Maß- und Wahrscheinlichkeitstheorie Skript Felsenstein W2022

Ida Hönigmann

December 22, 2022

1 Produkträume und Maße

Auf dem kartesischen Produkt von Grundmengen

$$\Omega = \bigtimes_{i=1}^{n} \Omega_i$$

wird eine Produkt-(Sigma)Algebra konstruiert, wobei $(\Omega_i, \mathcal{A}_i)$ Messräume sind. Es soll $(\Omega_i, \mathcal{A}_i)$ in $(\Omega, \mathcal{A})$ eingebettet werden.

Def (Projektion).

$$\pi_i : \Omega \mapsto \Omega_i \ mit \ \pi_i(\omega_1, ..., \omega_n) = \omega_i$$

Die Produktalgebra wird von den Projektionen erzeugt.

Definition 1 (Produktalgebra).

$$\mathcal{A} := \bigotimes_{i=1}^{n} \mathcal{A}_i := \sigma(\pi_1, ..., \pi_n)$$

$$d.h.\ \pi_1^{-1}(A_1) \cap \pi_2^{-1}(A_2) \cap ... \cap \pi_n^{-1}(A_n) \ für \ A_i \in \mathcal{A}_i$$

erzeugen diese Algebra (bzw. $A_1 \times A_2 \times ... \times A_n$).

Wenn $\mathcal{A}_i = \sigma(\mathcal{E}_i)$ erzeugt wird von $\mathcal{E}_i$, wird $\bigotimes \mathcal{A}_i$ von $\bigtimes \mathcal{E}_i$ erzeugt.

Satz 1 (1.PR). *$\mathcal{A}_i = \sigma(\mathcal{E}_i)$ mit $\mathcal{E}_i \subset 2^{\Omega_i}$. Ω_i sei aus $\mathcal{E}_i$ monoton erreichbar ($E_{i,k} \nearrow \Omega_i$)[a].*

$$\mathcal{E} = \{\bigtimes_{i=1}^{n} E_i, E_i \in \mathcal{E}_i\}$$

Dann gilt

$$\bigotimes_{i=1}^{n} \mathcal{A}_i = \sigma(\mathcal{E}).$$

[a] Man kann fordern, dass $\Omega_i \in \mathcal{E}_i$, dann ist es erfüllt.

Beweis Satz 1.PR. $\sigma(\mathcal{E})$ ist die kleinste Sigma-Algebra, sodass die Projektionen messbar sind: $\tilde{\mathcal{A}}$ sei Sigma-Algebra auf $\Omega = \bigtimes_i \Omega_i$.

π_i ist $\tilde{\mathcal{A}} - \mathcal{A}_i$ messbar $\iff \sigma(\mathcal{E}) \subseteq \tilde{\mathcal{A}}$.

$\implies$: Alle π_i seien $\tilde{\mathcal{A}} - \mathcal{A}_i$ messbar, d.h. $\pi_i^{-1}(\mathcal{E}_i) \subseteq \tilde{\mathcal{A}}$, wobei für $E_i \in \mathcal{E}_i$

$$\pi_i^{-1}(E_i) = \Omega_1 \times \Omega_2 \times ... \times E_i \times \Omega_{i+1} \times ... \times \Omega_n$$

Da

$$\underbrace{\bigtimes_{i=1}^{n} E_i}_{\in \mathcal{E}} = \bigcap_{i=1}^{n} \pi_i^{-1}(E_i) \subseteq \tilde{\mathcal{A}} \implies \sigma(\mathcal{E}) \subseteq \tilde{\mathcal{A}}.$$

$\Longleftarrow$: Es gelte $\sigma(\mathcal{E}) \subseteq \tilde{\mathcal{A}}$. Jedes Ω_i ist aus $\mathcal{E}_i$ monoton erreichbar mit $E_{i,k} \nearrow \Omega_i, k \to \infty$.

$$F_k = E_{1,k} \times E_{2,k} \times ... \times E_i \times ... \times E_{n,k} \nearrow \pi_i^{-1}(E_i)$$
$$\lim F_k = \bigcup_{k=1}^{\infty} F_k = \pi_i^{-1}(E_i) \in \sigma(\mathcal{E}) \subset \tilde{\mathcal{A}}$$

Urbild vom Erzeuger in $\tilde{\mathcal{A}}$, $\pi_i^{-1}(\mathcal{E}_i) \subseteq \tilde{\mathcal{A}} \forall i$, alle π_i sind $\tilde{\mathcal{A}} - \mathcal{A}_i$ messbar.

Alle Projektionen $\sigma(\mathcal{E})$ messbar, also $\sigma(\pi_1, ..., \pi_n) \subset \sigma(\mathcal{E})$. Nach obigem setze $\tilde{\mathcal{A}} = \sigma(\pi_1, ..., \pi_n) \implies \sigma(\mathcal{E}) \subseteq \sigma(\pi_1, ..., \pi_n)$ also $\sigma(\mathcal{E}) = \sigma(\pi_1, ..., \pi_n)$. □

Beberkung. *$\{A_1 \times ... \times A_n | A_i \in \mathcal{A}_i\}$ ist keine Sigma-Algebra, da nicht Vereinigungs-stabil.*

Die komponentenweise Behandlung im Produktraum ist anwendbar auf n-dim. Funktionen.

Folgerung.

$$f_i : \Omega_0 \mapsto \Omega_i, f = (f_1, ..., f_n) =: \otimes_i f_i$$
$$f : \Omega_0 \mapsto \bigtimes_i \Omega_i = \Omega$$

Dann gilt:

$$f \textit{ ist } (\Omega_0, \mathcal{A}_0) \mapsto \underbrace{\bigotimes_{i=1}^{n} (\Omega_i, \mathcal{A}_i)}_{\textit{Produktraum}} \textit{ messbar} \iff f_i \mathcal{A}_0 - \mathcal{A}_i \textit{ messbar}$$

Beweis Folgerung. $\implies$: Da $f_i = \pi_i \circ f$ und $\pi_i \otimes_j \mathcal{A}_j - \mathcal{A}_i$ messbar ist f_i als Verkettung messbar.

