Das Ergebnis ist die Zahl $x \in (0,1]$, auf die der feste Zeiger nach Stillstand des Rades zeigt.
Dreht man das Glücksrad jetzt n mal (also bei n–maliger Wiederholung des Experiments), so erhalten wir als Ergebnis des Versuchs (Versuchsausgang) n Zahlen $(x_1,...,x_n) \in (0,1]^n$. Die Folge $x_1,...,x_n$ erscheint zunächst regellos. Unser Ziel besteht nun darin, dennoch Regeln und Gesetzmäßigkeiten zu erkennen.
Zu gegebenen $(x_1,...,x_n)$ und $A \subset (0,1]$ betrachtet man die **relative Häufigkeit**

(1.1) $\text{rel. } H[A] := \frac{1}{n} \sum_{m=1}^{n} \mathbf{1}_A \, (x_m)$,

(1.2) wobei $\mathbf{1}_A$ die Indikatorfunktion von A ist, definiert durch $\mathbf{1}_A \, (x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$.

Also ist rel. $H[A]$ ein Maß dafür, wie häufig bei oftmaliger Wiederholung ein Ergebnis in A liegt. Wir werden uns bei der Definition eines **Wahrscheinlichkeitsmaßes** (kurz **W–Maß**) $P[\cdot]$ an der Definition und den Eigenschaften der " relativen Häufigkeit " rel. $H[\cdot]$ orientieren.

Die Bedeutung der W–Theorie gründet sich auf eine wesentliche Erfahrungstatsache (**Stabilität relativer Häufigkeiten**, **empirisches Gesetz der großen Zahlen**): rel $H[A]$ führt bei verschiedenen langen Versuchsreihen zu annähernd gleichen Werten. Dies führt dazu, daß man A eine Wahrscheinlichkeit $P[A] \in [0,1]$ zuordnen möchte, die unabhängig von der speziellen Versuchsreihe ist, hier verbunden mit der Prognose: In $P[A] \cdot 100\%$ aller Fälle liegt x_m in A.

Dabei kann $P[A]$ empirisch ermittelt werden und/oder durch spezielle Kenntnis der Versuchsanordnung erraten werden. Das letztere gilt für das Glücksrad.

(1.3) Für ein Intervall $I \subset (0,1]$ sei $a{+}I := \{ x \in (0,1] \,;\, x - a \in I \bmod 1 \}$.

a+I entsteht aus I durch Verschiebung der Skala um a.
Ist das Glücksrad völlig symmetrisch und ist der Antrieb so, daß kein Einfluß der Ausgangslage auf das Ergebnis festgestellt werden kann, so ist auch kein Einfluß einer Verschiebung der Skala zu vermuten, also wird folgendes eintreten:

(1.4) $\text{rel. } H[a{+}I] \approx \text{rel. } H[I]$ für große n .

Die beobachtbare Unabhängigkeit der rel. H. von n bei großem n motiviert einen Übergang zu idealen Werten $P[A]$, für die (1.4) exakt gilt, also

(1.5) $P[a{+}I] = P[I]$ für alle Intervalle $I \subset (0,1]$

Es gibt weitere Eigenschaften, die $P[\cdot]$ von rel. $H.[\cdot]$ erben sollte (wobei hier Ω für den vorliegenden Raum $(0,1]$ geschrieben wird):

(1.6) (1) $P[A] \geq 0$ (**Positivität**)

(2) $P[\Omega] = 1$ (**Normiertheit**)

(3′) $P[A_1 \cup A_2] = P[A_1] + P[A_2]$, falls A_1, A_2 disjunkt (**endliche Additivität**)

(3) $P[\bigcup_{m=1}^{\infty} A_m] = \sum_{m=1}^{\infty} P[A_m]$, falls (A_m) paarw. disjunkt (**σ–Additivität**).

Die Eigenschaft (3) folgt für $P[\cdot] := \text{rel. } H.[\cdot]$ aus (3'), denn es gibt nur endlich viele k mit $\bigcup_{m=1}^{n} \{x_m\} \cap A_k \neq \emptyset$, also mit rel. $H[A_k] > 0$. Für allgemeine Mengenfunktionen P folgt (3) nicht aus (3'), ist aber sehr fruchtbar für die mathematische Theorie.

<u>**Definition**</u> Eine Mengenfunktion P mit (1.6) heißt **W–Maß**.

Die Forderungen (1.5) und (1.6) erfüllt allein das Lebesgue–Maß λ^1 auf (0,1] (vgl. Barner/Flohr II (83) S. 198, Bauer (78) S. 45). $\lambda^1[A]$ ist für alle Borel–Mengen A in $(0,1]$, sogar für alle λ^1–mb Mengen A in (0,1] erklärt. Nach Vitali (1909) existiert keine Mengenfunktion P mit (1.5) und (1.6), die für alle Teilmengen von $(0,1]$, also auf der Potenzmenge $\mathfrak{P}((0,1])$ definiert ist. Also muß der Definitionsbereich von P spezifiziert werden. Dies führt für das Glücksradbeispiel zu folgendem W–Modell $(\Omega_1, \mathfrak{F}_1, P_1)$:

$\Omega_1 := (0,1]$

$\mathfrak{F}_1$:= System der Borel–Mengen in (0,1]

$P_1 := \lambda^1|_{\mathfrak{F}_1}$

Das Mengensystem $\mathfrak{F}_1$ hat gegenüber dem System der λ_1–mb. Mengen den Vorteil, daß $\mathfrak{F}_1$ einerseits nicht von dem zugrundegelegten Maß abhängt under andererseits alle für die Praxis interessanten Mengen enthält. Es ist sogar schwierig und nur mit Hilfe des Auswahlaxiom möglich, ein $A \in \mathfrak{P}(\Omega_1) \setminus \mathfrak{F}_1$ zu konstruieren.

1.7 Definition Das W–Maß P_1 heißt (stetige) **Gleichverteilung** U((0,1]) (=: U(0,1)) auf $(0,1]$.

Allgemein wird die Gesamtheit aller bei einem Versuch möglichen **Ergebnisse** ω durch den **Merkmalraum** Ω beschrieben. Die Problematik, daß ein W–Maß P i. a. nicht auf der Potenzmenge $\mathfrak{P}(\Omega)$ erklärt werden kann, wird allgemein dadurch gelöst, daß eine σ–Algebra $\mathfrak{F}$ als Definitionsbereich von P zugrundegelegt wird.

1.8 Definition $\mathfrak{F} \subset \mathfrak{P}(\Omega)$ ist **σ–Algebra** (in Ω), wenn gilt:

(1) $\emptyset \in \mathfrak{F}$ $\quad$ ($\Longleftrightarrow \Omega \in \mathfrak{F}$ wegen (2))

(2) $A \in \mathfrak{F} \Longrightarrow A^c \in \mathfrak{F}$, wobei $A^c := \Omega \setminus A$

(3) $A_m \in \mathfrak{F}, m \in \mathbb{N} \Longrightarrow \bigcup_{m=1}^{\infty} A_m \in \mathfrak{F}$

Die Eigenschaften (1) und (3) sichern gerade, daß die Forderungen (2) und (3) in (1.6) wohldefiniert sind. Die Eigenschaft (2) in 1.8 ermöglicht die Negationen von Eigenschaften $\mathfrak{E}$, die ein $A \in \mathfrak{F}$ gemäß $A = \{ \omega \in \Omega \,;\, \omega \text{ erfüllt die Eigenschaft } \mathfrak{E} \}$ beschreiben.

Die Wahl von P_1 ergab sich im wesentlichen aus Symmetriegründen.

Da eine Modellbildung nie rein logisch zu begründen ist, kann sie nur motiviert werden durch Eigenschaften des empirischen Experiments (hier : Unabhängigkeit von den Skalen) und den entsprechenden Eigenschaften des mathematischen Modells (hier: Translationsinvarianz).
Eine Aufgabe der W–theorie ist es, ein Arsenal von Modellen bereitzustellen und ihre Eigenschaften zu studieren. Grundsätzlich ist die Modellbildung abhängig vom Kenntnisstand über das empirische Experiment und vom Grad der Komplexheit, den man zulassen will. So kann man etwa bei der Notierung von Lebenszeiten noch nach Geschlecht, Wohnort oder Lebensweise differenzieren.

Die **Mathematische Statistik** bietet Verfahren an, mit denen man aufgrund von beobachteten Daten (Ergebnissen) entscheiden kann, ob ein gewähltes W–Modell zur Beschreibung des vorliegenden Experimentes geeignet ist.
Die Problematik der Modellbildung, die auch bei anderen mathematischen Disziplinen auftaucht (vgl. Geometrie; euklidische– bzw. Kugel–Geometrie), kann durch das eingangs vorgestellte 4 – Phasen – Schema sehr gut verdeutlicht werden. Ein Versuch, dieses Beispiel in das Schema einzuarbeiten, könnte wie folgt aussehen :

1. Phase :
   Beschreibung des empirischen Sachverhalts durch empirische Begriffe
   (Symmetrie, relative Häufigkeit).
2. Phase :
   Abbildung des in Phase 1 gewonnen empirischen Begriffssystems auf ein mathematisches (wahrscheinlichkeitstheoretisches) Modell
   (Definition eines W–Maßes, W–Raum, Modellannahmen, usw.)
3. Phase :
   Rein logische Untersuchung des in Phase 2 entworfenen Modells
   (Gesetz der großen Zahlen, usw.)
4. Phase :
   Umkehrung der Phase 2, d. h. empirische Interpretation der in Phase 3 erhaltenen Resultate
   (Aussagen über (Un–) Sicherheiten eines zukünftigen Ereignisses, Prognosen, usw.)

Solange das mathematische Modell noch nicht erprobt ist, schließt sich noch eine Phase an.

5. Phase :
   Vergleich der in Phase 4 gefundenen Resultate mit der Wirklichkeit
   (mit Hilfe der Mathematischen Statistik).

Erweiterung des Beispiels:

Bei n–maliger Wiederholung des Experiments ist ein mögliches Ergebnis: $x = (x_1,...,x_n) \in (0,1]^n$.
Ist $I = \times_1^n I_m \subset (0,1]^n$ ein n–dim Intervall, $a = (a_1,...,a_n) \in \mathbb{R}^n$, so sei

(1.9) $\qquad a+I := \{ (x_1,...,x_n) \in (0,1]^n \,;\, x_m - a_m \in I_m \bmod 1 \}$

Eine naheliegende Forderung an ein W–Maß P_n auf $(0,1]^n$ zur Beschreibung des Experiments ist:

(1.10) $\qquad P_n[a+I] = P_n[I]$

D.h.: verschiebt man die Skala vor dem m–ten Versuch um a_m, so soll das keinen Einfluß ausüben. Es gibt wieder genau ein W–Maß P_n auf den Borel–Mengen in $(0,1]^n$ mit (1.10) und zwar das Lebesgue–Maß λ^n, (Barner/Flohr II (83) S. 198/211, Bauer (78) S.45).
Das Wahrscheinlichkeitsmodell (**W–Modell**) $(\Omega_n,\mathfrak{F}_n,P_n)$ sieht nun wie folgt aus:

$$\Omega_n = (0,1]^n$$

$$\mathfrak{F}_n = \text{System der Borel–Mengen in } (0,1]^n$$

$$P_n = \lambda^n\big|_{\mathfrak{F}_n} =: U((0,1]^n) \quad \text{(Gleichverteilung auf } (0,1]^n\text{).}$$

Eine Analyse auf $(\Omega_n, \mathfrak{F}_n, P_n)$

Mit $x \in (0,1]^n = \Omega_n$ sei eine Größe (Gewinn) $X(x)$ verknüpft, sodaß $X : \Omega_n \mapsto \mathbb{R}$ Borel–meßbar ist und $\int_{\Omega_n} X^2(x)\, dx < \infty$ $\left[\Rightarrow \int_{\Omega_n} |X(x)|\, dx < \infty \quad \text{wegen} \quad |X| \le 1 + X^2 \right]$. Sei

$$E[X] := \int_{\Omega_n} X(x)\, dx \quad (\in \mathbb{R}),$$ **Mittelwert**, **Erwartungswert** und

$$\mathrm{Var}[X] := E[\{X - E[X]\}^2] \quad (< \infty),$$ **Varianz** von X bzgl. $(\Omega_n, \mathfrak{F}_n, P_n)$.

$\mathrm{Var}[X]$ ist ein Maß für die **Streuung** um den Mittelwert; im Falle $\mathrm{Var}[X]=0$ gilt:
$X(x) = \text{const}$ für λ^n–f.a. x mit $\text{const} = E[X]$.

Der Satz von **Fubini** (Barner & Flohr II (83), S. 304; Bauer (78), S. 117/118) besagt:

(1.11) $$E[X] = \int_0^1 (\ldots \int_0^1 (\int_0^1 X(x_1,\ldots,x_n)\, dx_{\pi(n)})\, dx_{\pi(n-1)}) \ldots dx_{\pi(1)}$$
für Permutationen $\pi : \{1,\ldots,n\} \mapsto \{1,\ldots,n\}$.

Wir betrachten jetzt Gewinne $G(t)$, $t \in (0,1]$, die auf jeder Stufe anfallen, und setzen:

$$G_m(x) := G(x_m) \quad \left[= g(x_1) \cdot \ldots \cdot g(x_{m-1}) \cdot G(x_m) \cdot g(x_{m+1}) \cdot \ldots \cdot g(x_n) \text{ mit } g \equiv 1 \right]$$

$$\overline{G}_n(x) := 1/n \sum_{m=1}^{n} G(x_m) \quad 1/n \sum_{m=1}^{n} G_m(x) \quad \text{auf } (0,1]^n.$$

Dann gilt offenbar mit (1.11):

(1.11a) $$E[G_m] = \int G(t)\, dt = E[G] =: a.$$

Man erwartet, daß die Bildung des arithmetischen Mittels einen stabilisierenden Effekt hat. Dies soll gezeigt werden. Mit (1.11a) folgt:

(1.12a) $$E[\overline{G}_n] = \frac{1}{n} \sum_1^n E[G_m] = a\,.$$

Betrachtet man die Streuung um den Mittelwert, so folgt ebenso wie in (1.11a):

$$\mathrm{Var}[\overline{G}_n] := \int_0^1 \ldots \int_0^1 \{\tfrac{1}{n} \sum_1^n (G(x_m) - a)\}^2 \overbrace{dx_1 \ldots dx_n}^{= \, dx}$$
$$= \frac{1}{n^2} \left\{ \sum_1^n \int_0^1 (G(t)-a)^2\, dt + \sum_{k \neq m} \int_0^1 (G(t)-a)\, dt \cdot \int_0^1 (G(y)-a)\, dy \right\},$$

indem man die Doppelsumme $\{\sum_1^n \ldots\}^2$ aufspaltet in zwei Summen über Terme mit gleichem Index bzw. mit verschiedenen Indizes. Nach Def. von a gilt $\int_0^1 (G(t)-a)\, dt = 0$, somit folgt:

(1.12b) $$\mathrm{Var}[\overline{G}_n] = \frac{1}{n} \int_0^1 (G(t)-a)^2\, dt =: \frac{1}{n} \mathrm{Var}[G] \longrightarrow 0 \quad (\text{für } n \longrightarrow \infty).$$

Diese Beziehung kann man offenbar als stabilisierenden Effekt deuten.

Sei $A_\varepsilon := \{ x \in \Omega_n ; |\overline{G}_n(x) - a| > \varepsilon \}$ für $\varepsilon > 0$, dann ist

$$\mathrm{Var}[\overline{G}_n] \geq E[\,\mathbf{1}_{A_\varepsilon} \cdot \{\overline{G}_n - a\}^2]$$

$$\geq E[\mathbf{1}_{A_\varepsilon} \cdot \varepsilon^2] = \varepsilon^2 \lambda^n [A_\varepsilon] \quad \text{, also}$$

(1.13) $\qquad P_n[A_\varepsilon] \leq \frac{1}{\varepsilon^2} \cdot \frac{1}{n} \cdot \mathrm{Var}[G]$ (**Tschebyscheff – Ungleichung**).

Wegen $P_n[A_\varepsilon^c] = 1 - P_n[A_\varepsilon]$ folgt:

1.14 Theorem (Ein schwaches Gesetz der großen Zahlen): Sei $\int_0^1 G^2(t)\, dt < \infty$, dann gilt:

$$P_n[\,\{ x \in \Omega_n ; |\overline{G}_n(x) - E[G]\,| \leq \varepsilon \}\,] \xrightarrow[n\to\infty]{} 1 \quad \text{für alle } \varepsilon > 0.$$

Dies ist ein mathematisches Gegenstück zum empirischen Gesetz der großen Zahlen. Der Satz kann auch als ein **Ergodensatz** angesehen werden. In der Sprache der Ergodentheorie ist $\overline{G}(x)$ ein **zeitliches Mittel** (entlang der Folge von Ergebnissen $x_1,\dots,x_n$) und $E[G]$ ein **räumliches Mittel** (über alle möglichen Werte von G).

1.15. Spezialfall

Sei $\quad G := \mathbf{1}_A$ mit A Borel–Menge in $(0,1]$, dann ist

$$\overline{G}_n(x) = \frac{1}{n} \Sigma_1^n \mathbf{1}_A(x_i) = \text{rel. } H[A]$$

$$E[G] = \int_0^1 \mathbf{1}_A(t)\, dt = \lambda^1[A] =: p = P_1[A]$$

$$\mathrm{Var}[G] = \int_0^1 (\mathbf{1}_A(t) - p)^2\, dt = \int_0^1 (\mathbf{1}_A(t) - 2p\, \mathbf{1}_A(t) + p^2)\, dt = p - 2p^2 + p^2 = p \cdot (1-p)\,,$$

also ist (vgl. (1.13)):

$$P_n[\{ x \in \Omega ; |\tfrac{1}{n} \Sigma_1^n \mathbf{1}_A(x_i) - p| \leq \varepsilon \}] \geq 1 - \frac{1}{n\varepsilon^2}\, p \cdot (1-p)\,. \;[]$$

1.16. Bemerkung

Jedes reelle zufallsabhängige Phänomen läßt sich durch $(\Omega_1, \mathfrak{A}_1, P_1)$ und eine Funktion $G : \Omega_1 \longrightarrow \mathbb{R}$ modellieren. Beweis vgl. 7.11. []

Beispiel zu 1.16 (a) **Werfen eines Würfels.**

Wähle $\qquad G(y) = i$ für $y \in (\frac{i-1}{6}, \frac{i}{6}]$ $i = 1,\dots,6$. Dann ist $P_1[\{y \in \Omega_1 ;\ G(y) = i\}] = 1/6$.

(b) **Werfen eines Reißnagels**. Hier liegen keine Symmetrien vor. Wenn man das Ergebnis, daß die Nagelspitze nach oben bzw. nach unten durch 0 und 1 beschreibt, so kann man $G(y) = 0$ für $y \in (0,p]$ und $G(y) = 1$ für $y \in (p,1]$ wählen. Bei $n = 26\,000$ Würfen hat man beobachtet, daß in 56,52 % der Würfe die Spitze nach oben lag. Also liegt die Wahl von $p = 0{,}5652$ nahe.

Die Wahl von p aufgrund von Beobachtungen ist bereits ein Beispiel für eine statistische Schätzung. $\square$

Das Drehen eines Glücksrades läßt sich mit dem Computer simulieren durch Angabe einer (Pseudo–) Zufallszahl $x \in (0,1]$. Mit Bemerkung (1.16) ist die Simulation des Glücksrades Grundlage für die **Simulation eines beliebigen zufallsabhängigen Phänomens**.

1.17 Bemerkung (**Monte–Carlo–Methode**). Gemäß 1.14 kann ein Integral $\int_0^1 G(y)dy$ näherungsweise experimentell berechnet werden durch $\frac{1}{n}\sum_1^n G(x_i)$.

1.18 Definition Sei $\Omega_0 = \mathbb{R}$, $\mathfrak{F}_0$ = das System der Borel–Mengen in $\mathbb{R}$ und

$$\varphi(t) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} t^2\right), \quad t \in \mathbb{R}, \quad \left(\Rightarrow \int_{-\infty}^{\infty} \varphi(t)\, dt = 1\right).$$

Dann heißt $P_0 =: N(0,1)$ mit $P_0[A] := \int_A \varphi(t)\, dt$, $A \in \mathfrak{F}_0$ **standardisierte Normalverteilung**.

Dabei ist $x \longrightarrow \varphi(x)$ die **Gaußsche Glockenkurve**.

——————————————————+——————————+—————|

α β

1.19 Zentraler Grenzwertsatz .

Sei $(\Omega_n, \mathfrak{F}_n, P_n)$, G und $\bar{G}_n$ wie oben mit $a := E[G]$ $(\in \mathbb{R})$ und $0 < \sigma^2 := \mathrm{Var}[G] < \infty$.

Dann gilt für $-\infty \le \alpha < \beta \le +\infty$:

$$P_n[\{x \in \Omega_n;\ \alpha < \sqrt{n}\,(\bar{G}_n(x) - a)/\sigma \le \beta\}] \longrightarrow \int_\alpha^\beta \varphi(t)\, dt = N(0,1)[(\alpha,\beta]] \qquad (n \to \infty).$$

1.19 zeichnet die Normalverteilung aus und ist eine Verschärfung von 1.14 (vgl. Übung).

$N(0,1)[(\alpha,\beta]]$ ist die Fläche unter der Glockenkurve über dem Intervall $(\alpha,\beta]$.

Die Gestalt von G geht in den Grenzwert offenbar nur über die Größen a und σ ein.

<u>1.20 Bemerkung</u> zur Normierung. Sei $Y(t) := (G(t) - a)/\sigma$, $t \in (0,1]$, dann ist

$$\sqrt{n}\,(\overline{G}_n(x) - a)/\sigma = \sqrt{n}\,\tfrac{1}{n}\,\Sigma_1^n\, Y(x_m) =: \sqrt{n}\,\overline{Y}_n(x)$$

mit $E[Y] = 0$ und $\mathrm{Var}[Y] = 1$. Also kann beim Beweis von 1.19 o.E. a=0 und $\sigma^2=1$ angenommen werden. Es folgt:

$$E[\sqrt{n}\,(\overline{G}_n - a)/\sigma] = E[\sqrt{n}\,\overline{Y}_n] = \sqrt{n}\,E[\overline{Y}_n] = \sqrt{n}\,E[Y] = 0 \quad \textbf{(Zentrierung)}$$
$$\mathrm{Var}[\sqrt{n}\,(\overline{G}_n - a)/\sigma] = n\,E[\overline{Y}_n^2] = n\,\mathrm{Var}[\overline{Y}_n] = n\,\tfrac{1}{n}\,\mathrm{Var}[Y] = 1 \quad (\textbf{Normierung}).$$

Zentrierung und Normierung bezeichnet man zusammenfassend als **Standardisierung**. $\square$

Jedes $f \in C(\overline{\mathbb{R}}) = C(\overline{\mathbb{R}},\mathbb{R})$ kann identifiziert werden mit einem $f \in C(\mathbb{R})$, für das $\lim_{y\to\pm\infty} f(y)$ in $\mathbb{R}$ existiert. Jedes solche f ist als stetige Funktion auf einem Kompaktum glm. stetig und beschränkt.

<u>1.21 Lemma</u> Für $f \in C(\overline{\mathbb{R}})$ gilt unter den Voraussetzungen von 1.19 mit a=0, $\sigma^2=1$:

$$E[f(\sqrt{n}\,\overline{G}_n)] \longrightarrow I(f) := \int_{-\infty}^{\infty} f(t)\cdot\varphi(t)\,dt \quad (n\to\infty)\,.$$

Theorem 1.19 folgt offenbar dann, wenn die Aussage von Lemma 1.21 auch für $f = \mathbf{1}_I$, I Intervall, gezeigt ist. Dies wird dann in einem weiteren Schritt geschehen.

Beweis von 1.21: (i) Zunächst erfolgt eine erneute Zentrierung:

$$E[f(\sqrt{n}\,\overline{G}_n)] - I(f) = E[\tilde{f}(\sqrt{n}\,\overline{G}_n)] \quad \text{mit } \tilde{f} = f - I(f) \in C(\overline{\mathbb{R}}).$$

(ii) Sei $h(y) := \frac{1}{\varphi(y)} \int_{-\infty}^{y} \tilde{f}(t)\cdot\varphi(t)\,dt$. Wegen

(1.22) $\qquad \varphi'(y) = -\,y\cdot\varphi(y)$

folgt als Darstellung von $\tilde{f}$:

(1.22a) $\qquad \tilde{f}(y) = h'(y) - y\cdot h(y)$.

(iii) $h' \in C(\overline{\mathbb{R}})$ wegen $\lim_{y\to\pm\infty} y\cdot h(y) = -\lim_{y\to\pm\infty} \tilde{f}(y)$ [vgl. Übung].

(iv) Sei $r(\varepsilon) := \max_{y,\,t\in\overline{\mathbb{R}},\ \mathrm{dist}(y,t)\le\varepsilon} |h'(y) - h'(t)|$, $\varepsilon \ge 0$, dann ist r als Stetigkeitsmodul einer stetigen Funktion auf einem Kompaktum beschränkt und stetig; insbesondere gilt:

$$r(\varepsilon) \downarrow 0 \text{ für } \varepsilon \downarrow 0 \quad \text{(vgl. Übung)}.$$

(v) Sei $G_m(x) := G(x_m)$, $W_n := \sqrt{n}\,\overline{G}_n = \Sigma_1^n\, G_m/\sqrt{n}$, $\qquad Z_{2,n} = \Sigma_2^n\, G_m/\sqrt{n}$.

Dann ist $\qquad W_n = G_1/\sqrt{n} + Z_{2,n}$ und es ist zu zeigen mit (1.22a):

(1.22b) $\qquad \big[E[\tilde{f}(W_n)] =\big]\ E[h'(W_n) - W_n\cdot h(W_n)] \to 0$.

(vi) Wegen der Invarianz von W_n geg. Permutationen der Koordinaten und mit Fubini (1.11) folgt:

$$E[W_n \cdot h(W_n)] = n^{-\frac{1}{2}} \Sigma_1^n E[G_m \cdot h(W_n)] = n^{-\frac{1}{2}} \Sigma_1^n E[G_1 \cdot h(W_n)] = \sqrt{n}\, E[G_1 \cdot h(W_n)].$$

(vii) Eine Taylorentwicklung von h um $Z_{2,n}$ liefert:

$$h(W_n) = h(Z_{2,n}) + n^{-\frac{1}{2}} G_1 \cdot h'(Z_{2,n}) + n^{-\frac{1}{2}} G_1 \cdot R_n \quad \text{mit } |R_n| \le r(|G_1|/\sqrt{n}).$$

(viii) Es folgt:

$$\sqrt{n}\, E[G_1 \cdot h(W_n)] = \sqrt{n}\, E[G_1 \cdot h(Z_{2,n})] + E[G_1^2 \cdot h'(Z_{2,n})] + E[G_1^2 \cdot R_n]$$

mit $E[G_1 \cdot h(Z_{2,n})] = E[G_1] \cdot E[h(Z_{2,n})] = 0 \cdot E[h(Z_{2,n})] = 0$ nach Fubini,

$E[G_1^2 \cdot h'(Z_{2,n})] = E[G_1^2] \cdot E[h'(Z_{2,n})] = 1 \cdot E[h'(Z_{2,n})]$ nach Fubini

und $|E[G_1^2 \cdot R_n]| \le E[G_1^2 \cdot r(|G_1|/\sqrt{n})] \to 0 \ (n\to\infty)$.

Die letztere Konvergenz folgt nach dem Lebesgueschen Satz von der majorisierten Konvergenz (Barner, Flohr II (83) S. 276 , Bauer (78) S. 77) auf dem Raum $(\Omega_1,\mathfrak{F}_1,P_1)$ wegen

$r(|G_1|/\sqrt{n}) \to 0$ auf Ω_1 und $|G_1^2 \cdot r(|G_1|/\sqrt{n})| \le \text{const} \cdot G_1^2$; dabei ist $\text{const} \cdot G_1^2$ nach Vorauss. eine integrierbare Majorante..

(ix) Für die nicht (für $n\to\infty$) verschwindenen Terme in (1.22b) ergibt sich

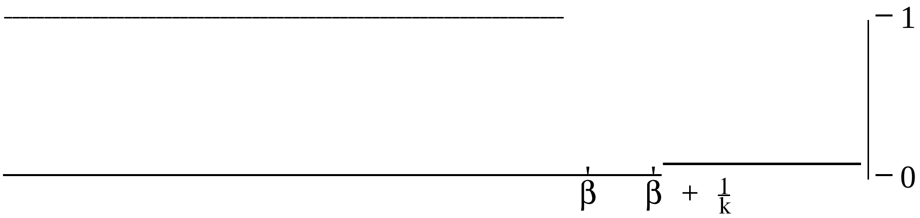
$$|E[h'(W_n) - h'(Z_{2,n})]| \le E[r(|G_1|/\sqrt{n})] \to 0$$

wieder wegen maj. Konvergenz. $\square$

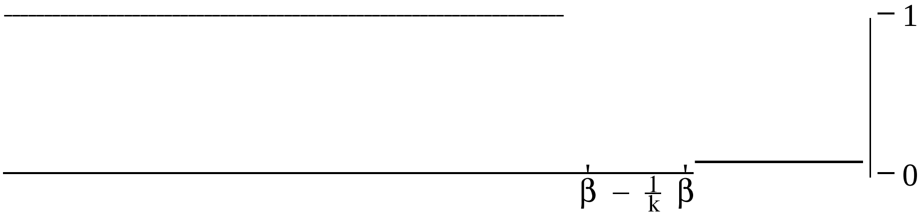
Beweisskizze zu Theorem 1.19. Der Rest des Beweises ist eine Übungsaufgabe.

Betrachte zunächst $\alpha = -\infty$ und definiere zu β Folgen von Funktionen $\bar{f}_k \in C(\bar{\mathbb{R}})$ mit

$\mathbf{1}_{(-\infty,\beta]} \le \bar{f}_k$ gemäß



und $\underline{f}_k \in C(\bar{\mathbb{R}})$ mit $\underline{f}_k \le \mathbf{1}_{(-\infty,\beta]}$ gemäß



Dann gilt für $k\to\infty$: $I(\bar{\underline{f}}_k) \longrightarrow I(\mathbf{1}_{(-\infty,t]})$. $\square$

§2 W–Räume

Der Gesamtheit aller bei einem Versuch möglichen **Ergebnisse** ω wird der **Merkmalraum** Ω zugeordnet. Typische Beispiele für Ω sind:

$\{0,1\}$, $\{1,...,n\}$, $\{0,...,n\}$, $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $[0,\infty)$, $\mathbb{R}$, $C([0,\infty))$,

$\Omega=\Omega_1\times...\times\Omega_n$, $\Omega=\Omega_1\times\Omega_2\times...$ mit Ω_i wie oben.

2.1 Def. Ein System $\mathfrak{F} \subset \mathfrak{P}(\Omega)$ heißt **σ–Algebra** (in Ω) und $(\Omega,\mathfrak{F})$ **Meßraum,** wenn gilt:

1) $\emptyset \in \mathfrak{F}$

2) $A \in \mathfrak{F} \Rightarrow A^c \in \mathfrak{F}$

3) $A_i \in \mathfrak{F}$, $i\in\mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}$

2.2 Folgerungen. 4) $\Omega \in \mathfrak{F}$ $(\emptyset^c=\Omega)$

5) $A_1,...,A_n \in \mathfrak{F} \Rightarrow \cup_1^n A_i \in \mathfrak{F}$ $(A_i=\emptyset\ \forall\ i>n)$

6) $A_i \in \mathfrak{F}$, $i\in\mathbb{N} \Rightarrow \cap_1^n A_i \in \mathfrak{F}$, $\cap_1^\infty A_i \in \mathfrak{F}$ (De Morgan)

Jedes $A \in \mathfrak{F}$ heißt **Ereignis**, $\{\omega\}$ **Elementarereignis**. Man sagt:
Das Ereignis A tritt ein, falls das beobachtete Ergebnis ω in A liegt. Die mengentheoretischen Operationen und Beziehungen lassen sich in aussagenlogische Verknüpfungen übersetzen:

$A^c \cong$ das Ereignis A tritt nicht ein

$A\cap B \cong$ das Ereignis A und das Ereignis B treten ein

$A\cup B \cong$ das Ereignis A oder das Ereignis B treten ein

$A \subset B \cong$ das Ereignis A impliziert das Ereignis B.

2.3 Beispiele. 1) $\mathfrak{P}(\Omega)$ ist die feinste,

2) $\{\emptyset,\Omega\}$ ist die gröbste,

3) $\{\emptyset, \Omega, A, A^c\}$ ist die von $A\subset\Omega$ erzeugte σ–Algebra.

2.4 Def. Sei $\mathfrak{C} \subset \mathfrak{P}(\Omega)$

$\sigma(\mathfrak{C}) := \bigcap \{\mathfrak{F}, \mathfrak{F}\ \sigma\text{– Algebra}, \mathfrak{F} \supset \mathfrak{C}\}$ heißt die **von $\mathfrak{C}$ erzeugte σ–Algebra**.

$\mathfrak{C}$ heißt **Erzeugendensystem**.

Die Def. ist sinnvoll, denn einmal gilt $\mathfrak{P}(\Omega) \supset \mathfrak{C}$ und zum anderen:

(2.5) $\mathfrak{F}_i$ σ–Algebra in Ω, $i\in I$ (I bel.) $\Rightarrow \bigcap_i \mathfrak{F}_i$ ist σ–Algebra (leicht)

2.6 Folgerung. Ist Ω abzb, $\mathfrak{C}$ das System der Elementarereignisse $\{\{\omega\}, \omega\in\Omega\}$, dann gilt $\sigma(\mathfrak{C}) = \mathfrak{P}(\Omega)$. (Bew. $A = \bigcup_{\omega\in A} \{\omega\}$.) Für abzb. Ω ist $\mathfrak{P}(\Omega)$ die natürliche σ–Algebra.

2.7 Def. Ist $(\Omega,\mathfrak{T})$ ein top. Raum, so heißt $\sigma(\mathfrak{T})$ die **Borelsche σ–Algebra.** Ist $\mathfrak{T}_d$ die übliche Top. auf $\mathbb{R}^d$, so heißt $\mathfrak{B}_d = \sigma(\mathfrak{T}_d)$ das System der **Borelschen Mengen in $\mathbb{R}^d$**.

2.8 Bemerkung. (ohne Beweis) Ist $\mathfrak{A}_d = (\mathfrak{B}_d)_{\lambda^d}$ das System der λ^d–mb Mengen, so gilt:

$$\mathfrak{B}_d \subsetneqq \mathfrak{A}_d \subsetneqq \mathfrak{P}(\mathbb{R}^d).$$

<u>**Weitere in $\mathfrak{F}$ erlaubte Operationen**</u>:

(2.9) Disjunkte Vereinigung: $\boxed{A+B :=}\ A\cup B$, falls $A\cap B=\emptyset$.

(2.10) Zerlegung durch Fallunterscheidung: $\Omega = \sum_i B_i \Rightarrow A \subset \sum B_i \Rightarrow$

$A = \bigcup_i A\cap B_i$ (1. Distributivgesetz) (Sind die B_i disjunkt, so auch die $A\cap B_i$!)

(2.11) Differenz $A\setminus B := A\cap B^c$ ($\cong$ A aber nicht B)

(2.12) symmetrische Differenz $A\Delta B := (A\setminus B) + (B\setminus A) = (A\cup B) \setminus (A\cap B)$ ($\cong$ genau eins der Ereignisse A, B tritt ein)

(2.13) mengentheoretischer Limes

$\overline{\lim}\, A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \{\omega\in\Omega;\ \omega \in A_n \text{ für } \infty \text{ viele n}\}$

$\underline{\lim}\, A_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m = \{\omega\in\Omega;\ \omega \in A_n \text{ für alle bis auf endlich viele n}\}$.

Stets gilt: $\underline{\lim}\, A_n \subset \overline{\lim}\, A_n$. Im Falle $\underline{\lim}\, A_n = \overline{\lim}\, A_n =: A$ setzt man $\lim A_n := A$ oder $A_n \to A$

(2.14) Monotone Folgen

(A_n) isoton $:\Leftrightarrow A_n\uparrow\ :\Leftrightarrow A_1\subset A_2\subset\ldots\Leftrightarrow: A_n \uparrow \cup A_m$

(A_n) antiton $:\Leftrightarrow A_n\downarrow\ :\Leftrightarrow A_1\supset A_2\supset\ldots\Leftrightarrow: A_n \downarrow \cap A_m$

Es gilt: $A_n\uparrow \Rightarrow A_n \to \cup A_n$, $A_n\downarrow \Rightarrow A_n \to \cap A_m$

(2.15) Folgerung. (A_m) bel. Dann $\bigcup_1^n A_m \uparrow \bigcup_1^\infty A_m$ und $\bigcap_1^n A_m \downarrow \bigcap_1^\infty A_m$

2.16 Reduktionssatz Zu $(A_i, i\in\mathbb{N})\ \exists\ (B_i, i\in\mathbb{N})$ mit den Eigenschaften:

(B_i) (paarw.) disjunkt, $B_i\subset A_i$, $\bigcup_1^n A_i = \bigcup_1^n B_i$, $1\le n\le\infty$, nämlich:

$B_1 := A_1$, $B_n := A_n \setminus (A_1\cup\ldots\cup A_{n-1}) = A_n\cap A_1^c\cap\ldots\cap A_{n-1}^c$

(2.17) Indikatorfunktionen: Seien (B_i) disjunkt, (A_i) beliebig:

$$1_{\cup B_n} = \Sigma 1_{B_n},\ 1_{\cap A_n} = \Pi\, 1_{A_n},\ 1_{\overline{\lim}\, A_n} = \overline{\lim}\, 1_{A_n},\ 1_{\underline{\lim}\, A_n} = \underline{\lim}\, 1_{A_n}.$$

2.18 Def. Kolmogoroffsche Axiome. Ist $\mathfrak{F}$ σ–Algebra in Ω, so heißt $P : \mathfrak{F} \mapsto \mathbb{R}$ **W–Maß** oder **W–Verteilung** und $(\Omega,\mathfrak{F},P)$ **W–Raum**, wenn P (1) positiv (2) normiert (3) σ–additiv ist.

2.19 Eigenschaften eines W–Maßes

(4) $P[\emptyset]=0$ (**Nulltreue**) (Bew. $P[\emptyset] = \sum P[\emptyset]$)

(5) P ist (**endlich**) **additiv** (Bew. setze $A_i=\emptyset\ \forall\ i>n$)

(6) P ist **isoton:** $A \subset B \Rightarrow P[A] \le P[B]$ (vgl. (7))

(7) P ist **subtraktiv**: $A \subset B \Rightarrow P[B\setminus A] = P[B] - P[A]$

(8) $P[A^c] = 1 - P[A]$ (wegen (7))

(9) $0 \le P[A] \le 1$ (wegen (6))

(10) $P[A \cup B] = P[A] + P[B] - P[A \cap B]$ (leicht)

(11) **Siebformel** von Sylvester–Poincaré:

$$P[\bigcup_{i=1}^n A_i] = \sum_{\emptyset\neq T\subset\{1,...,n\}} (-1)^{|T|-1} P[\bigcap_{i\in T} A_i] \qquad \text{(Übung)}$$

(12) P ist **sub–σ–additiv**: $P[\cup A_n] \le \sum P[A_n]$ (mit 2.16)

(13) $\cup B_n = \Omega \Rightarrow A \subset \cup B_n \Rightarrow P[A] = \sum_n P[A \cap B_n]$ (mit 2.10)

§ 3 Diskrete W–Räume

3.1 Def. $(\Omega,\mathfrak{F},P)$ heißt **diskreter W–Raum**, falls Ω abzb. und $\mathfrak{F} = \mathfrak{P}(\Omega)$.

3.2 Def. Sei Ω abzb. Eine Abb. $f : \Omega \mapsto [0,\infty)$ heißt **Z–Dichte** (Zähldichte) auf Ω, falls

$\sum_{\omega\in\Omega} f(\omega) = 1$. ($\sum_{\omega\in\Omega}$ ist unabh. von der Reihenfolge der Summation.)

Ist P W–Maß auf $\mathfrak{P}(\Omega)$, so heißt $\omega \mapsto P(\{\omega\}) =: P(\omega)$ Z–Dichte von P.

Interpretation: Jedes $\omega\in\Omega$ erhält ein Gewicht. Es besteht eine 1–1 Beziehung zwischen den Z–Dichten und den W–Maßen auf einem abzb. Raum.

3.3 Satz. Sei Ω abzb.

(a) Jedes W–Maß P auf $\mathfrak{P}(\Omega)$ ist durch seine Z–Dichte f eindeutig bestimmt.

(b) Zu jeder Z–Dichte f auf Ω gibt es ein W–Maß P auf $\mathfrak{P}(\Omega)$ mit Z–Dichte f, nämlich

(*) $P[A] = \sum_{\omega\in A} f(\omega)$, $A \in \mathfrak{P}(\Omega)$.

Bew. a) $A = \bigcup_{\omega\in A} \{\omega\} \Rightarrow$ (*) ist notwendig. b) z.z. P ist σ–additiv (Umordnungssatz). $\square$

3.4 Def. Sei $\Omega = \mathbb{N}_0$.

($f(k) =$) $\pi(\lambda;k) = e^{-\lambda} \lambda^k/k!$ definiert die **Poissonverteilung** $\pi(\lambda)$ zu $\lambda>0$,

($f(k) =$) $Geo(\rho;k) = (1-\rho)^k\rho$ definiert die **geometrische Verteilung** $Geo(\rho)$ zu $0<\rho<1$.

Die Bedeutung von $\pi(\lambda)$ wird in 5.17, 5.19, die von Geo(p) in 12.17, 12.18 erklärt.

3.5 Beispiel. Zu $x_1,...,x_n \in \Omega$ ist das W–Maß $P[A] := \mathrm{Emp}[A|x_1,...,x_n] = \frac{1}{n}|\{i;\ x_i \in A\}|$ auf $\mathfrak{P}(\Omega)$ definiert. Ist Ω abzählbar, so hat P die Z–Dichte $f(\omega) = \frac{1}{n}|\{i;\ x_i=\omega\}|$. []

Ein Spezialfall von 3.5 ergibt sich in:

3.6 Def. $(\Omega,\mathfrak{F},P)$ heißt **Laplacescher W–Raum** und P (**diskrete**) **Gleichverteilung** $U(\Omega)$ auf Ω,, wenn
1) Ω endlich 2) $\mathfrak{F} = \mathfrak{P}(\Omega)$ 3) $P(\omega) = \frac{1}{|\Omega|}$.