$\Longleftarrow$: $A \in \otimes \mathcal{A}_i$ Menge aus der Produktalgebra mit $A = \times_{i=1}^n A_i \leftarrow$ erzeugen $\otimes_{i=1}^n \mathcal{A}_i$.

$$f^{-1}(A) = \bigcap_{i=1}^{n} f_i^{-1}(A_i) \in \mathcal{A}_0$$

Diese Rechtecke erzeugen $\sigma(f) = \sigma(f_1, ..., f_n) \subseteq \mathcal{A}_0$ also ist $f \mathcal{A}_0 - \otimes \mathcal{A}_i$ messbar. □

Anwendung auf Borel-Algebra: $\mathcal{B}^k = \otimes_{i=1}^k \mathcal{B}$ erzeugt von den Rechtecken $\times_i (a_i, b_i]$. Die Lebesgue-Mengen werden nicht von den Produkten erzeugt: $\otimes_{i=1}^k \mathcal{L} \subsetneq \mathcal{L}_k$.

Nicht alle Nullmengen von $\mathcal{B}^k$ sind durch die Produkte mit allen Nullmengen erzeugbar.

Def (Schnitt). *Besondere Mengen (für Integralberechnungen) sind die Schnitte:*

$$A \in \mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$$
$$A_{x_1} := \{x_2 \in \Omega_2 | (x_1, x_2) \in A\}$$
$$A_{x_2} := \{x_1 \in \Omega_1 | (x_1, x_2) \in A\}$$

Die Schnitte sind messbar.

Satz 2 (2.PR). *$A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ messbar bez. Produktalgebra, dann ist $A_{x_1} \in \mathcal{A}_2$ und $A_{x_2} \in \mathcal{A}_1$ für alle x_1 bzw. x_2.*

Proof. x_1 sei fest. Betrachte $\mathcal{M} = \{A \subseteq \Omega_1 \times \Omega_2 | A_{x_1} \in \mathcal{A}_2\}$. $\mathcal{M}$ ist Sigma-Algebra: $\Omega \in \mathcal{M}$, $(A_{x_1})^c = (A^c)_{x_1}$ und $\cup A_{x_i,1} = (\cup A_i)_{x_1}$.

Für die Erzeuger Mengen $A_1 \times A_2 \in \mathcal{E}$ mit $A_i \in \mathcal{A}_i$

$$(A_1 \times A_2)_{x_1} = \begin{cases} A_2, & x_1 \in A_1 \\ \emptyset, & x_1 \notin A_1 \end{cases}$$

also $(A_1 \times A_2)_{x_1} \in \mathcal{A}_2$, $\mathcal{E} \subseteq \mathcal{M}$ und $\sigma(\mathcal{E}) = \mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq \mathcal{M}$, alle Mengen, alle x_1 erfüllen die Messbarkeit-Bedingungen. □

Die Abbildungen $x_1 \mapsto \mu_2(A_{x_1})$ sind messbare Abbildungen.

Satz 3 (3.PR). *$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$, auf $\mathcal{A}_2$ sei μ_2 ein sigma-endliches Maß auf $(\Omega_2, \mathcal{A}_2, \mu_2)$. Dann ist $x_1 \mapsto \mu_2(A_{x_1})$ eine $\mathcal{A}_1$ messbare Abbildung (entsprechendes gilt auch für $x_2 \mapsto \mu_1(A_{x_2})$).*

Proof. Für ein $A \in \mathcal{A}$ sei $f_A(x_1) = \mu_2(Ax_1)$

1. μ_2 sei endlich. Betrachte $\mathcal{D} = \{E \in \mathcal{A} | f_E ist \mathcal{A}_1$ Borel-messbar$(\mathbb{R}^+, \mathcal{B}^+)\}$

   $\mathcal{D}$ ist ein Dynkin-System:

   - $\Omega \in \mathcal{D}$, da $\Omega_{x_1} = \Omega_2 \in \mathcal{A}_2 \forall x_1 : f_\Omega \equiv \mu_2(\Omega_2)$ konstant
   - $A \subseteq \mathcal{B}$ und f_A, f_B messbar
     $A_{x_1} \subseteq B_{x_1}$ und $\mu_2(B_{x_1} \setminus A_{x_1}) = \mu_2(B_{x_1}) - \mu_2(A_{x_1}) = f_B(x_1) - f_A(x_1)$ ist messbar als Differenz messbarer Funktionen.
   - $A_i \in \mathcal{A}$ und disjunkt, $B = \bigcup_i A_i$
     Wenn $A_i \in \mathcal{D}$, $B_{x_1} = \bigcup A_{i,x_1}$

     $$\mu_2(B_{x_1}) = f_B(x_1) = \sum_i f_{A_i}(x_1) = \mu_2(\bigcup A_{x_1})$$

     Eine abzählbare Summe messbarer Funktionen ist messbar (Darstellung jeder messbarer Funk $f = \sum_j c_j 1_{c_j}$).

   Jede Menge aus $\mathcal{E} = \{A_1 \times A_2 | A_i \in \mathcal{A}_\rangle\}$ (dem Erzeuger von $\mathcal{A}$) ist auch in $\mathcal{D}$, weil $f_{A_1 \times A_2}(x_1) = \mu_2((A_1 \times A_2)_{x_1}) = \mu_2(A_2)1_{A_1}(x_1)$. $c1$ ist messbar.

   $\mathcal{E}$ ist ein durchschnitt-stabiler Erzeuger und $\mathcal{E} \subseteq \mathcal{D} \subseteq \mathcal{A} = \sigma(\mathcal{E})$ also $\mathcal{A} = \mathcal{D}$, alle solchen Funktionen sind messbar bez. $\mathcal{A}_\infty$.

2. Ist μ_2 sigma-endlich, es gibt eine Folge $A_{2,i} \nearrow \Omega_2$ mit endlichem Maß, daher auch eine disjunkte Folge (D_n), die eine Zerlegung von Ω_2 sind und $\mu_2(B) = \mu_2(\bigcup_n (D_n \cap B)) = \sum_n \mu_2(B \cap D_n)$.

   $$f_B(x_1) = \mu_2(B_{x_1}) = \sum_n \underbrace{\mu_2(B_{x_1} \cap D_n)}_{<\infty} = \sum_n f_{(B_{x_1} \cap D_n)}(x_1)$$

   also Summe messbarer Funktionen.

□

Mit diesen Funktionen wird das Produktmaß erklärt.

Def (Produktmaß)**.**

$$\mu\left(\bigtimes_{i=1}^k A_i\right) = \prod_{i=1}^k \mu_i(A_i)$$

Wenn $\Omega = \Omega_1^k$ mit $\mu_i = \mu$ gilt

$$\mu\left(\bigtimes_{i=1}^k A_i\right) = \prod_{i=1}^k \mu(A_i).$$

Dieses Maß ist eindeutig definiert, wenn μ_1, μ_2 sigma-endlich sind, dann sind die Funktionen $f_B(x_1)$ bzw. $f_B(x_2)$ die "Dichten" bezüglich den Randmaßen μ_1 bzw. μ_2. μ_1, μ_2 werden auch als marginale Maße bezeichnet.

Satz 4 (4.PR)**.** *1. μ_2 sei sigma-endlich. Dann definiert*

$$\mu(A) = \int_{\Omega_1} \mu_2(A_{x_1}) d\mu_1(x_1)$$

das Produktmaß auf $(\Omega, \mathcal{A})$.

2. Sind beide Maße μ_1,μ_2 sigma-endlich, dann ist μ eindeutig und

$$\mu(A) = \int_{\Omega_1} \mu_2(A_{x_1}) d\mu_1(x_1) = \int_{\Omega_2} \mu_1(A_{x_2}) d\mu_2(x_2)$$

Proof. $f_A(x_1) = \mu_2(A_{x_1})$ ist eine $\mathcal{A}_1$-messbare Funktion $\forall A$ und durch die Additivität $f_{\cup A_i} = \sum f_{A_i}$ ist μ ein Maß auf $(\Omega, \mathcal{A})$.

Für $A = A_1 \times A_2$ gilt

$$\mu(A_1 \times A_2) = \int_{\Omega_1} \underbrace{\mu_2((A_1 \times A_2)_{x_1})}_{\mu_2(A_2)1_{A_1}(x_1)} d\mu_1 = \mu_2(A_2) \underbrace{\int_{A_1} d\mu_1}_{\mu_1(A_1)}$$

ist Produktmaß.

μ ist auf dem (durchschnitts-stabilen) Erzeuger definiert, da die Maße sigma-endlich sind, ist μ eindeutig (Eindeutigkeitssatz). Vice versa gelten alle Gleichungen analog bez. μ_2. □

Beispiel. *Insbesonders bei endlichem Maß anwendbar. X,Y stochastische Größen, dann wird eindeutig eine zweidim. Verteilung auf* $\mathbb{R}^2$ *durch* $P[(X, Y) \in A \times B] := P[X \in A]P[Y \in B]$ *Produktverteilung.*

Verallgemeinerung der Maße von Schnitten ist die Schnittfunktion.

Definition 2 (Schnittfunktion)**.**

$$\begin{aligned} f : \Omega_1 \times \Omega_2 \mapsto \Omega' &\ \textit{Dann heißt} \\ x_2 \mapsto f_{x_1}(x_2) = f(x_1, x_2) &\ x_1\textit{-Schnitt} \\ x_1 \mapsto f_{x_2}(x_1) = f(x_1, x_2) &\ x_2\textit{-Schnitt} \end{aligned}$$

von f. (messbare Funktion $\Omega_1 \times \Omega_2 \to (\Omega', \mathcal{A}')$*).*

Satz 5 (5.PR)**.** *f sei messbar* $(\Omega, \mathcal{A}) \mapsto (\Omega', \mathcal{A}')$*. Die Schnittfunktionen sind messbar:*

$$\begin{aligned} f_{x_1} &\ \textit{ist messbar } (\Omega_2, \mathcal{A}_2) \mapsto (\Omega', \mathcal{A}') \\ f_{x_2} &\ \textit{ist messbar } (\Omega_1, \mathcal{A}_1) \mapsto (\Omega', \mathcal{A}') \end{aligned}$$

Proof. $A' \in \mathcal{A}'$

$$f_{x_1}^{-1}(A') = \{x_2 \in \Omega_2 | f(x_1, x_2) \in A'\} = \{x_2 \in \Omega_2 | (x_1, x_2) \in f^{-1}(A')\} = (\underbrace{f^{-1}(A')}_{\text{mb } \in \mathcal{A}})_{x_1}$$

jede Schnittmenge ist messbar.

Analog für $f_{x_2}^{-1}(A') \in \mathcal{A}_1$. □

Mit den Funktionenschnitten lässt sich auch ein mehrdim. Integral "zerteilen".