Folgerung: $P[A] = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}$ (Laplace)

Beipiele 1) Werfen einer Münze
2) Werfen eines Würfels
3) Ziehen einer Karte aus einem Skatspiel

"**Gegenbeispiel**": Werfen eines Reißnagels.

In Laplaceschen W–Räumen wird die W. durch Abzählen von Mengen ermittelt. Hilfsmittel stellt die **Kombinatorik** bereit:

(3.7) Geg. A mit $|A|=n$; dann gilt: $|A^r| = n^r$.
Sei $1 \le r \le n$ und $B=\{(b_1,...,b_r) \in A^r;\ b_i \ne b_j\ \forall\ i \ne j\}$. Dann gilt: $|B| = (n)_r = n(n-1)...(n-r+1)$.

(3.8) Geg. $A = \{0,1\}$, $1 \le k \le r \in \mathbb{N}$, $B = \{(b_1,...,b_r) \in A^r;\ |\{i,\ b_i=1\}| = k\}$.Dann gilt:
$|B| = \binom{r}{k}$ = Anzahl aller k–element. Teilmengen einer r–element.Menge.

Verallgemeinerung:

(3.9) Geg. $A = \{1,...,n\}$, $k_m \in \mathbb{N}_0$, $\sum_{m=1}^{n} k_m = r$,
$B = \{(b_1,...,b_r) \in A^r;\ |\{i;\ b_i=m\}| = k_m,\ 1 \le m \le n\}$. Dann gilt:
$|B| = \frac{r!}{k_1!k_2!...k_n!} =: \binom{r}{k_1,k_2,...,k_n}$ (**Multinomialkoeffizient**).

3.10 Beispiel: Geburtagszwillinge. Es werden r Personen nach ihrem Geburtstag befragt. Wir wählen $\Omega = \{1,2,...,365\}^r$ und nehmen an, daß das Experiment näherungsweise durch einen Laplaceschen W–Raum beschrieben werden kann. Es interessiert die W., daß (mindestens) zwei Personen am gleichen Tag Geburtstag haben. Sei also

$$A = \{ (\omega_1,...,\omega_r) \in \Omega;\ \exists\ i \ne j \text{ mit } \omega_i=\omega_j \} .$$

Es ist $|\Omega| = 365^r$, $|A^c| = (365)_r$, somit folgt: $P[A] = 1 - P[A^c] = 1 - (365)_r/365^r$.

Zahlenwerte (nach Stirling) für (r;P[A]): (30; 0,71), (40; 0,89), (50; 0.97), (60; 0,99)
P[A] erhöht sich noch bei einer Abweichung von der Gleichverteilung.

§4 Zufallsvariablen und ihre Verteilungen.

Ziel: Eine Zva X ist eine Größe mit Werten in einem Raum Ω', sodaß $P[X\in A']$ erklärt ist (für geeignete $A' \subset \Omega'$).

Der Merkmalraum Ω wird so gewählt, daß $\omega \in \Omega$ alle (interessierenden) Informationen über das Experiment enthält. Jeder Aspekt dieses Experiments ist also durch ω bestimmt. Eine Zva X soll aber gerade einen Teilaspekt des Experiments beschreiben. Unser Ziel wird also erreicht, wenn man X als eine Funktion $X : \Omega \mapsto \Omega'$ definiert. Es überrascht, daß eine Zufallsvariable als eine eindeutige Funktion definiert werden soll. Die Unsicherheit über $X(\omega)$ liegt dann jedoch in (dem Eintreten von) $\omega \in \Omega$.

Beispiel. Es werde n–mal gewürfelt. (i) X ist die Summe der gewürfelten Zahlen.
(ii) X_m ist das Ergebnis des m–ten Wurfes.

4.1 Def. Zu $f : \Omega \mapsto \Omega'$ heißt

$f^{-1} : \mathfrak{P}(\Omega') \mapsto \mathfrak{P}(\Omega)$ mit $f^{-1}(A'):=\{\omega\in\Omega;\ f(\omega)\in A'\}$ die zu f gehörige **Urbildfunktion**. und

$f : \mathfrak{P}(\Omega) \mapsto \mathfrak{P}(\Omega')$ mit $f(A):=\{\omega'\in\Omega;\ \exists\ \omega\in A \text{ mit } f(\omega)=\omega'\}$ **Bildfunktion**.

4.2 Lemma. Sei $f:\Omega\mapsto\Omega'$bel., A', $A_i' \subset \Omega'$, I bel. Indexmenge. Dann:

(a) $f^{-1}(\bigcap_{i\in I} A_i') = \bigcap_{i\in I} f^{-1}(A_i')$
(b) $f^{-1}(\bigcup A_i') = \bigcup f^{-1}(A_i')$; sind dabei die A_i' disjunkt, so auch die $f^{-1}(A_i')$.
(c) $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\Omega') = \Omega$
(d) $f^{-1}(A'^c) = (f^{-1}(A'))^c$
(e) $(g\circ f)^{-1} = f^{-1}\circ g^{-1}$ falls $g:\Omega'\mapsto\Omega''$

Schreibweisen: $\{\omega\in\Omega;\ X(\omega)\in A'\} = X^{-1}(A') =: \{X\in A'\}$, $P[\{X\in A'\}] =: P[X\in A']$.
Beispiele: $\{X=x\}$, $\{X\leq x\}$, $P[X=x]$, $P[X\leq x]$.

4.3 Def. Sei $(\Omega,\mathfrak{F},P)$ ein W–Raum, $(\Omega',\mathfrak{F}')$ Meßraum. Dann heißt $X:\Omega\mapsto\Omega'$ eine **Zufallsvariable** (**Zva**) auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega',\mathfrak{F}')$, falls $\{X\in A'\}\in\mathfrak{F}\ \forall\ A'\in\mathfrak{F}'$.

Für $\Omega'\subset\mathbb{R}$ heißt X **reelle Zva**, für $\Omega'\subset\overline{\mathbb{R}}$: **erweitert reelle Zva**, für $\Omega'\subset\mathbb{R}^d$: **d–dimensionaler Zufallsvektor**, für Ω' abzb.: **diskrete Zva**..

4.4 Bemerkung: Ist $(\Omega,\mathfrak{F},P)$ diskret, so ist jede Abb. $X:\Omega\mapsto\Omega'$ Zva. Ist Ω' abzb., $\mathfrak{F}' = \mathfrak{P}(\Omega')$, so gilt:

X (diskrete) Zva $\Leftrightarrow$ $\{X=x\} \in \mathfrak{F}\ \forall\ x\in\Omega'$ (vgl. 2.6).

4.5 Satz. Sei X Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega',\mathfrak{F}')$. Dann ist durch

$$P'[A'] := P[X\in A'] = P[X^{-1}(A')] \text{ ein W–Maß } P' \text{ auf } \mathfrak{F}' \text{ definiert.}$$

Schreibweise: $P'= PX^{-1}$.

Bew. mit 4.2b,c. []

4.6 Def. PX^{-1} heißt **Verteilung** von X.

Schreibweise: $X \sim P' \Leftrightarrow PX^{-1} = P'$ (wenn etwa P nicht interessiert);

z.B. $X \sim \pi(\lambda)$ heißt $P[X=k] = e^{-\lambda}\lambda^k/k!$.

4.7 Bemerkung: Die Verteilung einer diskreten Zva ist durch die Werte $P[X=x]$, $x\in\Omega'$, festgelegt. (Bew. 3.3a).

4.8 Bemerkung. Die Aussage 1.16 bedeutet gerade: Zu jedem W–Maß Q auf $(\mathbb{R}^1,\mathfrak{B}_1)$ ex Zva X auf $(\Omega_1,\mathfrak{F}_1,P)$ wie in §1 mit Werten in $(\mathbb{R},\mathfrak{B}_1)$, so daß $X \sim Q$.

4.9 Beispiel. $Q = U(\{1,...,6\})$. Sei $(\Omega_1,\mathfrak{F}_1,P_1)$ wie in §1, $X(\omega) = i$ für $\omega \in (\frac{i-1}{6},\frac{i}{6}]$, i=1,...,6. Dann gilt: $\Omega'=\{1,...,6\}$, $P_1X^{-1}[\{i\}] = P_1[X=i] = \frac{1}{6}$, $i\in\Omega'$, d.h. $X \sim U(\{1,...,6\})$.

§5 Elementare bedingte W. und stochastisch unabhängige Zva.

Problem: Wie hängen die Schulnoten in Mathematik und Deutsch eines zufällig heraus–gegriffenen Schülers voneinander ab?

Eine Untersuchung bei n Schülern ergibt als Ergebnis (Stichprobe) Tupel

$$\omega_1,...,\omega_n \in \{1,...,6\}^2 =: \Omega.$$

1. Wähle $i,j \in \{1,...,6\}$ fest und betrachte $A:=\{i\}\times\{1,...,6\} \subset \Omega$, $B:=\{1,...,6\}\times\{j\} \subset \Omega$.
2. Sondere nun die Tupel $\omega_m \in B$ aus. Deren Anzahl ist $n\cdot \text{rel.H}[B]$ und betrachte in der reduzierten Folge die rel. Häufigkeit von A, also

$$\text{rel.H}[A|B] := \frac{1}{n\cdot\text{rel.H}[B]}\sum_{m=1}^{n} 1_A(\omega_m)1_B(\omega_m) = \text{rel.H}[A\cap B]/\text{rel.H}[B].$$

Dieser Wert kann nun etwa mit rel.H[A] verglichen werden. Die Größe rel.H [A|B] dient als Vorbild für die folgende:

5.1 Def. Sei $(\Omega,\mathfrak{F},P)$ W–Raum, $A,B\in\mathfrak{F}$. Dann heißt $P[A|B] := \frac{P[A\cap B]}{P[B]}$ bedingte W. von A unter der Bedingung B. Ist $P[B]=0$, so sei $P[A|B]=0$ (oder beliebig $\in [0,1]$).

Sprechweise: P heißt auf B konzentriert, falls $P[B]=1$.

5.2 Satz. Sei $(\Omega,\mathfrak{F},P)$ W–Raum, $B\in\mathfrak{F}$ mit $P[B]>0$. Dann ist $P[\cdot|B]$ ein W–Maß auf $\mathfrak{F}$, das auf B konzentriert ist.

Bew. einfach mit (2.10). []

Beispiel. Die W., daß bei 2 Münzwürfen zweimal Zahl erscheint, unter der Bed., daß die erste Münze Zahl zeigt, ist 1/2. Die W., daß bei 2 Münzwürfen zweimal Zahl erscheint, unter der Bed., daß mindestens einmal Zahl dabei ist, ist 1/3. $\square$

Problem. Seien X,Y diskrete Zva auf $(\Omega.\mathfrak{F},P)$. Wann gilt

$(*)$ $P[Y=j\,|\,X=i] = P[Y=j]$?

Interpretation: Wann liefert die Kenntnis von X keine Information über Y.
Sei $P[X=j]>0$, dann gilt: $(*) \Leftrightarrow P[X=i,Y=j] = P[X=i]\cdot P[Y=j]$
Die rechte Aussage ist symmetrisch in X und Y und ist auch sinnvoll, falls $P[X=i]=0$. Deshalb wird sie für die Definition von "X und Y unabhängig" genommen.

5.3 Def. Sei J endlich, X_i, $i\in J$, Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i.\mathfrak{F}_i)$. Dann heißt die Familie

$(X_i, i\in J)$ (stochastisch) unabhängig $:\Leftrightarrow$

(5.4) $$P[X_i\in B_i,\ i\in J] = \mathbf{\Pi}_{i\in J}\, P[X_i\in B_i] \quad \forall\ B_i\in\mathfrak{F}_i$$

5.5 Folgerung. $(X_i,i\in J)$ unabh. $\Leftrightarrow$ $(X_i,\ i\in I)$ unabh. $\forall\ I\subset J$. (Bew. Setze $B_i=\Omega_i$ für $i\in J\setminus I$).

5.6 Def. Seien nun J bel. und X_i wie in 5.3. dann heißt die Familie

$(X_i,\ i\in J)$ stoch. unabh. $:\Leftrightarrow$ $(X_i,\ i\in I)$ unabh. $\forall\ I\subset J$ mit $2\le|I|<\infty$.

5.7 Folgerungen. (a) Die Def. 5.3 und 5.6 sind konsistent gemäß 5.5.
(b) Ist $J=\mathbb{N}$, so gilt: $(X_i,\ i\in\mathbb{N})$ unabh. $\Leftrightarrow$ $(X_i,\ 1\le i\le n)$ unabh. $\forall$ n.

Bew. von b)"$\Leftarrow$" Zu I mit $|I|<\infty$ wähle n, so daß $I\subset\{1,...,n\}$. $\square$

5.8 Satz. Seien X_i, $1\le i\le n$, diskrete Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in Ω_i. Dann gilt:

$(X_i,\ 1\le i\le n)$ unabh. $\Leftrightarrow$ $P[X_1=k_1,...,X_n=k_n] = P[X_1=k_1]\cdot...\cdot P[X_n=k_n]$ $\forall\ k_i\in\Omega_i$, $1\le i\le n$.

Bew. von "$\Leftarrow$" mit $B_i = \Sigma_{k\in B_i}\{k\}$. $\square$

5.9 Bemerkung. Seien Ω_i endlich, $1\le i\le n$, $\Omega=\Omega_1\times...\times\Omega_n$, $X_i:\Omega\to\Omega_i$ die Projektionen.

Dann sind äquivalent:

(1) $(\Omega,\mathfrak{F},P)$ ist Laplacescher W–Raum.
(2) $(\Omega,\mathfrak{F},P)$ ist diskreter W–Raum, so daß
(a) $X_1,...,X_n$ unabh. (b) $X_i\sim U(\Omega_i)$, $1\le i\le n$.

Bew. "(1)$\Rightarrow$(2)" $P[X_1=k_1,...,X_n=k_n] = P[\{(k_1,...,k_n)\}] = \frac{1}{|\Omega|} = \Pi_i\,\frac{1}{|\Omega_i|}$.

Durch Summation über k_j mit $j\neq i$ erhält man: $P[X_i=k_i] = \frac{1}{|\Omega_i|}$.

ad "(2)$\Rightarrow$(1)": $P[\{(k_1,...,k_n)\}] = P[X_1=k_1,...,X_n=k_n] = \Pi_i\, P[X_i=k_i] = \Pi_i\,\frac{1}{|\Omega_i|} = \frac{1}{|\Omega|}$ $\square$

Beispiele: Geburtstagszwillinge 3.10, n–maliges Werfen einer Münze oder eines Würfels.

Sprech– und Schreibweise: X_i, $i\in I$, heißen **identisch verteilt**, wenn $X_i \sim Q$ für ein W–Maß Q.

X_i , $i\in I$, sind **iid**, wenn sie unabhängig und identisch verteilt (independent, identically distributed) sind.

5.10 Beispiel: Multinomialverteilung.

Seien $X_1,\dots,X_n$ iid Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $\{\alpha_0,\dots,\alpha_d\}$, wobei $n, d\in\mathbb{N}$,

mit $P[X_m=\alpha_i] = p(\alpha_i) =: p_i$, $0\le i\le d$, für eine Z–Dichte p auf $\{\alpha_0,\dots,\alpha_n\}$.

$Y_i(\omega) := |\{m;\ 1\le m\le n\},\ X_m(\omega) = \alpha_i\}|$, $0\le i\le d$, also $Y_0 = n - (Y_1+\dots+Y_d)$

$Y := (Y_1,\dots,Y_d)$, also $Y : \Omega \to K_{d,n} := \{(k_1,\dots,k_d) \in \mathbb{N}_0^d;\ k_1+\dots+k_d\le n\}$.

Zu $(k_1,\dots,k_d)\in K_{d,n}$, $k_0:= n - (k_1+\ \ +k_d)$, betrachte

$B := \{(x_1,\dots,x_n) \in \{\alpha_0,\dots,\alpha_d\}^n,\ \alpha_i$ tritt k_i–mal auf, $0\le i\le d\}$.

Dann gilt $|B| = \dfrac{n!}{k_0!k_1!\dots k_d!}$ nach (3.9). Es ist

$\{Y = (k_1,\dots,k_d)\} = \bigcup_{(x_1,\dots,x_n)\in B} \{X_1=x_1,\dots,X_n=x_n\} \in \mathfrak{F}$, also ist Y Zva nach 4.4 und

$P[Y=(k_1,\dots,k_d)] = \sum_{(x_1,\dots,x_n)\in B} P[X_1=x_1]\dots P[X_n=x_n] =$

$$= \sum_{(x_1,\dots,x_n)\in B} p_0^{k_0}\, p_1^{k_1}\ \ p_d^{k_d} = |B|\cdot \prod_0^d p_i^{k_i} =$$

(5.11) $$b(n,(p_1,\dots,p_d);(k_1,\dots,k_d)) := n! \prod_0^d \frac{p_i^{k_i}}{k_i!} \quad \text{mit} \quad \begin{array}{l} p_0 = 1-(p_1+\dots+p_d) \\ k_0 = n-(k_1+\dots+k_d) \end{array}.$$

5.12 Def. Das W–Maß auf $K_{d,n}$ mit der Z–Dichte $b(n,\vec{p};\cdot)$ heißt **Multinomialverteilung** $b(n,\vec{p})$ für $\vec{p} = (p_1,\dots,p_d) \in [0,1]^d$ mit $\sum_1^d p_i \le 1$.

Spezialfall: d=1, $\alpha_0=0$ ($\hat{=}$ Mißerfolg), $\alpha_1=1$ ($\hat{=}$ Erfolg), $K_{1,n}=\{0,\dots,n\}$, $p_1=:p=P[X_m=1]$, $p_0=1-p=P[X_m=0]$, $Y \equiv$ Anzahl der Erfolge bei n Versuchen, d.h. $Y = \sum_{m=1}^n X_m$, $P[Y=k] =$

(5.13) $$b(n,p;k) = \binom{n}{k}p^k(1-p)^{n-k}$$

5.14 Def. Das W–Maß auf $\{0,\dots,n\}$ mit der Z–Dichte $b(n,p;\cdot)$ heißt **Binomialverteilung** $b(n,p)$ für $p\in[0,1]$. $n\in\mathbb{N}$.

5.15 Folgerung. Seien $Z_1,\dots,Z_n$ iid Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega',\mathfrak{F}')$, $B\in\mathfrak{F}'$,

$X_m := \mathbf{1}_{\{Z_m\in B\}}$ $(= \mathbf{1}_B\circ Z_m)$, dann sind $X_1,\dots,X_n$ iid Zva mit $P[X_m = 1] = P[Z_1 \in B]=:p$ und

$$\sum_1^n X_m \sim b(n,p).$$

Approximation von b(n,p) durch $\pi(\lambda)$ für große n und kleine p.

5.16 Hilfssatz. $(1 + \frac{z}{n} + o(\frac{1}{n}))^n \to e^z \quad (n\to\infty)$ für $z\in\mathbb{C}$.

5.17 Gesetz der seltenen Ereignisse.

$b(n,p_n;k) \to \pi(\lambda;k)$, $k\in\mathbb{N}_0$, falls $n\to\infty$, $np_n\to\lambda > 0$.

Bew. durch Induktion nach k: k=0: $b(n,p_n;0) = (1-p_n)^n \to e^{-\lambda} = \pi(\lambda,0)$ nach 5.16.

$k\to k+1$: $b(n,p_n;k+1) = \frac{n-k}{k+1}\frac{p_n}{1-p_n}\, b(n,p_n;k) \to \frac{\lambda}{k+1}\pi(\lambda;k) = \pi(\lambda;k+1)$. $\square$

Beispiel: n=100, p=0,01, λ=1 dann $|b(n,p;k)-\pi(\lambda;k)| \le 0{,}002$.

5.18 Hilfssatz. Jede beschränkte Lösung $f:[0,\infty)\mapsto\mathbb{R}$ von $f(t+s)=f(t)f(s)$ $s,t\ge0$
hat die Gestalt $f(t) = e^{-\lambda t}$ für ein $\lambda\in[0,\infty]$.
(**Bew**. z.B. Hinderer S. 201).

5.19 **Beispiel: der Poissonprozeß.**

Geg. $T_n : \Omega \mapsto (0,\infty)$, $n\in\mathbb{N}$ mit $0 < T_1(\omega) < T_2(\omega) <$...und $T_n(\omega) \to \infty$ $\forall\omega\in\Omega$.

Für ein Intervall $I \subset [0,\infty)$ sei $N_I(\omega) := |\{n\in\mathbb{N};\, T_n(\omega) \in I\}|$.

Annahmen: $\forall\, 0\le a<b$ sei $N_{(a,b]}$ Zva auf $(\Omega,\mathfrak{F},P)$, sodaß gilt:

(1) die Zva $N_{(a+h,b+h]}$, $h\ge0$, sind identisch verteilt,

(2) für disjunkte Intervalle $I_1,\dots,I_n$ $(2\le n<\infty)$ sind $N_{I_1},\dots, N_{I_n}$ unabhängig.

Beh.: $\exists\, \lambda>0$ sodaß $N_{(a,b]} \sim \pi(\lambda(b-a))$.

Anwendungen: T_n ist der Zeitpunkt

a) des n–ten Telephonanrufs in einer Telephonzentrale,
b) der Ankunft des n–ten Kunden an einem Schalter,
c) der Aussendung des n–ten α–Teilchens einer radioaktiven Substanz,
d) des n–ten Ausfalls einer Maschine in einer Anlage.
e) der n–ten Schadensmeldung bei einer Versicherung.

Bew. der Beh.: Sei $f(t):=P[N_{(0,t]}=0]=P[T_1>t]$.

(i) $\exists\, \lambda\in[0,\infty]$ mit $f(t) = e^{-\lambda t}$, denn $f(t+s) = P[N_{(0,t]}=0,\, N_{(t,t+s]}=0] \overset{(2)}{=} P[N_{(0,t]}=0]\cdot P[N_{(t,t+s]}=0]$
$\overset{(1)}{=} P[N_{(0,t]}=0]\cdot P[N_{(0,s]}=0] = f(t)\cdot f(s)$; jetzt 5.18.

(ii) $0<\lambda<\infty$, denn wegen $\{T_1<\infty\} = \cup_n \{T_1 \le n\}$ und $P[T_1<\infty]>0$ (sogar =1) folgt mit 2.19(12):

$\exists$ n mit $P[T_1 \le n] > 0$, d.h. $\exists$ n mit $f(n)<1$.

Wegen $\{T_1>0\} = \cup_n \{T_1 > \frac{1}{n}\}$ und $P[T_1>0]>0$ (sogar =1) folgt ebenso: $\exists$ n mit $f(\frac{1}{n})>0$.

(iii) Zu $(0,t] = \cup_{i=1}^{n} (\frac{i-1}{n}t, \frac{i}{n}t]$ setze $X_i^{(n)} := 1_{\{N(\frac{i-1}{n}t,\, \frac{i}{n}t] > 0\}}$; dann gilt

$P[X_i^{(n)}=1] = P[X_1^{(n)}=1] = 1 - f(\frac{t}{n}) = 1 - e^{-\lambda t/n} =: p_n$.

Sei $Y^{(n)} := \sum_{i=1}^{n} X_i^{(n)}$ die Anzahl der Intervalle, in denen mindestens ein Zeitpunkt T_m liegt. Dann ergibt sich $Y^{(n)} \sim b(n,p_n)$ gemäß 5.15; dabei gilt: $np_n \to \lambda t$,

also nach 5.17: $b(n,p_n;k) \to \pi(\lambda t;k)$ für $n\to\infty$, d.h. $P[Y^{(n)}=k] \to \pi(\lambda t;k)$, $k\in\mathbb{N}_0$.

Andererseits gilt: $Y^{(n)}(\omega) \to N_{(0,t]}(\omega)$ $\forall\omega\in\Omega$; daraus folgt (Übung):

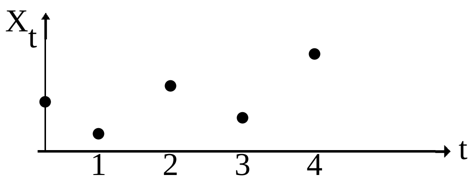
$P[Y^{(n)}=k] \to P[N_{(0,t]}=k]$ für $n\to\infty$. Insgesamt: $P[N_{(0,t]}=k] = \pi(\lambda t;k)$ $\forall$ $k\in\mathbb{N}_0$. $\Box$

$\{N_t, t\geq 0\}$ mit $N_t:=N_{(0,t]}$ heißt **Poissonprozeß mit Intensität λ** und ist ein wichtiges Beispiel für einen **stochastischen Prozeß mit stetigem Zeitparameter**.

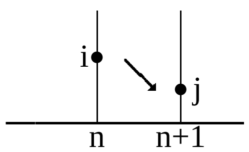
Die **Existenz** eines W–Raums mit den obigen Eigenschaften (Annahmen) in 5.19 wird später gezeigt.

Markoff–Ketten.

Wir betrachten jetzt eine Folge von diskreten Zva $(X_n, 0 \leq n < N)$ $(N \leq \infty)$ auf $(\Omega,\mathfrak{F},P)$ mit Werten in der abzählbaren Menge S und interpretieren X_n als Zustand eines Systems z.Zt. n.



Dabei wird p_{ij} die W. bestimmen, daß das System nach j geht, wenn es zuvor in i gewesen ist.



5.20 Definition. $\mathbb{P} = (p_{ij})_{i,j\in S}$ heißt eine **Übergangsmatrix** (Ü–Matrix) und die p_{ij} **Übergangs–wahrscheinlichkeiten**, falls $p_{ij} \geq 0$, $i,j \in S$, $\sum_{j\in S} p_{ij} = 1$ (d.h. $j \mapsto p_{ij}$ ist eine Z–Dichte) $\forall$ $i \in S$.

Die Ü–Matrizen bilden eine Halbgruppe bzgl. der Matrixmultiplikation

(5.21) $\mathbb{P} = \mathbb{P}' \cdot \mathbb{P}'' \Leftrightarrow p_{ik} = \sum_{j\in S} p'_{ij}\, p''_{jk}$.

5.22 Bezeichnungen $(p_{ij}^{(0)}) = \mathbb{P}^0$ ist die Einheitsmatrix,

$(p_{ij}^{(n)}) = \mathbb{P}^n$ ist das n–fache Produkt von $\mathbb{P}$ mit sich selbst.

5.23 Lemma. Sei $\mathbb{P} = (p_{ij})$ eine Ü–Matrix. Dann sind äquivalent:

i) $P[X_0 = i_0,...,X_n = i_n] = P[X_0 = i_0] \cdot p_{i_0 i_1} \cdot ... \cdot p_{i_{n-1} i_n}$, $\forall\, i_0,...,i_n \in S,\ 1 \le n < N$;

ii) $P[X_0 = i_0,...,X_{n+1} = i_{n+1}] = P[X_0 = i_0,...,X_n = i_n] \cdot p_{i_n i_{n+1}}$ $\forall\, i_0,...,i_{n+1} \in S,\ 1 \le n+1 < N$;

iii) $P[X_{n+1} = i_{n+1} \mid X_0 = i_0,...,X_n = i_n] = p_{i_n i_{n+1}}$

$\forall\, i_0,...,i_{n+1} \in S,\ 1 \le n+1 < N$, mit $P[X_0 = i_0,...,X_n = i_n] > 0$.

Beweis: Übung.

5.24 Definition. $(X_n,\ 0 \le n < N)$ ist eine (zeithomogene) **Markoff–Kette**, falls eine Ü–Matrix $\mathbb{P} = (p_{ij})$ existiert, sodaß eine der drei äquivalenten Bedingungen (i) – (iii) in 5.23 gelten. Die Eigenschaft (iii) heißt **Markoff–Eigenschaft** (Unabhängigkeit der Zukunft X_{n+1} von der Vergangenheit $(X_0,...,X_{n-1})$ bei bekannter Gegenwart X_n). Die Verteilung von X_0 heißt **Startverteilung**.

Eine Markoff–Kette ist eine Verallgemeinerung einer Folge von iid Zva mit Werten in S; in diesem Fall kann man setzen: $p_{ij} = p_j := P[X_0=j]$.

5.25 Satz (Variante der Markoff–Eigenschaft):

a) $P[X_{n+1} = j \mid (X_0,...,X_{n-1}) \in B,\ X_n = i] = p_{ij}$

$\forall\, B \subset S^n,\ i,j \in S,\ 1 \le n+1 < N$, mit $P[(X_0,...,X_{n-1}) \in B,\ X_n = i] > 0$.

b) $P[X_{n+1} = j_1,...,X_{n+m} = j_m \mid (X_0,...,X_{n-1}) \in B,\ X_n = i] = p_{ij_1} \cdot ... p_{j_{m-1} j_m}$

$\forall\, B \subset S^n,\ i,j_1,...,j_m \in S,\ n \le n+m < N$, mit $P[(X_0,...,X_{n-1}) \in B,\ X_n = i] > 0$.

Beweis: Übung. Ein wichtiger Spezialfall ist der Fall $B = S^n$.

5.26 Lemma. $P[X_n = i,\ X_{n+m} = k] = P[X_n = i] \cdot p_{ik}^{(m)}$.

Beweis. Für m=1 haben wir mit 5.25a ($B = S^n$):

$P[X_n = i,\ X_{n+1} = k] = P[X_n = i] \cdot P[X_{n+1} = k \mid X_n = i] = P[X_n = i] \cdot p_{ik}$.

Für den Schritt $m \to m+1$ schreiben wir mit einer Zerlegung durch Fallunterscheidung [(2.10)]:

$P[X_n = i, X_{n+m+1} = k] = \Sigma_j P[X_n = i, X_{n+m} = j, X_{n+m+1} = k]$
$= \Sigma_j P[X_n = i, X_{n+m} = j] \cdot P[X_{n+m+1} = k \mid X_n = i, X_{n+m} = j]$
$= \Sigma_j P[X_n = i] \cdot p_{ij}^{(m)} p_{jk} = P[X_n = i]\ p_{ik}^{(m+1)}$.
Dabei wurde die Induktionsvoraussetzung und wieder 5.25a benutzt. $\square$

5.27 Satz (Chapman–Kolmogoroff–Gleichungen). Für $P[X_n = i] > 0$ gilt:

(a) $P[X_{n+m} = k \mid X_n = i] = p_{ik}^{(m)}$;

(b) $P[X_{n+m} = k \mid X_n = i] = \Sigma_j P[X_t = j \mid X_n = i]\ P[X_{n+m} = k \mid X_t = j]$.

Aufgrund der Eigenschaft 5.27a heißen die $p_{ij}^{(m)}$ **m–Schritt–Übergangswahrscheinlichkeiten**.

Beweis: a) Dividiere in Lemma 5.26 durch $P[X_n = i]$.

b) Wegen $\mathbb{P}^m = \mathbb{P}^{t-n} \cdot \mathbb{P}^{n+m-t}$ gilt auch

(5.27c) $\quad p_{ik}^{(m)} = \Sigma_j\ p_{ij}^{(t-n)}\ p_{jk}^{(n+m-t)}$.

Nun folgt die Behauptung aus (a). $\square$

5.28 Korollar. $P[X_m = k] = \Sigma_i P[X_0 = i]\ p_{ik}^{(m)}$.

Insbesondere gilt: $P[X_1 = k] = \Sigma_i P[X_0 = i]\ p_{ik}$. Die Frage, wann die Verteilungen von X_0 und X_1 übereinstimmen, führt zu folgender Definition.

5.29 Def. Ein W–Maß auf S mit Z–Dichte $\{\nu_i\}$ heißt **stationäre Verteilung** , falls gilt:

(5.30) $\quad \nu_k = \Sigma_{i\in S}\ \nu_i\ p_{ik}\ \ (\forall\ k\in S)$.

5.31 Satz. Ist die Startverteilung eine stationäre Verteilung, so sind die X_n, $0 \le n < N$, identisch verteilt.

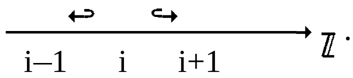
Beweis: Nach Voraussetzung definiert $P[X_0=i] = \nu_i$ eine stationäre Verteilung. Der Beweis folgt nun mit Induktion aus:

$$P[X_{n+1} = k] = \Sigma_i P[X_n = i]\ P[X_{n+1} = k \mid X_n = i] = \Sigma_i\ \nu_i\ p_{ik}.\ \square$$

5.32 Bemerkung. Ist S endlich, so existiert immer eine stationäre Verteilung.
Der Beweis wird in W–Theorie II (stochastische Prozesse) gebracht. $\square$

5.33 Satz. Sind (Y_n, $1 \le n < N$) diskrete iid Zva mit Werten in S' und existiert eine Funktion $f : S \times S' \mapsto S$, sodaß $X_n = f(X_{n-1}, Y_n)$, $1 \le n < N$, gilt, so ist (X_n) eine Markoff–Kette mit den Ü–W. $p_{ij} = P[f(i,Y_1) = j]$, falls X_0 konstant ist [oder allgemeiner falls auch $\{X_0, Y_n, 1 \le n < N\}$ unabhängig sind].

Beweis: Wir zeigen die Eigenschaft 5.23 (i):
$P[X_0 = i_0,\dots,X_n = i_n] = P[X_0 = i_0, f(X_0,Y_1)=i_1,\dots,f(X_{n-1},Y_n)=i_n]$
$= P[X_0 = i_0, f(i_0,Y_1)=i_1,\dots,f(i_{n-1},Y_n)=i_n] = P[X_0 = i_0] \cdot P[f(i_0,Y_1)=i_1] \cdot \dots \cdot P[f(i_{n-1},Y_n)=i_n]$;
denn es ist $\{X_0 = i_0\}$ entweder Ω oder $\emptyset$, und man kann offenbar $\{f(i_{m-1},Y_m)=i_m\} = \{Y_m \in B_m\}$ schreiben für gewisse $B_m \subset S'$. []

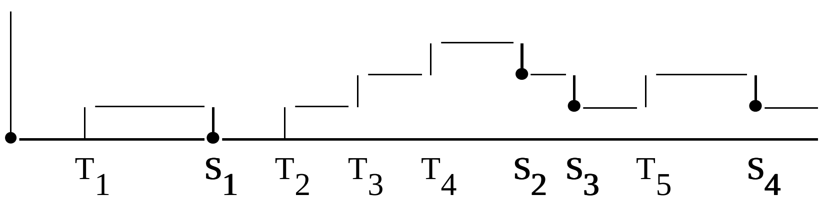
5.34 Beispiele für die Situation von 5.33.
(1) **Summen von iid Zva** erhält man, falls $S = \mathbb{Z} \supset S'$ und $f(i,y) = i + y$. Für den Fall $S' = \{-1.1\}$ heißt $\{X_n\}$ eine (1–dim.) **Irrfahrt** 

.

(2) **Ein Warteschlangenmodell.** Es sei $S = \mathbb{N}_0 \supset S'$ und $f(i,y) = (i-1)^+ + y$. In einem Bedienungssystem mögen zu Beginn X_0 Kunden auf Bedienung warten. Beschreibt Y_n die Anzahl der Ankünfte während der Bedienung des n–ten Kundens, so soll X_n die Länge der Warteschlange nach der n–ten Bedienung beschreiben. [Dabei können die Zeitpunkte der Ankünfte neuer Kunden durch einen Poisson–Prozeß beschrieben werden.] Somit haben wir

(5.35a) $X_{n+1} = (X_n - 1)^+ + Y_{n+1} =: f(X_n, Y_{n+1})$, insbesondere $X_{n+1} \ge X_n - 1$.

Ist $X_n = 0$, so muß man warten, bis wieder ein Kunde ankommt. Dann ist man ider Situation, daß die Warteschlange zu Beginn der Bedienung gleich eins ist.



$S_n \;\hat{=}$ Bedienungsbeendigungen
$T_n \;\hat{=}$ Ankunftszeiten

Es ergibt sich nach Satz 5.33:

(5.35b) $p_{ij} = P[(i-1)^+ + Y_1 = j]$

Wir möchten in dieser Situation eine notwendige Bedingung für die Existenz einer stationären Verteilung herleiten. Liegt eine stationäre Verteilung mit Z–Dichte (v_i) vor, so gilt:

(5.36) $\nu_j\, p_{jj-1} = \Sigma_{i=0}^{j-1} \nu_i \Sigma_{k=j}^{\infty} p_{ik}$ für $j \geq 1$.

Zum Beweis schreiben wir unter Benutzung von $p_{ik} = 0$ für $k < j < i$:

$\Sigma_{k<j} \nu_k = \Sigma_{k<j} \Sigma_i \nu_i p_{ik} = \nu_j p_{jj-1} + \Sigma_{i<j} \nu_i \Sigma_{k<j} p_{ik}$ sowie

$\Sigma_{i<j} \nu_i = \Sigma_{i<j} \nu_i \Sigma_k p_{ik} = \Sigma_{i<j} \nu_i \Sigma_{k<j} p_{ik} + \Sigma_{i<j} \nu_i \Sigma_{k\geq j} p_{ik}$.

Durch Vergleich dieser Identitäten folgt (5.**36**). Daraus erhalten wir
wegen $p_{jj-1} = 1 - \Sigma_{k=j}^{\infty} p_{jk}$:

(5.36)' $\nu_j = \Sigma_{i=0}^{j} \nu_i \Sigma_{k=j}^{\infty} p_{ik}$ für $j \geq 0$.

Damit ergibt sich

$$
\begin{aligned}
1 &= \Sigma_{j=0}^{\infty} \nu_j = \Sigma_{j=0}^{\infty} \Sigma_{i=0}^{j} \nu_i \Sigma_{k=j}^{\infty} p_{ik} \\
&= \Sigma_{i=0}^{\infty} \nu_i \Sigma_{k=i}^{\infty} p_{ik} \Sigma_{j=i}^{k} 1 = \Sigma_{i=0}^{\infty} \nu_i \Sigma_{k=i}^{\infty} (k-i+1)\, p_{ik} \\
&= \nu_0 \Sigma_{k=0}^{\infty} (k+1)\, P\lceil Y_1 = k\rceil + \Sigma_{i=1}^{\infty} \nu_i \Sigma_{k=i}^{\infty} (k-i+1)\, P\lceil i-1 + Y_1 = k\rceil \\
&= \nu_0 + \Sigma_{i=0}^{\infty} \nu_i \Sigma_{\ell=1}^{\infty} \ell\, P\lceil Y_1 = \ell\rceil = \nu_0 + E\lceil Y_1\rceil, \text{ also}
\end{aligned}
$$

(5.37) $\nu_0 = 1 - E\lceil Y_1\rceil$ mit $E\lceil Y_1\rceil = \Sigma_{\ell=1}^{\infty} \ell\, P\lceil Y_1 = \ell\rceil$.

Dabei ist $E\lceil Y_1\rceil$ die mittlere Anzahl der Ankünfte während einer Bedienung. Somit ergibt sich als notwendige Bedingung für die Existenz einer stationären Verteilung $E\lceil Y_1\rceil \leq 1$. Wäre $p_{jj-1} = P\lceil Y_1 = 0\rceil = 0$, so müßte $P\lceil Y_1 = 1\rceil = 1$ gelten; dieser Fall ist aber leicht zu übersehen und nicht interessant. Sei nun $p_{jj-1} = P\lceil Y_1 = 0\rceil > 0$. Ist dann $E\lceil Y_1\rceil = 1$, so ist $\nu_0 = 0$. Daraus folgt aber nach (5.36): $\nu_j = 0 \ \forall j$; dann definiert aber (ν_j) keine stationäre Verteilung. Also erhalten wir,
falls $P\lceil Y_1 = 1\lceil < 1$, als

notwendige Bedingung: $E\lceil Y_1\rceil < 1$.

In der Vorlesung W–theorie II (stochastische Prozesse) wird gezeigt, daß diese Bedingung auch hinreichend ist. Dies wird in den Übungen für den Speziallfall gezeigt, daß die Y_n geometrisch verteilt sind. Die Angabe von ν_0 in (5.37) erlaubt mit Hilfe von (5.36) eine rekursive Berechnung der Z–Dichte der stationären Verteilung.

KAPITEL II MASSTHEORIE.

§6 Mengensysteme.

Sprechweise: Ist $(\Omega,\mathfrak{F})$ ein Meßraum, so heißt jedes $A\in\mathfrak{F}$ eine ($\mathfrak{F}$–) **meßbare Menge.**

6.1 Bemerkung. Sind $\mathfrak{C}$ und $\mathfrak{C}'$ Mengensysteme in Ω, so gilt:
$\sigma(\mathfrak{C}) = \sigma(\mathfrak{C}') \Leftrightarrow \mathfrak{C} \subset \sigma(\mathfrak{C}'), \mathfrak{C}' \subset \sigma(\mathfrak{C})$.

Die σ–Algebra $\mathfrak{B}_d$ der Borelschen Mengen (vgl. Def. 2.7).

6.2 Bemerkung. $\mathfrak{B}_d$ enthält alle abgeschlossenen Mengen (als Komplemente offener Mengen), somit alle Einpunktmengen sowie alle abzb. Mengen (als abzb. Vereinigung von Einpunktm.).

6.3 Def. Zu $a = (a_1,...,a_d)$, $b = (b_1,...,b_d) \in \mathbb{R}^d$ definiere das **halboffene Intervall** $(a,b]$ als

$(a,b] := \times_{i=1}^{d} (a_i, b_i] \subset \mathbb{R}^d$ (:= $\emptyset$, falls $a_i \geq b_i$ für ein i); ferner

$J_d := \{(a,b]; a,b\in\mathbb{R}^d\}$, $\tilde{J}_d := \{(a,b]; a,b\in\mathbb{Q}^d\}$, dabei ist $\tilde{J}_d$ abzb..

6.4 Lemma.(a) $\mathfrak{B}_d = \sigma(J_d) = \sigma(\tilde{J}_d)$; insbesondere hat $\mathfrak{B}_d$ ein abzb. Erzeugendensystem.

(b) $(a,b]\cap(a',b'] = \times_{i=1}^{d} (\max(a_i,a_i'),\min(b_i,b_i')] \in J_d$.