Satz 6 (6.PR Satz von Fubini (-Tonelli))**.** *Produktraum* $(\Omega, \mathcal{A}, \mu) = \otimes_i(\Omega_i, \mathcal{A}_i, \mu_i)$*, f messbar* $\Omega \mapsto \mathbb{R}$*,* μ_i *sigma-endlich. Das zweidimensionale Integral von f ist aufspaltbar*

$$\int f d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f_{x_1}(x_2) d\mu_2 \right) d\mu_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f_{x_2}(x_1) d\mu_1 \right) d\mu_2$$

wenn eine der folgenden Bedingungen gilt:

1. $f \geq 0$*: Dann ist* $x_2 \mapsto \phi_2(x_2) = \int_{\Omega_1} f_{x_2} d\mu_1$ *messbar* $\mathcal{A}_2$ *und* $x_1 \mapsto \phi_1(x_1) = \int_{\Omega_2} f_{x_1} d\mu_2$ *messbar* $\mathcal{A}_1$
2. *f ist integrierbar,* $\int f d\mu < \infty$*: Dann sind* f_{x_1} μ_2*-integrierbar* $[\mu_1]$ *f.ü. und* f_{x_2} μ_1*-integrierbar* $[\mu_2]$ *f.ü.*
3. $\int_{\Omega_1} \int_{\Omega_2} |f_{x_1}| d\mu_2 d\mu_1 < \infty$ *oder* $\int_{\Omega_2} \int_{\Omega_1} |f_{x_2}| d\mu_1 d\mu_2 < \infty$*: Daraus folgt f integrierbar.*

Proof. 1. $f \geq 0 :$ 4 Schritte des Integralaufbaus

$f = 1_A, A \in \mathcal{A}$

$$x_2 \mapsto \phi_2(x_2) = \int_{\Omega_1} \underbrace{(1_A)_{x_2}}_{1_{A_{x_2}}} d\mu_1 = \mu_1(A_{x_2})$$

ϕ_2 ist messbar (laut Satz PR5) und nach Satz PR7 gilt

$$\int f d\mu = \mu(A) = \int_{\Omega_2} \mu_1(A_{x_2}) d\mu_2(x_2) = \int_{\Omega_2} \phi_2 d\mu_2$$

Für einfache Funktionen $f = \sum_{i=1}^n \alpha_i 1_{A_i}$ ergibt sich das aus der Linearität:

$$f_{x_2} = \sum \alpha_i 1_{(A_i)_{x_2}}(x_1)$$

Schritt 3: $f_2 \nearrow f, f_n...$ einfache Funktionen, dann gilt $(f_n)_{x_2} \nearrow f_{x_2}$ (wieder aus der Darstellung $f = \sum_{i=1}^\infty \alpha_i 1_{A_i}$ + monotone Konvergenz ablesbar)

2. f integrierbar, betrachte Positiv- und Negativteil

$$max(0, f)_{x_2} = max(0, f_{x_2}) = (f^+)_{x_2}$$
$$\text{also } (f^+)_{x_2} = (f_{x_2})^+ \text{ und } f^-_{x_2} = (f^-)_{x_2} \text{ mit } |f|_{x_2} = |f_{x_2}|$$

Wegen $\int |f| d\mu < \infty$ gilt wegen oben für $|f|$:

$$\int |f| d\mu = \int_{\Omega_1} \int_{\Omega_2} |f|_{x_1} d\mu_2 d\mu_1 = \int_{\Omega_1} \underbrace{\int_{\Omega_2} |f_{x_1}| d\mu_2}_{\phi_1} d\mu_1 < \infty$$

Daher muss ϕ_1 integrierbar sein (wie auch ϕ_2). Integrierbarkeit erfordert auch $(f^+), (f^-)$ integrierbar sind, aufgespalten $f = f^+ - f^-$ und $f_{x_i} = f^+_{x_i} - f^-_{x_i}$. Nach 1. für f^+, f^- getrennt ergibt

$$\int f d\mu = \int f^+ - \int f^- = \int_{\Omega_1} \int_{\Omega_2} f^+_{x_1} d\mu_2 d\mu_1 - \int \int f^- d\mu_2 d\mu_1$$

3. impliziert $\int_{\Omega_2} f^+_{x_1} d\mu_2 < \infty, \int_{\Omega_2} f^-_{x_1} d\mu_2 < \infty$ und somit f integrierbar und 2.

□

Durch Iteration gilt die Fubini-Schnitt Konstruktion auch für mehrdimensionale $k \in \mathbb{N}$ Integrale: Wenn $f : \Omega \mapsto \mathbb{R}$ messbar mit $\Omega = \Omega_1 \times ... \times \Omega_n$

$$\int f d\mu = \int_{\Omega_1} \left(\int_{\Omega_2 \times ... \times \Omega_n} f_{x_1} d\mu_2 \otimes ... \otimes \mu_k \right) d\mu_1(x_1)$$

Viele Folgerungen: Doppelreihen-Satz $\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$ Kriterien für Konvergenz

Folgerung (Maße mit Dichten bezüglich dem Produktmaß)**.** $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$, $\mu = \mu_1 \otimes \mu_2$
ν sei absolut stetig bzgl. μ ($\nu \ll \mu$), es existiert ein $f = \frac{d\nu}{d\mu} \geq 0$ mit $\nu(A) = \int_A f d\mu$ (f integrierbar)
μ sei sigma-endlich, ν erzeugt ein $\nu_1 \ll \mu_1$ auf $(\Omega_1, \mathcal{A}_1)$ und ein $\nu_2 \ll \mu_2$ auf $(\Omega_2, \mathcal{A}_2)$ mit den Dichten ϕ_1, ϕ_2.

$$A \in \mathcal{A}_1 : \nu_1(A_1) = \int_{A_1} \phi_1(x_1) d\mu_1 = \int_{A_1} \int_{\Omega_2} f_{x_1}(x_2) d\mu_2 d\mu_1$$

Beispiel. *2 dim SG mit Gleichverteilung auf dem Einheitskreis $P[(X,Y) \in K] = 1$. Dichte f bezgl. $\lambda^2 = \lambda_1 \otimes \lambda_1$. $f(x,y) = \frac{1}{\pi} 1_K$.*

$$\phi_1(x) = \int f_X(y) d\lambda = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} d\lambda(y) = \frac{2\sqrt{1-x^2}}{\pi}$$

(Symmetrie: $\phi_2(y) = \frac{2\sqrt{1-y^2}}{\pi}$)
Randverteilung ν_1 $P[X \in A] = \int_A \frac{2}{\pi} \sqrt{1-y^2} d\lambda(y)$
ν ist nicht das Produktmaß $\nu \neq \nu_1 \otimes \nu_2$ außer $f(x,y) = f_1(x) f_2(y)$.