Bew. a) (i) $\tilde{J}_d \subset J_d \subset \sigma(\mathfrak{T}_d)$ wegen $(a,b] = \cap_{m=1}^{\infty} (a, b+\frac{\vec{1}}{m})$.

(ii) $\mathfrak{T}_d \subset \sigma(\tilde{J}_d)$; denn sei C offen; dann $\forall\, c\in C\ \exists\, I^{(c)} \in \tilde{J}_d$ mit $c\in I(c) \subset C$, also
$C = \cup_{c\in C} I^{(c)} = \cup_{c\in C'} I^{(c)}$ für ein abzb. Menge C', denn $\tilde{J}_d$ ist abzb.. Jetzt 6.1.

Teil (b) ist klar $\square$

6.5 Bemerkung. Andere Erzeugendensysteme von $\mathfrak{B}_d$ sind (vgl. 6.15):

(i) $J_d' := \{(-\vec{\infty}, a] ; a\in\mathbb{R}^d\}$ (ii) $\tilde{J}_d' := \{ (-\vec{\infty}, a] ; a\in\mathbb{Q}^d\}$.

Produkt–σ–Algebren.

6.6 Def. Geg. Ω_i , $\mathfrak{C}_i \subset \mathfrak{P}(\Omega_i)$, σ–Algebren $\mathfrak{F}_i$ in Ω_i, $i\in\mathbb{N}$. Dann sei

$\times_1^n \mathfrak{C}_i := \{ \times_1^n C_i; C_i\in\mathfrak{C}_i\}$ für $n\in\overline{\mathbb{N}}$ (z.B $J_d = \times_1^d J_1$),

$\otimes_1^n \mathfrak{F}_i := \sigma(\times_1^n \mathfrak{F}_i)$, $n\in\overline{\mathbb{N}}$ ($\times_1^n \mathfrak{F}_i$ ist i.a. keine σ–Algebra).

$\otimes_1^n \mathfrak{F}_i$ heißt die (von den $\mathfrak{F}_i$ erzeugte) **Produkt–σ–Algebra** auf $\times_1^n \Omega_i$ und
$(\times_1^n \Omega_i, \otimes_1^n \mathfrak{F}_i)$ heißt **Produktmeßraum**.

6.7 Spezialfall. Für den Fall Ω_i abzb., $1\leq i\leq n<\infty$ gilt: $\otimes_1^n \mathfrak{P}(\Omega_i) = \mathfrak{P}(\times_1^n \Omega_i)$.

Denn $\times_1^n \Omega_i$ ist abzb, und $\times_1^n \mathfrak{P}(\Omega_i)$ enthält die Einpunktmengen
$\{(\omega_1,...,\omega_n)\} = \{\omega_1\}\times\{\omega_2\}\times...\times\{\omega_n\}$; jetzt 2.6.

Gilt $|\Omega_i| \geq 2$ so ist $\mathop{\times}_{i=1}^{\infty} \Omega_i$ überabzb. und $\mathfrak{P}(\mathop{\times}_{i=1}^{\infty} \Omega_i) \neq \bigotimes_1^{\infty} \mathfrak{P}(\Omega_i)$ (ohne Beweis).

Durch Abbildungen definierte σ–Algebren.

Schreibweise. Zu $f{:}\Omega \mapsto \Omega'$, $\mathfrak{C}' \subset \mathfrak{P}(\Omega')$ setze $f^{-1}(\mathfrak{C}') := \{f^{-1}(A'); A' \in \mathfrak{C}'\}$.

6.8 Satz. Geg. $f{:}\Omega \mapsto \Omega'$, $\mathfrak{F}'$ σ–Algebra auf Ω', $\mathfrak{C}' \subset \mathfrak{P}(\Omega')$. Dann gilt:

(a) $f^{-1}(\mathfrak{F}')$ ist σ–Algebra auf Ω;

(b) $f^{-1}(\sigma(\mathfrak{C}')) = \sigma(f^{-1}(\mathfrak{C}'))$.

Bew. a) folgt mit 4.2. b) "$\supset$" folgt aus (a). ad "$\subset$" Wir führen einen Schluß von $\mathfrak{C}'$ auf $\sigma(\mathfrak{C}')$ durch: Setze $\mathfrak{A}' := \{A' \subset \Omega';\ f^{-1}(A') \in \sigma(f^{-1}(\mathfrak{C}'))\}$. Dann gilt $\mathfrak{C}' \subset \mathfrak{A}'$, und es ist z.z. $\sigma(\mathfrak{C}') \subset \mathfrak{A}'$. Dies folgt aber, wenn $\mathfrak{A}'$ σ–Algebra ist. Dies ergibt sich wiederum mit 4.2. $\Box$

6.9 Beispiel. Seien $(\Omega_i, \mathfrak{F}_i)$ Meßräume, $i \in \mathbb{N}$. $\Omega = \mathop{\times}_1^n \Omega_i$ für ein $n \in \overline{\mathbb{N}}$, $\pi_i : \Omega \mapsto \Omega_i$ Projektion. Dann gilt etwa $\pi_1^{-1}(\mathfrak{F}_1) = \{A \times \mathop{\times}_2^n \Omega_i,\ A \in \mathfrak{F}_1\}$ (Zylindermengen).

Spur–σ–Algebren.

Schreibweise. Zu $\mathfrak{C} \subset \mathfrak{P}(\Omega)$, $B \subset \Omega$ setzte $B \cap \mathfrak{C} := \{B \cap C;\ C \in \mathfrak{C}\}$.
Sei $i : B \to \Omega$ die natürliche Injektion. Dann gilt: $i^{-1}(C) = B \cap C$, $i^{-1}(\mathfrak{C}) = B \cap \mathfrak{C}$.

6.10 Def. Ist $(\Omega, \mathfrak{F})$ ein Meßraum, $\emptyset \neq B \subset \Omega$. So heißt

$B \cap \mathfrak{F}$ die **Spur–σ–Algebra** von $\mathfrak{F}$ auf B.

Gemäß 6.8a ist $B \cap \mathfrak{F} = i^{-1}(\mathfrak{F})$ wirklich eine σ–Algebra.

(6.11) Für den Fall $B \in \mathfrak{F}$ gilt: $B \cap \mathfrak{F} = \{A \in \mathfrak{F};\ A \subset B\} \subset \mathfrak{F}$.

Aus 6.8b folgt:

(6.12) $B \cap \sigma(\mathfrak{C}) = \sigma(B \cap \mathfrak{C})$.

Beispiel. $(\Omega, \mathfrak{F}) = (\mathbb{R}^1, \mathfrak{B}_1)$, $\mathfrak{C} = J_1$, B Intervall.

Zusammenhang zwischen ⊗ und σ(...).

6.13 Satz. Geg. $\mathfrak{C}_i \subset \mathfrak{P}(\Omega_i)$, $1 \leq i \leq n \in \mathbb{N}$, sodaß $\forall$ i Folgen $\{C_{ik}, k \in \mathbb{N}\}$ existieren mit $C_{ik} \uparrow \Omega_i$.

Dann gilt: $\sigma(\mathop{\times}_1^n \mathfrak{C}_i) = \bigotimes_1^n \sigma(\mathfrak{C}_i)$ $\left[= \sigma(\mathop{\times}_1^n \sigma(\mathfrak{C}_i))\right]$.

Bew."$\subset$" klar. ad "$\supset$" Sei $\pi_i : \mathop{\times}_1^n \Omega_j \mapsto \Omega_i$ die Projektion, also $\pi_1^{-1}(C_1) = C_1 \times \mathop{\times}_2^n \Omega_i$. Aus $C_{jk} \uparrow \Omega_j$ folgt $\pi_1^{-1}(C_1) \in \sigma(\mathop{\times}_1^n \mathfrak{C}_i)$, also $\pi_i^{-1}(\mathfrak{C}_i) \subset \sigma(\mathop{\times}_1^n \mathfrak{C}_i)$ und damit gemäß 6.8b $\pi_i^{-1}(\sigma(\mathfrak{C}_i)) = \sigma(\pi_i^{-1}(\mathfrak{C}_i)) \subset \sigma(\mathop{\times}_1^n \mathfrak{C}_i)$. Aus $A_1 \times \ldots \times A_n = \cap_i \pi_i^{-1}(A_i)$ für $A_i \in \sigma(\mathfrak{C}_i)$ folgt die Beh. $\Box$

Anwendung auf $\mathfrak{B}_n$.

6.14 Korollar. $\mathfrak{B}_n = \otimes_1^n \mathfrak{B}_1$.

Bew. Wähle $\mathfrak{C}_i = J_1$ in 6.13 und benutze 6.4.

6.15 Korollar. $\mathfrak{B}_n = \sigma(J'_n)$.

Bew. Der Fall n=1 ist klar wegen $(-\infty,a] = \cup_n (-n,a]$ und $(a,b] = (-\infty,b] \setminus (-\infty,a]$.

Der Fall n>1 folgt nun aus 6.13 mit $\mathfrak{C}_i = J'_1$ und 6.14. $\square$

Die σ–Algebra $\overline{\mathfrak{B}}$ auf $\overline{\mathbb{R}}$.

6.16 Def. $\overline{\mathfrak{B}} := \sigma(\{[-\infty,a];\ a \in \mathbb{R}^1\})$.

(6.17) $\overline{\mathfrak{B}} = \{B, B \cup \{+\infty\}, B \cup \{-\infty\}, B \cup \{-\infty,+\infty\};\ B \in \mathfrak{B}_1\}$ (Übung)

Aus (6.17) oder (6.12) mit $\mathfrak{B}_1 = \sigma(J'_1)$ folgt sofort:

(6.18) $\mathbb{R}^1 \cap \overline{\mathfrak{B}} = \mathfrak{B}_1$.

Dynkinsysteme.

6.19 Def. $\mathfrak{C} \subset \mathfrak{P}(\Omega)$ heißt **∩–stabil**, falls: $A \cap B \in \mathfrak{C}\ \forall A, B \in \mathfrak{C}$.

Beispiele: J_d (vgl. 6.4b), J'_d, $\mathfrak{T}_d$.

6.20 Def. $\mathfrak{D} \subset \mathfrak{P}(\Omega)$ heißt **Dynkinsystem**, falls

(1) $\emptyset \in \mathfrak{D}$

(2) $A \in \mathfrak{D} \Rightarrow A^c \in \mathfrak{D}$

(3) $A_i \in \mathfrak{D}$, $i \in \mathbb{N}$, disjunkt $\Rightarrow \cup_1^\infty A_i \in \mathfrak{D}$.

6.21 Weitere Eigenschaften.

(4) $\Omega \in \mathfrak{D}$

(5) Sind $A_1, A_2 \in \mathfrak{D}$ und disjunkt, so gilt: $A_1 \cup A_2 \in \mathfrak{D}$

(6) $A, B \in \mathfrak{D}$, $B \subset A \Rightarrow A \setminus B \in \mathfrak{D}$ (wegen $A \setminus B = (A^c \cup B)^c$.

6.22 Lemma. Für $\mathfrak{D} \subset \mathfrak{P}(\Omega)$ gilt: $\mathfrak{D}$ σ–Algebra $\Leftrightarrow \mathfrak{D}$ ist ∩–stabiles Dynkinsystem.

Bew. "⇒" klar. "⇐" folgt aus dem Reduktionssatz 2.16.

6.23 Def. Für $\mathfrak{C} \subset \mathfrak{P}(\Omega)$ sei $\delta(\mathfrak{C}) := \bigcap \{\mathfrak{D};\ \mathfrak{D}$ ist Dynkinsystem, $\mathfrak{D} \supset \mathfrak{C}\}$, also das kleinste Dynkinsystem, das $\mathfrak{C}$ enthält..

Der folgende Satz wird ein wichtiges Beweisprinzip liefern.

6.24 Satz. Ist $\mathfrak{C} \subset \mathfrak{P}(\Omega)$ $\cap$–stabil, so ist $\delta(\mathfrak{C})$ σ–Algebra.

Bew. Wegen 6.22 genügt es z.z.: $\delta(\mathfrak{C})$ ist $\cap$–stabil.
Setze $\mathfrak{D}_B := \{A \subset \Omega;\ A \cap B \in \delta(\mathfrak{C})\}$. **Ziel:** $\delta(\mathfrak{C}) \subset \mathfrak{D}_B\ \ \forall\, B \in \delta(\mathfrak{C})$. Es gilt zunächst:
(*) $\mathfrak{D}_B$ ist Dynkinsystem $\forall\, B \in \delta(\mathfrak{C})$ (wegen $A^c \cap B = B \setminus A \cap B$ und $(\cup A_i) \cap B = \cup (A_i \cap B)$)
Wir führen einen Schluß von $\mathfrak{C}$ auf $\delta(\mathfrak{C})$ durch. Es gilt:

$\mathfrak{C}$ $\cap$–stabil $\Leftrightarrow$ $\mathfrak{C} \subset \mathfrak{D}_B\ \ \forall\, B \in \mathfrak{C}$ $\Rightarrow \delta(\mathfrak{C}) \subset \mathfrak{D}_B\ \forall\, B \in \mathfrak{C}$ wegen (*)

($\Leftrightarrow A \cap B \in \delta(\mathfrak{C})\ \ \forall\, B \in \mathfrak{C}\ \forall\, A \in \delta(\mathfrak{C})$) $\Leftrightarrow \mathfrak{C} \subset \mathfrak{D}_A\ \forall\, A \in \delta(\mathfrak{C})$

$\Rightarrow \delta(\mathfrak{C}) \subset \mathfrak{D}_A\ \ \forall A \in \delta(\mathfrak{C})$ wegen (*) $\Leftrightarrow$ Ziel. $\square$

6.25 Dynkin–Argument: Ist $\mathfrak{C}$ $\cap$–stabil und $\mathfrak{D}$ ein Dynkinsystem mit $\mathfrak{D} \supset \mathfrak{C}$,
dann gilt auch: $\mathfrak{D} \supset \sigma(\mathfrak{C})$.

Bew. Es gilt $\mathfrak{D} \supset \delta(\mathfrak{C})$ und nach 6.24 $\delta(\mathfrak{C}) \supset \sigma(\mathfrak{C})$ $\left[;\text{ dabei gilt stets } \delta(\mathfrak{C}) \subset \sigma(\mathfrak{C})\right]$. $\square$

Beweisprinzipien für σ–Algebren.

Sei $\mathfrak{C} \subset \mathfrak{P}(\Omega)$. Zu zeigen sei: alle $A \in \sigma(\mathfrak{C})$ haben eine bestimmte Eigenschaft E, falls alle $A \in \mathfrak{C}$ die Eigenschaft E haben. Man setze:
$\mathfrak{D} := \{A \subset \Omega;\ A \text{ hat Eigenschaft E}\}$; also z.z. $\mathfrak{D} \supset \sigma(\mathfrak{C})$. Dann gilt:

(1) $\mathfrak{C} \subset \mathfrak{D}$, $\mathfrak{D}$ σ–Algebra $\Rightarrow \mathfrak{D} \supset \sigma(\mathfrak{C})$;

(2) $\mathfrak{C} \subset \mathfrak{D}$, $\mathfrak{C}$ $\cap$–stabil, $\mathfrak{D}$ Dynkin–System $\Rightarrow \mathfrak{D} \supset \sigma(\mathfrak{C})$.

Prinzip (1) wurde bereits im Beweis von 6.8b angewandt; (2) ist gerade 6.25.

6.26 Beispiel. (Spezialfall des Eindeutigkeitssatzes für W–Maße).
Seien P_1 und P_2 W–Maße auf $(\mathbb{R}^n, \mathfrak{B}_n)$. Dann heißt $F_i(x) := P_i\left[(-\vec{\infty}, x]\right]$, $x \in \mathbb{R}^n$, die **Verteilungsfunktion** zu P_i. Es gilt:

Stimmen F_1 und F_2 überein, so sind auch P_1 und P_2 identisch.

Ein W–Maß auf $\mathfrak{B}_n$ ist also durch seine Verteilungsfunktion bestimmt.
Zum Beweis wählt man $\mathfrak{C} = J_n'$ und $\mathfrak{D} = \{B \in \mathfrak{B}_n;\ P_1[B] = P_2[B]\}$. Nach 6.15 gilt: $\mathfrak{B}_n = \sigma(J_n')$ und J_n' ist $\cap$–stabil. Ferner kann leicht gezeigt werden, daß $\mathfrak{D}$ ein Dynkinsystem ist. Nach Voraussetzung gilt auch: $J_n' \subset \mathfrak{D}$. Somit folgt mit 6.25: $\mathfrak{B}_n \subset \mathfrak{D}$, also $\mathfrak{B}_n = \mathfrak{D}$. $\square$

§7 Maße.

Rechnen in $\overline{\mathbb{R}}$.

$x \pm \infty := \pm\infty \ \ \forall\, x\in\mathbb{R}^1$; nicht erklärt sind: $\infty - \infty, -\infty + \infty$, $\quad a\cdot(\pm\infty) := (\pm\infty)\cdot a := \begin{cases} \pm\infty & 0<a\le\infty \\ 0 & \text{für } a=0 \\ \mp\infty & -\infty\le a<0 \end{cases}$,

$\frac{1}{0} := \infty$, $\frac{1}{\pm\infty} := 0$, $\frac{a}{b} := a\cdot\frac{1}{b}$, insbesondere $\frac{0}{0} = 0$.

Rechnen in $[0.\infty]$.

In $[0,\infty]$ sind "+" und "$\cdot$" assoziativ, kommutativ und distributiv.

Für $a_i\in[0,\infty]$, $i\in I$, I abzb., sei $\sum_{i\in I} a_i := \sup\{\sum_J a_j\ ;\ |J|<\infty, J\subset I\} = \sup \sum_{J_n} a_j$ .falls $J_n\uparrow I$.

Dann gilt: $\sum_1^\infty a_i$ absolut konvergent $\Leftrightarrow$ $\sum_1^\infty |a_i| \ \left[= \sup_n \sum_1^n |a_i|\right] < \infty$.

Weitere Eigenschaften:

Isotonie: $0\le a_i\le b_i\le\infty \Rightarrow \sum_I a_i \le \sum_I b_i$;

Semi–Linearität: $a_i, b_i, c \in [0,\infty] \ \Rightarrow\ \sum c a_i = c\cdot\sum a_i$, $\ \sum(a_i+b_i) = \sum a_i + \sum b_i$.

Umordnungsatz: $a_i \in [0,\infty]$, $I = \cup_{j\in J} I_j$, (I_j) disjunkt $\Rightarrow$ $\sum_I a_j = \sum_J \sum_{I_j} a_i$.

Zum **Bew.** des Umordnungssatzes betrachte die Fälle:
(i) $\sum a_i < \infty$ (ii) $\sum_J \sum_{I_j} a_i < \infty$ (ii) $\exists$ i mit $a_i=\infty$.

7.1 Def. Geg. $(\Omega,\mathfrak{F})$ Meßraum. Eine Abb. $\mu : \mathfrak{F} \mapsto \overline{\mathbb{R}}$ heißt **Maß** und $(\Omega,\mathfrak{F},\mu)$ **Maßraum**, wenn gilt: μ ist
(1) **nulltreu**: $\mu[\emptyset]=0$
(2) **positiv**: $\mu[A]\ge 0 \ \ \forall\, A\in\mathfrak{F}$;
(3) **σ–additiv**: $\mu[\cup_1^\infty A_i] = \sum_1^\infty \mu[A_i] \ \ \forall\, A_i\in\mathfrak{F}$ (paarw.) disjunkt.

7.2 Lemma. Sei $(\Omega,\mathfrak{F},\mu)$ Maßraum, A_i, $i\in\mathbb{N}$, A, B $\in\mathfrak{F}$; dann ist μ :

(4) **additiv**: $A_1,\dots,A_n \in \mathfrak{F}$, disjunkt $\Rightarrow \mu[\cup_1^n A_i] = \sum_1^n \mu[A_i]$;

(5) **isoton**: $\mu[A] \le \mu[B]$ falls $A\subset B$;
(6) **subtraktiv**: $\mu[B\setminus A] = \mu[B] - \mu[A]$, falls $A\subset B$, $\mu[B]<\infty$;
(7) **modular**: $\mu[A] + \mu[B] = \mu[A\cup B] + \mu[A\cap B]$;
(8) **subadditiv:** $\mu[A\cup B] \le \mu[A]+\mu[B]$;
(9) **σ–subadditiv**: $\mu[\cup A_i] \le \sum_i \mu[A_i]$:

(10) $\quad\mu[\cup A_i] = 0$, falls $\mu[A_i]=0 \ \ \forall$ i;

(11) $\quad\mu[\cap A_i] = 1$, falls μ W–Maß und $\mu[A_i]=1 \ \ \forall$ i;

(12) **stetig von unten**: $\mu[A_i] \uparrow \mu[A]$, falls $A_i \uparrow A$;

(13) **stetig von oben**: $\mu[A_i] \downarrow \mu[A]$, falls $A_i \downarrow A$ und $\exists$ k mit $\mu[A_k] < \infty$;

(14) **$\emptyset$–stetig**: $\mu[A_i] \downarrow 0$, falls $A_i \downarrow \emptyset$ und $\exists$ k $\mu[A_k] < \infty$.

Bew. von (4) – (11) vgl. §2.
ad (12): Sei $(A_i) \uparrow$, also: $A_i = \cup_1^i A_j$. Wähle B_i wie im Reduktionssatz 2.16; dann folgt:
$\mu[A_i] = \mu[\cup_1^i B_j] = \Sigma_1^i \mu[B_j] \uparrow \Sigma_1^\infty \mu[B_j] = \mu[\cup_1^\infty B_j] = \mu[A]$.
ad (13): o.E. k=1; betrachte jetzt $A_i' := A_1 \setminus A_i$ und verwende (12). []

7.3 Def. Sei $(\Omega,\mathfrak{F},\mu)$ Maßraum, $\mathfrak{C} \subset \mathfrak{F}$. Dann heißt μ
(i) **endlich**, falls $\mu[\Omega] < \infty$,
(ii) **σ–endlich auf $\mathfrak{C}$**, falls $\exists$ $A_i \in \mathfrak{C}$, $i \in \mathbb{N}$, mit $A_i \uparrow \Omega$ und $\mu[A_i] < \infty$ $\forall$ i,
(iii) **σ–endlich**, falls μ σ–endlich auf $\mathfrak{F}$.

7.4. Beispiele. (1) $\mu \equiv 0$. (2) W–Maße sind normierte Maße.
(3) Ist Ω abzb., so heißt $\mu[A] := |A|$, $A \subset \Omega$, **Zählmaß** auf $\mathfrak{P}(\Omega)$. Dann ist μ σ–endlich .
(4) Ist Ω abzb., $p : \Omega \mapsto [0,\infty]$, so ist $\mu[A] = \Sigma_{i \in A}\, p(i)$ ein Maß auf $\mathfrak{P}(\Omega)$ ($\Leftrightarrow$ Umordnungssatz).

7.5 Eindeutigkeitssatz für Maße. Sei $\mathfrak{C} \subset \mathfrak{P}(\Omega)$ $\cap$–stabil. Dann gilt:
(a) Zwei W–Maße P_1 und P_2 auf $\sigma(\mathfrak{C})$, die auf $\mathfrak{C}$ übereinstimmen, stimmen auf $\sigma(C)$ überein.
(b) Seien μ_1 und μ_2 zwei Maße auf $\sigma(\mathfrak{C})$, so daß μ_1 σ–endlich auf $\mathfrak{C}$ ist. Stimmen dann μ_1 und μ_2 auf $\mathfrak{C}$ überein, so auch auf $\sigma(\mathfrak{C})$.

Bew. a) vgl. Beispiel 6.26.
b) Reduktion auf (a) mit $P_{in}[B] := \mu_i[B \cap A_n]/\mu_i[A_n]$, wobei (A_n) wie in 7.3(ii). []

7.6 Def. Ein Maß μ auf $(\mathbb{R}^d,\mathfrak{B}_d)$ heißt **Borel–Maß**, falls $\mu((a,b]) < \infty$ $\forall$ $(a,b] \in J_d$.
[Ist $(E,\mathfrak{T})$ ein top. Raum, so heißt ein Maß μ auf $\sigma(\mathfrak{T})$ Borel–Maß, falls $\mu[K] < \infty$ $\forall$ K kompakt].

7.7 Anwendungen von 7.5.
(a) Jedes Borel–Maß auf $\mathfrak{B}_d$ ist durch seine Werte auf J_d festgelegt.
(b) Jedes W–Maß P auf $\mathfrak{B}_d$ ist durch seine Werte auf J_d' festgelegt, also durch die Funktion

$$F(x)= := P[(-\vec{\infty},x]],\ x \in \mathbb{R}^d.$$

$F : \mathbb{R}^d \mapsto [0,1]$ heißt die zu P gehörige **Verteilungsfunktion** (vgl. Bsp. 6.26).

(c) Jedes W–Maß auf $\mathfrak{F}_1 \otimes \mathfrak{F}_2$ ist durch seine Werte auf $\mathfrak{F}_1 \times \mathfrak{F}_2$ bestimmt.

Die zu der Gleichverteilung U(0,1) auf (0,1) gehörige Verteilungsfunktion F ist offenbar:

$F(x) = x,\ 0 \le x \le 1,\ F(x) = 0,\ x \le 0,\ F(x) = 1,\ x \ge 1.$

7.8 Def. $F : \mathbb{R} \to [0,1]$ heißt (1–dim.) **Verteilungsfunktion** (Vf), falls

(i) F isoton,

(ii) F rechtsstetig,

(iii) $F(-\infty) := \lim_{x\to-\infty} F(x) = 0$, $F(+\infty) := \lim_{x\to+\infty} F(x) = 1$.

7.9 Lemma. Ist P W–Maß auf $\mathfrak{B}_1$, so ist $F(x) = P[(-\infty,x]]$ Vf im Sinne von 7.8.

Der **Beweis** folgt aus der Stetigkeit von oben und unten eines W–Maßes.

7.10 Def. Ist F (1–dim) Vf, so heißt

$F^{-} : (0,1) \mapsto \mathbb{R}$, $F^{-}(u) := \inf\{x\in\mathbb{R};\ F(x)\ge u\}$ **verallgemeinerte Inverse** von F

7.11 Satz. Sei F (1–dim) Vf. Dann gehört zur Verteilung der Zva F^{-} auf $(\Omega,\mathfrak{F},P) := ((0,1),(0,1)\cap\mathfrak{B}_1,U(0,1))$ gerade die Verteilungsfunktion F.

Bew. F^{-} ist Zva (als isotone Funktion: vgl. 8.4). die Aussage folgt nun aus [vgl. Übung]:

(7.12) $\qquad F^{-}(u)\le x \iff F(x) \ge u\ ,0<u<1.$ □

Der Beweis von 7.11 löst zugleich das in 1.16 und 4.8 angesprochene Problem. Zu einem beliebigen W–Maß Q auf $\mathfrak{B}_1$ mit Vf F wähle als Zva X auf $(\Omega,\mathfrak{F},P)$ gerade $X = F^{-}$. Dann gilt nach 7.11: $X \sim Q$. Will man $(0,1]$ statt (0,1) haben, so kann X(1) beliebig definiert werden wegen $\lambda_1[\{1\}] = 0$.

7.13 Fortsetzungssatz. Zu jeder 1–dim. Vf F existiert (genau ein) W–Maß auf $\mathfrak{B}_1$, das F als Vf hat.

Bew. Das gesuchte W–Maß ist gemäß 7.11 gerade die Verteilung von F^{-} aufgefaßt als Zva auf $((0,1),(0,1)\cap\mathfrak{B}_1,U(0,1))$. Dabei wurde die Existenz von U(0,1), also des Lebesgue–Maßes benutzt. □

Gemäß 7.13 ist die durch $F(x) = P[(-\infty,x]]$, $x\in\mathbb{R}$, def. Abb. $P \mapsto F$ eine Bijektion zwischen der Menge der W–Maße auf $\mathfrak{B}_1$ und der Menge der (1–dim) Vf. Es gilt dann:

$P[(a,b]) = F(b) - F(a),\quad P[(b,\infty)] = 1 - F(b).$

7.14 Verallgemeinerung auf Borel–Maße. $G : \mathbb{R}^1 \mapsto \mathbb{R}^1$ heißt **maßdefinierende Fkt,** falls G isoton und rechtsstetig ist. Ist μ Borel–Maß auf $\mathfrak{B}_1$, $c\in\mathbb{R}$ beliebig, so wird durch $G(x):=c+\mu[(0,x]]$ für $x\ge0$ und $G(x):=c-\mu[(x,0]]$ für $x<0$ eine maßdefinierende Fkt definiert mit

(7.15) $\qquad \mu[(a,b]] = G(b)-G(a)\ ,\ \forall a,b,\ a<b.$

Ist G eine maßdefinierende Fkt, so existiert genau ein Borelmaß μ mit (7.15) (Übung,).

7.16 Beispiel: $G(x) = x$ bei λ^1 (hier ist $c = 0$ zu wählen).

7.17 Bemerkung. Ist $\mu[(-\infty.0]] < \infty$, so ist in 7.14 $c = \mu[(-\infty.0]]$, also $G(x)=\mu[(-\infty,x]]$ wählbar.

7.18 Bemerkung. In der Situation von 7.14 gilt wegen $(x-\frac{1}{n}, x] \downarrow \{x\}$:

$$\mu[\{x\}] = G(x)-G(x-0)\,, \quad \mu[(a,b)] = G(b-0) - G(a) \quad \forall\, a,b,\ a<b.$$

7.19 Def. Ein Maß μ heißt **stetig**, falls $\mu[\{x\}]=0 \ \ \forall x$.

7.20 Folgerung: Ein W–Maß P [bzw. ein Borelmaß μ] auf $(\mathbb{R}^1,\mathfrak{B}_1)$ ist genau stetig, wenn die Vf. [bzw. eine im Sinne von 7.15 zu μ gehörende maßdefinierende Fkt.] stetig ist.

§8 Meßbare Funktionen.

8.1 Def. Seien $(\Omega,\mathfrak{F})$ und $(\Omega',\mathfrak{F}')$ Meßräume.

Dann heißt $f : \Omega \mapsto \Omega'$ $(\mathfrak{F}-\mathfrak{F}'-)$ meßbar (mb) (auch $\mathfrak{F}$–mb), falls $f^{-1}(\mathfrak{F}') \subset \mathfrak{F}$.

Schreibweise: $f^{-1}(A') = \{f\in A\}$, $\mu[\{f\in A\}] = \mu[f\in A]$.

Zufallsvariablen sind gerade mb Funktionen auf einem W–Raum. Jede konstante Fkt $f\equiv c$ ist mb, wegen $f^{-1}(A') = \Omega$ bzw. $\emptyset$ falls $c \in A'$ bzw. $c \notin A'$. Eine Indikatorfkt 1_A ist genau dann mb, wenn A mb. Gemäß Beispiel 6.9 sind Projektionen mb.

8.2 Satz. Geg. Meßräume $(\Omega,\mathfrak{F})$, $(\Omega',\mathfrak{F}')$ mit $\mathfrak{F}'=\sigma(\mathfrak{E}')$, Dann gilt: $f : \Omega \mapsto \Omega'$ mb $\Leftrightarrow f^{-1}(\mathfrak{E}') \subset \mathfrak{F}$.

Bew. für "$\Leftarrow$" 6.8b. $\square$

8.3 Korollar. Sei $(\Omega,\mathfrak{F})$ Meßraum, $f:\Omega\mapsto\overline{\mathbb{R}}$. Dann ist f genau dann $\mathfrak{F}$–$\overline{\mathfrak{B}}$–mb, wenn gilt:

$\{f\leq x\} \in \mathfrak{F} \quad \forall x\in\mathbb{R}$.

Bew. 8.2 und 6.16 mit $f^{-1}([-\infty,x]) = \{f \leq x\}$. $\square$

8.4 Anwendung. Jede monotone Fkt $f : \mathbb{R} \mapsto \overline{\mathbb{R}}$ ist mb..
Denn falls z.B. f isoton , gilt: $\forall x\ \exists y$ mit $\{f\leq x\} = (-\infty,y]$ oder $(-\infty,y)$ $\square$.

8.5 Anwendung. Sei $(\Omega,\mathfrak{F})$ Meßraum und $f, g : \Omega \mapsto \overline{\mathbb{R}}$ $\mathfrak{F}$–mb.

Dann liegen in $\mathfrak{F}$: $\{f<g\}$, $\{f\leq g\}$, $\{f=g\}$, $\{f\neq g\}$.

Bew. $\{f<g\} = \cup_{r\in\mathbb{Q}} \{f\leq r\}\setminus\{g\leq r\}$, $\{f\leq g\}=\{g<f\}^c$, $\{f=g\}=\{f\leq g\}\cap\{g\leq f\}$. $\square$

8.6 Satz. Stetige Funktionen sind mb.; d.h. sind $(\Omega,\mathfrak{T})$, $(\Omega',\mathfrak{T}')$ top. Räume und ist $f : \Omega \mapsto \Omega'$ stetig, so ist f $\sigma(\mathfrak{T})$–$\sigma(\mathfrak{T}')$–mb.

Bew. $f^{-1}(\mathfrak{T})\subset\mathfrak{T}'$, jetzt 8.2. $\square$

8.7. Satz. Die Hintereinanderschaltung von mb Abb. ist mb.; d.h. sind $f : \Omega \mapsto \Omega'$ $\mathfrak{F}$–$\mathfrak{F}'$–mb und $g : \Omega' \mapsto \Omega''$ $\mathfrak{F}'$–$\mathfrak{F}''$–mb, so ist $g\circ f$ $\mathfrak{F}$–$\mathfrak{F}''$–mb.

Bew. $(f\circ g)^{-1}=g^{-1}\circ f^{-1}$ (4.2e). $\square$

8.8 Anwendung. Geg $f : \Omega \mapsto \Omega'$ mit $f(\Omega) \subset \Omega^o \subset \Omega'$; dann gilt:

f $\mathfrak{F}$–$\mathfrak{F}'$–mb $\Leftrightarrow$ $f : \Omega \mapsto \Omega^o$ $\mathfrak{F}$–$\Omega^o\cap\mathfrak{F}'$–mb.

Bew. Sei $i:\Omega^o\mapsto\Omega'$ Injektion, dann $i^{-1}(\mathfrak{F}') = \Omega^o\cap\mathfrak{F}'$. Betrachte nun $i\circ f$. $\square$

8.9 Beispiel. $\Omega'=\overline{\mathbb{R}}$, $\Omega^o=\mathbb{R}$ mit (6.18).

8.10 Satz. Geg. $(\Omega,\mathfrak{F})$, $(\Omega_i,\mathfrak{F}_i)$, $f_i : \Omega \mapsto \Omega_i$, $i\in\mathbb{N}$, $f=(f_1,...,f_n)$ oder $f=(f_1,f_2,...)$. Dann gilt für $n\in\overline{\mathbb{N}}$:

f ist $\mathfrak{F}$–$\bigotimes_1^n \mathfrak{F}_i$–mb $\Leftrightarrow$ f_i ist $\mathfrak{F}$–$\mathfrak{F}_i$–mb $1\le i\le n$.

Bew. $f^{-1}(\times_1^n A_i) = \cap f_i^{-1}(A_i)$. Für "$\Leftarrow$" benutze 8.2; für "$\Rightarrow$" setze $A_j=\Omega_j$ $\forall\, j\neq i$. $\square$

8.11 Bemerkung. Die Abb. $(x,y) \mapsto g(f_1(x),f_2(y))$ von $\Omega_1\times\Omega_2$ nach Ω'' ist $\mathfrak{F}_1\otimes\mathfrak{F}_2$–$\mathfrak{F}''$–mb, wenn

$f_i : \Omega_i \mapsto \Omega_i'$ $\mathfrak{F}_i$–$\mathfrak{F}_i'$–mb, i=1,2, und $g:\Omega_1'\times\Omega_2'\mapsto\Omega''$ $\mathfrak{F}_1'\otimes\mathfrak{F}_2'$–$\mathfrak{F}''$–mb.

Bew. $(x,y) \mapsto x \mapsto f_1(x)$ und $(x,y) \mapsto y \mapsto f_2(y)$ sind mb. Jetzt 8.10 und 8.7. $\square$

8.12. Lemma. Stückweise mb Fkt sind mb: D.h. seien $(\Omega,\mathfrak{F})$, $(\Omega'.\mathfrak{F}')$ Meßräume, I abzb und

$\Omega = \cup_{i\in I} A_i$ mit $A_i\in\mathfrak{F}$. Für $f : \Omega \mapsto \Omega'$ gilt dann: f $\mathfrak{F}$–$\mathfrak{F}'$–mb $\Leftrightarrow$ $f|_{A_i}$ $A_i\cap\mathfrak{F}$–$\mathfrak{F}'$–mb $\forall i\in I$.

Bew. Übung

8.13 Satz. Seien $f,g : \Omega \mapsto \overline{\mathbb{R}}$ $\mathfrak{F}$–mb, $a,b\in\mathbb{R}$. Dann sind $af+b$, $f\cdot g$, f/g und $f\pm g$, falls definiert, $\mathfrak{F}$–mb. .

Bew. Es genügt z.z. (a)$f\cdot g$ mb (b) $f+g$ mb (c) $q:\overline{\mathbb{R}}\mapsto\overline{\mathbb{R}}$ mit $q(x)=1/x$ mb; denn z.B. gilt: $f/g=f\cdot q\circ g$.
ad (a) Definiere $A_e:=\{|f|<\infty,\ |g|<\infty\}$ und $A_i:=\{f\cdot g=i\}$, $i=0,\pm\infty$. Dann gilt $\Omega = \cup_i A_i$ mit $A_i\in\mathfrak{F}$. Letzteres sieht man für $i=0,\pm\infty$ mit Hilfe einer weiteren Zerlegung durch Fallunterscheidung. Nun gilt $f\cdot g|_{A_e} = h(f|_{A_e},g|_{A_e})$ mit h stetig. Jetzt 8.6, 8.7, 8.10, 8.12.

ad (b): wie (a). ad (c): verwende 8.3 oder 8.12 und 8.6. $\square$

8.14 Satz. Seien $f_n:\Omega\mapsto\overline{\mathbb{R}}$ $\mathfrak{F}$–mb, $n\in\mathbb{N}$. Dann sind $\sup f_n$, $\inf f_n$, $\underline{\lim}\, f_n$, $\overline{\lim}\, f_n$ und

(im Falle der Existenz) $\lim f_n$ $\mathfrak{F}$–mb.

Bew. $\{\sup f_n\le x\} = \cap\, \{f_n\le x\}$, $\inf f_n = -\sup(-f_n)$, $\underline{\lim}\, f_n = \sup_n \inf_{m\ge n} f_m$. $\square$

8.15 Def. Für $x \in \bar{\mathbb{R}}$ sei $x^{\pm} := \max(0, \pm x) \geq 0$, und für $f : \Omega \mapsto \bar{\mathbb{R}}$ sei $f^{\pm}(x) := (f(x))^{\pm}$, also $f = f^{+} - f^{-}$.

8.16 Lemma. Für $f : \Omega \mapsto \bar{\mathbb{R}}$ gilt: f $\mathfrak{F}$–mb $\Leftrightarrow$ f^{+}, f^{-} $\mathfrak{F}$–mb $\Rightarrow$ $|f|$ $\mathfrak{F}$–mb.

Bew. Die Abb $\alpha_{\pm}(x) = x^{\pm}$ sind monoton, also mb., und $f^{\pm} = \alpha_{\pm} \circ f$, $|f| = f^{+} + f^{-}$. []

8.17 Def. Sei $(\Omega, \mathfrak{F})$ Meßraum.

$\mathfrak{M}_0 := \{f : \Omega \to \mathbb{R}\,;\ f \text{ mb},\ f(\Omega) \subset [0,\infty),\ |f(\Omega)| < \infty\}$. Jedes $f \in \mathfrak{M}_0$ heißt **primitiv.**

$\mathfrak{M}_+ := \{f : \Omega \to \bar{\mathbb{R}}\,;\ f \text{ mb},\ f(\Omega) \subset [0,\infty]\}$.

Für $f \in \mathfrak{M}_0$ heißt $\quad f = \Sigma_{a \in f(\Omega)}\, a \cdot \mathbf{1}_{\{f=a\}}$ **Normaldarstellung von f**.

$\Pi_n : \mathfrak{M}_+ \mapsto \mathfrak{M}_0$ sei def. durch

$$\Pi_n(f) := \Sigma_{i=1}^{n2^n} \frac{i-1}{2^n}\, \mathbf{1}_{\{\frac{i-1}{2^n} \leq f < \frac{i}{2^n}\}} + n\, 1_{\{f \geq n\}}\,.$$

8.18 Lemma. Für $\Pi_n : \mathfrak{M}_+ \mapsto \mathfrak{M}_0$ gilt:

(a) $\quad 0 \leq \Pi_n(f) \leq n$,

(b) $\quad \Pi_n(f) \leq f$, $\Pi_n(f) \uparrow f$ $(n \to \infty)$,

(c) $\quad \Pi_n(f) \leq \Pi_n(g)$ falls $f \leq g$.

§9 Der Integralbegriff.

Sei $(\Omega, \mathfrak{F}, \mu)$ ein Maßraum.

<u>**Integrale primitiver Funktionen**</u>.

Die Definition des Integrals $\int f d\mu$ soll so erfolgen, daß es linear ist und μ fortsetzt im Sinne von

(9.1) $\int 1_A d\mu := \mu[A]$, $A \in \mathfrak{F}$.

Dies führt über die Normaldarstellung notwendigerweise zu:

9.2 Def. Für $f \in \mathfrak{M}_0$ sei: $\int f d\mu := \Sigma_{a \in f(\Omega)}\, a \cdot \mu[f=a] \in [0,\infty]$.

9.3 Lemma. (a) $\int (af) d\mu = a \cdot \int f d\mu$, $\quad 0 \leq a < \infty$, $f \in \mathfrak{M}_0$.

(b) $\int (f+g) d\mu = \int f d\mu + \int g d\mu$, $\quad f, g \in \mathfrak{M}_0$.

(c) $f, g \in \mathfrak{M}_0$, $f \leq g$ $\Rightarrow$ $\int f d\mu \leq \int g d\mu$.

(d) $\int (\Sigma_1^n a_i 1_{A_i})\, d\mu = \Sigma_1^n a_i\, \mu[A_i]$, $\quad 0 \leq a_i < \infty$, $A_i \in \mathfrak{F}$.

Beweis. a) leicht; b),c) mit $\int h(f,g)\, d\mu = \sum_{a,b} h(a,b)\cdot\mu[f{=}a,g{=}b]$, dieser Ausdruck ist linear und isoton in h; d) folgt aus (9.1), sowie (a) und (b). $\square$

Integrale nicht–negativer mb Funktionen.

Die Definition des Integrals soll so erfolgen, daß es auf $\mathfrak{M}_+$ stetig von unten ist. Dies führt in Hinblick auf 8.18b notwendigerweise zur:

9.4 Def. Für $f\in\mathfrak{M}_+$ sei: $\int f d\mu := \lim_n \int \Pi_n(f) d\mu \in [0.\infty]$.

(Nach 8.18(c) und 9.3c ist die Folge $\int\Pi_n(f)d\mu$ isoton.) Der folgende grundlegende Satz zeigt, daß die Definition unabhängig ist von der speziellen approximierenden Folge .

9.5 Satz. Geg. sei $f\in\mathfrak{M}_+$ und $(f_n)\subset\mathfrak{M}_0$ mit $f_n\uparrow f$. Dann gilt: $\lim_n \int f_n d\mu = \int f d\mu$.

Bew.(i) $1_{A_n}\uparrow 1_A \Rightarrow \int 1_{A_n} d\mu \uparrow \int 1_A d\mu$ wegen der Stetigkeit von unten von μ.

(ii) $B_n\uparrow\Omega$, $g\in\mathfrak{M}_0 \Rightarrow \int g 1_{B_n} d\mu \uparrow \int g d\mu$ wegen 9.3d sowie $A\cap B_n\uparrow A$ und (i).

(iii) $g\in\mathfrak{M}_0$, $g \le \lim f_n \Rightarrow \int g d\mu \le \lim \int f_n d\mu$. Bew.:

Zu $\alpha<1$ def $B_n=\{f_n \ge \alpha g\}\in\mathfrak{F}$, dann $B_n\uparrow\Omega$ also nach (ii): $\lim \int g 1_{B_n} d\mu = \int g d\mu$. Andererseits:

$f_n \ge \alpha g 1_{B_n} \in\mathfrak{M}_0\ \forall n \Rightarrow \int f_n d\mu \ge \alpha\int g 1_{B_n} d\mu\ \forall n \Rightarrow \lim \int f_n d\mu \ge \alpha\cdot\lim \int g 1_{B_n} d\mu = \alpha\int g d\mu$.

(iv) $\Pi_k(f) \le f = \lim f_n\ \forall k \Rightarrow \int\Pi_k(f)d\mu \le \lim \int f_n d\mu\ \forall k$ wegen (iii) $\Rightarrow \lim_k \int\Pi_k(f)d\mu \le \lim_n \int f_n d\mu$.

Ebenso "$\ge$" Vertausche $\Pi_k(f)$ und f_k. $\square$

9.6 Folgerung. Die Definitionen von $\int f d\mu$ für $f\in\mathfrak{M}_0$ und $f\in\mathfrak{M}_+$ sind konsistent.

Bew. Zu $f\in\mathfrak{M}_0$ wähle $f_n=f\ \forall n$ in 9.5.

9.7 Lemma. Sei $f,g\in\mathfrak{M}_+$, $0\le a<\infty$. Dann gilt:

(a) $\int(af)d\mu = a\cdot\int f d\mu$, (b) $\int(f+g)d\mu = \int f d\mu + \int g d\mu$ (c) $f\le g \Rightarrow \int f d\mu \le \int g d\mu$.

Bew. a) $\mathfrak{M}_o \ni a\cdot\Pi_n(f) \uparrow a\cdot f$.

b) $\mathfrak{M}_0 \ni \Pi_n(f)+\Pi_n(g) \uparrow f+g$. c) 8.18c. $\square$

Integrierbare und quasi–intb. Funktionen.

In allgemeinen Fall soll die Definition des Integrals wieder so erfolgen, daß es linear ist. Dies führt wegen der Zerlegung $f = f^+ - f^-$ zur

9.8 Def. Sei $f : \Omega \mapsto \overline{\mathbb{R}}$ mb.

(a) f heißt (μ–) **quasi–intb** $:\Leftrightarrow f \in \mathfrak{L}'_\mu :\Leftrightarrow \int f^+ d\mu < \infty$ oder $\int f^- d\mu < \infty \Leftrightarrow: \int f d\mu$ existiert.

Für $f \in \mathfrak{L}'_\mu$ definiert man: $\int f d\mu \ (= \int (f^+ - f^-) d\mu) := \int f^+ d\mu - \int f^- d\mu$.

(b) f heißt (μ–) **intb** $:\Leftrightarrow f \in \mathfrak{L}_\mu :\Leftrightarrow \int f^+ d\mu < \infty$ und $\int f^- d\mu < \infty \Leftrightarrow: \int f d\mu$ ist endlich.

9.9 Bemerkung. $\mathfrak{M}_+ \subset \mathfrak{L}'_\mu$; $\mathfrak{M}_0 \subset \mathfrak{L}_\mu \Leftrightarrow \mu$ ist endlich.

9.10 Satz. Seien $f,g : \Omega \mapsto \overline{\mathbb{R}}$ mb, $a,M \in \mathbb{R}$. Dann gilt:

(a) $f \in \mathfrak{L}_\mu \Leftrightarrow \int |f| d\mu < \infty$,

(b) $f \in \mathfrak{L}'_\mu \Rightarrow |\int f d\mu| \le \int |f| d\mu$,

(c) $f \le M$, μ endlich $\Rightarrow f \in \mathfrak{L}'_\mu$, $\int f d\mu \le M \cdot \mu[\Omega]$,

(d) $f \le g$, $f,g \in \mathfrak{L}'_\mu \Rightarrow \int f d\mu \le \int g d\mu$ (**Isotonie**),

(e) $|f| \le g$, $g \in \mathfrak{L}_\mu \Rightarrow f \in \mathfrak{L}_\mu$ (**Majorantenkriterium**),

(f) $f \in \mathfrak{L}'_\mu \Rightarrow af \in \mathfrak{L}'_\mu$ und $\int (af) d\mu = a \cdot \int f d\mu$ (**Linearität**),

(g) $f,g \in \mathfrak{L}'_\mu$ so daß f+g und $\int f d\mu + \int g d\mu$ definiert sind $\Rightarrow$

$f+g \in \mathfrak{L}'_\mu$ und $\int (f+g) d\mu = \int f d\mu + \int g d\mu$ (**Linearität**).

Bew. a) $\int |f| d\mu = \int f^+ d\mu + \int f^- d\mu$;

b) o.E. $\int |f| d\mu < \infty$; $|a-b| \le a+b$ für $a,b \ge 0$;

d) $f \le g \Rightarrow f^+ \le g^+$, $f^- \ge g^-$;

c) $\int M d\mu = M\mu[\Omega]$, jetzt (d);

e) wegen (a) und (d);

f) $a \ge 0 \Rightarrow (af)^\pm = a \cdot f^\pm$, $a < 0 \Rightarrow (af)^\pm = -a \cdot f^\mp$

g) $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$, Rest: Übung. []

9.11 Bemerkung. Jede Summe/Reihe ist Integral bzgl. des Zählmaßes.

Sei $(\Omega,\mathfrak{F}) = (\mathbb{N}_0, \mathfrak{P}(\mathbb{N}_0))$, μ Zählmaß auf Ω.

Fall 1: $0 \le f < \infty$ auf $\mathbb{N}_0$. Setze $f_n := f \cdot \mathbf{1}_{\{0,\dots,n\}} = \Sigma_{0 \le m \le n} f(m) 1_{\{m\}}$, dann ist f_n primitiv und $f_n \uparrow f$, also folgt mit 9.5: $\int f_n d\mu \uparrow \int f d\mu$;

andererseits gilt: $\int f_n d\mu = \Sigma_{0 \le m \le n} f(m) \mu[\{m\}] = \Sigma_{0 \le m \le n} f(m) \uparrow \Sigma_{0 \le m < \infty} f(m)$.

Fall 2: $f \ge 0$, $\exists k$ mit $f(k) = \infty$, dann folgt: $\int f d\mu \ge M \cdot \mu[\{k\}] = M \ \forall M$. Also gilt: $\int f d\mu = \infty = \Sigma f(m)$.

Fall 3: $f \in \mathfrak{L}'_\mu \Leftrightarrow \Sigma f^+(m) < \infty$ oder $\Sigma f^-(m) < \infty \Rightarrow \int f d\mu = \Sigma f^+(m) - \Sigma f^-(m) =: \Sigma f(m)$.

Bei der Beziehung $\pi/4 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \ldots$ ist die rechte Seite im Sinne unserer maßtheoretischen Def. der Reihe nicht erklärt.

Fall 4: $f \in \mathfrak{L}_\mu \Leftrightarrow \Sigma|f(m)| < \infty \Leftrightarrow \Sigma f(m)$ konvergiert absolut.

Integrale über meßbare Mengen.

9.12 Def. Sei $f : \Omega \mapsto \bar{\mathbb{R}}$, $A \in \mathfrak{F}$. Man sagt $_A\int f d\mu$ existiert, falls $f \cdot 1_A \in \mathfrak{L}'_\mu$, und definiert dann:

$$_A\int f d\mu := \int f \cdot 1_A d\mu.$$

9.13 Bemerkung: $_A\int f d\mu$ hängt von f nur über $f\big|_A$ ab.

9.14 Satz. Seien $f : \Omega \mapsto \bar{\mathbb{R}}$ mb, $A,B \in \mathfrak{F}$, dann gilt:

(a) $\int f d\mu$ ex [bzw. ist endlich] $\Rightarrow$ $_A\int f d\mu$ ex [bzw. ist endlich] ;

(b) $f \le g$ auf A, $_A\int f d\mu$ und $_A\int g d\mu$ ex $\Rightarrow$ $_A\int f d\mu \le {_A\int} g d\mu$;

(c) $f \in \mathfrak{L}'_\mu$, A,B disjunkt $\Rightarrow$ $_{A\cup B}\int f d\mu = {_A\int} f d\mu + {_B\int} f d\mu$.

Bew. a) $(f \cdot 1_A)^{\pm} \le f^{\pm}$; b) $f \cdot 1_A \le g \cdot 1_A$ und 9.10d; c) $1_{A\cup B} = 1_A + 1_B$ und 9.10g.

Beziehungen zwischen Integral und Integrand.

9.15 Sprechweise: $N \in \mathfrak{F}$ heißt (μ–) Nullmenge, wenn $\mu[N]=0$.

Sei $K(\omega)$ eine Aussage über $\omega \in \Omega$. Man sagt: K ist μ–f.s. richtig, wenn gilt:

$$N_K := \{\omega;\ K(\omega) \text{ ist falsch}\} \subset N \text{ für eine Nullmenge N.}$$

Meist gilt: $N_K \in \mathfrak{F}$, also ist $N_K = N$ wählbar.

Beispiel: (i) $f \ge 0$ μ–f.s.. (ii) Ist $\mu \equiv 0$, so ist jede Aussage μ–f.s. richtig.

Gemäß 7.2 ist jede abzählbare Vereinigung von Nullmengen wieder eine Nullmenge.

9.16 Satz. Seien $f,g : \Omega \mapsto \bar{\mathbb{R}}$ mb, $A \in \mathfrak{F}$.

(a) Ist $f \ge 0$. dann gilt: $\int f d\mu = 0 \Leftrightarrow f=0$ μ–f.s..

(b) Falls $f=g$ μ–f.s., gilt: $f \in \mathfrak{L}'_\mu \Leftrightarrow g \in \mathfrak{L}'_\mu \Rightarrow \int f d\mu = \int g d\mu$.

(c) $\int |f| d\mu < \infty \Rightarrow |f| < \infty$ μ–f.s..

(d) $\mu[|f| \ge \varepsilon] \le \varepsilon^{-\alpha} \cdot \int |f|^\alpha d\mu$ $\forall\ \varepsilon > 0, \alpha > 0$ (**Markoffsche Ungleichung**).

(e) Ist N Nullmenge, so gilt: $_{\Omega\setminus N}\int f d\mu$ ex $\Leftrightarrow f \in \mathfrak{L}'_\mu \Rightarrow {_{\Omega\setminus N}\int} f d\mu = \int f d\mu$.

(f) Ist N Nullmenge, so gilt: $_N\int f d\mu = 0$.

Bew. d) $|f|^{\alpha} \geq \varepsilon^{\alpha} 1_{\{|f|\geq\varepsilon\}}$.

a) "$\Rightarrow$" $N:=\{f>0\} = \cup_n \{f\geq\frac{1}{n}\}$, wobei gemäß (d) mit $\alpha=1$: $\mu[f\geq\frac{1}{n}] = 0$; "$\Leftarrow$" $\Pi_n(f) \leq n\cdot 1_N$.

f) $f^{\pm}\cdot 1_N = 0$ μ–f.s., verwende jetzt (a)"$\Leftarrow$".

b) Mit $N := \{f\neq g\}$ gilt $\int_{\Omega\setminus N} f^{\pm}d\mu = \int_{\Omega\setminus N} g^{\pm}d\mu$ und nach (f) auch: $\int_N f^{\pm}d\mu = \int_N g^{\pm}d\mu$. Jetzt 9.14c.

c) Mit (d) folgt: $\mu[|f|=\infty] \leq \mu[|f|\geq n] \leq \frac{1}{n}\cdot\int|f|d\mu \to 0$ $(n\to\infty)$.

e) $f = f\cdot 1_{\Omega\setminus N}$ μ–f.s., verwende jetzt (b). $\square$

9.17 Anwendung (a) (vgl. 9.7a): $\int(\infty f)d\mu = \infty\cdot\int fd\mu$ für $f\in\mathfrak{M}_+$. Bew. mit 9.16a,c.

(b) Aussage 9.16b liefert die Rechtfertigung für das Argument der **Nullmengenelimination**. Wenn etwa $\mu = P$ ein w–Maß ist und ein Integrand eine Eigenschaft außerhalb einer Nullmenge N hat, so kann o. E. angenommen werden, daß die Eigenschaft überall gilt. Andernfalls geht man zum Raum $\Omega\setminus N$ über oder genauer zum Raum $(\Omega',\mathfrak{F}',P') = (\Omega\setminus N, \Omega'\cap\mathfrak{F}, P|_{\mathfrak{F}'})$.

9.18 Bemerkung. Sei $f:\Omega\mapsto\overline{\mathbb{R}}$ μ–f.s. erklärt, d.h. es gibt eine Nullmenge, so daß $f:\Omega\setminus N\mapsto\overline{\mathbb{R}}$ erklärt ist. Dann sagt man:

$f\in\mathfrak{L}'_\mu$ [bzw. $\mathfrak{L}_\mu$] $:\Leftrightarrow$ $\bar{f}\in\mathfrak{L}'_\mu$ [bzw. $\mathfrak{L}_\mu$] mit $\bar{f}:= f\cdot 1_{\Omega\setminus N}$ (also :=0 auf N) und setzt $\int fd\mu := \int\bar{f}d\mu$.

Gemäß 9.16b hängt $\int fd\mu$ nicht von der speziellen Definition von $\bar{f}$ auf N ab.

9.19 Beispiel (vgl. 9.10g). Sei $f,g\in\mathfrak{L}'_\mu$ und $\int fd\mu+\int gd\mu$ definiert, dann ist auch $f+g$ μ–f.s. erklärt und es gilt: $f+g\in\mathfrak{L}'_\mu$ sowie $\int(f+g)d\mu = \int fd\mu + \int gd\mu$. [Übung].

9.20 Satz. Sei $f,g\in\mathfrak{L}'_\mu$, μ σ–endlich. Dann gilt:

(a) $f\leq g$ μ–f.s. $\Leftrightarrow$ $\int_A fd\mu \leq \int_A gd\mu$ $\forall A\in\mathfrak{F}$;

(b) $f=g$ μ–f.s. $\Leftrightarrow$ $\int_A fd\mu = \int_A gd\mu$ $\forall A\in\mathfrak{F}$.

Bew. Es genügt (a)"$\Leftarrow$" zu zeigen. Geg. sei $A_n\uparrow\Omega$ mit $\mu[A_n]<\infty$. Dann hat man

$\{f>g\} = \cup_{n\in\mathbb{N},\, r,s\in\mathbb{Q}, r>s} \{f\geq r>s\geq g\}\cap A_n =: \cup A_{r,s,n}$ wobei gilt:

$-\infty < r\cdot\mu[A_{r,s,n}] \leq \int_{A_{r,s,n}} fd\mu \leq$ (nach Vor.) $\int_{A_{r,s,n}} gd\mu \leq s\cdot\mu[A_{r,s,n}] < \infty$.

Also folgt: $\mu[A_{r,s,n}]=0$ $\forall r,s,n$ und somit $\mu[f>g]=0$. $\square$

<u>**Maßtheoretische Induktion.**</u>

Soll gezeigt werden, daß $\int fd\mu$ eine bestimmte Eigenschaft hat $\forall f\in\mathfrak{L}'_\mu$; dann führt oft das folgende Vorgehen zum Ziel:

Schritt 1: $f = 1_A$, $A\in\mathfrak{F}$: Dieser Fall ist meist leicht zu behandeln.

Schritt 2: $f = \Sigma_1^n a_i 1_{A_i} \in \mathfrak{M}_0$, $0\le a_i<\infty$. $A_i\in\mathfrak{F}$: Verwende die Linearität des Integrals.

Schritt 3: $f \in \mathfrak{M}_+$, also f mb und $f \ge 0$: Wähle $(f_n)\subset\mathfrak{M}_0$ mit $f_n\uparrow f$ und verwende die Stetigkeit des Integrals auf $\mathfrak{M}_+$ von unten (vgl. Def.9.4 oder 9.5 oder 10.1).

Schritt 4: $f\in\mathfrak{L}'_\mu$: Zerlege f in $f=f^+-f^-$ und verwende die Linearität des Integrals.

Manchmal kann noch der folgende Schritt vorausgehen:

Schritt 0: $f = 1_C$, $C \in \mathfrak{C}$ mit $\mathfrak{F} = \sigma(\mathfrak{C})$ und $\mathfrak{C}$ $\cap$–stabil.

Von Schritt 0 nach Schritt 1 kommt man mit dem Dynkinargument 6.25.

9.21 Beispiel. $\int f\, d(c\mu) = c \cdot \int f d\mu$, $c\ge 0$.

Schreibweise: $\int f d\mu =: \int f(\omega)\mu[d\omega] = \int d\mu\, f$ und $_a\int^b \ldots := {}_{(a,b]}\int \ldots$ bzw. $_a\int^\infty \ldots := {}_{(a,\infty)}\int \ldots$.

§10 Eigenschaften des Integrals.

Sei $(\Omega,\mathfrak{F},\mu)$ Maßraum; betrachtet werden nur mb Funktionen von Ω nach $\bar{\mathbb{R}}$.

Konvergenzsätze.

10.1 Satz von der monotonen Konvergenz (B. Levi). Ist (f_n) eine isotone Folge mb Fkt. mit $f_n\ge g$ $\forall$ n für ein $g \in \mathfrak{L}_\mu$ (z.B. $g\equiv 0$), dann gilt:

$$\lim \int f_n d\mu = \int(\lim f_n)d\mu \quad [\text{mit } \lim f_n\in \mathfrak{L}'_\mu] \;;$$

das Integral ist also insbesondere auf $\mathfrak{M}_+$ von unten stetig.

Bew. (i) $f_n^-\le g^- \Rightarrow f_n$, $f:=\lim f_n \in \mathfrak{L}'_\mu$.

(ii) O.E. g=0, betrachte sonst $f_n-g\ge 0$.

(iii) $f_k \le f \Rightarrow \int f_k d\mu \le \int f d\mu \Rightarrow \lim \int f_k d\mu \le \int f d\mu$,

(iv) Zeige mit 8.18: $\Pi_k(f_k) \uparrow f$ $(k\to\infty)$. Dann folgt mit 9.5:

$$\int f d\mu = \lim_k \int \Pi_k(f_k) d\mu \le \lim \int f_k d\mu. \; \square$$

10.2 Bemerkung. 10.1 gilt entsprechend für antitone Folgen.

10.3 "Gegenbeispiel". $f_n\equiv -1/n$, $\mu[\Omega]=\infty$; dann $\int f_n d\mu=\infty$, aber $\int \lim f_n d\mu = 0$.

10.4 Korollar. Sei $f_n \geq 0$ und mb , $n \in \mathbb{N}$; dann gilt: $\int(\Sigma f_n)d\mu = \Sigma \int f_n d\mu$.

Eine Verallgemeinerung wird in 11.13 als Folgerung aus dem Satz von Fubini gegeben.

10.5 Satz. Sei $f \geq 0$ und mb und sei $\nu[A] := {}_A\int f d\mu$,. $A \in \mathfrak{F}$. Dann gilt:

(a) ν ist ein Maß.

(b) Gilt $f<\infty$ μ–f.s., dann ist mit μ auch ν σ–endlich.

Bew. a) mit (2.17) und 10.4.
b) Sei o.E. $f<\infty$ auf Ω. Wähle $(A_n) \subset \mathfrak{F}$ mit $A_n \uparrow \Omega$, $\mu[A_n]<\infty$ und setze $B_n := A_n \cap \{f \leq n\}$. Dann folgt: $B_n \uparrow \Omega$ und $\nu[B_n] \leq n \cdot \mu[A_n] <\infty$. []

10.6 Lemma von Fatou. Seien f_n, $n \in \mathbb{N}$, mb.

(a) Existiert $g \in \mathfrak{L}_\mu$ (z.B. $g \equiv 0$) mit $f_n \geq g$, $n \in \mathbb{N}$, so gilt: $\int \underline{\lim}\, f_n d\mu \leq \underline{\lim} \int f_n d\mu$ [mit $\underline{\lim}\, f_n \in \mathfrak{L}'_\mu$].

(b) Existiert $h \in \mathfrak{L}_\mu$ (z.B. $h \equiv 0$) mit $f_n \leq h$, $n \in \mathbb{N}$, so gilt: $\int \overline{\lim}\, f_n d\mu \geq \overline{\lim} \int f_n d\mu$ [mit $\overline{\lim}\, f_n \in \mathfrak{L}'_\mu$].

Beweis: a) $v_n := \inf_{k \geq n} f_k \uparrow \underline{\lim}\, f_n$ und $f_n \geq v_n \geq g$; somit folgt mit monot. Konv. 10.1:

$\underline{\lim}\, f_n \in \mathfrak{L}'_\mu$ und $\underline{\lim} \int f_n d\mu \geq \lim \int v_n d\mu = \int \underline{\lim}\, f_n d\mu$.

Teil b) wie oder mit (a). []

10.6 (c) Folgerung. Der f.s. Limes einer L_1–beschränkten Folge ist intb.; d.h.: gilt $f_n \to f$ f.s. und $\sup_n \int |f_n| d\mu < \infty$, so folgt: $\int |f| d\mu < \infty$.

Zum **Beweis** verwende Fatou 10.6(a) für $|f_n|$. []

10.7 Satz von der majorisierten Konvergenz (Lebesgue).

Sei (f_n) eine (μ–f.s.) konvergente Folge mb. Fkt mit $|f_n| \leq g$. $n \in \mathbb{N}$, für ein $g \in \mathfrak{L}_\mu$; dann gilt:

$$\lim \int f_n d\mu = \int \lim f_n d\mu \quad [\text{mit } \lim f_n \in \mathfrak{L}_\mu].$$

Bew. Wegen 10.6 gilt: $\int \underline{\lim} \ldots \leq \underline{\lim} \int \ldots \leq \overline{\lim} \int \ldots \leq \int \overline{\lim} \ldots$. []

Maße mit Dichten.

10.8 Def. Sei $f \geq 0$ und mb; f heißt dann f **μ–Dichte** des Maßes ν auf $(\Omega, \mathfrak{F})$, falls $\nu[A] = {}_A\int f d\mu$, $A \in \mathfrak{F}$.

Eine λ^d–Dichte heißt **Lebesgue–Dichte**.

Schreibweise: $d\nu = f d\mu$ oder $f = \frac{d\nu}{d\mu}$ oder $\nu[d\omega] = f(\omega)\mu[d\omega]$.

10.9 Satz. Sei μ σ–endlich und $d\nu_i = f_i d\mu$, i=1,2; dann gilt:

$$\nu_1 = \nu_2 \iff f_1 = f_2 \quad \mu\text{–f.s.} .$$

Bew. mit 9.16b für "$\Leftarrow$" und 9.20b für "$\Rightarrow$". $[]$

10.10 Satz. Sei $d\nu = f\, d\mu$ und $g : \Omega \mapsto \overline{\mathbb{R}}$ mb.; dann gilt:

$$\int g\, d\nu = \int g \cdot f\, d\mu \quad [\text{wobei } g \in \mathfrak{L}'_\nu \iff g \cdot f \in \mathfrak{L}'_\mu].$$

Bew. Der Fall $g = 1_A$ ist klar; jetzt maßtheoretische Induktion. $[]$

10.11 Bemerkung. Aus $d\varphi = g d\nu$ und $d\nu = f d\mu$ folgt $d\varphi = g \cdot f d\mu$.

Bew. $\varphi[A] = \int 1_A g d\nu = \int 1_A g f d\mu$. $[]$

10.12 Satz. Sei $\mathfrak{F} = \sigma(\mathfrak{C})$, $\mathfrak{C}$ $\cap$–stabil, ν σ–endlich auf $\mathfrak{C}$ sowie $f \geq 0$ und mb; dann gilt:

$$d\nu = f d\mu \iff \nu[C] = {}_C\!\int f d\mu \quad \forall\, C \in \mathfrak{C}.$$

Bew. "$\Leftarrow$" mit 10.5 und dem Eindeutigkeitssatz 7.5. $[]$

10.13 Def. Das Maß ν auf $(\Omega, \mathfrak{F})$ heißt **μ–stetig** und μ heißt **dominierendes Maß** zu ν $:\iff$ $\boldsymbol{\nu \ll \mu}$ $:\iff$ Jede μ–Nullmenge ist eine ν–Nullmenge.

10.14 Satz von Radon–Nikodym. Ist μ σ–endlich , so besitzt ein Maß ν genau dann eine μ–Dichte, wenn es μ–stetig ist.

Bew. "dann" Bauer (1978) S. 91. "genau dann" nach 9.16 f. $[]$

Integranden mit Parameter.

10.15 Satz. Sei $B \subset \mathbb{R}^d$ und für $f : \Omega \times B \mapsto \mathbb{R}^1$ gelte:

(i) $f(\cdot, t)$ ist mb $\forall\, t \in B$,

(ii) $f(\omega, \cdot)$ ist stetig $\forall\, \omega \in \Omega$ und

(iii) $\exists\, g \in \mathfrak{L}_\mu$ mit $|f(\cdot, t)| \leq g \quad \forall\, t \in B$.

Dann ist $t \mapsto \int f(\omega, t) \mu[d\omega]$ stetig auf B.

Bew. Bei $f(\cdot, t_n) \to f(\cdot, t_0)$ für $t_n \to t_0$ liegt maj. Konv. 10.7 vor. $[]$

10.16 Satz. Sei $B \subset \mathbb{R}^1$, B offen und für $f : \Omega \times B \mapsto \mathbb{R}^1$ gelte:

(i) $f(\cdot, t) \in \mathfrak{L}_\mu$ $\forall\, t \in B$,

(ii) $f(\omega, \cdot)$ ist db auf B $\forall\, \omega \in \Omega$ und

(iii) $\exists\, g \in \mathfrak{L}_\mu$ mit $|\frac{\partial}{\partial t} f(\cdot, t)| \leq g \quad \forall\, t \in B$.

Dann gilt: $$\frac{d}{dt} \int f(\omega, t)\, \mu[d\omega] = \int \frac{\partial}{\partial t} f(\omega, t)\, \mu[d\omega].$$

Bew. Für eine Nullfolge $(h_n)\subset\mathbb{R}^1$ setze: $\Delta_n(\omega) := \left[f(\omega,t+h_n)-f(\omega.t)\right]/h_n$.
Dann ist $\frac{\partial}{\partial t}f(\cdot,t) = \lim_{h_n\to 0} \Delta_n$ mb und es gilt nach dem Mittelwertsatz: $|\Delta_n|\le g$. Also liegt für $h_n\to 0$ maj. Konv. 10.7 vor bei $\lim_{h_n\to 0} \int \Delta_n \, d\mu$. []

Ein Beispiel wird sich etwa bei der Laplace–Transformierten mit $\Omega = B = (0,\infty)$ und $f(\omega,t) = e^{-t\omega}$ ergeben.

Bildmaße.

Erinnerung: $PX^{-1} = P\left[X\in\cdot\right]$ heißt Verteilung der Zva X (§4).

10.17 Def. Sei $T : \Omega \mapsto \Omega'$ $\mathfrak{F}$–$\mathfrak{F}'$–mb. Dann heißt das durch $\mu T^{-1}\left[A'\right] := \mu\left[T\in A'\right]$, $A'\in\mathfrak{F}'$, definierte Maß auf $\mathfrak{F}'$ das **Bildmaß** von μ unter T.

10.18 Bemerkung. μT^{-1} ist ein Maß (Bew. mit 4.5) und es gilt: $\mu(T'\circ T)^{-1} = (\mu T^{-1})T'^{-1}$.

10.19 Transformationsatz. Seien $T : \Omega \mapsto \Omega'$ $\mathfrak{F}$–$\mathfrak{F}'$–mb und $f : \Omega' \mapsto \bar{\mathbb{R}}$ $\mathfrak{F}'$–mb, $A'\in\mathfrak{F}'$; dann gilt:

$\int_{\{T\in A'\}} f\circ T \, d\mu = \int_{A'} f \, d\mu T^{-1}$, insbesondere $\int f\circ T \, d\mu = \int f \, d\mu T^{-1}$,

wobei jeweils das linke Integral genau dann existiert, wenn das rechte existiert.

Bew. Es ist $1_{\{T\in A'\}} = 1_{A'}\circ T$, also o.E. $A'=\Omega'$. Sei $f=1_{B'}$ für $B'\in\mathfrak{F}'$; dann
$\int 1_{B'}\circ T \, d\mu = \mu\left[T\in B'\right] = \mu T^{-1}\left[B'\right] = \int 1_{B'} d\mu T^{-1}$. Jetzt maßtheoretische Induktion. []

Berechnung von Lebesgue–Integralen.

10.20 Transformationssatz für Lebesgue–Integrale.. Sei $d\in\mathbb{N}$, $A\in\mathbb{R}^{d\times d}$ invertierbar, $b\in\mathbb{R}^d$ (Spaltenvektor), $T(x):=Ax+b$, $x\in\mathbb{R}^d$, sowie $B\in\mathfrak{B}_d$ und $f : \mathbb{R}^d \mapsto \bar{\mathbb{R}}$ mb. Dann gilt:

(a) $(\lambda^d)T^{-1} = |\det(A)|^{-1}\cdot\lambda^d$;

(b) $\int_B f\circ T \, d\lambda^d = |\det(A)|^{-1}\cdot\int_{T(B)} f d\lambda^d$, falls eines der Integrale existiert.

Bew. Übung oder Barner/Flohr (1983, S.312).

Schreibweise: $\int f d\lambda^d = \int f(x)\lambda^d\left[dx\right] =: \int f(x)dx$.

§11 Produktmaße.

Seien $(\Omega_i,\mathfrak{F}_i,\mu_i)$, i=1,2, Maßräume, $\Omega:=\Omega_1\times\Omega_2$, $\mathfrak{F}:=\mathfrak{F}_1\otimes\mathfrak{F}_2$.

11.1 Def. Für $A\subset\Omega$ und $x\in\Omega_1$ heißt $\mathbf{A_x}:=\{y\in\Omega_2; (x,y)\in A\}$ der **x–Schnitt** von A.

11.2 Lemma. Für $x\in\Omega_1$ ist $i_x:\Omega_2\mapsto\Omega$ definiert durch $i_x(y)=(x,y)$ eine $\mathfrak{F}_2$–$\mathfrak{F}$–mb Abbildung und es gilt:

$A_x=i_x^{-1}(A)$.

Bew. i_x ist mb nach 8.10. $\square$

Aus 11.2 und 4.2 folgt:

11.3 Lemma. Für $x\in\Omega_1$, A, B, $A_i \subset \Omega$ gilt:

$(\Omega_1\times\Omega_2)_x=\Omega_2$, $(\cap A_i)_x=\cap (A_i)_x$, $(\cup A_i)_x=\cup (A_i)_x$, $(A\setminus B)_x=A_x\setminus B_x$.

11.4 Lemma. (a) Für $A\in\mathfrak{F}$ ist $A_x\in\mathfrak{F}_2$ $\forall$ $x\in\Omega_1$.

(b) Ist $f:\Omega\mapsto\Omega'$ $\mathfrak{F}$–$\mathfrak{F}'$–mb , so ist $f(x,\cdot)$ $\mathfrak{F}_2$–$\mathfrak{F}'$–mb $\forall$ $x\in\Omega_1$.

Bew. folgt aus 11.2 und $f(x,\cdot) = f\circ i_x$. $\square$

11.5 Bemerkung: Aus der $\mathfrak{F}_2$–Meßbarkeit von $f(x,\cdot)$ und der $\mathfrak{F}_1$–Meßbarkeit von $f(\cdot,y)$ $\forall x,y$ folgt i.a. nicht die $\mathfrak{F}$–Meßbarkeit von f.

11.6 Vorstufe zu Fubini (Integration bzgl eines Parameter). Sei μ_2 σ–endlich.

(a) Für $A\in\mathfrak{F}$ ist $x\mapsto\mu_2[A_x]$ $\mathfrak{F}_1$–$\bar{\mathfrak{B}}$–mb.

(b) Ist $f:\Omega\mapsto\bar{\mathbb{R}}$ $\mathfrak{F}$–mb und existiert $h(x):=\int f(x,y)\mu_2(dy)$ $\forall$ $x\in\Omega_1$, so ist h $\mathfrak{F}_1$–$\bar{\mathfrak{B}}$–mb.

Offenbar ist 11.6b auch eine Aussage über Intgranden mit Parametern.

Bew. a): (i) Sei μ_2 endlich und $\mathfrak{D}:=\{A\in\mathfrak{F};\ x\mapsto\mu_2[A_x]\ \text{mb}\}$. Dann ist $\mathfrak{D}$ Dynkinsystem, denn z.B. gilt mit 11.3: $\mu_2[(A^c)_x] = \mu_2[\Omega_2]-\mu_2[A_x]$. Ferner hat man $\mathfrak{D}\supset\mathfrak{F}_1\times\mathfrak{F}_2$; denn $(A_1\times A_2)_x = A_2$ $\forall$ $x\in A_1$ bzw. $=\emptyset$ $\forall$ $x\notin A_2$, also $\mu_2[(A_1\times A_2)_x] = 1_{A_1}(x)\cdot\mu_2[A_2]$. Jetzt Dynkin–Argument.

(ii) Sei $(B_n)\subset\mathfrak{F}_2$ mit $B_n\uparrow\Omega_2$ und $\mu_2[B_n]<\infty$; dann gilt $\mu_2[B_n\cap A_x]\uparrow\mu_2[A_x]$, wobei $\mu_2[B_n\cap\cdot]$ endliches Maß und $\mu_2[B_n\cap A_x]$ nach Teil (i) mb ist.

b) Sei $f=1_A$ für ein $A\in\mathfrak{F}$. Dann ist $h(x) = \int 1_A(x,y)\mu_2[dy] = \int 1_{A_x}(y)\mu_2[dy] = \mu_2[A_x]$ nach Teil (a) mb. Jetzt maßtheoretische Induktion. $\square$

11.7 Satz über das Produktmaß. Seien μ_1,μ_2 σ–endlich. Dann ist

$(\mu_1\times\mu_2)[A] := \int \mu_2[A_x]\ \mu_1[dx]$, $A\in\mathfrak{F}$, das einzige Maß μ auf $\mathfrak{F}$ mit

(11.8) $\mu[A_1\times A_2] = \mu_1[A_1]\cdot\mu_2[A_2]$, $A_i\in\mathfrak{F}_i$, i=1,2.

$\mu_1\times\mu_2$ ist wieder σ–endlich und heißt das durch μ_1 und μ_2 bestimmte **Produktmaß.**

Bew. (i) Nach 11.6a ist die Def. sinnvoll. (ii) $\mu_1\times\mu_2$ ist Maß; die σ–Additivität folgt aus:

$\int\mu_2[(\cup A_i)_x]\ \mu_1[dx] = \int\mu_2[\cup (A_i)_x]\ \mu_1[dx] = \int(\Sigma\mu_2[(A_i)_x])\mu_1[dx] = \Sigma\int\mu_2[(A_i)_x]\ \mu_1[dx]$ für disjunkte A_i, wobei 10.4 benutzt wurde. (iii) $\mu_1\times\mu_2$ erfüllt (11.8) wegen $\mu_2[(A_1\times A_2)_x] = 1_{A_1}(x)\cdot\mu_2[A_2]$. (iv) $\mu_1\times\mu_2$ ist σ–endlich auf $\mathfrak{F}_1\times\mathfrak{F}_2$. Denn wähle A_i^n mit $A_i^n\uparrow\Omega_i$ und $\mu_i[A_i^n]<\infty$, i=1,2; dann folgt $A_1^n\times A_2^n\uparrow\Omega$ und $(\mu_1\times\mu_2)[A_1^n\times A_2^n]=\mu_1[A_1^n]\cdot\mu_2[A_2^n]<\infty$.

(v) Ist μ ein Maß mit (11.8), so folgt $\mu=\mu_1\times\mu_2$ aus dem Eindeutigkeitssatz 7.5. []

11.9 Satz von Fubini. Seien μ_1 und μ_2 σ–endlich.

(a) **(Tonelli)** Sei $f:\Omega_1\times\Omega_2\mapsto\overline{\mathbb{R}}$ $\mathfrak{F}$–mb und f≥0. Dann gilt:

(11.10) $$\int f\, d(\mu_1\times\mu_2) = \int\left[\int f(x,y)\mu_2[dy]\right]\mu_1[dx] = \int\left[\int f(x,y)\mu_1[dx]\right]\mu_2[dy]\ .$$

(b) Sei $f\in\mathfrak{L}'_{\mu_1\times\mu_2}$. Dann ist $x\mapsto\int f(x,y)\mu_2[dy]$ μ_1–f.s. erklärt und aus $\mathfrak{L}'_{\mu_1}$ sowie $y\mapsto\int f(x,y)\mu_1[dx]$ μ_2–f.s. erklärt und aus $\mathfrak{L}'_{\mu_2}$ und es gilt (11.10).

Bew. a) Setze $\mu'[A] := \int\mu_1[A_y]\mu_2[dy]$, dann zeigt man wie in 11.7, daß μ' ein Maß auf $\mathfrak{F}$ ist und die Eigenschaft (11.8) hat. Also gilt nach 11.7: $\mu'=\mu_1\times\mu_2$.

Damit gilt (11.10) für $f=1_A$, $A\in\mathfrak{F}$. Nun folgt (11.10) für f≥0 mit maßtheoretischer Induktion.

b) Sei nun $f\in\mathfrak{L}'_{\mu_1\times\mu_2}$; dann folgt (11.10) für $f^{\pm}$. Setze $h_{\pm}(x) = \int f^{\pm}(x,y)\mu_2[dy]$. Dann ist nach (a) und nach Voraussetzung $\int f d\mu_1\times\mu_2 = \int f^+ d\mu_1\times\mu_2 - \int f^- d\mu_1\times\mu_2 = \int h_+ d\mu_1 - \int h_- d\mu_2$ erklärt.

Aus 9.19 folgt, daß h_+-h_- μ_1–f.s. erklärt ist und gilt: $\int h_+ d\mu_1 - \int h_- d\mu_1 = \int(h_+-h_-)d\mu_1$.

Dort wo h_+-h_- erklärt ist, gilt $h_+-h_- = \int f(\cdot,y)\mu_2[dy]$. []

11.11 Hinweis. Zum Nachweis von $f\in\mathfrak{L}'_{\mu_1\times\mu_2}$ [bzw. $f\in\mathfrak{L}_{\mu_1\times\mu_2}$] verwende 11.9a für $f^{\pm}$ [bzw. für |f|].

Erinnerung: Eine hinreichende Bedingung für die im Satz von Fubini verlangte Produktmeßbarkeit von f wurde in 8.11 gegeben.

11.12 Spezialfall. Für $A_i\in\mathfrak{F}_i$ gilt:

$$\int_{A_1\times A_2} f_1(x)f_2(y)(\mu_1\times\mu_2)[d(x,y)] = \int_{A_1} f_1(x)\mu_1[dx]\cdot\int_{A_2} f_2(y)\mu_2[dy]$$

für $f_i\geq 0$, i=1,2, oder für $f_i\in\mathfrak{L}_{\mu_i}$, i=1,2.

Bew. Wegen $1_{A_1\times A_2}(x,y) = 1_{A_1}(x)\cdot 1_{A_2}(y)$ sei o.E. $A_i = \Omega_i$. Betrachte nun erst $|f_i|$. $\square$

11.13 Anwendung. Sei μ_1 σ–endlich, $f_n : \Omega_1 \mapsto \bar{\mathbb{R}}$ $\mathfrak{F}_1$–mb, $n\in\mathbb{N}_0$, sodaß $\Sigma\int f_n^+ d\mu_1<\infty$ oder $\Sigma\int f_n^- d\mu_1<\infty$.

Dann ist Σf_n μ_1–f.s. erklärt $[$als $\Sigma f_n^+ - \Sigma f_n^-]$ und es gilt:

$$\Sigma\int f_n\, d\mu_1 = \int (\Sigma f_n)\, d\mu_1 .$$

Bew. Definiere $\Omega_2=\mathbb{N}_0$, $f(x,n) := f_n(x)$, μ_2 als das Zählmaß auf $\mathbb{N}_0$ und verwende 11.9 mit 11.11 sowie 9.11. $\left[\text{Etwa nach 8.12 mit } A_i = \Omega_1\times\{i\} \text{ ist } (\omega_1,n) \mapsto f_n(\omega_1) \text{ mb.}\right]$ $\square$

11.14 Verallgemeinerung auf d Räume mit vollständiger Induktion gemäß $\overset{d}{\underset{1}{\times}}\mu_i := \overset{d-1}{\underset{1}{\times}}\mu_i \times \mu_d$,also

$\overset{d}{\underset{1}{\times}}\mu_i[A_1\times...\times A_{n-1}\times A_n] := (\overset{d-1}{\underset{1}{\times}}\mu_i \times \mu_d)[A_1\times...\times A_{n-1}\times A_n] = \overset{d-1}{\underset{1}{\times}}\mu_i[A_1\times...\times A_{n-1}]\cdot\mu_d[A_n]$

$= \mathbf{\Pi}_{i=1}^{d}\mu_i[A_i]$.