Def (Ordinatenmenge, Graph)**.** *Die Punkte unter einer positiven Funktion $f \geq 0$ heißt Ordinatenmenge $O_f = \{(x,y) | 0 \leq < \leq f(x)\}$ und der Graph ist $\Gamma_f = \{(x, f(x)) | x \in \Omega\}$.*

Für sigma-endliches Maß μ und f messbar, dann ist O_f eine bezüglich $\mathcal{A} \otimes \mathcal{B}$ messbare Menge und $\int f d\mu = (\mu \otimes \lambda)(O_f)$. Der Graph ist eine Nullmenge, $(\mu \otimes \lambda)(\Gamma_f) = 0$

1.1 ∞-dim. Produkträume

Die Konstruktion der ∞-dim. Produkt-Messräume ist der endl. dim. Konstruktion entsprechend. I sei Index-Menge,

$$\otimes_{i\in I}\mathcal{A}_i := \sigma(\pi_i, i \in I) \qquad \pi_i\text{... Projektion auf } \Omega_i$$

beispielsweise $\otimes_{i\in I}\mathcal{B}$ auf $\Omega = \mathbb{R}^I$ (Funktionenraum).

Def (Zylinder,Pfeiler)**.** *$Z \subseteq \Omega^I$ heißt Zylinder, wenn $Z = \pi_j^{-1}(C) = C \times \Omega^{j^C}$ wobei $J \subseteq I$ eine endliche Teilindexmenge ist und C die endlich dim. Basis des Zylinders ist; und Pfeiler, wenn C ein Rechteck $\times_{i\in J}A_i$ ist.*

Auf den Pfeilern lässt sich das n-dim. Produktmaß erklären. $P^n(C) = \pi_{j\in J}P(A_j)$

Im abzählbar unendlichen Fall ist die Vorgangsweise ähnlich, wie im endl. dim. Fall:

Satz 7 (7.PR)**.** *X_i sei eine Folge von SGn auf $(\Omega_i, \mathcal{A}_i)$ mit gegebener gemeinsamer Verteilung*

$$P_i(A) = P[X_j \in A_j, j \leq i, X_i \in A]$$

für $A_j \in \mathcal{A}_j$ und $A \in \mathcal{A}_i$. D.h. die gemeinsame Verteilung von $(X_1, ..., X_n)$ sind auf $(\times_i\Omega_i, \otimes_i\mathcal{A}_i)$ mit

$$P[(X_1, ..., X_n) \in A_n, X_i \in \mathbb{R}, i > n] = P[(X_1, ..., X_n) \in A_n]$$

also für Zylinder $Z = \pi_{1,...,n}^{-1}(C_n)$, $C_n \in \otimes_i\mathcal{A}_i$ gilt $P[Z] = P_n(Cn)$.
Wenn diese Wahrscheinlichkeitsverteilung verträglich sind, d.h. die Randverteilungen (bzw. alle Teilmengen endl.) eindeutig sind, lässt sich obiges Prinzip verallgemeinern.

Satz 8 (8.PR Satz von Kolmogoroff)**.** *$(\Omega_i, \mathcal{A}_i), i \in I$ sei eine Familie von Messräumen und P_j sind endl. dim. Verteilungen auf $(\mathbb{R}^J, \mathcal{B}_{|J|}), J \subset I, J$ endlich. Diese Verteilungen seien verträglich ($P_j = P_k\pi_{k,j}^{-1}, J \subset K, K$ ebenfalls endlich).*
Dann existiert ein eindeutiges Maß P auf $(\mathbb{R}^I, \mathcal{B}_I = \otimes_{i\in I}\mathcal{B})$ mit genau diesen Randverteilungen ($P(A) = P_J(\pi_j^{-1}(A)), A \in \otimes_{i\in J}\mathcal{A}_i$).

Das Produktmaß $\mu_1\otimes\mu_2$ auf $\mathcal{A}_1\otimes\mathcal{A}_2$ muss nicht vollständig sein, wenn $(\Omega_i, \mathcal{A}_i, \mu_i), i = 1, 2$ vollständige Maßräume sind.

Beispiel. *$(\mathbb{R}^2, \mathcal{L}_2, \lambda_2)$ ist ein vollständiger Maßraum. $\lambda_2(\mathbb{R} \times \{1\}) = 0$ und für beliebiges $A \subseteq \mathbb{R}\lambda_2(A \times \{1\}) = 0$. Wenn $A \notin \mathcal{L}$, dann sollte aber trotzdem der Schnitt von $A \times \{1\}$, $(A \times \{1\})_1 = A \in \mathcal{L}$ messbar sein, das ist ein Widerspruch. Es folgt also $\mathcal{L} \otimes \mathcal{L} \neq \mathcal{L}_2$.*

Satz 9 (9.PR)**.** *$(\Omega_i, \mathcal{A}_i, \mu_i), i = 1, 2$ seinen sigma-endliche Maßräume. Wenn die messbaren Funktionen $f_i : (\Omega_i, \mathcal{A} : i) \to (\mathbb{R}, \mathcal{B})$ entweder*

- *$f_i \geq 0$ oder*
- *f_i ist integrierbar, $i = 1, 2$*

ist, dann gilt

$$\int f_1 \cdot f_2 d\mu_1 \otimes d\mu_2 = \int f_1 d\mu_1 \cdot \int f_2 d\mu_2$$

Im 2. Fall ist $f_1 \cdot f_2$ auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$ integrierbar.