11.15 Folgerung. Für das d–dim. Lebesgue–Maß auf $\mathfrak{B}_d = \overset{d}{\underset{1}{\otimes}}\mathfrak{B}_1$ gilt: $\lambda^d = \overset{d}{\underset{1}{\times}}\lambda^1 := \lambda^{d-1}\times\lambda^1$.

Bew. Jedes d–dim. Intervall in J_d läßt sich schreiben als $A\times B$ mit $A\in J_{d-1}$, $B\in J_1$. Dann gilt: $\lambda^d[A\times B] = \lambda^{d-1}[A]\cdot\lambda^1[B]$. Aus dem Satz 11.7 über das Produktmaß folgt nun $\lambda^d = \lambda^{d-1}\times\lambda^1$. $\square$

Ist λ^d bisher noch nicht definiert worden, sondern nur λ^1, so kann 11.15 auch als Definition von λ^n benutzt werden.

11.16 Bemerkung. Sei Ω_i abzb., μ_i das Zählmaß auf Ω_i, $1\le i\le d<\infty$.

Dann ist $\mu_1\times...\times\mu_d$ das Zählmaß auf $\Omega_1\times...\times\Omega_d$.

Bew. O.E. d=2. $\mu_1\times\mu_2[\{(x,y)\}] = \mu_1\times\mu_2[\{x\}\times\{y\}] = \mu_1[\{x\}]\cdot\mu_2[\{y\}] = 1.$ $\square$

11.17 Satz. Seien μ_i, ν_i σ–endliche Maße auf den Meßräumen $(\Omega_i,\mathfrak{F}_i)$ mit $d\nu_i=f_i d\mu_i$, $1\le i\le d<\infty$;

dann gilt: $(\overset{d}{\underset{1}{\times}}\nu_i)(d(x_1,...,x_d)) = \Pi_{i=1}^{d} f_i(x_i)\ (\overset{d}{\underset{1}{\times}}\mu_1)(d(x_1,...,x_d))$, d.h.

$d(\overset{d}{\underset{1}{\times}}\nu_i) = \Pi_{i=1}^{d} f_i\circ\pi_i\ d(\overset{d}{\underset{1}{\times}}\mu_1)$ mit $\pi_i : \overset{d}{\underset{1}{\times}}\Omega_j \mapsto \Omega_i$ als Projektion.

Bew. O.E. d=2. Wir verwenden 10.12 mit $\mathfrak{C} = \mathfrak{F}_1\times\mathfrak{F}_2$. Mit 11.12 erhält man:

$(\nu_1\times\nu_2)[A_1\times A_2] = \nu_1[A_1]\cdot\nu_2[A_2] = {}_{A_1}\!\int f_1 d\mu_1\cdot{}_{A_2}\!\int f_2 d\mu_2 = {}_{A_1\times A_2}\!\int f_1(x)f_2(y)(\mu_1\times\mu_2)[d(x,y)]$ $\square$

KAPITEL III. WAHRSCHEINLICHKEITSTHEORIE.

§ 12 W–Maße mit Dichten.

12.1 Def. Seien X_i Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$, $1\le i\le d\le\infty$.
[Dann ist $X=(X_1,...,X_d)$ gemäß 8.10 Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\times_1^d \Omega_i, \otimes_1^d \mathfrak{F}_i)$].
Die Verteilung von X, also $PX^{-1}=P[(X_1,...,X_d)\in\cdot\,]$, heißt **gemeinsame Verteilung** der X_i, $1\le i\le d$. Die Verteilung von X_i, also PX_i^{-1}, heißt (1–dim.) **Randverteilung** von PX^{-1}.

12.2. Bemerkung PX_i^{-1} ist bestimmt durch PX^{-1}; denn z.B. für i=1 gilt:

$$P[X_1\in B]=P[(X_1,X_2,...)\in B\times\Omega_2\times...].$$

Aber die 1–dim. Randverteilungen bestimmen nicht die gemeinsame Verteilung. Dazu das

12.3 Beispiel. Sei $X\sim b(1,\frac{1}{2})$, dann $1-X\sim b(1,\frac{1}{2})$, aber die Verteilung von (X,X) ist nicht gleich der von (X,1–X).

12.4 Sprechweise. Eine Begriff, der für W–Maße definiert ist, soll auch als Begriff für Zva aufgefaßt werden und dann mit dem für die Verteilung der Zva definierten Begriff identifiziert werden. So heißt eine [Lebesgue–] Dichte oder die Verteilungsfunktion der Verteilung einer Zva auch [Lebesgue–] Dichte oder Verteilungsfunktion dieser Zva. Eine Dichte der gemeinsamen Verteilung von gewissen Zva heißt auch **gemeinsame Dichte** dieser Zva.

12.5 Folgerung. Seien X_i Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$, μ_i σ–endliche Maße auf $(\Omega_i,\mathfrak{F}_i)$, $1\le i\le d<\infty$. Dann ist $f : \times_1^d \Omega_i \mapsto [0,\infty)$ genau dann gemeinsame Dichte der X_i, wenn gilt:
f ist $\otimes_1^d \mathfrak{F}_i$–mb und
$P[(X_1,...,X_d)\in B] = {}_B\!\int f\, d \times_1^d \mu_i \quad \forall\, B \in \otimes_1^d \mathfrak{F}_i$ [oder $\forall\, B \in \times_1^d \mathfrak{F}_i$ gemäß 10.12]

12.6 Bemerkung. Seien X_i , μ_i wie in 12.5, i=1,2. Haben dann X_1,X_2 die gemeinsame $\mu_1\times\mu_2$–Dichte f, so ist $\quad x \mapsto f_1(x) := \int f(x,y)\mu_2[dy]$ eine μ_1–Dichte von X_1.

Bew. f_1 ist nach 11.6 mb. Sei $A\in\mathfrak{F}_1$; dann gilt nach Fubini und wegen $1_{A\times\Omega_2}(x,y)=1_A(x)$:

$$P[X_1\in A] = P[(X_1,X_2)\in A\times\Omega] = \int 1_{A\times\Omega_2} f d\mu_1\times\mu_2 = {}_A\!\int f_1 d\mu_1. \square$$

12.7 Bemerkung. Sei $(\Omega,\mathfrak{F},P)$ diskreter W–Raum. Dann ist f genau dann eine Dichte von P bzgl. des Zählmaßes μ auf Ω, wenn f eine Zähldichte von P ist.

Bew. $\Sigma_{x\in A}\, f(x) = \Sigma_{x\in\Omega}\, \mathbf{1}_A(x)f(x) = \int \mathbf{1}_A f d\mu = {}_A\!\int f d\mu$ mit 9.11. $\square$

12.8 Satz. Sei P W–Maß auf $\mathfrak{B}_d$ mit Vf. F und $f : \mathbb{R}^d \mapsto [0,\infty)$ mb. Dann ist f genau dann eine Lebesgue–Dichte von P, wenn gilt: $F(x) = \int_{(-\vec{\infty},x]} f(y)dy \quad \forall\, x\in\mathbb{R}^d$.

Bew. mit 10.12 und 6.15. $\Box$

12.9 Satz. Sei $A\in\mathbb{R}^{d\times d}$ invertierbar, $b\in\mathbb{R}^d$. Hat der d–dim.Zufalls–(Spalten–)vektor X die λ^d–Dichte f, so hat AX+b die λ^d–Dichte $x \mapsto |\det(A)|^{-1}\cdot f(A^{-1}(x-b))$.
Für d=1 hat also aX+b die λ^1–Dichte $x \mapsto \frac{1}{|a|}\cdot f(\frac{x-b}{a})$, $(a\neq 0)$.

Bew. mit 10.20 als Übung.

12.10. Anwendung. Sei d=1. Hat X die Lebesgue–Dichte f mit f(x)=f(–x) und die Vf $F(t) = P[X\le t]$, so gilt: $\quad P(-X)^{-1} = PX^{-1}$ und $P[|X|>t] = 2F(-t) = 2[1-F(t)] \quad \forall\, t>0$.

Bew. mit 12.9 und:
$P[|X|>t] = P[-X>t]+P[X>t] = P[X\le -t]+P[-X\le -t]$ wobei $P[X=\pm t]=0$. $\Box$

12.11 Beispiel. Ist $f : \mathbb{R} \mapsto [0,\infty)$ mb mit $\int f(x)\,dx = 1$; so gehört zu f gemäß 10.5 genau ein W–Maß auf $\mathfrak{B}_1$, dessen Lebesgue–Dichte f ist.

Die **stetige Gleichverteilung** U(a,b) auf (a,b) ist definiert als das W–Maß auf $\mathfrak{B}_1$ mit der Lebesgue–Dichte $\frac{1}{b-a}\cdot 1_{(a,b)}$.

12.12 Bemerkung. Hat X die Dichte $f\cdot 1_B$, z.B. $B = [0,\infty)$, so gilt $X\in B$ f.s. $\left(P[X\in B^c] = \int_{B^c} f\cdot 1_B d\mu = 0\right)$.

12.13 Hilfsmittel. $\Gamma(\nu) := \int_0^\infty x^{\nu-1}e^{-x}dx = 2^{-(\nu-1)}\int_0^\infty y^{2\nu-1}e^{-y^2/2}dy,\ \nu>0;$

$\Gamma(\frac{1}{2})=\sqrt{\pi}$, $\Gamma(n) = (n-1)!$, $n\in\mathbb{N}$, $\Gamma(\nu+1) = \nu\Gamma(\nu)$, $\nu>0$.

12.14 Def. Das W–Maß auf $(\mathbb{R},\mathfrak{B}_1)$ mit der Lebesgue–Dichte

$$\varphi_{a,\sigma^2}(x) := (2\pi\sigma^2)^{-\frac{1}{2}}\exp\{-(x-a)^2/2\sigma^2\} \quad ,x\in\mathbb{R},$$

heißt **Normalverteilung** oder **N(a,σ²)** für $a\in\mathbb{R}$, $\sigma^2>0$. **N(0,1)** heißt **standardisierte Normalverteilung**.

Die Def. ist sinnvoll, da $\varphi\ge 0$, stetig, mb., und $\int\varphi(x)dx=1$ nach 12.13 und Variablensubstitution.

12.15 Eigenschaften. (a) $\varphi_{0,1}(x) = \varphi_{0,1}(-x)$; also $X\sim N(0,1) \Leftrightarrow -X\sim N(0,1)$.

(b) Für $X\sim N(a,\sigma^2)$ gilt: $\alpha X+\beta \sim N(\alpha a+\beta,\alpha^2\sigma^2)$ für $\alpha\neq 0$,
speziell: $X = \sigma\tilde{X}+a$ mit $\tilde{X} := \frac{X-a}{\sigma} \sim N(0,1)$.

(c) Ist Φ die Vf zu N(0,1) und $X\sim N(a,\sigma^2)$, so gilt:
$P[|X-a|>k\cdot\sigma] = 2\Phi(-k) = 2[1-\Phi(k)]$; insbesondere:
$P[|X-a|>2\sigma] \le 0.05$ und $P[|X-a|>3\sigma] \le 0,003$ (**3–σ–Regel**) .

Bew. a) 12.10. b) 12.9. c) Sei $\tilde{X}=(X-a)/\sigma$; dann $P[|X-a|>k\sigma] = P[|\tilde{X}|>k]$; vgl. jetzt 12.10 und Vertafelung von Φ. Der Übergang von X zu $\tilde{X}$ entspricht der Transformation in 1.20. □

12.16 Def. Das W–Maß auf $(\mathbb{R},\mathfrak{B}_1)$ mit der Lebesgue–Dichte

$$\gamma_{\alpha,\nu}(x) := \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\alpha x} \cdot 1_{(0,\infty)}(x) ,$$

heißt **Γ–Verteilung** oder $\boldsymbol{\Gamma_{\alpha,\nu}}$ für $\nu>0$, $\alpha>0$.

$\Gamma_{1/2,n/2} =: \boldsymbol{\chi^2_n}$ heißt **χ²–Verteilung** mit n Freiheitsgraden.

$\Gamma_{\alpha,1} =:$ **Exp(α)** heißt **Exponentialverteilung**, wobei $\gamma_{\alpha,1}(x) = \alpha e^{-\alpha x} 1_{(0,\infty)}(x)$.

$\Gamma_{\alpha,n}$,$n\in\mathbb{N}$, heißt **Erlang–Verteilung**.

12.17 Eigenschaften. (a) $X \sim N(0,1) \Rightarrow X^2 \sim \chi^2_1$.

(b) $X_1,\dots,X_n$ iid, $X_i\sim N(0,1) \Rightarrow \Sigma_1^n X_i^2 \sim \chi^2_n$.

(c) $X \sim \mathrm{Exp}(\alpha) \Leftrightarrow$ X hat die Vf. $t \mapsto (1-e^{-\alpha t})\cdot 1_{(0,\infty)}(t)$.

(d) Sei X Zva mit Werten in $(0,\infty)$; dann gilt: $X \sim \mathrm{Exp}(\alpha)$ für ein $\alpha\in(0,\infty) \Leftrightarrow$

$P[X>s+t\,|\,X>s] = P[X>t] \quad \forall\, s,t>0$ (**Gedächtnislosigkeit**).

(e) X und Y unabhängig, $X \sim \mathrm{Exp}(\alpha)$, $Y \sim \mathrm{Exp}(\beta) \Rightarrow \min(X,Y) \sim \mathrm{Exp}(\alpha+\beta)$.

(f) $X_1,\dots,X_n$ iid, $X_i\sim \mathrm{Exp}(\alpha) \Rightarrow \Sigma_1^n X_i \sim \Gamma_{\alpha,n}$.

Bew. a) Sei $t\geq 0$, $\varphi:=\varphi_{0,1}$.

$P[X^2\leq t] = P[-\sqrt{t}\leq X\leq\sqrt{t}] = \int_{-\sqrt{t}}^{\sqrt{t}}\varphi(x)dx = 2\cdot\int_0^{\sqrt{t}}\varphi(x) = \int_0^t\varphi(\sqrt{y})/\sqrt{y}\,dy$ (Variablensubstitution). Nach 12.8 hat X^2 die λ–Dichte $y \mapsto y^{-\frac{1}{2}}\varphi(y^{\frac{1}{2}})1_{(0,\infty)}(y) = \gamma_{\frac{1}{2},\frac{1}{2}}(y)$.

b) 13.13d c) leicht.

d) "⇒" mit $P[X>t]=e^{-\alpha t}$. "⇐" Für $f(t):=P[X>t]$ gilt nach Vor. $f(s+t)=f(s)\cdot f(t)$; wegen $0<X<\infty$ gilt $1=f(0)\geq f(t)\geq f(\infty)=0$. Jetzt 5.18 .

e) $P[\min(X,Y)\geq t] = P[X\geq t, Y\geq t] = P[X\geq t]\cdot P[Y\geq t]$. f) 13.13d. □

In Beispiel 5.19 mit dem Poissonprozeß wurde für den ersten Zeitpunkt T_1 gezeigt: $P[T_1>t] = e^{-\lambda t}$. Also gilt in dieser Situation nach 12.17(c): $T_1 \sim \mathrm{Exp}(\lambda)$.

12.18 Bemerkung. (a) Das diskrete Analogon zur Exponentialvert. ist die geometrische Vert..
Z.B. $X \sim \mathrm{Geo}(p) \Rightarrow P[X\geq n] = \Sigma_{m=n}^\infty (1-p)^m p = (1-p)^n = e^{-\alpha n}$ mit $\alpha=-\log(1-p) > 0$.

(b) Das diskrete Analogon zur Γ–Vert. ist die **negative Binomialvert. Nb(ν,p)**, $\nu>0$, $0<p<1$, mit der Z–Dichte $\mathrm{Nb}(\nu,p;k) := \binom{\nu+k-1}{k}p^\nu(1-p)^k =: (-1)^k\binom{-\nu}{k}p^\nu(1-p)^k$, $k\in\mathbb{N}_0$, also $\mathrm{Nb}(1,p) = \mathrm{Geo}(p)$.

§13 Stochastische Unabhängigkeit

13.1 Satz. Für $1\le i\le n<\infty$ seien X_i Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$.

(a) Sind Q_i W–Maße auf $(\Omega_i,\mathfrak{F}_i)$, dann gilt:

$X_1,...,X_n$ unabh.; $X_i\sim Q_i$, $1\le i\le n$ $\Leftrightarrow$ $\times_1^n Q_i$ ist die gemeinsame Vert. der X_i, $1\le i\le n$.

(b) Sind μ_i σ–endliche Maße, f_i μ_i–Dichten eines W–Maßes Q_i auf $(\Omega_i,\mathfrak{F}_i)$, dann gilt:

$X_1,...,X_n$ unabh.; f_i ist μ_i–Dichte von X_i, $1\le i\le n$ $\Leftrightarrow$

$f(x_1,...,x_n) := \Pi_1^n f_i(x_i)$ ist gemeinsame μ–Dichte der X_i, $1\le i\le n$, mit $\mu = \times_1^n \mu_i$.

(c) Sind $\mathfrak{C}_i \subset \mathfrak{P}(\Omega_i)$ $\cap$–stabil mit $\sigma(\mathfrak{C}_1\times...\times\mathfrak{C}_n) = \mathfrak{F}_1\otimes...\otimes\mathfrak{F}_n$ (vgl. 6.13), dann gilt:

$X_1,...,X_n$ unabh. $\Leftrightarrow$ $P[X_1\in C_1,...,X_n\in C_n] = \Pi_1^n P[X_i\in C_i]$ $\quad\forall\, C_i\in\mathfrak{C}_i$, $1\le i\le n$.

Bew. Es gilt:

(13.2) $\qquad P[X_1\in B_1,...,X_n\in B_n] = P(X_1,...,X_n)^{-1}[B_1\times...\times B_n]$

a) $X_1,...,X_n$ unabh.; $X_i\sim Q_i$ $\forall i$ $\Leftrightarrow$ $P[X_1\in B_1,...,X_n\in B_n] = \Pi_i Q_i[B_i]$ $\forall\, B_i$;

dabei folgt "$\Leftarrow$" mit $B_j=\Omega_j$ $\forall\, j\neq i$. Die gleiche Beziehung wurde in 5.1 für einen Spezialfall gezeigt.

Gemäß (13.2) und 11.7 [für n=2 direkt und für n>2 durch Induktion] ist das gleichbedeutend mit: $P(X_1,...,X_n)^{-1} = \times_1^n Q_i$

b) "$\Rightarrow$" folgt aus (a)"$\Rightarrow$", da gemäß 11.17 gilt:

(*) $\quad$ f ist μ–Dichte von $\times_1^n Q_i$.

"$\Leftarrow$" Nach Vorauss. und (*) gilt: $P(X_1,...,X_n)^{-1} = \times_1^n Q_i$. Jetzt (a)"$\Leftarrow$".

c) "$\Leftarrow$" $\mathfrak{C}_1\times...\times\mathfrak{C}_n$ ist $\cap$–stabil. Nach dem Eindeutigkeitssatz folgt nun : $P(X_1,...,X_n)^{-1} = \times_1^n PX_i^{-1}$; jetzt kann (a)"$\Leftarrow$" angewendet werden. $\square$

13.3 Folgerung. Bei unabh. Zva ist die gemeinsame Verteilung durch die 1–dim. Randverteilungen bestimmt.

13.4 Folgerung: Sind X_i, i=1,2, unabh. Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$; dann gilt:

$$P[(X_1,X_2)\in B] = \int PX_2^{-1}[B_x]\ PX_1^{-1}[dx] = \int P[(x,X_2)\in B]\ PX_1^{-1}[dx] \quad \text{für } B\in\mathfrak{F}_1\otimes\mathfrak{F}_2.$$

Als Beispiel erhält man für reelle Zva: $P[X_1+X_2\le t] = \int P[x+X_2\le t]\ PX_1^{-1}[dx]$.

Hat X_2 die λ^1–Dichte f_2 oder haben X_1 und X_2 die λ^1–Dichte f_1 bzw. f_2, so ergibt sich:

$$P[X_1+X_2\le t] = \int\left[\int \mathbf{1}_{(-\infty,t]}(x+y)\ f_2(y)\ dy\right] PX_1^{-1}[dx] = {}_{-\infty}\!\int^t \left[\int f_2(z-x)\ PX_1^{-1}[dx]\right] dz \quad \text{sowie}$$

$$P[X_1+X_2\le t] = {}_{-\infty}\!\int^t f_2\star f_1(z)\ dz \quad \text{mit} \quad f_2\star f_1(z) = \int f_2(z-x)\ f_1(x)\ dx.$$

13.5 Konstruktion eines W–Modells:
Es existieren n unabhängig Zva mit vorgegebenen Verteilungen. Genauer:
Gegeben seien W–Räume $(\Omega_i,\mathfrak{F}_i,Q_i)$, $1\le i\le n$. **Gesucht** ist ein W–Raum $(\Omega,\mathfrak{F},P)$ und Zva X_i auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$, so daß $X_1,\dots,X_n$ unabh. sind mit $X_i \sim Q_i$, $1\le i\le n$.
Lösung: Wähle $\Omega := \times_1^n \Omega_i$, $\mathfrak{F} := \otimes_1^n \mathfrak{F}_i$, $P = \times_1^n Q_i$, X_i als Projektion $X_i : \Omega \mapsto \Omega_i$.
Dann ist X_i mb (vgl. 6.9), $(X_1,\dots,X_n) = \mathrm{Id}_\Omega$ und $P(X_1,\dots,X_n)^{-1} = P = \times_1^n Q_i$. Jetzt 13.1a.
Diese Konstruktion der Zva kann auch für n=1 nützlich und bequem sein. $\square$

13.6 Satz. Funktionen von unabh. Zva sind wieder unabh.:
Seien X_i unabh. Zva auf $(\Omega.\mathfrak{F},P)$ mit Werten in $(\Omega_i,\mathfrak{F}_i)$, $i\in I$, I abzb. und $I \supset \cup_{j\in J} I_j$; dann sind die Zva $g_j\circ(X_i,i\in I_j)$, $j\in J$, unabh., wobei $g_j : \times_{i\in I_j} \Omega_i \mapsto \Omega'_j$ $\quad \otimes_{i\in I_j}\mathfrak{F}_i$–$\mathfrak{F}'_j$–mb, $j\in J$, ist.

Bew. Sei o.E. J endlich, vgl. Def. 5.6..
Wegen $\{g_j\circ(X_i,i\in I_j)\in B_j\} = \{(X_i,i\in I_j)\in g_j^{-1}(B_j)\}$ sei o.E. g_j die Identität.
Aus schreibtechnischen Gründen wird hier nur der Fall $J=\{0,1\}$, $I_0=\{0\}$, $I_1=\mathbb{N}$ behandelt.
Also z.z.: $X_i,i\in\mathbb{N}_0$, unabh. $\Rightarrow$ X_0, $(X_i,i\in\mathbb{N})$ unabh. Zva mit Werten in Ω_0 und $\times_1^\infty \Omega_i$.
Nach Voraussetzung gilt:

(*) $\quad P[X_0\in A_0,(X_1,\dots,X_n)\in A_1\times\dots\times A_n] = \Pi\, P[X_i\in A_i] = P[X_0\in A_0]\,P[(X_1,\dots,X_n)\in A_1\times\dots\times A_n]$.

Setze $\mathfrak{C} := \times_1^\infty \mathfrak{F}_i$; $\mathfrak{C}$ ist $\cap$–stabil, enthält $\times_1^\infty \Omega_i$ und erzeugt $\otimes_1^\infty \mathfrak{F}_i$. Wegen

(13.7) $\qquad A_1\times A_2\times\dots = \cap_n\, A_1\times\dots\times A_n\times\Omega_{n+1}\times\Omega_{n+2}\times\dots$.

und (*) gilt nun:

$$P[X_0\in A_0,\ (X_i,i\in\mathbb{N})\in C] = P[X_0\in A_o]\cdot P[(X_i,i\in\mathbb{N})\in C] \quad \text{für } A_0\in\mathfrak{F}_0\ (=:\mathfrak{C}_0),\ C\in\mathfrak{C}.$$

Jetzt kann 13.1c angewendet werden, wenn gilt: $\sigma(\mathfrak{F}_0\times\mathfrak{C}) = \mathfrak{F}_0\otimes\sigma(\mathfrak{C})$ $(= \mathfrak{F}_0 \otimes \otimes_1^\infty \mathfrak{F}_i = \otimes_o^\infty \mathfrak{F}_i)$.
Dies ist aber sehr einfach zu sehen. $\square$

13.8 Satz. Seien X und Y unabhängige Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\mathbb{R}^d,\mathfrak{B}_d)$ bzw. $(E,\mathfrak{E})$; sei f λ^d–Dichte von X, $b:E\mapsto\mathbb{R}^d$ mb, $A:E\mapsto\mathbb{R}^{d\times d}$ mb, so daß die Matrixinverse $A(Y)^{-1}$ f.s. existiert. Dann hat $A(Y)X+b(Y)$ die λ^d–Dichte

$$z \mapsto \int |\det A(y)|^{-1} f(A(y)^{-1}(z-b(y)))\, PY^{-1}[dy].$$

Bew. mit 12.9 und Fubini als Übung.

13.9 Spezialfälle. Sei $E=\mathbb{R}^d$, g eine λ^d–Dichte von Y $[\Rightarrow Y\neq 0 \text{ f.s.}]$ und h die Dichte von $Z:=A(Y)X+b(Y)$ aus 13.8:

(a) $Z=X\pm Y$; dann $h(z) = \int f(z\mp y)g(y)dy$.
(b) $d=1$, $Z=X\cdot Y$; dann $h(z) = \int f(x/y)\cdot g(y)\cdot |y|^{-1} dy$.
(c) $d=1$, $Z=X/Y$ $[:=1_{\{Y\neq 0\}}X/Y]$; dann $h(z) = \int [f(xy)\cdot g(y)\cdot |y|]\,dy$.

13.10 Definition. (a) Sind μ und ν W–Maße auf $(\mathbb{R}^d,\mathfrak{B}_d)$ und ist $T : \mathbb{R}^d\times\mathbb{R}^d \mapsto \mathbb{R}^d$ die durch $T(x,y)=x+y$ def. Abb., so heißt das W–Maß $(\mu\times\nu)T^{-1}=:\mu\star\nu$ die **Faltung** von μ und ν,
(b) Sind f und g λ^d–Dichten von W–Maßen, so heißt $x\mapsto f\star g(x) := \int f(x-y)g(y)dy$ **Faltung** von f und g. Gilt speziell: $d=1$, $f=g=0$ auf $(-\infty.0)$, so ist $f\star g(x) := {}_0\int^x f(x-y)g(y)dy$.
(c) Sind f und g Z–Dichten auf $\mathbb{N}_0$, so heißt $f\star g(k) := \sum_{j=0}^k f(k-j)g(j)$ **Faltung** von f und g.

13.11 Eigenschaften. Die Faltung ist kommutativ und assoziativ.

13.12 Satz. (a) Sind X_1 und X_2 unabh. d–dim. Zva auf $(\Omega.\mathfrak{F},P)$, so gilt:

$$P(X_1+X_2)^{-1}= PX_1^{-1}\star PX_2^{-1}.$$

(b1) Sind X_1 und X_2 unabhängige Zva mit Werten in $\mathbb{R}^d$ und mit λ^d–Dichten f_i, i=1,2, so hat X_1+X_2 die λ^d–Dichte $f_1\star f_2$.
(b2) Seien Q_i W–Maße auf $\mathbb{R}^d$. Hat (nur) Q_1 die λ^d Dichte f_1, so hat $Q_1\star Q_2$ die λ^d–Dichte $x \mapsto \int f_1(x-y)Q_2[dy]$.
(c) Sind X_1 und X_2 unabhängige Zva mit Werten in $\mathbb{N}_0$ und mit Z–Dichten f_i, i=1,2, so hat X_1+X_2 die Z–Dichte $f_1\star f_2$.

Bew. a) $P(X_1+X_2)^{-1}=: PT(X_1,X_2)^{-1}=(P(X_1,X_2)^{-1})T^{-1}=(PX_1^{-1}\times PX_2^{-1})T^{-1}$.
(b1) ergibt sich aus Folgerung 13.4. (b2) Wähle (gemäß 13.5) unabh. Zva X und Y mit Dichte f bzw. $Y\sim Q_2$ und verwende 13.8 mit $A(Y)X + b(Y) = X + Y$.
c) $\{X+Y=k\} = \cup_{i=0}^k \{X=k-i,Y=i\}$. $\square$

13.13 Abgeschlossenheit gegenüber der Faltung.
(a) $b(n,p)\star b(m,p) = b(n+m,p)$;
(b) $\pi(\lambda_1)\star\pi(\lambda_2) = \pi(\lambda_1+\lambda_2)$;
(c) $N(a,\sigma^2)\star N(b,\tau^2) = N(a+b,\sigma^2+\tau^2)$;
(d) $\Gamma_{\alpha,\mu}\star\Gamma_{\alpha,\nu} = \Gamma_{\alpha,\mu+\nu}$, speziell: $\chi^2_n\star\chi^2_m = \chi^2_{n+m}$.

Bew. durch Rechnung oder später mit Hilfe der Transformierten (vgl. § 15). Dabei gilt (a) auch wegen $b(n,p) = b(1,p)*...*b(1,p)$ und (d) für $\mu,\nu \in \mathbb{N}$ wegen $\Gamma_{\alpha,n} = \Gamma_{\alpha,1}*...*\Gamma_{\alpha,1}$.

§14 Erwartungswert und Varianz.

Seien stets X, $X_1,\dots,X_n$,Y erweitert relle Zva auf $(\Omega,\mathfrak{F},P)$.

Erwartungswert.

14.1 Def. Ist X P–quasi–intb., so heißt $EX = E[X] := \int XdP$ **Erwartungswert** (Erw.) von X. Dabei sagt man: X quasi–intb. $\Leftrightarrow$: EX existiert; X intb. $\Leftrightarrow$: EX ist endlich.

Gemäß § 9 kennen wir bereits die folgenden Eigenschaften.

14.2 Korollar. (a) Jede nach oben oder nach unten beschränkte Zva. ist quasi–intb.;

(b) EX ex. $\Leftrightarrow E[X^+]<\infty$ oder $E[X^-]<\infty \Rightarrow EX = E[X^+] - E[X^-]$;

(c) $E1_A = P[A]$, $A\in\mathfrak{F}$;

(d) EX, EY ex., $X\le Y \Rightarrow EX \le EY$;

(e) EX ex. $\Rightarrow |EX| \le E[|X|]$;

(f) X=Y f.s., EX ex. $\Rightarrow$ EY =EX;

(g) EX ex., $\alpha\in\mathbb{R}$, oder $X\ge 0$, $0\le\alpha\le\infty \Rightarrow E[\alpha X] = \alpha EX$;

(h) EX, EY ex., EX + EY ist def. $\Rightarrow$ X+Y ist f.s. def., $E[X+Y] = EX + EY$;

(i) EX endlich $\Leftrightarrow E[|X|] <\infty$;

(j) $a\le X$, $a<X$, $X\le b$, bzw.$X<b$ für gewisse $a,b\in\mathbb{R} \Rightarrow a\le EX$, $a<EX$, $EX\le b$, bzw. $EX<b$.

14.3 Korollar. Sei Z Zva mit Werten in $(\Omega',\mathfrak{F}')$, $g:\Omega'\mapsto\bar{\mathbb{R}}$ mb, μ Maß auf $(\Omega',\mathfrak{F}')$. Die folgenden Identitäten gelten mit der Maßgabe, daß die linke Seite genau dann existiert, wenn die rechte Seite existiert.

(a) $E[g(Z)] = \int gdQ$, falls Z~Q; insbesondere:
$E[X^k] = \int x^k Q[dx]$,$k\in\mathbb{N}$, falls X~Q.

(b) Sei Ω' abzb., dann: $E[g(Z)] = \sum_{i\in\Omega'} g(i)\, P[Z=i]$.

(c) Z habe die μ–Dichte f, dann: $E[g(Z)] = \int g\cdot f\, d\mu$;insbesondere:
$E[X^k] = \int x^k f(x)dx$, $k\in\mathbb{N}$, falls X die Lebesgue–Dichte f hat.

$E[g(Z)]$ hängt also von Ω,P,Z nur über $Q = PZ^{-1}$ ab.

Bew. a) folgt aus dem Transformationssatz 10.19, b) ist ein Spezialfall von (c) mit μ als Zählmaß. c) folgt aus (a) und 10.10. $\square$

14.4 Beispiel. $X \sim U(a,b) \Rightarrow EX = \frac{1}{b-a}\cdot \int_a^b x dx = \frac{a+b}{2}$.

$X \sim N(0,1) \Rightarrow EX = \int x\,\varphi(x)\, dy = 0$ mit der Dichte φ zu N(0,1), da $x\cdot\varphi(x)$ ungerade ist.

14.5 Def. Die **Cauchy–Verteilung** ist die Verteilung auf $(\mathbb{R}^1,\mathfrak{B}_1)$ mit der λ^1–Dichte:

$$x \mapsto \frac{1}{\pi}\cdot\frac{1}{1+x^2}\quad \left[= \frac{1}{\pi}\cdot\frac{d}{dx}\arctan(x)\right].$$

14.6 Beispiel. Ist X Cauchy–verteilt, so gilt: $E[X^+] = E[X^-] = \infty$, d.h. EX ex. nicht. Nach 12.10 und 14.3a gilt nämlich: $E[X^-] = E[(-X)^-] = E[X^+] = {}_0\int^\infty \frac{1}{\pi}\cdot\frac{x}{1+x^2}dx \ge \frac{1}{2\pi}\cdot {}_1\int^\infty \frac{1}{x}\,dx = \infty$. □

14.7 Satz. (a) Ist $X\ge 0$, so gilt: $EX = {}_0\int^\infty P[X>t]\,dt$.

(b) Hat X Werte in $\mathbb{N}_0$, so gilt: $EX = \sum_{n=0}^\infty P[X>n]$.

Bew. a) Sei $B = \{(t,x),\ 0<x<t\}$ $(\in \mathfrak{B}_2)$, $X\sim Q$, dann folgt mit Fubini:

$${}_0\int^\infty P[X>t]\,dt = \int Q[B_t]\lambda^1[dt] = (\lambda^1\times Q)[B] = \int \lambda^1[B_x]Q[dx] = \int x^+ Q[dx] = E[X^+].$$

b) wie oder mit (a). □

14.8 Beispiel. Für $X \sim \mathrm{Geo}(p)$ folgt (vgl. 12.18a): $EX = \sum_{n=0}^\infty (1-p)^{n+1} = \frac{1-p}{p}$.

14.9 Korollar zu Fubini. Seien Z_i unabh. Zva mit Werten in $(\Omega_i,\mathfrak{F}_i)$, i=1,2, $h:\Omega_1\times\Omega_2 \mapsto \bar{\mathbb{R}}$ mb, sodaß $E[h(Z_1,Z_2)]$ existiert, dann gilt:

$$E[h(Z_1,Z_2)] = E[\bar{h}(Z_1)] \quad \text{mit} \quad \bar{h}(z) = E[h(z,Z_2)],$$

insbesondere: $P[(Z_1,Z_2)\in B] = E[p(Z_1)]$ mit $p(z) = P[(z,Z_2)\in B] = P[Z_2\in B_z]$.

14.10 Korollar. Gilt $X_i\ge 0$, $1\le i\le n$, oder $E[|X_i|]<\infty$, $1\le i\le n$ und sind $X_1,\dots,X_n$ unabhängig,

so gilt: $E[\Pi_1^n X_i] = \Pi_1^n E[X_i]$, insbesondere ex. $E[\Pi_1^n X_i]$.

Bew. o.E. n=2, jetzt 11.12 mit 13.1a. □

Ungleichungen.

14.11 Markoffsche Ungleichung. (vgl. 9.16 d):

$$P[|X|\ge\varepsilon] \le \varepsilon^{-p} E[|X|^p] \quad \forall p>0.$$

14.12 Schwarzsche Ungleichung (Spezialfall der Hölderschen Ungleichung):

Im Falle $E[X^2]<\infty$, $E[Y^2]<\infty$ gilt:

$$|E[X\cdot Y]| \le \{E[X^2]\cdot E[Y^2]\}^{\frac{1}{2}}$$

mit "=" $\Leftrightarrow$ $X=\alpha Y$ f.s. für ein $\alpha\in\mathbb{R}$ oder $Y=0$ f.s..

Bew. Wegen $|E[X\cdot Y]| \le E[|X|\cdot|Y|]$ sei o.E.: $X\ge 0, Y\ge 0$; jetzt wie üblich. □

14.13 Jensensche Ungleichung. Sei I bel Intervall in $\mathbb{R}$ (z.B. $I=\mathbb{R}$) mit $X\in I$ f.s., $E[|X|]<\infty$, $g:I\mapsto\mathbb{R}$ konvex. Dann gilt: $EX\in I$ und

$$g(EX) \le E[g(X)].$$

Interpretation: Bei konvexer Verlustfkt. (konkaver Nutzenfunktion) empfiehlt sich Risikoaversion.

Bew. $EX \in I$ folgt aus 14.2j. Der Fall $EX \in \partial I$ impliziert nach 9.16a: X = const f.s., ist also klar. Liege nun EX im Innern von I, dann ex. Stützgerade $y=\alpha x+\beta$ an g im Punkt EX; also $g(x) \geq \alpha x+\beta$ und $g(EX)=\alpha EX+\beta$. $\Rightarrow E[g(X)] \geq E[\alpha X+\beta] = \alpha EX+\beta = g(EX)$. []

14.14 Anwendung. Für p>0 sei $\|X\|_p := E[|X|^p]^{1/p}$, dann gilt für q<p: $\|X\|_q \leq \|X\|_p$.

Bew. $I=[0,\infty)$, $g(x) = x^{p/q}$ ist konvex. Sei $\|X\|_q < \infty$; dann folgt nach 14.13:
$E[|X|^q]^{p/q} \leq E[|X|^{q \cdot p/q}]$. Zum Fall $\|X\|_q = \infty$ vgl.14.15b. []

Momente

14.15 Lemma. Sei p>0, $a \in \mathbb{R}$. Dann gilt:
(a) $E[|X-a|^p] < \infty \Leftrightarrow E[|X|^p] < \infty \Leftrightarrow E[|X-b|^p] < \infty \;\; \forall\, b \in \mathbb{R}$.
(b) $E[|X|^p] < \infty \Rightarrow E[|X|^q] < \infty \;\; \forall\, 0 \leq q \leq p$.

Bew. a) Für $\alpha, \beta \in \mathbb{R}$ gilt: $|\alpha+\beta|^p \leq (2 \cdot |\alpha| \vee |\beta|)^p \leq 2^p(|\alpha|^p + |\beta|^p)$.
b) $|X|^q \leq 1 + |X|^p$ für $0 \leq q \leq p$. []

14.16 Def. Im Falle der Existenz heißen $E[X^k]$ **k–tes Moment** und bei endlichem Erw. a=EX: $E[(X-a)^k]$ **k–tes zentrales Moment**, $V[X] := E[(X-a)^2]$ **Varianz** von X.

14.17 Bemerkung. (a) $E[X^2] < \infty \Leftrightarrow$ X hat endliche Varianz. (Bew. 14.15).
(b) $V[X]=0 \Leftrightarrow X=\text{const}\ (\in \mathbb{R})$ f.s. (Bew. 9.16.a). []

14.18 Def. Ist $E[X^2] < \infty$, $E[Y^2] < \infty$, so heißen
$\mathrm{Kov}[X,Y] := E[(X-EX)(Y-EY)]$ **Kovarianz** von X und Y,
$\mathrm{Kor}[X,Y] := \mathrm{Kov}[X,Y] \cdot \{V[X] \cdot V[Y]\}^{-\frac{1}{2}}$ **Korrelationskoeffizient** von X und Y.
X und Y heißen **unkorreliert**, **positiv korreliert** bzw. **negativ korreliert**, falls
$\mathrm{Kor}[X,Y] = 0, > 0$ bzw. < 0.

14.19 Satz. Sei $E[X^2] < \infty$, $E[Y^2] < \infty$, dann gilt: (a) $\mathrm{Kov}[X,X] = V[X]$;
(b) $V[X] = E[X^2] - E[X]^2$ (**Steinersche Gleichung**),
$\mathrm{Kov}[X,Y] = E[XY] - E[X] \cdot E[Y]$;
(c) $|\mathrm{Kor}[X,Y]| \leq 1$ und $|\mathrm{Kor}[X,Y]| = 1 \Leftrightarrow \exists\, \alpha \neq 0, \beta \in \mathbb{R}$, mit $X = \alpha Y + \beta$ f.s., $V[X] \neq 0 \neq V[Y]$.

Bew. a),b) leicht, c) mit der Schwarzschen Ungleichung. []

14.20 Beispiel. Betrachte n–malige Wiederholung eines 2–stufigen Experiments (etwa eine repräsentative Umfrage nach 2 Merkmalen). Als Ergebnis (Stichprobe) erhält man: $(x_1,y_1),\ldots,(x_n,y_n) \in \mathbb{R}^2$ (etwa x_i, y_i Körpergröße oder IQ von Vater und Sohn bei Paar i).

Wähle $\Omega_1 \supset \cup \{x_i\}$, $\Omega_2 \supset \cup \{y_i\}$, [o.E. Ω_i endlich]

$X:\Omega_1\times\Omega_2 \mapsto \Omega_1$ und $Y:\Omega_1\times\Omega_2\mapsto\Omega_2$ als Projektionen sowie:

$P[A] := \mathrm{Emp}[A\,|\,(x_1,y_1),\ldots,(x_n,y_n)] = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_A((x_i,y_i))$ für $A\subset\Omega_1\times\Omega_2$.

Dann $E[g(X,Y)] = \sum_{x,y} g(x,y)\, P[\{(x,y)\}] = \frac{1}{n}\sum_i g(x_i,y_i)$,

also $EX = \frac{1}{n}\sum x_i =: \bar{x}$, $\bar{x}$ heißt **Stichprobenmittel** $EY = \frac{1}{n}\sum y_i =: \bar{y}$,

$V[X] = \frac{1}{n}\sum (x_i-\bar{x})^2$, dabei heißt $\frac{1}{n-1}\sum(x_i-\bar{x})^2$ **Stichprobenstreuung**,

[der Faktor $\frac{1}{n-1}$ statt $\frac{1}{n}$ wird in der Statistik erklärt] $V[Y] = \frac{1}{n}\sum(y_i-\bar{y})^2$,

$\mathrm{Kov}[X,Y] = \frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y})$, $\mathrm{Kor}[X,Y] =: r_{xy}$ heißt **empirischer Korrelationswert**.

$r_{xy} \in [-1,+1]$ wird als Maß für den linearen Zusammenhang interpretiert.

Als kritische Werte werden $\pm\{\frac{4}{n+2}\}^{\frac{1}{2}}$ angesehen.

14.21 Satz. Sei $E[X_i^2]<\infty$, $i=1,\ldots,n$. Dann gilt:

(a) $\sum_1^n X_i$ hat endliche Varianz $V[\sum X_i] = \sum_i V[X_i] + 2\cdot\sum_{i<j}\mathrm{Kov}[X_i,X_j]$;

(b) $X_1,\ldots,X_n$ unabh. $\Rightarrow X_1,\ldots,X_n$ paarweise unkorreliert $\Rightarrow$

$$V[\sum X_i] = \sum_i V[X_i] \quad \text{(\textbf{Gleichung von Bienaymé})}.$$

Bew. a) Sei $n=2$. $(X_1+X_2)^2 \le X_1^2+X_2^2+2\,|X_1X_2|$; jetzt benutze die Schwarzsche Ungleichung. Dann o.E. $E[X_i]=0$, betrachte sonst $X_i' := X_i - E[X_i]$. b) mit 14.10. $\square$

14.22 Beispiele.

(a) $X\sim U(a,b) \Rightarrow E[X]=\frac{a+b}{2}$, $E[X^k] = \frac{1}{b-a}\cdot {}_a\int^b x^k dx$, $V[X] = E[X^2] - E[X]^2 = \frac{1}{12}(b-a)^2$.

(b) $X\sim\Gamma_{\alpha,\nu} \Rightarrow E[X]=\nu/\alpha$, $V[X] = \nu/\alpha^2$. Allgemeiner gilt:

$E[X^k] = (k+\nu-1)(k+\nu-2)\ldots(\nu+1)\nu\cdot\alpha^{-k}$;

$E[X^k] = \frac{1}{\Gamma(\nu)}\, {}_0\int^\infty x^k \alpha^\nu x^{\nu-1} e^{-\alpha x} dx = \frac{1}{\Gamma(\nu)}\,\alpha^{-k}\, {}_0\int^\infty z^{k+\nu-1} e^{-z} dz = \frac{1}{\Gamma(\nu)}\,\alpha^{-k}\,\Gamma(k+\nu)$.

Jetzt benutze die Eingeschaft der Γ–Funktion (12.13) .

(c) $X\sim N(0,1) \Rightarrow E[X] = 0$, $V[X] = E[X^2] = 1$. Allgemeiner gilt:

$E[X^{2k-1}] = \int x^{2k-1}\varphi_{0,1}(x)dx = 0$ da $x\mapsto x^{2k-1}\varphi_{0,1}(x)$ ungerade;

$E[X^{2k}] = E[(X^2)^k]$ mit $X^2\sim\Gamma_{\frac{1}{2},\frac{1}{2}}$ (vgl. 12.17a), also nach (b)

$E[(X^2)^k] = (k-\frac{1}{2})(k-\frac{1}{2}-1)\ldots\frac{1}{2}\cdot(\frac{1}{2})^{-k}$, somit: $E[X^{2k}] = (2k-1)\cdot(2k-3)\cdot\ldots\cdot 3\cdot 1$.

$X \sim N(a,\sigma^2) \Rightarrow EX = a$, $V[X] = E[(X-a)^2] = \sigma^2$. Allgemeiner gilt:
$(X-a)/\sigma \sim N(0,1) \Rightarrow E[(X-a)^{2k-1}] = 0$, $E[(X-a)^{2k}] = (2k-1)\cdot(2k-3)\cdot\ldots\cdot 3\cdot\sigma^{2k}$, $k\in\mathbb{N}$.

(d) $X \sim b(n,p) \Rightarrow EX = np$, $V[X] = np(1-p)$.

Bew. Gemäß 5.15 o.E. $X = \sum_1^n X_i$ mit (X_i) iid $\sim b(1,p)$, dann $E[X_i]=E[X_i^2]=0\cdot(1-p)+1\cdot p=p$ also $V[X_i]=p-p^2=p(1-p)$. Jetzt 14.2h und 14.21b.

(e) $X \sim \pi(\lambda) \Rightarrow EX = V[X] = \lambda$.

Bew. mit $k\pi(\lambda;k) = \lambda\pi(\lambda;k-1)$ und $E[X(X-1)] = \lambda^2$ wegen $k(k-1)\pi(\lambda;k) = \lambda^2\pi(\lambda;k-2)$.

§15 Transformierte.

Erwartungswert komplexer Zva.

15.1 Def. Sei $(\Omega,\mathfrak{F},P)$ ein W–Raum, $i=\sqrt{-1}$.

(a) $Z:\Omega\mapsto\mathbb{C}$ heißt eine **komplexe Zva**, falls $\Re(Z)$ und $\Im(Z)$ reelle Zva sind ($\Leftrightarrow$ $(\Re(Z),\Im(Z))$ ist 2–dim. Zufallsvektor).

(b) Sind für die komplexe Zva Z die Erwartungswerte von $\Re(Z)$ und $\Im(Z)$ endlich, so sagt man:"EZ ex. in $\mathbb{C}$" und nennt $EZ := E[\Re(Z)] + iE[\Im(Z)]$ den **Erwartungswert** von Z.

15.2 Bemerkungen. Durch Zerlegung in Real– und Imaginärteil zeigt man.

(a) Der Erwartungswert bleibt auch für komplexe Zva **linear**.

(b) Ex. EZ in $\mathbb{C}$, so gilt: $E[\bar{Z}] = \overline{EZ}$.

(c) EZ ex in $\mathbb{C} \Leftrightarrow E[|Z|] <\infty \Rightarrow |EZ| \le E[|Z|]$.
Bew. "$\Leftrightarrow$" $|Z| \le |\Re(Z)| + |\Im(Z)| \le 2|Z|$; "$\Rightarrow$" vgl. Übung, dort Z d–dim.

(d) Der **Satz von der majorisierten Konvergenz** gilt auch für komplexwertige Zva unter der Vorauss. $|Z_n|\le Y$ mit $E[|Y|] < \infty$.

(e) Der **Satz von Fubini** gilt auch für komplexwertige Zva, deren Erw. in $\mathbb{C}$ ex.

(f) Sind Z_1 und Z_2 unabh. komplexe Zva mit $E[|Z_i|] <\infty$, i=1.2, dann gilt:
$$E[Z_1\cdot Z_2] = E[Z_1]\cdot E[Z_2].$$

Laplace–Transformierte.

15.3 Def. (a) Ist Q ein W–Maß auf $[0.\infty)\cap\mathfrak{B}_1$, so heißt $\psi:[0,\infty)\mapsto[0,1]$ mit
$\psi(t) := \int e^{-tx}Q[dx]$ die **Laplace–Transformierte** (LT) von Q.

(b) Ist X eine Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $[0,\infty)$ [oder mit Werten in $[0,\infty]$ und $X\neq\infty$ f.s.], so heißt die LT ψ von PX^{-1}, also (gemäß 14.3a) $\psi(t) = E[e^{-tX}]$ auch LT von X.

15.4 Def. (a) Ist Q ein W–Maß auf $\mathbb{N}_o$, so heißt $g:[0,1]\mapsto[0,1]$ mit

$$g(s) := \Sigma_{n=0}^{\infty} Q[\{n\}]s^n$$ **erzeugende Funktion von Q.**

(b) Ist X eine Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $\mathbb{N}_0$ [oder mit Werten in $\bar{\mathbb{N}}_0$ und $X\neq\infty$ f.s.], so heißt die erzeugende Funktion von PX^{-1}, also

$$g(s) = \Sigma_{n=0}^{\infty} P[X=n]s^n = E[s^X]$$ auch **erzeugende Funktion von X**.

15.5 Bemerkung. Ist Q ein W–Maß auf $\mathbb{N}_0$ mit erzeugender Funktion g, so ist $\psi(t) = g(e^{-t})$ die LT von Q wegen $\int e^{-tx}Q[dx] = \sum e^{-tn}Q[\{n\}]$ und umgekehrt entsprechend. Deswegen o.E. hier nur Behandlung der LT.

15.6 Satz. Sei ψ die LT zum W–Maß Q und der Zva X. Dann gilt:

(a) $P[X=0]=1 \Leftrightarrow \psi \equiv 1$,
$P[X=0]<1 \Leftrightarrow \psi$ streng antiton $\Leftrightarrow P[X=0] = \psi(+\infty) := \lim_{t\to\infty} \psi(t) < \psi(t) < \psi(0) = 1,\ 0<t<\infty$.

(b) $\psi(z) := \int e^{-zx}Q[dx]$ ist holomorph in $G = \{z\in\mathbb{C};\ \Re(z)>0\}$ und stetig in $\bar{G} = \{z\in\mathbb{C};\ \Re(z)\geq 0\}$, also ist ψ bel. oft db. in $(0,\infty)$ und es gilt: $\psi^{(k)}(0+) = (-1)^k E[X^k]$.

(c1) Sind X,Y unabh. Zva mit LT ψ und χ, so ist $\psi\cdot\chi$ LT von X+Y.

(c2) Ist ψ_k LT von Q_k, k=1,2, so ist $\psi_1\cdot\psi_2$ LT von $Q_1\star Q_2$.

Bew. a) Für s<t ist $\psi(s)-\psi(t) = \int_{(0,\infty)}(e^{-sx}-e^{-tx})Q[dx]$, also $\psi(s)-\psi(t)=0$ falls $Q[(0,\infty)]=0$ und $\psi(s)-\psi(t)>0$ falls $Q[(0,\infty)]>0$ gemäß 9.16a mit $\mu = Q|_{(0,\infty)\cap\mathfrak{B}_1}$.

Ferner mit maj. Konv.: $\lim_{t\to\infty} \int e^{-tx}Q[dx] = \int\lim... = \int 1_{\{0\}}dQ = Q[\{0\}]$.

b) Die Stetigkeit von ψ auf $\bar{G}$ folgt aus 10.15 mit $\mathbb{C}\cong\mathbb{R}^2$.

Für $z\in G$ und $|\zeta|\leq\delta<\Re(z)=:r$ gilt:

$$|(e^{-(z+\zeta)x}-e^{-zx})/\zeta| = e^{-rx}|(e^{-\zeta x}-1)/\zeta| \leq e^{-rx}\cdot\Sigma_{n>0}\,\delta^{n-1}x^n \leq e^{-(r-\delta)x}/\delta\ .$$

Nun folgt wie in 10.16 mit maj.Konv.:

$$\lim_{\zeta\to 0} \int(e^{-(z+\zeta)x}-e^{-zx})/\zeta\ Q[dx] = \int\lim... = \int(-x)e^{-zx}Q[dx].$$

Da auch $x\mapsto x^k e^{-(r-\delta)x}$ beschränkt ist auf $[0,\infty)$, folgt mit vollst. Induktion:

$\psi^{(k)}(z) = \int(-x)^k e^{-zx}Q[dx]$, also mit monot. Konv.:

$\psi^{(k)}(0+) = (-1)^k\lim_{t\to 0} \int x^k e^{-tx}Q[dx] = (-1)^k\int x^k Q[dx]$.

c1) $E[e^{-t(X+Y)}] = E[e^{-tX}\cdot e^{-tY}]E[e^{-tX}]\cdot E[e^{-tY}]$ mit 13.6 und 14.10.

c2) folgt aus (c1) Wähle X,Y unabh, mit $X\sim Q_1$, $Y\sim Q_2$. []

15.7 Beispiele. (a) $\pi(\alpha)$ hat die erzeugende Funktion $g(s) = e^{\alpha(s-1)}$.

(b) $b(n,p)$ hat die erzeug. Fkt. $g(s) = (1-p+ps)^n$.

(c) $\Gamma_{\alpha,\nu}$ hat die LT $\psi(t) = (\frac{\alpha}{\alpha+t})^{\nu}$.

Bew. a) $g(s) = e^{-\alpha} \sum \frac{1}{k!}\alpha^k \cdot s^k = e^{-\alpha} \cdot e^{\alpha s}$.

b) b(1,p) hat die erz. Fkt. $g(s) = (1-p)s^0+ps$, jetzt 15.6c mit 5.15.

c) $\psi(t) = \int e^{-tx}\,\Gamma_{\alpha\nu}[dx] = {}_0\int^{\infty} e^{-tx}\,\gamma_{\alpha\nu}(x)\,dx = \frac{\alpha^{\nu}}{\Gamma(\nu)}\cdot {}_0\int^{\infty} e^{-tx}x^{\nu-1}e^{-\alpha x}dx =$

$$\frac{\alpha^{\nu}}{\Gamma(\nu)}\cdot {}_0\int^{\infty} x^{\nu-1}e^{-(\alpha+t)x}dx = \frac{\alpha^{\nu}}{\Gamma(\nu)}\cdot(\alpha+t)^{-\nu}\cdot {}_0\int^{\infty} y^{\nu-1}e^{-y}dy = \frac{\alpha^{\nu}}{\Gamma(\nu)}\cdot(\alpha+t)^{-\nu}\Gamma(\nu).\ \square$$

Charakteristische Funktionen.

15.8 Def. (a) Ist Q ein W–Maß auf $\mathfrak{B}_d$, so heißt $\varphi : \mathbb{R}^d \mapsto \mathbb{C}$ mit

$$\varphi(t) = \int e^{it^{\top}x}\,Q[dx] = \int e^{i\cdot\sum t_k x_k}\,Q[dx] \quad (= \int\cos(t^{\top}x)Q[dx] + i\cdot\int\sin(t^{\top}x)Q[dx])$$

charakteristische Funktion (CF) oder Fourier–Transformierte von Q.

(b) Ist X d–dim. Zufallsvektor auf $(\Omega,\mathfrak{F},P)$, so heißt die CF φ von PX^{-1},also $\varphi(t) = E[\exp\{it^{\top}X\}]$ auch CF von X.

15.9 Satz. $X=(X_1,...,X_d)^{\top}$ habe die CF φ. Dann gilt:

(a) $|\varphi| \le 1 = \varphi(0)$ und φ ist glm. stetig;

(b) Ist $A\in\mathbb{R}^{n\times d}$, $b\in\mathbb{R}^n$, so hat AX+b die CF: $\mathbb{R}^n \ni t \mapsto \exp\{it^{\top}b\}\cdot\varphi(A^{\top}t)$.

(c1) Sind X und Y unabhängige d–dim. Zva mit CF φ bzw. χ, so ist $\varphi\cdot\chi$ CF von X+Y;

(c2) Sind φ_k CF von Q_k, k=1,2, so ist $\varphi_1\cdot\varphi_2$ CF. von $Q_1\star Q_2$.

(d) $\mathbb{R}^m \ni t \mapsto \varphi((t,0,...,0))$ ist CF von $(X_1,...X_m)$ für m<d.

(e) Sind Q_k W–Maße auf $\mathfrak{B}_1$ mit CF φ_k, $1\le k\le d$, so ist

$$(t_1,...,t_d) \mapsto \Pi_{k=1}^{d}\,\varphi_k(t_k) \quad \text{CF von } \mathop{\times}_1^d Q_k.$$

Bew. a) Mit 15.2c folgt: $|\varphi(t)| \le E[|\exp\{it^{\top}X\}|] = 1 = \varphi(0)$. Mit 15.2c,d folgt:
$|\varphi(t)-\varphi(t+h)| = |E[\exp\{it^{\top}X\}(1-\exp\{ih^{\top}X\})]| \le E[|1-\exp\{ih^{\top}X\}|] \to 0$ für $h\to 0$.
b) $E[\exp\{it^{\top}(AX+b)\}] = \exp[it^{\top}b\}\cdot E[\exp\{i(A^{\top}t)^{\top}X\}]$.
c) analog zu 15.6c bzw. mit Fubini 15.2f. d) folgt direkt oder aus (b).
e) $\exp\{it^{\top}x\} = \Pi_{k=1}^{d}\exp\{it_k x_k\}$. Jetzt Fubini 15.2f. $\square$

15.10 Beziehung zur LT.
Sei Q [z.B. $Q = \Gamma_{\alpha\nu}$] ein W–Maß auf $[0,\infty)$ mit LT ψ [z.B. $\psi(t) = (\frac{\alpha}{\alpha+t})^{\nu}$] und CF φ, G = $\{z\in\mathbb{C};\ \Re(z)>0\}$ und $\hat{\psi}$ eine stetige Fortsetzung von ψ auf $\bar{G}$, die in G holomorph ist [z.B. $\hat{\psi}(z) = (\frac{\alpha}{\alpha+z})^{\nu}$]. Eine solche Fortsetzung ist nach 15.6 auch $\psi(z) = \int e^{-zx}Q[dx]$. Nach dem Identitätssatz für holomorphe Funktionen gilt $\psi = \hat{\psi}$ auf $\bar{G}$ und somit: $\varphi(t) = \hat{\psi}(-it)$ [z.B. $\varphi(t) = (\frac{\alpha}{\alpha-it})^{\nu}$].

15.11 Beispiele. (a) $N(0,1)$ hat die CF $\varphi(t) = \exp\{-\frac{1}{2}t^2\}$,

$N(a,\sigma^2)$ hat die CF $\varphi(t) = \exp\{iat-\frac{1}{2}\sigma^2 t^2\}$.

(b) $\overset{d}{\underset{1}{\times}} N(0,\sigma^2)$ hat die λ^d–Dichte $\mathbb{R}^d \ni t \mapsto (2\pi\sigma^2)^{-d/2}\exp\{-\|t\|^2/2\sigma^2\}$ und

die CF $\mathbb{R}^d \ni t \mapsto \nu_{0,\sigma^2}(t) := \exp\{-\sigma^2\|t\|^2/2\}\}$.

Bis auf den Faktor σ^2 bzw σ^{-2} im Exponenten unterscheiden sich Dichte und CF also nur durch die Normierungskonstante. Dies wird im Beweis vom Eindeutigkeitssatz 15.16 benutzt werden.

(c) Die Cauchy–Vert. hat die CF $\varphi(t) = \exp\{-|t|\}$.

(d) $\Gamma_{\alpha,\nu}$ hat die CF $\varphi(t) = (\frac{\alpha}{\alpha-it})^{\nu}$.

Bew. a) Nach 15.9b genügt es, $N(0,1)$ zu betrachten. Da $\sin(tx)\exp\{-\frac{1}{2}x^2\}$ ungerade und $\cos(tx)\exp\{-\frac{1}{2}x^2\}$ gerade ist, gilt: $\varphi(t) = 2\cdot(2\pi)^{-\frac{1}{2}}\cdot {}_0\int^{\infty}\cos(tx)\,\exp\{-\frac{1}{2}x^2\}\,dx = \varphi(-t)$. Also o.E. $t>0$.

Aus 10.16 oder 15.22 folgt: $\varphi'(t) = -2\cdot(2\pi)^{-\frac{1}{2}}\,{}_0\int^{\infty}\sin(tx)\cdot x\cdot\exp\{-\frac{1}{2}x^2\}\,dx$.

Mit partieller Integration ergibt sich $\varphi'(t) = -t\,\varphi(t)$ bei $\varphi(0)=0$.

b) Mit $\|t\|^2 = \sum t_k^2$ und 11.17 und 15.9e.

c) Wie in a) folgt $\varphi(t) = \frac{1}{\pi}\cdot\int e^{itx}\cdot\{1+x^2\}^{-1}\,x = \varphi(-t) =: \frac{1}{\pi}\int f(x)\,dx$, also o.E. $t>0$.

Sei $f(z) := e^{itx}/(1+z^2) = g(z)/(z-i)$ mit $g(z) := e^{itz}/(z+i)$

und H_r der Halbkreis um 0 in der oberen Halbebene von $\mathbb{C}$ mit Radius r. Für $r>1$ liegt i innerhalb die Kurve $[-r.r]\cup H_r$. Nach der Cauchyschen Integralformel gilt daher

${}_{[-r.r]\cup H_r}\int f(z)dz = \frac{1}{2\pi i}\int\frac{2\pi i\cdot g(z)}{z-i} = 2\pi i\cdot g(i) = \pi\cdot e^{-t}$. Für $z\in H_r$ gilt:

$|f(z)| \le |\exp\{itz\}|\cdot\{|z^2|-1\}^{-1} = |\exp\{it\Re(z)\}|\cdot|\exp\{-t\Im(z)\}|\cdot\{|z^2|-1\}^{-1} \le \{r^2-1\}^{-1}$.

Es folgt: $|{}_{H_r}\int f(z)dz| \le \pi\, r\cdot\{r^2-1\}^{-1} \to 0$ (für $r\to\infty$) und somit:

$$\pi\cdot\varphi(t) = \lim_{r\to\infty}\ {}_{[-r,r]}\int f(x)dx = \pi e^{-t}.$$

d) $\Gamma_{\alpha,\nu}$ hat nach 15.7c die LT $(\frac{\alpha}{\alpha+t})^{\nu}$. Nun ist $(\frac{\alpha}{\alpha+z})^{\nu}$ eine holomorphe Fortsetzung auf $\{z\in\mathbb{C};\ \Re(z)>-\alpha\}$; jetzt 15.10. []

Das nächste Ziel sind die **Eindeutigkeitssätze**.

15.12 Parsevalsche Gleichung. Seien Q_i W–Maße auf $\mathfrak{B}_d$ mit CF φ_i, i=1,2. Dann gilt:

$$\int\varphi_2(x-y)Q_1[dx] = \int\exp\{-it^Ty\}\varphi_1(t)Q_2[dt]\ ,y\in\mathbb{R}^d,\ \text{speziell: } \int\varphi_2 dQ_1=\int\varphi_1 dQ_2.$$

Bew. Mit Fubini 15.2f gilt:

$\int\exp\{-it^Ty\}\,(\int\exp\{it^Tx\}Q_1[dx])\,Q_2[dt] = \int(\int\exp\{it^T(x-y)Q_2[dt]\})\,Q_1[dx]$. []

15.13 Def. $C_b(\mathbb{R}^d)$ sei die Menge der beschränkten, stetigen Funktionen auf $\mathbb{R}^d$.

15.14 Lemma. Gilt für zwei W–Maße Q_1,Q_2 auf $\mathfrak{B}_d$: $\int f dQ_1 = \int f dQ_2$, $f\in C_b(\mathbb{R}^d)$, so stimmen sie überein.

Bew. 17.2. □

15.15 Lemma. Ist Q W–Maß auf $\mathfrak{B}_d$ und $Q_\sigma := Q \star \bigtimes_1^d N(0,\sigma^2)$, dann gilt für $\sigma^2\to 0$:

$\int f dQ_\sigma \to \int f dQ \ \ \forall f\in C_b(\mathbb{R}^d)$ [d.h. Q_σ konvergiert schwach gegen Q].

Bew. Wähle $X_1,\ldots,X_d,Y$ unabhängig mit $X_k \sim N(0,1)$ und $Y\sim Q$. Dann gilt

$Y + \sigma\cdot(X_1,\ldots,X_d) \sim Q_\sigma$ gemäß 12.15b, 13.1a, 13.12a,

also $\int f dQ_\sigma = E[f(Y + \sigma\cdot(X_1,\ldots,X_d))]$; jetzt maj.Konv. . □

15.16 Eindeutigkeitssatz für CF. W–Maße auf $\mathfrak{B}_d$ mit gleicher CF stimmen überein.

Bew. Sei Q W–Maß mit CF φ, Q_σ wie in 15.15.

Idee: $\int f dQ$, $f\in C_b(\mathbb{R}^d)$, läßt sich durch φ ausdrücken, dann 15.14.

Q_σ hat nach 15.11b und 13.12b2 die λ^d–Dichte

$y \mapsto (2\pi\sigma^2)^{-d/2}\cdot\int \exp\{-\|y-x\|^2/2\sigma^2\}Q[dx] = (2\pi\sigma^2)^{-d/2}\cdot\int \nu_{0,1/\sigma^2}(x-y)Q[dx]$,

also nach 15.12 die λ^d–Dichte (wobei sich $(\sigma^2)^{-d/2}$ wegkürzt):

$(2\pi)^{-d}\cdot\int \exp\{-it^T y\}\cdot\varphi(t)\cdot\exp\{-\tfrac{1}{2}\sigma^2\|t\|^2\}dt =: g(\varphi,\sigma^2,y)$.

Gemäß 15.15 gilt nun eine **Inversionsformel**:

(15.17) $\int f dQ = \lim_{\sigma^2\to 0}\int f(y)g(\varphi,\sigma^2,y)dy \quad \forall f\in C_b(\mathbb{R}^d)$. □

15.18 Eindeutigkeitssatz für LT. W–Maße auf $[0,\infty)\cap\mathfrak{B}_1$ mit gleicher LT stimmen überein.

Bew. Nach 15.6 und dem Identitätssatz für holomorphe Funktionen ist $\psi(z) := \int e^{-zx}Q[dx]$ auf $\bar{G} = \{z\in\mathbb{C};\ \Re(z)\geq 0\}$ durch seine Werte $\psi(t)$, $t\in[0,\infty)$ bestimmt. Die CF φ ist wiederum durch ψ auf $\bar{G}$ bestimmt wegen $\varphi(t)=\psi(-it)$, $t\in\mathbb{R}^1$. □

15.19 Korollar. Seien X_k reelle Zva mit CF φ_k, $1\leq k\leq d$. Dann gilt:

$X_1,\ldots,X_d$ unabhängig $\Leftrightarrow$ $(X_1,\ldots,X_d)$ hat die CF $(t_1,\ldots,t_d)\mapsto \Pi_1^d\,\varphi_k(t_k)$.

Bew. Sei $X_k\sim Q_k$. Nach 15.9e ist $(t_1,\ldots,t_d)\mapsto\Pi_1^d\,\varphi_k(t_k)$ die CF von $\bigtimes_{k=1}^d Q_k$; jetzt 13.1a und 15.16. □

15.20 Korollar. Sind Q_k W–Maße mit CF φ_k [bzw. LT ψ_k bzw. erzeugender Fkt g_k], k=1,2,3; dann gilt: $Q_3 = Q_1\star Q_2 \Leftrightarrow \varphi_3 = \varphi_1\cdot\varphi_2$ [bzw. $\psi_3=\psi_1\cdot\psi_2$ bzw. $g_3=g_1\cdot g_2$].

Bew. 15.9c2 und 15.16 [bzw.15.6c2 und 15.18]. □

15.21 Anwendungen. (a) $\pi(\alpha)\star\pi(\beta) = \pi(\alpha+\beta)$ wegen $e^{\alpha(s-1)} \cdot e^{\beta(s-1)} = e^{(\alpha+\beta)(s-1)}$.

(b) $b(n,p)\star b(m,p) = b(n+m,p)$ wegen $(1-p+ps)^n \cdot (1-p+ps)^m = (1-p+ps)^{n+m}$.

(c) $\Gamma_{\alpha,\mu}\star\Gamma_{\alpha,\nu} = \Gamma_{\alpha,\mu+\nu}$ wegen $(\frac{\alpha}{\alpha+t})^\mu \cdot (\frac{\alpha}{\alpha+t})^\nu = (\frac{\alpha}{\alpha+t})^{\mu+\nu}$.

(d) $N(a,\sigma^2)\star N(b,\tau^2) = N(a+b,\sigma^2+\tau^2)$

wegen $\exp\{iat-\frac{1}{2}\sigma^2 t^2\} \cdot \exp\{ibt-\frac{1}{2}\tau^2 t^2\} = \exp\{i(a+b)t-\frac{1}{2}(\sigma^2+\tau^2)t^2\}$.

(e) Sind $X_1,\dots,X_n$ unabhängige, Cauchy–verteilte Zva, dann ist auch $\frac{1}{n}\sum_1^n X_k$ Cauchy–verteilt. Denn die CF von $\sum_1^n X_k$ ist $e^{-n|t|}$ nach 15.11c und 15.9c1; die CF von $\frac{1}{n}\sum_1^n X_k$ ist nach 15.9b $e^{-n|t/n|} = e^{-|t|}$.

(f) Hat X die CF φ, so gilt: φ reellwertig $\Leftrightarrow PX^{-1} = P(-X)^{-1}$. Denn mit 15.2b gilt:
φ reellwertig $\Leftrightarrow \varphi(t) = \overline{\varphi(t)} \Leftrightarrow E[e^{itX}] = E[e^{-itX}]$.

15.22 Satz. φ sei die CF der reellen Zva X mit $E[|X|^k]<\infty$ für ein $k\in\mathbb{N}$. Dann gilt:
(a) φ ist k–mal stetig db, wobei $\varphi^{(k)}(0) = i^k \cdot E[X^k]$;
(b) $\varphi(t) = \sum_{j=0}^k \frac{1}{j!} E[X^j] \cdot (it)^j + o(|t|^k)$ für $t\to 0$.

Bew. a) Sei $X\sim Q$ und $\varphi_j(t) := \int (ix)^j e^{itx} Q[dx]$. Beh. $\frac{d}{dt}\varphi_j(t) = \varphi_{j+1}(t)$, $0\le j<k$.
Es ist $\frac{1}{h}[\varphi_j(t+h)-\varphi_j(t)] = \int (ix)^j e^{itx} \cdot \frac{1}{h}[e^{ihx}-1]\ Q[dx] =: \int f_h dQ$.
Nun gilt für $y\in\mathbb{R}$: $|e^{iy}-1| = |\int_0^y ie^{it}dt| \le |y|$, also für den gesamten Integranden: $|f_h(x)| \le |x|^{j+1}$.
Also für $j<k$ mit maj.Konv.: $\varphi_j'(t) = \int \lim_{h\to 0} f_h\, dQ = \varphi_{j+1}(t)$, somit $\varphi^{(j)} = \varphi_j$.
Die Stetigkeit von φ_k folgt auch mit maj. Konv. . b) Standardbeweis. []

15.23 Lemma. Für ein W–Maß Q auf $\mathfrak{B}_1$ mit CF φ gilt:

$$Q[[-\tfrac{2}{\delta},\tfrac{2}{\delta}]^c] \le \frac{1}{\delta} \cdot \int_{-\delta}^{\delta}(1-\varphi(t))dt \quad \forall\delta>0.$$

Das Verhalten von φ bei 0 bestimmt also das Verhalten von Q bei $\pm\infty$. In 15.22 wurde umgekehrt aus dem Verhalten von Q bei $\pm\infty$, was sich in der Endlichkeit von Momenten ausdrücken kann, auf die Differenzierbarkeit von φ in 0 geschlossen.
Bew. von 15.23: Mit Fubini folgt $\int_{-\delta}^{\delta}(1-\varphi(t))dt = \int_{-\delta}^{\delta}\int(1 - \cos tx - i\cdot\sin tx)Q[dx]\ dt$

$= \int \int_{-\delta}^{\delta}(1 - \cos tx - i\cdot\sin tx)\ dt\ Q[dx] = \int 2\delta(1 - \frac{\sin\delta x}{\delta x})\ Q[dx]$.

Der Integrand ist nicht negativ und sogar $\ge \delta\cdot 1_{\{|x|>2/\delta\}} \cdot 2[1 - \frac{1}{|\delta x|}] \ge \delta \cdot 1_{\{|x|>2/\delta\}}$. []

§16 Mehrdimensionale Verteilungen.

16.1 Def. Sei $X=(X_1,...,X_n)^T$ n–dim. Zufallsvektor auf $(\Omega.\mathfrak{F},P)$. Sind alle Zva X_k P–intb., so heißt X P–intb. und $EX = (EX_1,...,EX_n)^T \in \mathbb{R}^n$ **Erwartungswertvektor** von X. Gilt $E[\|X\|^2] = \sum_1^n E[X_k^2] < \infty$, so heißt X **quadratisch intb.** und $\mathfrak{Cov}[X] := [\mathrm{Kov}[X_j,X_k]] \in \mathbb{R}^{n\times n}$ **Kovarianzmatrix** von X. (Im Falle n=1 gilt $\mathfrak{Cov}[X]=V[X]$.)

16.2 Satz. Sei X intb. n–dim Zva, $C \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Dann gilt:

(a) $E[CX+b] = CEX + b$, (b) $\|EX\| \le E[\|X\|]$.

Bew. a) leicht , b) Übung.

16.3 Satz. Sei X quadrat.intb.n–dim.Zva. Dann gilt:

(a) $\mathfrak{Cov}[X] = E[(X-EX)(X-EX)^T]$;

(b) Für $C\in\mathbb{R}^{m\times n}, b\in\mathbb{R}$, ist CX+b quadrat.intb. und es gilt:
$\mathfrak{Cov}[CX+b] = C\mathfrak{Cov}[X]C^T \in\mathbb{R}^{m\times m}$, speziell für $a\in\mathbb{R}^n$: $V[a^TX] = a^T\mathfrak{Cov}[X]a$.

(c) $\mathfrak{Cov}[X]$ ist eine symmetrische und nichtnegativ definite [positiv semi–definite] Matrix, somit genau dann positiv definit, wenn sie nicht singulär ist.

(d) **Standardisierung**: Sei der Rang von $\mathfrak{Cov}[X]$ gleich $\mathrm{rg}\mathfrak{Cov}[X]=d\le n$. Dann gibt es quadrat.intb. d–dim. Zva W sowie $\Sigma\in\mathbb{R}^{n\times d}$ mit $\mathrm{rg}\Sigma=d$, so daß
$EW = 0$, $\mathfrak{Cov}[W] = I_d$ (Einheitsmatrix), $X = \Sigma W + EX$.
(Für n=1, $\sigma^2:=V[X]>0$, a:=EX gilt: $X=\sigma W+a \Leftrightarrow W=\frac{X-a}{\sigma}$.)

Bew. a) klar, b) mit (a), c) $\mathfrak{Cov}[X]$ nichtnegativ definit wegen (b) und $V[a^TX]\ge 0$.
d) Ziel: Zerlegung: $\mathfrak{Cov}[X]=\Sigma\Sigma^T$.Wegen (c) existiert orthogonale Matrix $U\in\mathbb{R}^{n\times n}$ mit $\mathfrak{Cov}[X] = U\Lambda U^T$, wobei $\Lambda = [\delta_{jk}\lambda_k]$ und $\lambda_k\ge 0$ die Eigenwerte von $\mathfrak{Cov}[X]$ sind. O.E. $\lambda_1,...,\lambda_d>0$ und $\lambda_{d+1}=...=\lambda_n=0$. Nun Zerlegung von Λ mithilfe von $\Lambda^{\frac12}$. Setze dazu $\Lambda^{\pm\frac12} := [\delta_{jk}\lambda_k^{\pm\frac12}] \in \mathbb{R}^{d\times d}$ und $B:=\begin{bmatrix}\Lambda^{\frac12}\\ 0\end{bmatrix} \in \mathbb{R}^{n\times d}$, $B^- := [\Lambda^{-\frac12}\ 0] \in \mathbb{R}^{d\times n}$ (verallgem. Inverse),
$\tilde X := U^T(X-EX)$, dann $\tilde X_k=0$ f.s. für $d<k\le n$; $\Sigma := UB \in \mathbb{R}^{n\times d}$, $W := B^-\tilde X$. Rest als Übung. []

Erinnerung (15.22): Ist X reelle Zva mit $E[X^2]<\infty$ und CF φ, dann:

$$\varphi(t) = 1 + itEX - \tfrac12 t^2(V[X]+E[X]^2) + o(t^2).$$

16.4 Satz. Sei X quadrat.–intb.Zufallsvektor mit CF F. Dann gilt:

$$\varphi(t) = 1 + it^TEX - \tfrac12 t^T\mathfrak{Cov}[X]t - \tfrac12(t^TEX)^2 + o(\|t\|^2).$$

Bew. Übung.

Erinnerung (12.15b): $X \sim N(a,\sigma^2) \Leftrightarrow X = \sigma W + a$ für ein $W \sim N(0,1)$.

16.5 Def. Für $d\in\mathbb{N}$ seien $W_1,...,W_d$ reelle iid Zva mit $W_k \sim N(0,1)$ und $W=(W_1,...,W_d)^T$. Die Verteilung von W ist die **standardisierte d–dim. Normalverteilung** $N(0,I_d) := \bigtimes_1^d N(0,1)$.
Der Zufallsvektor $X=(X_1,...,X_n)^T$ besitzt eine **Normalverteilung $N(a,\Sigma\Sigma^T)$**, und die reellen Zva $X_1,...,X_n$ haben als gemeinsame Verteilung eine Normalverteilung $N(a,\Sigma\Sigma^T)$, falls eine d–dim. Zva W wie oben existiert sowie $\Sigma\in\mathbb{R}^{n\times d}$, $a\in\mathbb{R}^n$ mit $X = \Sigma W + a$.

16.6 Satz. Seien X und W wie in 16.5. Dann gilt:
(0) $N(a,\Sigma\Sigma^T)$ ist wohl definiert, d.h. die Verteilung von X hängt nur von a und $\Sigma\Sigma^T$ ab.
(a) $EX=a$. $\mathfrak{Cov}[X]=\Sigma\Sigma^T$.
(b) $N(a,0)$ ist das auf a konzentrierte Einpunktmaß δ_a, wenn 0 die Nullmatrix ist.
(c) $\bigtimes_1^n N(a_k,\sigma_k^2) = N(a,\Gamma)$ mit $a=(a_1,...,a_n)^T$, $\Sigma=\left[\sigma_k\delta_{jk}\right]\in\mathbb{R}^{n\times n}$, $\Gamma=\Sigma\Sigma^T=\left[\sigma_k^2\delta_{jk}\right]$.
(d) Für $C\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$, $Y=CX+b$ gilt: $Y \sim N(CEX+b,C\mathfrak{Cov}[X]C^T)$.
(e) Für $J \subset \{1,...,n\}$ ist haben X_j, $j\in J$, als gemeinsame Verteilung wieder eine Normalverteilung.
(f) Zu einer symmetrischen nichtnegativ definiten Matrix $\Gamma\in\mathbb{R}^{n\times n}$ und $a\in\mathbb{R}^n$ ex. Zufallsvektor X mit $X\sim N(a,\Gamma)$. Dabei kann $\mathrm{rg}\Gamma=:d$ wie in 16.5 gewählt werden.
(g) Ist $\mathrm{rg}\mathfrak{Cov}[X] = n$, so hat X die λ^n–Dichte
$x \mapsto ((2\pi)^n\det\mathfrak{Cov}[X])^{-\frac{1}{2}}\cdot\exp\{-\frac{1}{2}(x-a)^T\mathfrak{Cov}[X]^{-1}(x-a)\}$.
(h) $N(a,\Gamma)$ hat die CF $t \mapsto \exp\{ia^Tt - \frac{1}{2}t^T\Gamma t\}$.
(i) $N(a_1,\Gamma_1)\star N(a_2,\Gamma_2) = N(a_1+a_2,\Gamma_1+\Gamma_2)$.
(j) $X_1,...,X_n$ unabh. $\Leftrightarrow$ $X_1,...,X_n$ paarweise unkorreliert.
(k) Gilt $X_1,...,X_n$ unabh., $X_k\sim N(a_k,\sigma^2)\ \forall k$, $U\in\mathbb{R}^{n\times n}$ orthogonal, so folgt für $(Y_1,...,Y_n)^T=UX$:
$Y_1,...,Y_n$ unabh., $Y_k\sim N(b_k,\sigma^2)\ \forall k$ mit $(b_1,...b_n)^T=Ua$.

Bew. 0) Sei ν CF von W; dann folgt mit 15.9b und 15.11b:
$E\left[e^{it^T(\Sigma W+a)}\right] = e^{it^Ta}\cdot\nu(\Sigma^Tt) = e^{it^Ta - \frac{1}{2}t^T\Sigma\Sigma^Tt}$. Jetzt Eindeutigkeitssatz 15.16 für CF.
a) folgt mit 16.2 und 16.3 aus $EW=0$, $\mathfrak{Cov}[W]=I_d$.
b) Wähle $\Sigma=0\in\mathbb{R}^{n\times d}$.
c) Wähle $d=n$, $\Sigma W+a = (\sigma_1W_1+a_1,...,\sigma_nW_n+a_n)^T$.Dann gilt $\Sigma W+a \sim \bigtimes_1^n N(a_k,\sigma_k^2)$ nach 12.15b und 13.1a.
d) $Y=C\Sigma W+Ca+b$, $C\Sigma(C\Sigma)^T=C\Sigma\Sigma^TC^T$.
e) folgt aus (d).
f) Wähle $\Sigma\in\mathbb{R}^{n\times d}$ wie im Bew. von 16.3d zu Γ anstelle von $\mathfrak{Cov}[X]$ mit $d = \mathrm{rg}\Gamma$, dann $\Gamma=\Sigma\Sigma^T$.
g) Wähle Σ wie in f) und 16.3d. Die λ^n–Dichte von W sind etwa nach 15.11b bekannt. Damit ergibt sich die λ^n–Dichte von $\Sigma W+a$ nach 12.9.

h) folgt nach 0).

i) wie im 1–dim. Fall 15.21d.

j) "$\Leftarrow$" $\mathfrak{Cov}[X] = [\delta_{jk}\sigma_k^2]$, jetzt (c) [und 13.1a]. "$\Rightarrow$" folgt aus 14.21.

k) Nach (c) gilt $X \sim \bigotimes_{k=1}^{n} N(a_k,\sigma_k^2) = N(a,\sigma^2 I_n)$ und nach (d)

$Y \sim N(Ua, U\mathfrak{Cov}[X]U^T) = N(Ua, U(\sigma^2 I_n)U^T) = N(Ua, \sigma^2 I_n)$; jetzt (j) und (c). []

Spezialfall n = d = 2.