Proof. Eigentlich klar, da $(f_1 \cdot f_2)_{x_1} = f_1(x_1) \cdot f_2(.)$ $\mathcal{A}_2$ messbar ist und nach Fubini

$$\int f_1 \cdot f_2 d\mu_1 \otimes d\mu_2 = \int_{\Omega_1}\int_{\Omega_2} f_1(x_1)f_2(x_2)d\mu_2(x_2)d\mu_1 = \int_{\Omega_1}\left(f_1(x_1)\int_{\Omega_2} f_2(x_2)d\mu_2(x_2)\right) = \int f_1 d\mu_1 \cdot \int f_2 d\mu_2$$

Genauso gilt $\int |f_1 f_2| d\mu_1 \otimes d\mu_2 = \int |f_1| d\mu_1 \int |f_2| d\mu_2$ und wenn die rechte Seite endlich ist, gilt $f_1 \cdot f_2 \in \mathcal{L}_1$ bezüglich dem Produktraum. $\square$

Die Betrachtung unabhängiger SGn $X_i, i = 1, ..., k$ erfolgt bequemer auf dem Produktraum. Auch wenn alle $\Omega_i = \Omega$, also alle SGn auf dem selben Raum definiert sind, übersiedelt man für die Erklärung der gemeinsamen Verteilung auf den Produktraum.

Wenn $X = (X_1, ..., X_k)$ ein Vektor unabhängiger SGn X_i ist, gilt

$$PX^{-1} = PX_1^{-1} \otimes ... \otimes PX_k^{-1}.$$

Wenn PX^{-1} ein Wahrscheinlichkeitsmaß mit Dichte ist, dann gilt

$$PX^{-1}(A) = \int_A f_X d\lambda_k = \int_A f(x_1, ..., x_k) d\lambda(x_1)...d\lambda(x_k)$$

und die Randverteilung von X_i ist mit Dichte

$$PX_i^{-1}(A_i) = \int_{A_i \times \mathbb{R}^{k-1}} f_X d\lambda_k = \int_{A_i} \left(\int_{\mathbb{R}} ... \int_{\mathbb{R}} f_X d\lambda ... d\lambda \right) d\lambda$$

und wieder erhält man die Randdichte

$$f_i(x_i) = \int_{\mathbb{R}} ... \int_{\mathbb{R}} f(x_1, ..., x_k) d\lambda(x_1)...d\lambda(x_{i-1}) d\lambda(x_{i+1})...d\lambda(x_k)$$

(Das ist nicht neu, aber jetzt wird kein Umweg über das Riemann-Integral benötigt.)

Wenn die X_i unabhängig sind ist die gemeinsame Dichte

$$f_X(x_1, ..., x_k) = \prod_{i=1}^k f_i(x_i).$$

Nach dem letzten Satz gilt für integrierbare f und g und unabhängige X und Y

$$\mathbb{E} f(X) g(Y) = \mathbb{E} f(X) \cdot \mathbb{E} g(Y)$$

d.h. auch $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$ und die SGn X und Y sind unkorreliert.

Die Umkehrung gilt natürlich nicht, da Unkorreliertheit nur bedeutet, dass es keinen linearen Zusammenhang gibt.

Wenn $Y = c$ f.s., dann sind X,Y unabhängig, wenn $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$, dass dann automatisch gilt. Auch wenn $X \sim A_{p_1}$ (Alternativ-verteilt) und $Y \sim A_{p_2}$ und

$$\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y = 0(1 - p_1 p_2) + 1 p_1 p_2 = p_1 p_2$$

d.h. $P(X = 1)P(Y = 1) = p_1 p_2$ und X und Y sind unabhängig.

Ansonsten ist nur bei Normalverteilung kein Unterschied zwischen Unkorreliertheit und Unabhängigkeit.

Beispiel. $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ *o.B.d.A* $\mu_x = \mu_y = 0$

Die gemainsame Dichte zerfällt (siehe EI24)

$$f(x, y) = \underbrace{\frac{1}{\sqrt{2\pi}\sigma_x} exp\left(-\frac{1}{2} \left(\frac{x}{\sigma_x} \right)^2 \right)}_{f_1(x)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sigma_x} \frac{\sigma_x}{\sigma_y \sqrt{1 - \rho^2}} exp\left(-\frac{1}{2} \left(\frac{\sigma_x^2 (y - m)^2}{(1 - \rho^2)\sigma_y^2} \right)^2 \right)}_{f_2(x,y)}$$

mit $m = \rho x \frac{\sigma_y}{\sigma_x}$ *d.h.*

$$\mathbb{E}XY = \int\int xy f(x, y) d\lambda_2 = \int \underbrace{\int y f_2(x, y) dy}_{m} \, x f_1(x) dx = \int m x f_1(x) dx = \rho \int x^2 f_1(x) dx = \rho \frac{\sigma_y}{\sigma_x} \sigma_x^2 = \rho \sigma_x \sigma_y$$

X,Y sind unkorreliert $\iff$ $\rho = 0$, *dann ist* $f(x, y) = f_1(x) \cdot \Phi(\frac{y}{\sigma_y})$ *(*Φ *Dichte der* $\mathcal{N}(0, 1)$*) und sind X,Y unabhängig.*

Zwei Normalverteilungen können nur linear abhängen. Ansonsten kann sogar eine vollständige Abhängigkeit (nicht linear) bei unkorrelierten SGn vorliegen.

Beispiel. $X \sim U_{-1,1}, Y = X^2$. *Dann gilt* $\mathbb{E}X = 0$ *und* $\mathbb{E}XY = \mathbb{E}X^3 = 0$ *und* X *und* Y *sind unkorreliert.*

Nur wenn für alle integrierbaren f,g $f(X)$ und $g(Y)$ unkorreliert sind, dann sind X und Y unabhängig.

Der Satz von Fubini ist ein wichtiges Werkzeug auch um bekannte Sätze der Integrationstheorie zu verallgemeinern, wie beispielsweise die partielle Integration.

Satz 10 (10.PR)**.** *μ_F und μ_G seinen Lebesgue-Stieltes Maße mit F bzw. G als Verteilungsfunktionen. $G_-(x) = \lim_{\tilde{x} \nearrow x} G(\tilde{x})$ ist der linksseitige Grenzwert von G. Dann gilt*

$$\int_{(a,b]} F d\mu_G + \int_{(a,b]} G_- d\mu_F = F(b)G(b) - F(a)G(a)$$

Proof. in der Übung.

Wenn F und G stetig differenzierbar sind, dann gilt $d\mu_G = G' d\lambda$ und $d\mu_F = F' d\lambda$ und

$$\int_{(a,b]} F \cdot G' d\lambda + \int_{(a,b]} G \cdot F' d\lambda = F(b)G(b) - F(a)G(a)$$

und in der üblichen Schreibweise für Stammfunktionen $\int FG' = FG - \int F'G$ □

Mit dem Hauptsatz der Diff- u. Integrationstheorie lässt sich auch die bedingte Verteilung auf stetige Verteilungen erweitern.

Für diskrete SGn X,Y (beispielsweise auf $\mathbb{N}$ verteilt) ist

$$P[X = k | Y = l] = \frac{P[X = k, Y = l]}{P[Y = l]}, k, l \in \mathbb{N}$$

die Punktwahrscheinlichkeit p_k der bedingten Verteilung $X|Y = l$.

Besitzt X eine stetig differenzierbare VF F mit $F' = f_X$ als Dichte, gilt

$$\lim_{\triangle \to 0} \frac{F(x + \triangle) - F(x)}{\triangle} = f_X(x) \text{ oder}$$
$$\frac{P[X \in [x - \triangle, x]]}{\triangle} \to f_X(x) \text{ für } \triangle \to 0 \text{ oder}$$
$$P[X \in [x - \triangle, x]] \sim f_X(x)\triangle$$

Mit dieser "infidezimalen Wahrscheinlichkeit" als Dichte erhält man, wenn auch Y die Dichte f_Y hat,

$$P[X \in [x - \triangle, x] | Y \in [y - \triangle, y]] = \frac{P[X \in [x - \triangle, x], Y \in [y - \triangle, y]]}{P[Y \in [y - \triangle, y]]} =$$
$$\frac{\int_{x-\triangle}^{x} \int_{y-\triangle}^{y} f_{X,Y}(s,t) ds dt}{\int_{[y-\triangle,y]} f_Y(t) d\lambda(t)} \sim \frac{f_{X,Y}(x,y)\triangle^2}{f_Y(y)\triangle} = \frac{f_{X,Y}(x,y)\triangle}{f_Y(y)}$$

Definition 3 (bedingte Dichte)**.** *$X : \Omega \to \mathbb{R}$ und $Y : \Omega \to \mathbb{R}$ sind SGn mit Dichte $f(x, y)$. f_X, f_Y sind die Randdichten von X und Y. Dann heißt für y mit $f_Y(y) > 0$*

$$f(x|y) := \frac{f(x, y)}{f_Y(y)}$$

die bedingte Dichte von X bedingt durch $Y = y$. $f(x|y)$ ist PY_-^{-1} f.s. definiert.

Da für festes y

$$\int_{\mathbb{R}} f(x|y) d\lambda(x) = \frac{\int_{\mathbb{R}} f(x, y) d\lambda(x)}{\int_{\mathbb{R}} f(x, y) d\lambda(x)} = 1$$

und $f(x|y) \geq 0$ ist $f(.|y)$ eine Wahrscheinlichkeitsdichte und tatsächlich eine Verteilung festgelegt.

Wenn $X|Y = yPY^{-1}$-f.s. einen endlichen Erwartungswert besitzt, dann heißt die Funktion

$$y \mapsto \mathbb{E}[X|Y = y] = \int x f(x|y) d\lambda(x)$$

bedingter Erwartungswert. Dann ist $h(Y) := \mathbb{E}[X|Y]$ eine PY^{-1}-f.s. messbare Funktion. $h(.)$ heißt auch Regressionsfunktion. $h(Y)$ ist als Prognose von X nach der Beobachtung von Y zu verstehen.

Beispiel. *(X, Y) sei bivariat normalverteilt $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. Die Darstellung der bivariaten Normalverteilungsdichte (PR18) führt auf die bedingte Dichte*

$$f(x|y) = \frac{1}{\sqrt{2\pi v}} exp\left(-\frac{1}{2}\left(\frac{x-m}{v}\right)^2\right)$$

$$mit\ m = \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y)\ und\ v^2 = \sigma_x^2(1 - \rho^2)$$

d.h. $X|Y \sim \mathcal{N}(m, v^2)$.

Die Regressionsfunktion ist linear

$$\mathbb{E}[X|Y] = \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y)$$

Auch diese Eigenschaft charakterisiert die Normalverteilung.

Für die Regressionsfunktion $H(Y) := \mathbb{E}[X|Y]$ gilt bei Normalverteilung $\mathbb{E}H(Y) = \mu_x$. Das ist generell der Fall:

$$\mathbb{E}H(Y) = \int \mathbb{E}[X|Y = y]dPY^{-1} = \int\int xf(x|y)dxf_Y(y)dy$$

$$= \int\int x\underbrace{f(x|y)f_Y(y)}_{f(x,y)}dydx = \int\int xf(x, y)dydx = \int xf_X(x)dx = \mathbb{E}X.$$

Die mittlere Prognose entspricht dem unbedingten Erwartungswert.