Im Fall n=2 gilt: $\mathfrak{Cov}[X] = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ mit $\rho = \mathrm{Kor}[X_1,X_2]$.

Dabei gitl: $\mathrm{rg}\mathfrak{Cov}[X] = 2 \Leftrightarrow |\rho|<1, \sigma_1^2,\sigma_2^2 >0$. Dies sei jetzt vorausgesetzt.

Für die λ^2–Dichte f von X ergibt sich die Darstellung (nachrechnen!) mit $t_i = (x_i - a_i)/\sigma_i$:

$$f(x_1,x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \cdot \exp\{-\frac{1}{2(1-\rho^2)}(t_1^2 - 2\rho t_1 t_2 + t_2^2)\} .$$

Dies ist eine 2–dim. Glockenfläche; die Kurven $f(x_1,x_2) = \mathrm{const}$ sind konzentrische Ellipsen um a [bzw. Kreise im Fall $\rho=0$].

Die folgende Beziehung ist interessant; dabei sei wie in 12.14 $\varphi_{\alpha,\sigma^2}$ die λ^1–Dichte von $N(\alpha.\sigma^2)$:

$$f(x_1,x_2) = \varphi_{a_1,\sigma_1^2}(x_1) \cdot \varphi_{a_2+\rho t_1\sigma_2,(1-\rho^2)\sigma_2^2}(x_2) .$$

Nach 16.6e ist φ_{a_1,σ_1^2} λ^1–Dichte von X_1.

Wären X_1,X_2 diskrete Zva mit $P[X_1=x_1]>0$, so wäre

(*) $P[X_2 \in B | X_1=x_1] = \sum_{x_2\in B} P[X_1=x_1, X_2=x_2]/P[X_1=x_1]$.

Für den hier vorliegenden Fall $P[X_1=x_1] = 0 \ \forall x_1$ braucht man eine allgemeinere Def. der bedingten W. ($\longrightarrow$ W.th. II). Mit dieser ergibt sich ein Ausdruck, den man aus (*) durch formales Ersetzen der Z–Dichte durch die Lebesgue–Dichte erhält:

$$P[X_2 \in B | X_1=x_1] = \int_B f(x_1,x_2)\, dx_2 / \varphi_{a_1,\sigma_1^2}(x_1) = N(a_2+\rho t_1\sigma_2,(1-\rho^2)\sigma_2^2)[B] .$$

Daraus ergibt sich für den bedingten Erwartungswert:

$$E[X_2 | X_1=x_1] = a_2+\rho t_1\sigma_2 =: g(x_1).$$

Dabei heißt g **Regressionsgerade.**

Die Bezeichnung Regressionsgerade ergibt sich aus einem historischen Beispiel [Galton 1886]: Seien X_1, X_2 Körpergröße von Vater und Sohn und es liege der stationäre Fall $a_1 = a_2 =: \alpha$, $\sigma_1^2 = \sigma_2^2 =: \sigma^2$ vor. Dann ist $g(x_1) = (1-\rho)\alpha + \rho x_1$; der bedingte Erwartungswert der Körpergröße des Sohnes bei gegebener Größe x_1 des Vaters ist also bei $\rho>0$ [Körpergröße von Vater und Sohn sind positiv korreliert] eine konvexe Linearkombination des globalen Mittelwerts α und der Körpergröße x_1 des Vaters. Ist etwa $x_1>\alpha$, so ist $\alpha < g(x_1) < x_1$. Dies wurde als Rückschritt (Regression) zum globalen Durchschnittswert α interpretiert.

Diese Interpretation ist allerdings fragwürdig, was man erkennt, wenn man die Rollen von Vater und Sohn vertauscht.

Eine Darstellung $X = \Sigma W + a$ ergibt sich mit $\Sigma := \begin{bmatrix} s_+\sigma_1 & s_-\sigma_1 \\ s_-\sigma_1 & s_+\sigma_2 \end{bmatrix}$ und $s_\pm := \sqrt{\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{1-\rho^2}}$; also:

$$X_1 = s_+\sigma_1 \cdot W_1 + s_-\sigma_1 \cdot W_2 + a_1$$

$$X_2 = s_+\sigma_2 \cdot W_1 + s_-\sigma_2 \cdot W_2 + a_2;$$ dabei gilt: $s_- = 0 \Leftrightarrow \rho = 0 \Leftrightarrow X_1, X_2$ sind unabhängig.

§17 Schwache Konvergenz von W–Maßen.

Sei (S,d) ein metrischer Raum, $\mathfrak{S}$ die Borel–σ–Algebra auf S.

17.1 Def. $\mathbb{P}(S)$ sei die Menge der W–Maße auf $\mathfrak{S}$, $\delta_x \in \mathbb{P}(S)$ sei das auf $\{x\}$ konzentrierte Einpunktmaß, $C_b(S)$ bzw. $C_b^L(S)$ sei die Menge der reellen, beschränkten und stetigen bzw. Lipschitz–stetigen Funktionen auf S,
$S(f) := \{x \in S;\ f \text{ ist in } x \text{ stetig}\}$ sei die Menge der Stetigkeitsstellen von $f : S \mapsto \mathbb{R}$.

17.2 Lemma. Für $\mu,\nu \in \mathbb{P}(S)$ gilt $\mu=\nu$, falls nur $\int f d\mu = \int f d\nu \quad \forall f \in C_b^L(S)$.

Bew. Sei $A \subset S$ abgeschlossen. Für $k \in \mathbb{N}$ definiere $f_k : S \mapsto [0,1]$ gemäß $f_k(x) := \exp\{-kd(x,A)\}$.
Dann gilt $f_k \in C_b^L(S)$ wegen $|f_k(x)-f_k(y)| \le k \cdot d(x,y)$ sowie $f_k \downarrow 1_A$. Mit maj.Konv. folgt:
$\mu[A] = \lim \int f_k d\mu = \lim \int f_k d\nu = \nu[A]$. Die Beh. folgt aus Eindeutigkeitssatz für Maße 7.5. []

17.3 Def. (a) Seien $\mu, \mu_n \in \mathbb{P}(S)$, $n \in \mathbb{N}$. Dann **konvergiert** μ_n **schwach** gegen μ :⇔ $\mu_n \xrightarrow{w} \mu$

["w" von weak] :⇔ $\int f d\mu_n \to \int f d\mu \quad \forall f \in C_b(S)$.

(b) Sind X_n Zva auf $(\Omega_n,\mathfrak{F}_n,P_n)$ mit Werten in S, $n \in \mathbb{N}$, so **konvergiert** (X_n) **in Verteilung** gegen $\mu \in \mathbb{P}(S)$:⇔ $X_n \xrightarrow{w} \mu$:⇔ $P_n X_n^{-1} \xrightarrow{w} \mu$. Gilt dabei $X \sim \mu$ für eine Zva X auf $(\Omega,\mathfrak{F},P)$, so konvergiert (X_n) in Verteilung gegen X (⇔: $X_n \xrightarrow{w} X$).

17.4 Beispiele. (a) $\delta_{x_n} \xrightarrow{w} \delta_x \Leftrightarrow x_n \to x$. S läßt sich also in $\mathbb{P}(S)$ einbetten vermöge $x \mapsto \delta_x$.

(b) Approximation von $\mu:=U(0,1)$ durch $\mu_n := U(\{\frac{k}{n};\ 1 \le k \le n\})$. Es gilt:
$\int f d\mu_n = \sum_{k=1}^n f(\frac{k}{n})\frac{1}{n} \to {}_0\int^1 f(x)dx \quad \forall f \in C_b(\mathbb{R})$, also $\mu_n \xrightarrow{w} \mu$.
(c) $N(x,\sigma^2) \xrightarrow{w} \delta_x$ für $\sigma^2 \to 0$. Wähle in 15.15 d=1, $Q=\delta_x$; dann gilt: $\delta_x \star N(0,\sigma^2) = N(x,\sigma^2)$.

17.5 Bemerkung. In 17.4 gilt <u>nicht</u>: $\mu_n[B] \to \mu[B] \quad \forall B \in \mathfrak{S}$.

Wähle in (a), (c) $B=\{x\}$, dann $\mu_n[B]=0$ aber $\mu[B]=1$. Wähle in (b) $B=\mathbb{Q}$.

Ist in (a) $S=\mathbb{R}$, $x_n > x\ \forall n$ und F_n Vf zu δ_{x_n}, F Vf zu δ_x, so gilt: $F_n(x)=0\ \forall n$, aber $F(x)=1$.

17.6 Eindeutigkeit des schwachen Limes. Aus $\mu_n \xrightarrow{w} \mu$ und $\mu_n \xrightarrow{w} \mu'$ folgt $\mu=\mu'$.

Bew. mit 17.2. []

17.7 Satz von Portmanteau. Seien μ, $\mu_n \in \mathbb{P}(S)$. Dann sind folgende Aussagen äquivalent:

(a) $\mu_n \overset{w}{\to} \mu$;

(b) $\underline{\lim}\, \mu_n[G] \geq \mu[G]$ $\forall$ G offen;

(c) $\overline{\lim}\, \mu_n[A] \leq \mu[A]$ $\forall$ A abgeschlossen;

(d) $\lim \mu_n[B] = \mu[B]$ $\forall B \in \mathfrak{S}$ mit $\mu[\partial B]=0$;

(e) $\int f d\mu_n \to \int f d\mu$ $\forall f \in C_b^L(S)$;

(f) $\int f d\mu_n \to \int f d\mu$ $\forall f: S \mapsto \mathbb{R}$ mb, beschränkt und μ–f.s. stetig

Bew. Schema: (a) $\Rightarrow$ (e) $\Rightarrow$ (c) $\Leftrightarrow$ (b) $\Rightarrow$ (d) $\Rightarrow$ (f) $\Rightarrow$ (a). Klar sind (a)$\Rightarrow$(e), (c)$\Leftrightarrow$(b) und (f)$\Rightarrow$(a).
"(e)$\Rightarrow$(c)" Sei A abgeschlossen und f_k wie im Bew. von 17.2. Dann

$$\overline{\lim}_n\, \mu_n[A] \leq \overline{\lim}_n \int f_k d\mu_n = \int f_k d\mu \downarrow \mu[A] \quad \text{für } k \to \infty.$$

"(b),(c) $\Rightarrow$ (d)" Sei B^o das Innere von B, $\overline{B}$ der Abschluß; dann
$\mu[B^o] \leq \underline{\lim}\, \mu_n[B^o] \leq \underline{\lim}\, \mu_n[B] \leq \overline{\lim}\, \mu_n[B] \leq \overline{\lim}\, \mu_n[\overline{B}] \leq \mu[\overline{B}] = \mu[B^o] + \mu[\partial B]$.

"(d)$\Rightarrow$(f)" Idee: Fasse $\int f d\mu$ als Erwarungsw. auf, und benutze dessen Darstellung vit Hilfe von Vfen. Ziel: Konvergenz der Vfen als Folgerung aus (d) mit $B = \{f \leq t\}$:
(*) $F_n(t) := \mu_n[f \leq t] \to F(t) := \mu[f \leq t]$ $\forall t \in S(F)$.

Wähle $N \in \mathfrak{S}$ mit $\mu[N]=0$ und $S(f)^c \subset N$. Dann gilt:
(**) $\partial\{f \leq t\} \subset \{f=t\} \cup N$;
denn für $f(x_n) \leq t$ [bzw. $>t$] und $x_n \to x$ folgt $f(x) \leq t$ [bzw. $\geq t$] oder $x \in N$.

Nach 7.18 gilt: $\mu[f=t] = \mu f^{-1}[\{t\}] = F(t) - F(t-0)$, also $\mu[f=t]=0$ für $t \in S(F)$.
Nun folgt (*) aus (**) und (d).
Die Unstetigkeitsstellen $S(F)^c$ einer Vf F sind abzb., somit gilt: $\lambda^1[S(F)^c]=0$.
O. E. sei $f \geq 0$; sonst ersetzte f durch $f - \|f\|$ mit $\|f\| = \sup|f|$. Aus 14.7 folgt nun
$\int f d\mu_n = {}_{[0,\infty)}\int (1-F_n(t))dt = {}_{[0,\|f\|]}\int (1-F_n(t))dt$, entsprechend für μ und F.
Die Beh. folgt nun mit maj. Konv.. $\square$

17.8 Korollar. Seien S' metr.Raum mit Borel–σ–Algebra $\mathfrak{S}'$, $h: S \mapsto S'$ $\mathfrak{S}$–$\mathfrak{S}'$–mb, $\mu, \mu_n \in \mathbb{P}(S)$, sodaß

$\mu_n \overset{w}{\to} \mu$ und h μ–f.s. stetig. Dann gilt: $\mu_n h^{-1} \overset{w}{\to} \mu h^{-1}$.

Bew. Z. z. $\forall f \in C_b(S)$: $\int f d\mu_n h^{-1} \to \int f d\mu h^{-1}$; d.h. gemäß 10.19 : $\int f \circ h\, d\mu_n \to \int f \circ h\, d\mu$.

Nun ist $f \circ h$ dort stetig, wo h stetig ist, also μ–f.s.. Jetzt 17.7f. $\square$

Der Fall $S = \mathbb{R}^1$.

17.9 Bemerkung. Sei F 1–dim. Vf, $F^-: (0,1) \mapsto \mathbb{R}$ die verallgem. Inverse von F gemäß 7.10. Dann gilt:
$S(F^-) = \{\alpha \in (0,1);$ F nimmt α höchstens einmal an$\}$, also:
$S(F^-) = (0,1) \Leftrightarrow$ F ist streng isoton. Bew. Übung.

17.10 Satz. Sei $S=\mathbb{R}^1$, $\mu,\mu_n \in \mathbb{P}(\mathbb{R}^1)$ mit Vf. F bzw. F_n, $n\in\mathbb{N}$. Dann sind äquivalent:

(a) $\mu_n \overset{w}{\to} \mu$;

(b) $F_n(s) \to F(s) \ \forall\, s\in S(F)$;

(b') $F(t-0) \le \underline{\lim}\, F_n(t-0) \le \overline{\lim}\, F_n(t) \le F(t) \ \forall\, t\in\mathbb{R}^1$;

(c) $F_n^-(\alpha) \to F^-(\alpha) \ \forall\, \alpha\in S(F^-)$;

(c') $F^-(\alpha) \le \underline{\lim}\, F_n^-(\alpha) \le \overline{\lim}\, F_n^-(\alpha+0) \le F^-(\alpha+0) \ \ \forall\, \alpha \in (0,1)$.

Bew. Schema: (a) $\Rightarrow$ (b') $\Rightarrow$ (b) $\Rightarrow$ (c') $\Rightarrow$ (c) $\Rightarrow$ (a).
"(a)$\Rightarrow$(b')" $(-\infty,t)$ ist offen, $(-\infty,t]$ ist abgeschlossen; jetzt 17.7b,c.
"(b')$\Rightarrow$(b)" klar. "(b)$\Rightarrow$(c')" Übung. "(c')$\Rightarrow$(c)" klar, da F^- isoton.
"(c)$\Rightarrow$(a)" F^-, F_n^- sind Zva auf $(\Omega.\mathfrak{F},P) := ((0,1), (0,1)\cap\mathfrak{B}_1, U(0,1))$ mit $F^-_{(n)} \sim \mu_{(n)}$ gemäß 7.11. Da F^- monoton, ist $S(F^-)^c$ abzb., also eine P–Nullmenge. Nun folgt mit Transform.–Satz 10.19 und maj. Konv.: $\int f\, d\mu_n = E[\, f\circ F_n^- \,] \to E[\, f\circ F^- \,] = \int f\, d\mu$ für $f\in C_b(\mathbb{R})$. $\square$

17.11 Zentraler Grenzwertsatz. Seien $X_1, X_2, \ldots$ iid Zva mit $a := EX_k \in \mathbb{R}$, $0 < \sigma^2 := V[X_k] < \infty$.

Dann gilt: $$W_n := \sqrt{\frac{n}{\sigma^2}} \cdot \left(\frac{1}{n}\sum_1^n X_k - a\right) \overset{w}{\to} N(0,1) .$$

17.12 Bemerkung: Die Existenz einer Folge von iid Zva mit $X_n \sim Q$ (zu vorgegebenem Q) wird erst später gezeigt. Sie ist hier auch nicht notwendig. Die Verteilung von W_n hängt nur von der gemeinsamen Verteilung von $X_1,\ldots,X_n$ ab, also von $\overset{n}{\underset{1}{\times}} Q$. Deswegen läßt sich die Aussage auch mit einem sogenannten **Dreiecksschema**: $X_k^{(n)}$, $1\le k\le n<\infty$ formulieren.
Dabei ersetzt man $X_1,\ldots,X_n$ jeweils durch identisch verteilte Zva $X_k^{(n)}$, $1\le k\le n<\infty$, wobei $X_1^{(n)},\ldots,X_n^{(n)}$ unabhängig sind auf $(\Omega_n,\mathfrak{F}_n,P_n)$. Ein Beispiel für $(\Omega_n,\mathfrak{F}_n,P_n)$ ist nach 13.5 $(\mathbb{R}^n,\mathfrak{B}_n,\overset{n}{\underset{1}{\times}} Q)$. $\square$

Bew. von 17.11: In 1.19 wurde in Hinblick auf 17.10b ein Spezialfall behandelt mit
$\tilde\Omega_n = (0,1)^n$, $\tilde{\mathfrak{F}}_n = (0,1)^n \cap \mathfrak{B}_n$, $\tilde P_n = \lambda^n\big|_{\tilde{\mathfrak{F}}_n} = \overset{n}{\underset{1}{\times}} U(0,1)$, $\tilde X_k^{(n)}(x_1,\ldots,x_n) := G(x_k)$,

also $\tilde X_k^{(n)} = G\circ Z_k$ wenn $Z_k : (0,1)^n \mapsto (0,1)$ die k–te Projektion ist. Damit sind aus jetziger Sicht die Z_k und damit die $\tilde X_k^{(n)}$ unabhängig und identisch verteilt mit $Z_k \sim U(0,1)$.

Sei F Vf der Verteilung Q der X_k und $G = F^{-}$; dann gilt nach 7.11: $U(0,1)G^{-1} = Q$.
Damit haben also die $\tilde{X}_k^{(n)}$, $1 \le k \le n$, die gleiche gemeinsame Verteilung wie die X_k, $1 \le k \le n$. $\square$

17.13 Lemma. Sei D dichte Teilmenge in (0,1), F, F_n, $n \in \mathbb{N}$, 1–dim. Vf, so daß:

$$F_n(t) \to F(t),\ F_n(t-0) \to F(t-0)\quad \forall\, t \in \{F^{-}(\alpha);\ \alpha \in D\}.\ \text{ Dann gilt: } \sup_t |F_n(t) - F(t)| \to 0.$$

Bew. Übung.

17.14 Satz. Sind $\mu, \mu_n \in \mathbb{P}(\mathbb{R})$ mit Vf F bzw. F_n und ist F stetig, so sind äquivalent:

(a) $\mu_n \xrightarrow{w} \mu$ (b) $\sup_t |F_n(t) - F(t)| \to 0$.

Bew. "(a)⇒(b)" Es gilt $S(F) = \mathbb{R}$, also mit 17.10b': $\lim F_n(t) = \lim F_n(t-0) = F(t) = F(t-0)$ $\forall\, t \in \mathbb{R}$. Jetzt 17.13. $\square$

17.15 Korollar zum zentralen Grenzwertsatz. Es sei $\Phi_{a,\sigma^2}(t) = \Phi(\frac{t-a}{\sigma})$ die Vf zu $N(a,\sigma^2)$ mit $\Phi := \Phi_{0,1}$.
Dann gilt unter den Voraussetzungen von 17.11 :

(a) $\sup_t |P_n[W_n \le t] - \Phi(t)| \to 0$;
(b) $\sup_t |P_n[\frac{1}{n}\Sigma_1^n X_k^{(n)} \le t] - \Phi_{a,\sigma^2/n}(t)| \to 0$;
(c) $\sup_t |P_n[\Sigma_1^n X_k^{(n)} \le t] - \Phi_{na,n\sigma^2}(t)| \to 0$.

Bew. a) mit 17.11 und 17.14; b) ersetzte in (a) t durch $\sqrt{n}(t-a)/\sigma$;
c) ersetzte in (a) t durch $(t-na)/\sigma\sqrt{n}$. $\square$

17.16 Beispiel (Moivre–Laplace). Sei Y_n Zva auf $(\Omega_n, \mathfrak{F}_n, P_n)$ mit $Y_n \sim b(n,p)$; dann gilt:

$$\sup_t |P_n[Y_n \le t] - \Phi((t-np)/\sqrt{n \cdot p(1-p)})| \to 0, \text{ und z.B. } P_n[Y_n \le np] \to \tfrac{1}{2} = \Phi(0).$$

Bew. O.E. $Y_n = \Sigma_1^n X_k^{(n)}$ mit $X_1^{(n)}, \ldots, X_n^{(n)}$ iid mit $X_k^{(n)} \sim b(1,p)$.
Moivre / Laplace zeigten : $(npq)^{\frac{1}{2}} \cdot b(n,p;k) \approx \varphi((npq)^{-\frac{1}{2}} \cdot (k - np))$, wobei φ die Dichte von N(0,1) ist. $\square$

Verteilungskonvergenz

17.17 Lemma. Sei S separabel; dann ist $\mathfrak{S} \otimes \mathfrak{S}$ die Borel–σ–Algebra $\mathfrak{B}(S \times S)$ auf (dem separablen metr. Raum) $S \times S$ und $S \times S \ni (x,y) \mapsto d(x,y)$ $\mathfrak{S} \otimes \mathfrak{S}$–mb.

Bew. "$\mathfrak{S} \otimes \mathfrak{S} \subset \mathfrak{B}(S \times S)$" gilt wegen 6.13 auch ohne Separabilität, da das Produkt von offenen Mengen offen ist. "$\mathfrak{B}(S \times S) \subset \mathfrak{S} \otimes \mathfrak{S}$" zeigt man wie "$\mathfrak{B}_d \subset \sigma(J_d)$" in 6.4. Dabei wird J_d durch eine abzählbare Basis von S ersetzt. Die Meßbarkeit von d folgt jetzt aus der Stetigkeit. $\square$

Konvergenz in Verteilung gegen eine Konstante kann durch einen äquivalenten Begriff ersetzt werden. Dazu:

17.18 Lemma. Für X_n, $n\in\mathbb{N}$, Zva auf $(\Omega_n,\mathfrak{F}_n,P_n)$ mit Werten in S und $x_0\in S$ sind äquivalent:

(a) $X_n \xrightarrow{w} x_0$;

(b) X_n konvergiert nach W. gegen x_0, d.h. $P_n[d(X_n,x_0)\geq\varepsilon] \to 0$ $(n\to\infty)$ $\forall\varepsilon>0$.

(c) $d(X_n,x_0) \xrightarrow{w} 0$.

Bew. "(a)⇒(b)" Mit 17.7c folgt:

$\overline{\lim}\, P_n[d(X_n,x_0)\geq\varepsilon] = \overline{\lim}\, P_n X_n^{-1}[\{x\in S;\, d(x,x_0)\geq\varepsilon\}] \leq \delta_{x_0}[\{x\in S;\, d(x,x_0)\geq\varepsilon\}] = 0.$

"(b)⇒(a)" Setze im Bew. des folgenden Störungslemma $Y_n=x_0$.

Offenbar gilt: "(b) ⇔ $d(X_n,x_0)$ konv. nach W. gegen 0"; verwende nun "(a) ⇔ (b)". $\square$

17.19 Bemerkung. Sind X_n, $n\in\mathbb{N}$, Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in S und $x_0\in S$, dann folgt jede der Aussagen 17.18(a) und (b) aus jeder der beiden folgenden Bedingungen:

(a) $X_n \to x_0$ P–f.s.;

(b) $E[d(X_n,x_0)^p] \to 0$ für ein $p>0$.

Bew. a) $\Omega^o:=\{X_n\to x_0\}=\{\overline{\lim}\, d(X_n,x_0)=0\}\in\mathfrak{F}$ und $\Omega^o\cap\cup_{m\geq n}\{d(X_m,x_0)>\varepsilon\}\downarrow\emptyset$ mit $P[\Omega^o] = 1$.

b) mit Markoffscher Ungleichung. $\square$

17.20 Störungslemma. S sei separabel und X_n,Y_n, $n\in\mathbb{N}$, seien Zva auf $(\Omega_n,\mathfrak{F}_n P_n)$ mit Werten in S. Gilt dann $Y_n \xrightarrow{w} \mu\in\mathbb{P}(S)$ und $d(X_n,Y_n) \xrightarrow{w} 0$, so folgt $X_n \xrightarrow{w} \mu$.

Bew. Gemäß 17.18 "(a)⇒(b)" konvergiert $d(X_n,Y_n)$ nach W. gegen 0. Sei $f\in C_b^L(S)$ mit Lipschitzkonstante L und $\|f\| = \sup|f|$. Dann gilt:

$|\int f\, dP_n X_n^{-1} - \int f\, dP_n Y_n^{-1}| = |E[f(X_n)-f(Y_n)]| \leq E[|f(X_n)-f(Y_n)|1_{\{d(X_n,Y_n)\geq\varepsilon\}}] +$

$E[|f(X_n)-f(Y_n)|1_{\{d(X_n,Y_n)<\varepsilon\}}] \leq 2\|f\|\cdot P[d(X_n,Y_n)\geq\varepsilon] + L\varepsilon$ $\forall\, \varepsilon>0$. Jetzt 17.7e. $\square$

17.21 Satz. S' sei metr. Raum (z.B. S' = S), S sei separabel, X_n,Y_n, $n\in\mathbb{N}_0$, seien Zva auf $(\Omega_n,\mathfrak{F}_n,P_n)$ mit Werten in S und $Y_0=y_0\in S$ konstant, $h:S\times S\mapsto S'$ sei mb. und in (x,y_0) stetig $\forall\, x\in S$. Gilt dann $X_n \xrightarrow{w} X_0$ und $Y_n \xrightarrow{w} y_0$, so folgt $h(X_n,Y_n) \xrightarrow{w} h(X_0,y_0)$.

Bew. (i) Es folgt leicht: $\tilde{Y}_n := (X_n,y_0) \xrightarrow{w} (X_0,y_0)$; setze $\tilde{X}_n := (X_n,Y_n)$. (ii) Die Metrik $d_2((x,y),(x',y'))$ $= d(x,x')+d(y,y')$ erzeugt die Produkttopologie auf $S\times S$, und $d_2(\tilde{X}_n,\tilde{Y}_n)=d(Y_n,y_0) \xrightarrow{w} 0$.

(iii) Mit 17.20 folgt nun: $(X_n,Y_n) \overset{w}{\to} (X_0,y_0)$. (iv) h ist $P_0(X_0,y_0)^{-1}$–f.s. stetig; jetzt 17.8. $\square$

17.22 Anwendungen (Cramér). Setze in 17.21: $S=S'=\mathbb{R}^d$,

$$X_n \overset{w}{\to} X_0\,,\ Y_n \overset{w}{\to} 0 \ \Rightarrow\ X_n \pm Y_n \overset{w}{\to} X_0 \quad \text{(setze } h(x,y) = x\pm y,\ y_0=0\text{);}$$

$$X_n \overset{w}{\to} X_0\,,\ Y_n \overset{w}{\to} 1 \ \Rightarrow\ X_n \cdot Y_n \overset{w}{\to} X_0 \text{ und } X_n/Y_n \overset{w}{\to} X_0 \quad \text{für } d=1$$

$$\text{(setze } h(x,y)=x\cdot y \text{ oder } h(x,y)=\frac{x}{y}\cdot 1_{\{y\neq 0\}}\,,\ y_0=1\text{).}$$

§18 Der Stetigkeitssatz.

18.1 Def. Ist S ein metrischer Raum, so heißt $\Gamma \subset \mathbb{P}(S)$ **straff**, falls gilt:

$\forall \varepsilon>0\ \exists$ (relativ–) kompakte Menge $K\subset S\ \forall \mu\in\Gamma$: $\mu[K] \geq 1-\varepsilon$.

18.2 Beispiel. $\{\delta_{x_n}, n\in\mathbb{N}\} \subset \mathbb{P}(\mathbb{R}^d)$ ist straff $\Leftrightarrow$ $\{x_n;\ n\in\mathbb{N}\}$ ist beschränkt.

Schreibweise: Für $\mu \in \mathbb{P}(\mathbb{R}^d)$ sei $\mu(a,b] := \mu[(a,b]]$, $\mu(a,b) := \mu[(a,b)]$.

18.3 Lemma. Sei $\Gamma \subset \mathbb{P}(\mathbb{R}^d)$, $\pi_j : \mathbb{R}^d \mapsto \mathbb{R}$ Projektion auf die j–te Koordinate. Dann gilt:

(a) Es sind äquivalent:

(i) Γ ist straff;

(ii) $\forall\ \varepsilon>0\ \exists\ a,b \in\mathbb{R}^d\ \forall\ \mu\in\Gamma$: $\mu(a,b] \geq 1-\varepsilon$;

(iii) $\Gamma\pi_j^{-1} := \{\mu\pi_j^{-1};\ \mu\in\Gamma\}$ ist straff $\forall\ 1\leq j\leq d$.

(b) $|\Gamma|<\infty \Rightarrow \Gamma$ ist straff;

(c) jede Teilmenge einer straffen Menge von W–Maßen ist straff.

Bew. In (a) ist (i)$\Leftrightarrow$(ii) klar. ad "(i) $\Leftrightarrow$ (iii) ": Für $a = (a_1,...,a_d)$, $b=(b_1,...,b_d)$ gilt:

$\mu\pi_j^{-1}(a_j.b_j]^c = \mu[\pi_j \notin (a_j,b_j]] \leq \mu(a,b]^c = \mu[\cup_k \{\pi_k \notin (a_k,b_k]\}] \leq \Sigma_k\ \mu\pi_k^{-1}(a_k.b_k]^c$.

(b) $\max_{1\leq k\leq n} \mu_k(a,b] \uparrow 1$ für $(a,b] \uparrow \mathbb{R}^d$. (c) ist klar. $\square$

18.4 Satz von Prohorov (oder **Helly** für d=1).

Für $\Gamma \subset \mathbb{P}(\mathbb{R}^d)$ sind folgende Aussagen äquivalent:

(a) Γ ist straff;

(b) Γ ist rel. folgenkompakt, d.h. $\forall\ (\mu_n)\subset\Gamma\ \exists\ (n')\subset\mathbb{N}$, $\mu\in\mathbb{P}(\mathbb{R}^d)$ mit $\mu_{n'} \overset{w}{\to} \mu$.

Bew. "(b)$\Rightarrow$(a)" mit 18.3a indirekt.

Ann. $\exists\ \varepsilon>0\ \forall\ n\ \exists\ \mu_n\in\Gamma$ mit $\mu_n(-\vec{n},\vec{n}]<1-\varepsilon$; dann ex. nach Vor. $(n')\subset\mathbb{N}$, $\mu\in\mathbb{P}(\mathbb{R}^d)$ mit $\mu_{n'} \overset{w}{\to} \mu$.

Wähle a,b mit $\mu(a,b)>1-\varepsilon$, dann gilt mit 17.7b einerseits $\underline{\lim}\ \mu_{n'}(a,b) \geq \mu(a,b) > 1-\varepsilon$

und andererseits $\mu_n(a,b] \leq 1-\varepsilon\ \ \forall\ n$ mit $(a,b] \subset (-\vec{n},\vec{n}]$ im Widerspruch zur Ann. .

"(a)⇒(b)" Der Beweis soll hier zunächst nur für den Fall **d=1** geführt werden. Der allgemeine Fall $d\in\mathbb{N}$ wird soll in einem Anhang zu § 18 behandelt werden.
Geg. sei $(\mu_n)\subset\Gamma$ mit Vf. F_n.

(i) Mit Hilfe eines Diagonalisierungsverfahrens erhält man $(n')\subset\mathbb{N}$ sodaß
$G(r):=\lim F_{n'}(r)$ ex. $\forall\, r\in\mathbb{Q}$. G ist dann isoton auf $\mathbb{Q}$.

(ii) $F(x):=\inf\{G(r);\, r\in\mathbb{Q},\, r>x\}$ ist isoton auf $\mathbb{R}$ und $0\le F\le 1$. Ferner ist F rechtsstetig.
Denn für $x_n\downarrow x$ hat man: $\forall\, r>x\ \exists\, n_0\ \forall\, n\ge n_0$: $x_n<r$, d.h.

$\forall\, r>x\ \overline{\lim}\, F(x_n)\le G(r)$, also $\overline{\lim}\, F(x_n)\le F(x)$.

(iii) Beh. $\forall\, \varepsilon>0\ \exists\, \eta<\infty$ mit: (1) $F(\eta)\ge 1-\varepsilon$, (2) $F(-\eta)\le\varepsilon$.
Wegen der Straffheit von (μ_n) (vgl. 18.3a,c) gilt: $\forall\, \varepsilon>0\ \exists\, \eta\in\mathbb{Q}\ \ \forall\, n$: $F_n(\eta)-F_n(-\eta)\ge 1-\varepsilon$.

Diese Eigenschaft überträgt sich über G auf F [vgl. (*)].
(iv) Gemäß (ii) und (iii) ist F eine Vf. Nach dem Fortsetzungssatz 7.13 ex. zu F ein W–Maß μ, sodaß F Vf zu μ ist.
(v) Beh. $\mu_{n'}\xrightarrow{w}\mu$. Bew. mit 17.10b:

Zu $x\in\mathbb{R}$, $\varepsilon>0$ ex. $r,s\in\mathbb{Q}$ mit $x-\varepsilon<r<x<s<x+\varepsilon$. Aus $F_{n'}(r)\le F_{n'}(x)\le F_{n'}(s)$ folgt:

(*) $F(x-\varepsilon)\le G(r)=\lim F_{n'}(r)\le \underline{\lim}\, F_{n'}(x)\le \overline{\lim}\, F_{n'}(x)\le \lim F_{n'}(s)=G(s)\le F(x+\varepsilon)$.

Ist nun $x\in S(F)$, so folgt für $\varepsilon\downarrow 0$: $\lim F_{n'}(x)=F(x)$.. []

18.5 Bemerkung. Die Eigenschaft (a) $\mu_n\xrightarrow{w}\mu$

ist äquivalent zu (b) $\forall\, (n')\subset\mathbb{N}\ \exists\, (n'')\subset(n')$ mit $\mu_{n''}\xrightarrow{w}\mu$.

Die schwache Konvergenz wurde mit Hilfe der Konvergenz in $\mathbb{R}$ definiert; für diese ist die Folgerung "(b)⇒(a)" aber bekannt. []

18.6 Stetigkeitssatz von Levy–Cramér. Seien $\mu,\mu_n\in\mathbb{P}(\mathbb{R}^d)$ mit CF φ bzw. φ_n, $n\in\mathbb{N}$.

Dann sind folgende Aussagen äquivalent:

(a) $\mu_n\xrightarrow{w}\mu$

(b) $\varphi_n(t)\to\varphi(t)\ \ \forall\, t\in\mathbb{R}^d$.

Bew. "(a)⇒(b)" folgt aus $\varphi_{(n)}(t)=\int\cos(t^Tx)\mu_{(n)}[dx]+i\cdot\int\sin(t^Tx)\mu_{(n)}[dx]$
mit $x\mapsto\cos(t^Tx),\ \sin(t^Tx)\in C_b(\mathbb{R}^d)\ \ \forall\, t\in\mathbb{R}^d$.

"(b)⇒(a)" (i) Beh. $\{\mu_n\}$ ist rel. folgenkompakt.

Nach 18.3a und 18.4 genügt es dazu z.z.: $\{\mu_n\pi_j^{-1}\}$ ist straff $\forall$ j. Bew. davon mit 15.23:

Wir betrachten zunächst den Fall d=1. Nach 15.9 ist φ stetig und $\varphi(0)=1$; also:
$\forall\ \varepsilon>0\ \exists\ \delta>0$ mit $\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi(t))dt<\varepsilon$. Aus (b) folgt: $\varphi_n(t)\to\varphi(t)\ \ \forall\ t\in\mathbb{R}^1$, also mit maj. Konv.:

$\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi_n(t))dt \to \frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi(t))dt$. Somit folgt:
$\exists\, n_0\ \forall\, n\geq n_0$: $\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi_n(t))dt < \varepsilon$ und mit 15.23: $\mu_n[-\frac{2}{\delta},\frac{2}{\delta}] \geq 1-\varepsilon$.
Durch Verkleinerung von δ kann erreicht werden, daß dies für alle $n\in\mathbb{N}$ gilt.
Für den Fall $d > 1$ ist nach 15.9 $\mathbb{R}\ni t \mapsto \varphi_n(...,0,t,0,...)$ CF zu $\mu_n\pi_j^{-1}$. Nun folgt:
$\varphi_n(t) \to \varphi(t)\ \forall\, t \Rightarrow$ CF zu $\mu_n\pi_j^{-1} \to$ CF zu $\mu\pi_j^{-1}\ \ \forall\, j \Rightarrow \{\mu_n\pi_j^{-1}\}$ ist straff $\forall$ j gemäß dem Fall d=1.
(ii) Beh. Wir zeigen 18.5b: Sei $(n')\subset\mathbb{N}$ geg; nach (i) ex. $(n'')\subset(n')$, $\tilde\mu\in\mathbb{P}(\mathbb{R}^d)$ mit $\mu_{n''} \xrightarrow{w} \tilde\mu$.
Gemäß "(a)⇒(b)" folgt: $\tilde\mu$ hat CF φ, also gilt $\tilde\mu=\mu$ nach dem Eindeutigkeitssatz 15.16. $\square$

18.7 Cramér–Wold–Trick: Reduktion auf d=1.
Seien X_n d–dim Zva auf $(\Omega_n,\mathfrak{F}_n,P_n)$, $n\in\mathbb{N}_0$. Dann sind folgende Aussagen äquivalent:
(a) $X_n \xrightarrow{w} X_0$;
(b) $c^\top X_n \xrightarrow{w} c^\top X_0\ \ \forall\, c\in\mathbb{R}^d$.

Bew. "(a)⇒(b)" folgt aus 17.8 mit $h(x)=c^\top x$. "(b)⇒(a)" Sei φ_n CF von X_n, dann hat nach 15.9b [setze dort $A = c^\top$] $c^\top X_n$ die CF: $\mathbb{R}\ni t\mapsto \varphi_n((c^\top)^\top t) = \varphi_n(t\cdot c)$. Nach Vorauss. (b) und 18.6"(a)⇒(b)" mit t=1 gilt: $\varphi_n(c)\to\varphi_0(c)\ \forall\, c\in\mathbb{R}^d$. Jetzt 18.6"(b)⇒(a)". $\square$

Anhang: **<u>Verallgemeinerungen auf Dim. d</u>** .

A18.1 Def. $F:\mathbb{R}^d\mapsto\mathbb{R}$ heißt **stetig von oben**, falls $F(x_n)\to F(x)$ für $x_n\to x$ und $x_n\geq x$ (komponentenweise). Für $a=(a_1,...,a_d)$, $b=(b_1,...,b_d)\in\mathbb{R}^d$, $a\leq b$, sei

$$\Delta_a^b F := \sum_{y=(y_1,...,y_d),\ y_i\in\{a_i,b_i\}} (-1)^{|\{i;y_i=a_i\}|}\cdot F(y)$$

Dabei summiert man gerade über die Menge $\times_{i=1}^{d}\{a_i,b_i\}$ der **Eckpunkte** des Intervalls $(a,b]$.

A18.2 Lemma. Ist F Vf zu $\mu\in\mathbb{P}(\mathbb{R}^d)$, so gilt für $a\leq b\in\mathbb{R}^d$: $\mu(a,b] = \Delta_a^b F$.

Für d=2 gilt $\Delta_a^b F = F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2)$.

Bew. $\mu(a,b] = \mu[\{x;\ x\leq b\}] - \mu[\cup_i \{x;\ x_i\leq a_i,\ x\leq b\}]$. Verwende jetzt für den zweiten Term die Siebformel 2.19(11). $\square$

A18.3 Fortsetzungssatz. Ist $F:\mathbb{R}^d\mapsto\mathbb{R}$ eine von oben stetige Funktion mit $\Delta_a^b F\geq 0\ \ \forall\, a\leq b$ und (*)

$\forall\,\varepsilon>0\ \exists\,\eta<\infty$: (1) $F(x)\geq 1-\varepsilon$, falls $x\geq\eta$ für alle j und
(2) $F(x)\leq\varepsilon$, falls $x_j\leq-\eta$ für ein j,

so existiert genau ein $\mu\in\mathbb{P}(\mathbb{R}^d)$ mit Vf F.

Bew. Die Eindeutigkeit folgt aus 7.7a. Die Existenz eines Borelmaßes μ mit $\mu(a,b] = \Delta_a^b F \;\; \forall\, a \le b$ folgt aus dem Fortsetzungssatz von Carathéodory [vgl. Billingsley (1978/86) Theorem 12.5]. Z.z. ist noch, daß μ ein W–Maß mit der Vf F ist.

Beh. $\mu(-\vec{\infty},b] = F(b) \; \forall\, b \in \mathbb{R}^d$. Zum Bew. wähle zu $\varepsilon>0$ ein η wie in (*) und $a \le -\vec{\eta}$. Dann gilt: $F(x) \le \varepsilon$ für alle $x \in \times_1^d \{-a_j, b_j\} \setminus \{b\}$, d.h. für alle Eckpunkte von $(a,b]$ außer b. Somit folgt nach A18.2 : $F(b) = \mu(-\vec{\infty},b] \ge \mu(a,b] = \Delta_a^b F \ge F(b) - (2^d-1)\varepsilon$.

Beh. $\mu[\mathbb{R}] = 1$; denn für $b \ge \vec{\eta}$ folgt nun mit (*)(1) und $F \le 1$: $1-\varepsilon \le \mu(-\vec{\infty},b] \le 1$. []

A18.4 Def. Ist F Vf zu $\mu \in \mathbb{P}(\mathbb{R}^d)$ und $\pi_k : \mathbb{R}^d \mapsto \mathbb{R}$ Projektion auf die k–te Koordinate, F_k die (1–dim.) Vf der k–ten Randverteilung $\mu\pi_k^{-1}$ von μ, dann sei: $S_d(F) := \times_{k=1}^d S(F_k) \subset \mathbb{R}^d$.

$S_d(F)$ hat gegenüber S(F) den Vorteil, eine Produktmenge zu sein; dies wird unten in (A18.7) benutzt. $S_d(F)$ kann genau so gut wie S(F) die schwache Konvergenz gegen μ charakterisieren.

A18.5 Lemma. Sei $\mu \in \mathbb{P}(\mathbb{R}^d)$ mit Vf F. Dann gilt:

(a) $\mu[\partial(-\vec{\infty},s]] = 0 \;\; \forall\, s \in S_d(F)$;

(b) $S_d(F)$ ist dicht in $\mathbb{R}^d$.

Bew. a) $\partial(-\vec{\infty},s] \subset \cup_{k=1}^d \{x \in \mathbb{R}^d;\, x_k = s_k\}$ und $\mu[\{x \in \mathbb{R}^d;\, x_k = s_k\}] = \mu\pi_k^{-1}[\{s_k\}] = 0$ für $s_k \in S(F_k)$, wobei F_k wie in 18.4 ist. b) $S(F_k)^c$ ist abzb.. []

A18.6 Satz. Sei $S = \mathbb{R}^d$, $\mu, \mu_n \in \mathbb{P}(\mathbb{R}^d)$ mit Vf F bzw. F_n, $n \in \mathbb{N}$. Dann sind äquivalent:

(a) $\mu_n \xrightarrow{w} \mu$;

(b) $\mu_n(a,b] \to \mu(a,b] \;\; \forall\, a,b \in S_d(F),\, a \le b$;

(c) $F_n(s) \to F(s) \;\; \forall\, s \in S_d(F)$.

Bew."(a)⇒(c)" mit 17.7d und A18.5a.

"(c)⇒(b)" Nach 7.22 gilt $\mu(a,b] = \Delta_a^b F$, wobei $\Delta_a^b F$ Linearkombination von Werten von F an den Eckpunkten $\times_{k=1}^d \{a_k,b_k\}$ von $(a,b]$, entsprechend für F_n und μ_n. Nun gilt aber:

(A18.7) $\qquad \times_{k=1}^d \{a_k,b_k\} \subset S_d(F)$ für $a,b \in S_d(F)$.

"(b)⇒(a)" Sei $\mathfrak{U} := \{(a,b];\, a,b \in S_d(F)\}$. Nach Vorauss. gilt $\mu_n[U] \to \mu[U] \;\; \forall\, U \in \mathfrak{U}$. Wegen (6.4b) (und im Grunde genommen (18.7)) ist $\mathfrak{U}$ $\cap$–stabil. Nach der Siebformel 2.19(11) folgt:

$\mu_n[\cup_{k=1}^m U_k] \to \mu[\cup_{k=1}^m U_k]$ für $U_k \in \mathfrak{U}$, $1 \le k \le m \in \mathbb{N}$. Sei nun G offen; dann gilt:

$G = \cup_{k=1}^{\infty} U_k$ für gewisse $U_k \in \mathfrak{U}$, da $S_d(F)$ dicht in $\mathbb{R}^d$ (vgl. Bew. von 6.4). Es folgt:
$\underline{\lim}\ \mu_n[G] \geq \lim \mu_n[\cup_1^m U_k] = \mu[\cup_1^m U_k]$; jetzt $m \to \infty$. $\square$

A18.8 Lemma. Sei $\Gamma \subset \mathbb{P}(\mathbb{R}^d)$, F_μ Vf. zu μ. Dann sind folgende Aussagen äquivalent:

(a) Γ ist straff;

(b) $\forall\ \varepsilon>0\ \exists\ \eta<\infty\ \forall\ \mu\in\Gamma$: (1) $F_\mu(x) \geq 1-\varepsilon$, falls $x_j \geq \eta$ für alle j und
(2) $F_\mu(x) \leq \varepsilon$, falls $x_j \leq -\eta$ für ein j.

Bew. "(a)⇒(b)" Zu $\varepsilon>0$ wähle η gemäß 18.3a (ii), sodaß $\forall\ \mu\in\Gamma\ \forall\ j$: $\mu\pi_j^{-1}(-\eta,\eta] > 1-\frac{\varepsilon}{d}$. Dann:
(1) $\forall\ \mu\in\Gamma$: $1-F_\mu(\vec{\eta}) = \mu[\cup_{j=1}^d \{\pi_j>\eta\}] \leq \sum_j \mu[\pi_j>\eta] \leq d\cdot\varepsilon/d = \varepsilon$;
(2) $\forall\ x\in\mathbb{R}^d\ \forall\ \mu\in\Gamma\ \forall\ j$: $\varepsilon/d \geq \mu[\pi_j\leq-\eta] \geq \mu[\pi_j\leq-\eta,\ \pi_k\leq x_k\ \forall\ k\neq j] = F_\mu(\ldots,x_{j-1},-\eta,x_{j+1},\ldots)$.
"(b)⇒(a)" zeigt man mit den Methoden vom Beweis zu A18.3. $\square$

Beweis von **18**.4 **Satz von Prohorov** "(a)⇒(b)" für $d\in\mathbb{N}$.
Ist (a) $\Gamma \subset \mathbb{P}(\mathbb{R}^d)$ straff, so ist (b) Γ rel. folgenkompakt, d.h. $\forall\ (\mu_n)\subset\Gamma\ \exists\ (n')\subset\mathbb{N}$, $\mu\in\mathbb{P}(\mathbb{R}^d)$ mit $\mu_{n'} \xrightarrow{w} \mu$.

Geg. $(\mu_n)\subset\Gamma$ mit Vf. F_n.

(i) Mit Hilfe eines Diagonalisierungsverfahrens erhält man $(n')\subset\mathbb{N}$ sodaß
$G(r) := \lim F_{n'}(r)$ ex. $\forall\ r\in\mathbb{Q}^d$. G ist dann (komponentenweise) isoton auf $\mathbb{Q}^d$.

(ii) $F(x) := \inf\{G(r);\ r\in\mathbb{Q}^d,\ r>x\}$ ist isoton auf $\mathbb{R}^d$, $0\leq F\leq 1$. Ferner ist F stetig von oben, denn sei $x_n \downarrow x$, dann $\forall\ r>x\ \exists\ n_0\ \forall\ n\geq n_0$: $x_n<r$, d.h. $\forall\ r>x\ \overline{\lim}\ F(x_n)\leq G(r)$, also $\overline{\lim}\ F(x_n)\leq F(x)$.

(iii) Beh. $\Delta_a^b F \geq 0$. Aus $\Delta_a^b F_n \geq 0\ \forall a\leq b$ folgt zunächst: $\Delta_r^s G \geq 0\ \forall r\leq s,\ r,s\in\mathbb{Q}^d$.
Für $\varepsilon>0$ gilt nun $|\Delta_r^s G - \Delta_a^b F| < \varepsilon$, wenn $r>a, r\in\mathbb{Q}^d$ nahe bei a, $s>b, s\in\mathbb{Q}^d$ nahe bei b und somit alle Eckpunkte von $(r,s]$ nahe bei den entsprechenden von $(a,b]$ liegen.

(iv) Beh. $\forall\ \varepsilon>0\ \exists\ \eta<\infty$: (1) $F(x) \geq 1-\varepsilon$, falls $x_j\geq\eta$ für alle j und
(2) $F(x) \leq \varepsilon$, falls $x_j \leq -\eta$ für ein j.

Diese Eigenschaft überträgt sich wegen der Straffheit von (F_n) (vgl. A18.8b) über G auf F.

(v) Nach dem Fortsetzungssatz A18.3 ex. zu F ein W–Maß μ mit Vf F. Z. z. ist noch:
(vi) $\mu_{n'} \xrightarrow{w} \mu$. Bew. mit A18.6c. Für $x\in S_d(F)$ gilt nach A18.5a
(*) $\lim_{\varepsilon\downarrow 0} F(x-\vec{\varepsilon}) = \mu(-\vec{\infty},x) = \mu(-\vec{\infty},x] = F(x) = \lim_{\varepsilon\downarrow 0} F(x+\vec{\varepsilon})$.
Zu $\varepsilon>0$ ex. $r,s\in\mathbb{Q}^d$ mit $x-\vec{\varepsilon}<r<x<s<x+\vec{\varepsilon}$. Aus $F_{n'}(r) \leq F_{n'}(x) \leq F_{n'}(s)$ folgt:
$F(x-\vec{\varepsilon}) \leq G(r) \leq \underline{\lim}\ F_{n'}(x) \leq \overline{\lim}\ F_{n'}(x) \leq G(s) \leq F((x+\vec{\varepsilon})$. Jetzt $\varepsilon\downarrow 0$ und (*). $\square$

§19 Grenzwertsätze.

Im folgenden Satz sei in Hinblick auf 18.7 o.E. d=1 (vgl. Beweis von 19.11)

19.1 Schwaches Gesetz der großen Zahlen. Seien $X_1,X_2,\ldots$ iid Zva mit a:= $EX_j \in \mathbb{R}$. Dann gilt:

(a) $\bar{X}_n := \frac{1}{n}\sum_{j=1}^n X_j \xrightarrow{w} a$;

(b) $\bar{X}_n$ konvergiert nach W. gegen a, d.h. $P[|\bar{X}_n - a| \geq \varepsilon] \to 0$ $(n\to\infty)$ $\forall\varepsilon>0$.

Bemerkung 17.12 gilt hier entsprechend.

Bew. von 19.1. (a) Sei φ CF von X_j, φ_n CF von $\bar{X}_n$, dann ist nach 15.9 $t\mapsto\varphi(t)^n$ CF von $\sum_{j=1}^n X_j^{(n)}$ und $\varphi_n(t) = (\varphi(\frac{t}{n}))^n$. Nach 15.22 gilt: $\varphi(t) = 1 + ita + o(t)$ für $t\to 0$,

also: $\varphi_n(t) = (1 + \frac{1}{n}ita + o(\frac{t}{n}))^n$ für t fest und $n\to\infty$. Nach 5.16 folgt $\varphi_n(t) \to e^{ita}$.

Dies ist die CF zu $\delta_a \in \mathbb{P}(\mathbb{R})$. Jetzt 18.6.

"(a)⇒(b)" Mit 17.7c folgt:

$\overline{\lim}\, P_n[|\bar{X}_n - a|)\geq\varepsilon] = \overline{\lim}\, P_n\bar{X}_n^{-1}[\{x\in S;\ |x,a|\geq\varepsilon\}] \leq \delta_a[\{x\in S;\ |x-a|\geq\varepsilon\}] = 0.$ □

Gilt auch $V[X_j^{(n)}]<\infty$, so folgt die gemäß 17.18 äquivalente Konvergenz nach W. direkt aus der Markoffschen (für p=2 ≅ Tschebyscheffschen) Ungleichung (vgl. 1.14).

19.2 Starkes Gesetz der großen Zahlen. Ist $X_1,X_2,\ldots$ eine Folge von reellen iid Zva auf $(\Omega.\mathfrak{F},P)$, so daß $a=E[X_j]\in\mathbb{R}$ existiert, dann gilt:

$$\bar{X}_n := \frac{1}{n}\sum_{j=1}^n X_j \to a \text{ f.s..}$$

19.2 bedeutet in der Situation der **Statistik**, wo P und a unbekannt sind, daß $\bar{X}_n$ ein **konsistenter Schätzer** für a ist.

Bew. in W–theorie II als Korollar zum **Martingalkonvergenzsatz**. Die Vorauss. von 19.2 ist erfüllbar nach der folgenden Verallgemeinerung von 13.5.

19.3 Satz. Seien $(\Omega'_n,\mathfrak{F}'_n,Q_n)$, $n\in\mathbb{N}$, W–Räume. Dann existieren $(\Omega,\mathfrak{F},P)$ und unabhängige Zva X_n auf $(\Omega,\mathfrak{F},P)$ mit Werten in Ω'_n und $X_n\sim Q_n$.

Bew. in W–Theorie II als Korollar zum **Satz von Ionescu–Tulcea.** Dieser wird wiederum mit dem **Fortsetzungssatz von Caratheodory** bewiesen.

19.4 Bemerkung. Das schwache Gesetz folgt aus dem starken Gesetz der großen Zahlen. Zu μ wähle dazu (X_n) wie in 19.3 mit $Q_n=\mu$. Dann folgt nach 19.2 und 17.19a: $\bar{X}_n \xrightarrow{w} a$. Dieses $\bar{X}_n$ hat aber die gleiche Verteilung wie das $\bar{X}_n$ aus 19.1. □

19.5 Borelsches starkes Gesetz der großen Zahlen. Sei (Y_n) Folge von iid Zva auf $(\Omega,\mathfrak{F},P)$ mit Werten in $(\Omega',\mathfrak{F}')$ und $A'\in\mathfrak{F}'$, $p:=P[Y_n\in A']$, $h_n:=\frac{1}{n}\sum_1^n 1_{\{Y_j\in A'\}}$. Dann gilt: $h_n \to p$ f.s. .

Bew. Wähle in 19.2 $X_n:=1_{\{Y_n\in A'\}}=1_{A'}\circ Y_n$. []

19.6 Def. Seien $X_1,\ldots,X_n$ reelle iid Zva auf $(\Omega,\mathfrak{F},P)$. Dann heißt

$F_n(t)(\omega):=\frac{1}{n}\sum_{j=1}^n 1_{(-\infty,t]}(X_j(\omega))$ **empirische Vf** zu $(X_1,\ldots,X_n)$.

Also $F_n(\cdot)(\omega)$ ist Vf zu $\mathrm{Emp}(\cdot\,|X_1(\omega),\ldots,X_n(\omega))$.

F_n ist "**Schätzer**" für Vf F der X_j auf der Basis der beobachteten Stichprobe $(X_1(\omega),\ldots,X_n(\omega))$. F_n hat in t Sprung der Höhe $\frac{m}{n}$, falls der Wert t in der Stichprobe m–mal angenommen wird, und ist zwichen den Sprüngen konstant.

19.7 Lemma. Sei (X_n) eine Folge reeller iid Zva mit Vf F und emp.Vf F_n zu $(X_1,\ldots,X_n)$. Dann gilt: (a) $E[F_n(t)]=F(t)$ [d.h. $F_n(t)$ ist **erwartungstreuer Schätzer** für $F(t)$];

(b) $F_n(t)\to F(t)$ f.s. [d.h. $F_n(t)$ ist **konsistenter Schätzer** für $F(t)$],

$F_n(t-0)\to F(t-0)$ f.s. $\forall\, t\in\mathbb{R}$.

Bew. a) $E[F_n(t)]=\frac{1}{n}\sum_1^n P[X_j\le t]=F(t)$.

b) Setze im Borelschen Gesetz 19.5: $A'=(-\infty,t]$ oder $(-\infty,t)$. []

19.8 Satz von Glivenko–Cantelli. Sei (X_n) eine Folge von reellen iid Zva auf $(\Omega,\mathfrak{F},P)$ mit Vf F und emp.Vf F_n zu $(X_1,\ldots,X_n)$. Dann gilt: $\sup_t|F_n(t)-F(t)|\to 0$ f.s..

Bew. $\sup_{t\in\mathbb{R}}\ldots=\sup_{t\in\mathbb{Q}}\ldots$ ist mb. Nach 19.7 ex. $\mathfrak{F}\ni\Omega^t\supset\{F_n(t)\to F(t),\ F_n(t-0)\to F(t-0)\}$ mit $P[\Omega^t]=1$. Sei $D:=(0,1)\cap\mathbb{Q}$, dann ist $T:=\{F^{-}(\alpha);\ \alpha\in D\}$ abzb.. Setze $\Omega_0:=\cap_{t\in T}\Omega^t$; dann $\Omega_0\in\mathfrak{F}$ und $P[\Omega_0]=1$. Auf Ω_0 gilt nun nach 17.13: $\sup_t|F_n(t)-F(t)|\to 0$. []

19.9 Korollar. Sei (X_n) eine Folge von reellen iid Zva auf $(\Omega,\mathfrak{F},P)$ mit $X_n\sim\mu\in\mathbb{P}(\mathbb{R}^1)$. Dann gilt:

$$\mathrm{Emp}(\cdot\,|X_1,\ldots,X_n)\overset{w}{\to}\mu \text{ f.s..}$$

Bew. Sei Ω_0 wie im Bew. von 19.8. Für $\omega\in\Omega_0$ gilt gemäß 17.10b:

$\mathrm{Emp}(\cdot\,|X_1(\omega),\ldots,X_n(\omega))\overset{w}{\to}\mu$. []

19.10 Anwendung: Weierstraßscher Approximationssatz. Sei f reelle, stetige Funktion auf $[0,1]$, $p_n^f(t):=\sum_{k=0}^n f(\frac{k}{n})\binom{n}{k}t^k(1-t)^{n-k}$ (Bernstein–Polynom). Dann gilt: $p_n^f\to f$ glm. auf $[0,1]$.

Bew. Seien X_n, $n\in\mathbb{N}$, iid auf $(\Omega,\mathfrak{F},P)$ mit $X_n\sim U(O,1)$, $S_n(t):= nF_n(t) = \sum_1^n 1_{[0,t]}(X_k)$.
Dann gilt $S_n(t)\sim b(n,t)$, also $E[f(F_n(t))] = E[f(\frac{1}{n}S_n(t))] = p_n^f(t)$.
Sei $\Omega_0:= \{\sup_{0\le t\le 1} |F_n(t)-t| \to 0\}$, dann gilt $P[\Omega_0]=1$ nach 19.8 und somit
$\sup_{0\le t\le 1} |p_n^f(t)-f(t)| = \sup_{0\le t\le 1} |E[f(F_n(t))-f(t)]| \le E[\sup_{0\le t\le 1} f(F_n(t))-f(t)|1_{\Omega_0}] \to 0$ wegen glm. Stetigkeit von f und maj. Konv.. []

Erinnerung: Ist $X = (X_1,...,X_d)$ d–dim. Zufallsvektor mit $E[\|X\|^2] = \sum_{k=1}^n E[X_k^2] < \infty$, so heißt X quadratisch intb. und $\mathfrak{Cov}[X] := \left[\mathrm{Kov}[X_j,X_k]\right] \in \mathbb{R}^{d\times d}$ Kovarianzmatris von X (vgl. 16.1).
Zu jeder Kovarianzmatrix $\Gamma \in \mathbb{R}^{d\times d}$ mit Rang rg $\Gamma = m \le d$ existiert $\Sigma\in \mathbb{R}^{d\times m}$ mit (vgl. 16.3b,d)

$$\text{rg } \Sigma= m \text{ und } \Gamma = \Sigma\Sigma^\top.$$

Für $m\in\mathbb{N}$ seien $W_1,...,W_m$ reelle iid Zva mit $W_k \sim N(0,1)$ und $W=(W_1,...,W_m)^\top$. Die Verteilung von W ist die **standardisierte d–dim. Normalverteilung** $N(0,I_m) := \bigtimes_1^m N(0,1)$ mit der Einheitsmatrix I_m.
Der Zufallsvektor $X=(X_1,...,X_d)^\top$ besitzt eine **Normalverteilung $N(a,\Gamma)$** mit $a\in\mathbb{R}^d$, $\Gamma \in \mathbb{R}^{d\times d}$, falls eine m–dim. Zva W wie oben existiert mit $X = \Sigma W + a$, $\Sigma\in\mathbb{R}^{d\times m}$, $\Gamma = \Sigma\Sigma^\top$. Dann ist Γ die Kovarianzmatrix von X (vgl. 16.5). Für $c \in \mathbb{R}^d$ hat $c^\top X = c^\top\Sigma W + c^\top a$ die Normalverteilung $N(c^\top a,c^\top\Sigma (c^\top\Sigma)^\top) = N(c^\top a,c^\top\Sigma\Sigma^\top c) = N(c^\top a,c^\top\Gamma c)$.

19.11 Zentraler Grenzwertsatz. Seien $X_1,X_2,...$ iid quadratisch intb. Zva mit
$EX_j =: a \in \mathbb{R}^d$, $\mathfrak{Cov}[X] =:\Gamma \in \mathbb{R}^{d\times d}$. Dann gilt:

$$\tilde{W}_n := \sum_1^n (X_k - a)/\sqrt{n} \xrightarrow{w} N(0,\Gamma) .$$

Bemerkung 17.12 gilt hier entsprechend.
Bew. von 19.11. a) Neuer Bew. für d=1. Sei $\sigma^2:= \mathrm{Var}[X_k]$. Setze $\tilde{X}_k:= X_k-a$; dann gilt: $(\tilde{X}_k)$ iid, $E[\tilde{X}_k]=0$, $E[\tilde{X}_k^2]=\sigma^2$. Sei φ CF von $\tilde{X}_k$, φ_n CF von $\tilde{W}_n= \sum_1^n \tilde{X}_k/\sqrt{n}$; dann folgt mit 15.9:
$\varphi_n(t) = (\varphi(t/\sqrt{n}))^n$. Da nach 15.22 $\varphi(t) = 1 + itE[\tilde{X}_1] - \frac{1}{2}t^2\sigma^2 + o(t^2)$ für $t\to 0$, folgt:

$$\varphi_n(t) = (1 - \frac{1}{n}\cdot\tfrac{1}{2}t^2\sigma^2 + o(\frac{t^2}{n}))^n \quad \text{für t fest, } n\to\infty.$$

Aus 5.16 folgt: $\varphi_n(t) \to \exp\{-\frac{1}{2}\sigma^2t^2\}$; dies ist die CF von $N(0,\sigma^2)$ nach 15.11 b mit $N(0,0) = \delta_0$. Jetzt ist der Stetigkeitssatz 18.6 anwendbar. Ist $\sigma^2 > 0$, so ergibt sich die Form 17.11 mit $W_n = \frac{1}{\sigma}\tilde{W}_n =: h(\tilde{W}_n)$ und 17.8.

b) Bew. für $d > 1$ mit dem Cramér–Wold–Trick. Sei $c \in \mathbb{R}^d$. Dann ist $c^\top \tilde{W}_n = \sum_1^n (c^\top X_k - c^\top a)/\sqrt{n}$; dabei sind $c^\top X_k$ reelle iid Zva mit $E[c^\top X_k] = c^\top a$ (vgl. 16.2), $\mathrm{Var}[c^\top X_k] = c^\top \Gamma c =: \sigma_c^2$ (vgl. 16.3).

Nach a) gilt: $c^\top \tilde{W}_n \overset{w}{\to} N(0,\sigma_c^2)$. Wähle nun $X \sim N(0,\Gamma)$. Dann hat $c^\top X$ die Normalverteilung $N(0,c^\top \Gamma c) = N(0,\sigma_c^2)$ wie oben gezeigt wurde, also $c^\top \tilde{W}_n \overset{w}{\to} c^\top X$ $\forall\, c \in \mathbb{R}^d$. Jetzt 18.7.

c) Ein direkter Beweis für bel. d kann analog zu a) geführt werden mit 16.4 und 16.6h. $\square$

Anwendung: Die Black–Scholes–Formel zur Bewertung von Optionen.

Das 1–Perioden–Modell:

Situation: Es bestehen zwei Anlagemöglichkeiten:
(i) Eine risikolose Anlage von ξ EUR in ein Sparkonto (festverzinsliches Wertpapier, bond) erbringt am Ende der Periode $R\cdot\xi$ DM mit $R > 1$.
Ann.: Das gleiche R gilt auch für $\xi<0$, also wenn Geld geliehen wird. (Modelleinschränkung).
(Zur Vereinfachung kann man sich zunächst R=1 denken.)

(ii) Eine risikobehaftete Anlage von ξ Anteilen in ein kursabhängiges Wertpapier (stock) [z.B. Aktien, Devisen oder Investmentfond] mit dem Kurs $X_m>0$ kostet z.Zt. m=0 $\xi\cdot X_0$ DM und hat z.Zt. m=1 den Wert $\xi\cdot X_1$.

Ann.: $X_0=x_0$ bekannt, $X_1\in\{\rho_-x_0, \rho_+x_0\}$ (Modelleinschränkung) mit $\rho_- < R < \rho_+$.

Die Bank bietet einen Vertrag an. Der Käufer zahlt z.Zt. m=0 einen Preis (Prämie) V und hat dafür z.Zt. m=1 einen Zahlungsanspruch $h(X_1)$.

Problem: Was ist zu geg. h ein **fairer Preis** V ?
Klassische Antwort: $V = E[R^{-1}\cdot h(X_1)]$, falls $P[X_1= \rho_\pm x_0]$ bekannt.

In der vorliegenden besonderen Sitation gibt es eine andere Antwort.
Die Bank kann sich gegen den Zahlungsanspruch absichern, in dem sie selbst mit dem Startkapital V in das Wertpapier (und das Sparkonto) investiert. Kauft sie ξ Anteile, so ist ihre Bilanz z.Zt. m=1 :

Einnahmen: $\xi\cdot X_1 + R\cdot(V-\xi\cdot x_0) = R\cdot V + \xi(X_1-R\cdot x_0)$ **Ausgaben:** $h(X_1)$.

In der vorliegenden Situation gibt es nach folgendem Lemma ein V, so daß die Bilanz in jedem der Fälle $X_1=\rho_-x_0$ oder $X_1=\rho_+x_0$ ausgeglichen ist. V wird dann als fairer Preis angesehen.

Setze $h_\pm := h(\rho_\pm x_0)$.

19.12 Lemma. Das Gleichungssystem $R\cdot V + \xi[\rho_\pm x_0-R\cdot x_0] = h_\pm$ in V und ξ hat eine eindeutige Lösung; dabei ist: $R\cdot V = p_-h_- + p_+h_+$ mit

$$p_-:= \frac{\rho_+ - R}{\rho_+ - \rho_-},\quad p_+:= \frac{R - \rho_-}{\rho_+ - \rho_-},\quad \text{und} \quad 0 < p_\pm < 1,\ p_-+p_+=1.$$

Würde X_1 mehr als zwei Werte annehmen, so hätte man für die beiden Unbekannten V und ξ mehr als zwei Bestimmungsgleichungen.

19.13 Bemerkung. Wählt man ein künstliches W–Maß P^*, so daß $P^*[X_1=\rho_\pm x_0] = p_\pm$, so gilt:

$$V = R^{-1}\cdot E^*[h(X_1)],\quad E^*[R^{-1}\cdot X_1] = x_0 \quad \text{wegen}$$

(19.14) $$p_-\cdot\rho_- + p_+\cdot\rho_+ = R.$$

[In dem häufigen Fall $h_- = 0 < h_+$ gilt: $\xi \cdot x_0 > V$, es wird also mehr in das Wertpapier investiert, als die Bank von dem Käufer erhält.]

Das n–Perioden Modell.

Modell: $\Omega := \{-,+\}^n = \{\omega = (i_1,\dots,i_n);\ i_m \in \{-,+\}\}$; , dann: $X_m(\omega) = \rho_{i_m} \cdot \dots \cdot \rho_{i_1} \cdot x_0 = \rho_{i_m} \cdot X_{m-1}(\omega)$.

Zu $\omega = (i_1,\dots,i_n)$ setze $\omega_m := (i_1,\dots,i_m)$, also $X_m(\omega) = X_m(\omega_m)$.

Es wird kein W–Maß vorausgesetzt!

Wir betrachten wieder die Situation: Die Bank bietet einen Vertrag an. Der Käufer zahlt z.Zt. m=0 einen Preis/Prämie V und hat dafür z.Zt. n einen Zahlungsanspruch $h(X_n)$.

Eine (call–)**Option** ist ein Vertrag, der dem Käufer das Recht einräumt, zum Fälligkeits– termin (maturity time) n für einen festen Wahrnehmungspreis (exercise/striking price) K unabhängig von dem vorliegenden Kurs X_n ein Wertpapier zu kaufen (Europäische Option). Dann ist

$h(X_n) := [X_n - K]^+$ der Zahlungsanspruch (contingent claim),

also der Gewinn des Käufers.

Problem: Welches ist ein fairer Preis V_0 für diese Option?

Die Bank braucht (im Prinzip) nicht tatenlos abzuwarten, ob der Kurs fällt, was günstig für die Bank ist, oder steigt, was günstig für den Käufer ist. Um den Verlust bei steigendem Kurs abzusichern (hedging), kann die Bank selbst Wertpapiere kaufen (in den stock investieren). Die damit verbundenen Unsicherheiten werden durch $\omega \in \Omega$ dargestellt.

Dabei kann eine Strategie verwendet werden, gegeben durch

$\xi_0 \in \mathbb{R}$, $\xi_m(\omega_m) \in \mathbb{R}$, $\omega_m \in \{-,+\}^m$, $1 \le m < n$.:

Die Bank startet mit dem Kapital V_0, gegeben durch den Preis für die Option.

Davon werden ξ_0 Anteile zum Kurs von x_0, also zum Preis von $\xi_0 x_0$, gekauft (in den stock investiert) und $V_0 - \xi_0 x_0$ zum Zinssatz r angelegt (in den bond investiert).

[Oft ist $V_0 - \xi_0 x_0 \le 0$, dann wird Geld geliehen zum selben Zinssatz wie beim Sparen!]

Z.Zt. m=1 verfügt die Bank beim Kurs X_1 über: $\xi_0 X_1 + R(V_0 - \xi_0 x_0) = RV_0 + \xi_0(X_1 - Rx_0)$.

Nun wird die Anlage umgeschichtet. Die Bank hält nun insgesamt ξ_1 Anteile am Wertpapier. Die Entscheidung über ξ_1 kann dabei über die bis dahin beobachtete Preisentwicklung abhängen, also $\xi_1 = \xi_1(i_1) = \xi_1(\omega_1)$, $i_1 \in \{-,+\}$.

Z.Zt. m=2 verfügt die Bank beim Kurs X_2 über:

$\xi_1 X_2 + R[RV_0+\xi_0(X_1-Rx_0) - \xi_1 X_1] = R^2 V_0 + R\xi_0(X_1-Rx_0) + \xi_1(X_2-RX_1)$;

usw. Z.Zt. m=n verfügt die Bank beim Kurs X_n über:

$$\hat{V}_n := R^n V_0 + \sum_{m=1}^{n} R^{n-m} \cdot \xi_{m-1} \cdot [X_m - RX_{m-1}] \quad \text{mit } \xi_m = \xi_m(\omega_m), 0 \le m < n.$$

19.15 Lemma. Zu einer beliebigen Funktion $h : \mathbb{R} \mapsto \mathbb{R}$ existieren $V_0 \in \mathbb{R}$, $\xi_0 \in \mathbb{R}$, $\xi_m : \{-,+\}^m \mapsto \mathbb{R}$, $1 \le m < n$,

sodaß $\forall \omega \in \Omega$ gilt: $\hat{V}_n(\omega) = h(X_n(\omega))$.

Dabei ist $R^n \cdot V_0 = \sum_{\omega=(i_1,\dots,i_n)\in\Omega} p_{i_1} \cdot p_{i_2} \cdot \ldots \cdot p_{i_n} \cdot h(X_n(\omega))$

mit $p_\pm$ wie in 19.12.

Wählt man im Fall $h(x) = [x-K]^+$ nun $V_0, \xi_0, \dots, \xi_n$ wie in 19.14, so hat die Bank aus dem Startkapital V_0 gerade soviel erwirtschaftet, daß sie dem Käufer seinen Zahlungsanspruch $(X_n-K)^+$ auszahlen kann, und zwar unter jeder möglichen Preisentwicklung $\omega \in \Omega$. Insofern ist das Startkapital V_0 gleichwertig mit der Option. Man kann damit die Option imitieren (dublizieren) oder risikolos absichern. V_0 wird dann als **fairer Preis** angesehen.

Bew. von 19.15. Setze $V_n(\omega) = h(X_n(\omega))$. Zu ω_k, $k<n$, wähle rekursiv $V = V_k(\omega_k)$ und $\xi = \xi_k(\omega_k)$ gemäß 19.12 als Lösung von

$$RV + \xi[X_{k+1}(\omega_k,\pm) - R \cdot X_k(\omega_k)] = V_{k+1}(\omega_k,\pm) \,. \,\Box$$

19.16 Darstellung von V_0:

$(\Omega,\mathfrak{P}(\Omega),P^*)$ sei künstliches W–Raum mit $P^*[\{(i_1,\dots,i_n)\}] = p_{i_1} \cdot p_{i_2} \cdot \ldots \cdot p_{i_n}$, also

$P^* = \bigotimes_{k=1}^{n} Q^*$, wobei $Q^*[\{\pm\}]=p_\pm$. Dann sind die Koordinatenvariablen $\pi_k : \Omega \mapsto \{-,+\}$, $1 \le k \le n$, iid mit $\pi_k \sim Q^*$, $X_m = \rho_{\pi_m} \cdot X_{m-1}$. Unter P^* gilt:

$$V_0 = R^{-n} \cdot E^*[\, R^{-n} \cdot h(X_n)]$$

$$E^*[R^{-(m+1)} \cdot X_{m+1} \,|\, (\pi_1,\dots,\pi_m)=\omega'_m] = R^{-m} \cdot X_m(\omega'_m), \quad 0 \le m < n.$$

Dabei ist $E^*[\ldots|A] := \int \ldots dP[\,\cdot\,|A] = E^*[\ldots \cdot 1_A]/P[A]$.

Die zweite Eigenschaft in (19.15) besagt gerade, daß $(R^{-m} \cdot X_m, 0 \le m \le n)$ ein **Martingal** ist, wobei $R^{-m} \cdot X_m$ der auf den Zeitpunkt 0 abdiskontierte Kurs z.Zt. m ist. P^* heißt deswegen auch **Martingalmaß**. Der Beweis folgt leicht aus (19.14). Aus (19.15) folgt nun:

$$V_0 + \sum_{m=1}^{n} \xi_{m-1} \cdot [R^{-m} \cdot X_m - R^{-(m-1)} \cdot X_{m-1}] = h(X_n).$$

Die linke Seite kann als zeitdiskretes stochastisches Integral aufgefaßt werden mit (ξ_m) als Integrand und dem Martingal $(R^{-m} \cdot X_m)$ als Integrator. Man nennt die obige Gleichung aus einen Martingaldarstellungssatz.

Zur weiteren Auswertung von V_o für Optionen setzen wir nun:

$$\rho_- = e^D,\ \rho_+ = e^U \text{ mit } e^D < R < e^U$$

$$Z_m(\omega) = 0 \text{ oder } 1 \text{ für } i_m = - \text{ bzw. } + \text{ also } Z_m = \mathbf{1}_{\{+\}}(\pi_m).$$

Unter P* sind damit die Z_m, $1 \leq m \leq n$, iid mit $Z_m \sim b(1,p_+)$, also $\sum_{m=1}^n Z_m \sim b(n,p_+)$ und es gilt weiter:

$$X_m = x_0 \cdot \exp\{\textstyle\sum_{k=1}^m [(U-D) \cdot Z_k + D]\}$$

Nach 19.15 gilt:

$$R^n \cdot V_0 = \sum_{\omega=(i_1,\dots,i_n)\in\Omega} p_{i_1} \cdot p_{i_2} \cdot \dots \cdot p_{i_n} [\rho_{i_1} \cdot \rho_{i_2} \cdot \dots \cdot \rho_{i_n} \cdot x_0 - K]^+$$

$$= \sum_{\omega\in\Gamma} p_{i_1} \cdot p_{i_2} \cdot \dots \cdot p_{i_n} [\rho_{i_1} \cdot \rho_{i_2} \cdot \dots \cdot \rho_{i_n} \cdot x_0 - K] \quad \text{mit}$$

$$\Gamma := \{\omega=(i_1,\dots,i_n) \in \Omega;\ \rho_{i_1} \cdot \rho_{i_2} \cdot \dots \cdot \rho_{i_n} \cdot x_0 > K\} = \{\omega \in \Omega;\ \textstyle\sum_{m=1}^n \log(\rho_{i_m}) > \log(K/x_0)\}.$$

Wegen $\log(\rho_{\pi_m}) = (U-D) \cdot Z_m + D$ gilt mit $a := [\log(K/x_0) - nD]/(U-D)$:

(19.17) $\quad \Gamma = \{ (U-D) \cdot \sum_{m=1}^n Z_m + nD > \log(K/x_0) \} = \{\sum_{m=1}^n Z_m > a\}$

Sei P' ein weiteres W–Maß auf Ω mit $P' = \bigotimes_{k=1}^n Q'$, wobei $Q'[\{\pm\}] = p'_\pm := p_\pm \cdot \rho_\pm / R$. Dabei ist Q' wegen (19.14) ein W–Maß. Dann ergibt sich für V_0 die Darstellung:

(19.18) $\quad V_0 = x_0 \cdot P'[\Gamma] - K \cdot R^{-n} \cdot P^*[\Gamma]$

$$= x_0 \cdot \sum_{k>a} \binom{n}{k} p'^{\,k}_+ \, p'^{\,n-k}_- \; - \; K \cdot R^{-n} \cdot \sum_{k>a} \binom{n}{k} p_+^k \, p_-^{n-k} \quad \text{mit}$$

$$a := [\log(K/x_0) - nD]/(U-D) \text{ und } p'_\pm := p_\pm \cdot \rho_\pm / R.$$

Dies ist die **Bewertungsformel für Optionen im sog. Binomialmodell.**

Konvergenz gegen ein zeitstetiges Modell.

Nun sollen die Periodenlängen immer kleiner und n immer größer gewählt werden, sodaß dabei der Fälligkeitstermin $0<T<\infty$ festgehalten bleibt. [Im Zeitalter der Computermärkte wirden eine Periodenlänge tatsächlich immer kleiner; beim DAX ist sie eine Minute.] Mit kleinerer Periodenlänge müssen auch kleinere Werte für die Sprunghöhen U und D angesetzt werden.

Wählt man als Periodenlänge T/n, so wird sich $U = \sigma\sqrt{T/n} = -D$ als ein geeigneter Ansatz erweisen, also $\rho_\pm = \exp\{\pm\sigma\sqrt{T/n}\} = e^{\pm U}$.

Dabei ist σ ein Maß für die Fluktuationen des Kurses. Je kleiner die Periodenlänge ist, desto weniger gravierend ist übrigens die Beschränkung für Z_n und damit X_n auf zwei Werte.

Ferner muß die Verzinsung für das Intervall $[0,T]$ konstant bleiben, also etwa folgendes gelten:
$R^n =: e^{rT}$ für ein $0<r<T$, d.h. $R = e^{rT/n}$.
Sei $\Gamma =: \{W_n > \log(K/x_0)\}$, also $W_n = \Sigma_{m=1}^{n} [(U-D)\cdot Z_m + D]$.

Unter P^* hat W_n die CF $\varphi_n^*(s) := (p_+ \cdot e^{isU} + p_- \cdot e^{-isU})^n$, und entsprechend gilt:

unter P' hat W_n die CF $\varphi_n'(s) := (p_+' \cdot e^{isU} + p_-' \cdot e^{-isU})^n$ [vgl. 15.9].

Dabei gilt mit $R = \exp\{rT/n\}$:
$p_+ = (R - \rho_-)/(\rho_+ - \rho_-) = (R - \exp\{-\sigma\sqrt{T/n}\})/(\exp\{\sigma\sqrt{T/n}\} - \exp\{-\sigma\sqrt{T/n}\}) =$

$$\frac{1 + rT/n + o(1/n) - (1 - \sigma\sqrt{T/n} + \tfrac{1}{2}\sigma^2 T/n + o(1/n))}{2\sigma\sqrt{T/n}\cdot(1 + o(1/\sqrt{n}))},$$ also mit $(1 + o(\varepsilon))^{-1} = 1 + o(\varepsilon)$:

$p_+ = \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{\sigma}\cdot\alpha_-\cdot\sqrt{T/n} + o(1/\sqrt{n})$ mit $\alpha_\pm := r \pm \tfrac{1}{2}\sigma^2$.

Der vorliegende Fall ist als komplizierter als beim zentralen Grenwertsatz; denn hier hängt p_+ und damit Q* auch noch von n ab. Es folgt:

$\varphi_n^*(s) = (p_+ \cdot [1 + isU - \tfrac{1}{2}s^2U^2 + o(1/n)] + p_- \cdot [1 - isU - \tfrac{1}{2} s^2U^2 + o(1/n)])^n$

$= (1 + isU\cdot(p_+-p_-) - \tfrac{1}{2}s^2U^2 + o(1/n))^n = (1 + is\alpha_-\cdot T/n - \tfrac{1}{2}s^2\sigma^2\cdot T/n + o(1/n))^n$

$= (1 + (is\alpha_-\cdot T - \tfrac{1}{2}s^2\sigma^2\cdot T)/n + o(1/n))^n \to \exp\{is\alpha_-\cdot T - \tfrac{1}{2}s^2\sigma^2\cdot T\}$.

Dies ist die CF zu $N(\alpha_-\cdot T, \sigma^2\cdot T)$, diese hat die Vf $x \mapsto \Phi((x - \alpha_-\cdot T)/\sigma\sqrt{T})$, wenn Φ die Vf zu N(0,1) ist (vgl. 12.15b). Dabei gilt: $1-\Phi(x)=\Phi(-x)$ gemäß 12.10.

Nach dem Stetigkeitssatz 18.6 und 17.10 folgt:
$P^*[\Gamma] \to 1 - \Phi([\log(K/x_0) - \alpha_-\cdot T]/\sigma\sqrt{T})) = \Phi([\log(x_0/K) + \alpha_-\cdot T]/\sigma\sqrt{T})$.

Ebenso zeigt man: $P'[\Gamma] \to \Phi([\log(x_0/K) + \alpha_+\cdot T]/\sigma\sqrt{T})$.

Insgesamt ergibt sich also aus 19.18 für $n\to\infty$:

(19.19) $V_0 \to x_0\cdot\Phi([\log(x_0/K) + \alpha_+\cdot T]/\sigma\sqrt{T}) - K\cdot e^{-rT}\cdot\Phi([\log(x_0/K) + \alpha_-\cdot T]/\sigma\sqrt{T})$.

Dies ist die **Black–Scholes–Formel** für die Bewertung einer Option in einem zeitstetigen Modell. Dabei ist $\alpha_\pm := r \pm \tfrac{1}{2}\sigma^2$,

r die Zinsrate im zeitstetigen Modell,
σ^2 die sog. Volatilität als Maß für die Kursfluktationen,
x_0 der Kurs z.Zt. 0,
T der Fälligkeitskeitstermin der Option,
K der Wahrnehmungspreis der Option.