1.2 Faltung von Maßen

Definition 4 (Faltungsmaß). *μ_1, μ_2 sind sigma-endliche Maße auf $(\mathbb{R}, \mathcal{B})$. Das Faltungsmaß ist durch*

$$\mu_1 * \mu_2(A) = \int \mu_1(A - y)\mu_2(dy)$$

definiert.

Das $\mu_1 * \mu_2$ tatsächlich ein Maß ist, ergibt sich daraus, dass es das von der Summe $S = X_1 + X_2$ erzeugte Maß ist. D.h. $\mu_1 * \mu_2 = (\mu_1 \otimes \mu_2)S^{-1}$

Proof. $S^{-1}(A) = \{(x, y) \in \mathbb{R}^2 : x + y \in A\}$, wobei $A \in \mathcal{B}$. Der Schnitt von $S^{-1}(A)$ ist $S^{-1}(A)_y = \{x|x \in A - y\} = A - y$

Die Darstellung des Produktmaßes mittels Schnitten wie zuvor,

$$\mu_1 \otimes \mu_2(B) = \int \mu_1(B_y)d\mu_2(y), B \in \mathcal{A}_1 \otimes \mathcal{A}_2$$

ergibt

$$\mu_1 * \mu_2(S^{-1}(A)) = \int \mu_1(A - y)d\mu_2(y) \text{ und}$$
$$(\mu_1 \otimes \mu_2)S^{-1} = \mu_1 * \mu_2.$$

Es ist auch in die andere "Richtung" darstellbar: $\mu_1 * \mu_2 = \int \mu_2(A - x)d\mu_1(dx)$ □

Neben der Kommutativität $\mu_1 * \mu_2 = \mu_2 * \mu_1$ besitzt die Faltung noch folgende Eigenschaften:

- $\mu_i, i = 1, 2, 3$ sind sigma-endliche Maße auf $(\mathbb{R}, \mathcal{B})$

$$(\mu_1 * \mu_2) * \mu_3 = \mu_1 * (\mu_2 * \mu_3)$$

Proof. $A \in \mathcal{B}_2$

$$(\mu_1 * \mu_2) * \mu_3(A) = \int \mu_1 * \mu_2(A-z)d\mu_3(z) = \int \left(\int \mu_2(A-z-x)d\mu_1(x) \right) d\mu_3 =$$

$$\int \int \mu_2(A-z-x)d\mu_3(z)d\mu_1 = \int \mu_2 * \mu_3(A-x)d\mu_1(x) = (\mu_2 * \mu_3) * \mu_1(A) = \mu_1 * (\mu_2 * \mu_3)(A)$$

□

- $\mu_1(\mathbb{R}) = \mu_2(\mathbb{R}) = 1 \implies \mu_1 * \mu_2(\mathbb{R}) = 1$

  Proof. Da $\mathbb{R} - y = \mathbb{R} \forall y \in \mathbb{R}$ gilt

  $$\mu_1 * \mu_2(\mathbb{R}) = \int \mu_2(\mathbb{R} - x)d\mu_1(x) = \int \mu_2(\mathbb{R})d\mu_1(x) = \mu_2(\mathbb{R}) \cdot \mu_1(\mathbb{R}) = 1$$

  Die Faltung von Wahrscheinlichkeitsmaßen ist auch ein Wahrscheinlichkeitsmaß. □

  Sind $\mu_1 = PX_1^{-1}$ und $\mu_2 = PX_2^{-1}$ von SGn X_1 bzw. X_2 induziert, dann ist $\mu_1 * \mu_2 = PX_1^{-1} * PX_2^{-1} = P(X_1 + X_2)^{-1}$ das von X_1, X_2 induzierte Maß, vorausgesetzt X_1, X_2 sind unabhängig, also für $X = (X_1, X_2), PX = PX_1^{-1} \otimes PX_2^{-1}$.

- Es existiert auch ein neutrales Element der Faltung μ_0 mit $\mu_0 * \mu = \mu = \mu * \mu_0$.

  Proof. Für $\mu_0 = \delta_0$, d.h. $\delta_0(A) = 1_A(0)$ gilt

  $$\mu_0 * \mu(A) = \int \mu_0(A-x)d\mu(x) = \int \underbrace{1_{A-x}(0)}_{1_A(x)} d\mu(x) = \int 1_A(x)d\mu(x) = \mu(A).$$

  □

Die Faltung bildet eine kommutative Halbgruppe.

Die bereits vorher erklärte Faltung von messbaren Funktionen hängt erwartungsgemäß mit der Faltung von Maßen zusammen. Bei Maßen mit Dichten ist die Dichte der Faltung genau die Faltung der Dichten.

Satz 11 (11.PR). *μ_1, μ_2 seien Maße mit Dichten bezüglich λ: $\mu_1 = \int f_1 d\lambda$ und $\mu_2 = \int f_2 d\lambda$ und f_1, f_2 sind reellwertig.*
Dann ist

$$\mu_1 * \mu_2(A) = \int_A \left(\int_{\mathbb{R}} f_1(s-y)f_2(y)d\lambda(y) \right) d\lambda(s)$$

also die Dichte von $\mu_1 * \mu_2$ ist die Faltung der Dichten $f_1 * f_2$.

Proof.

$$\mu_1 * \mu_2(A) = \int_{\mathbb{R}} \mu_1(A-y)d\mu_2(y) = \int_{\mathbb{R}} \int_{A-y} f_1(x)d\lambda(x)f_2(y)d\lambda(y)$$

Die Transformation $T_y(x) = x - y$ ist als lineare Transformation für λ translationsinvariant. Für jedes y gilt

$$\lambda T_y^{-1}(.) = \lambda(T_y^{-1}(.)) = \lambda(.)$$

Das innere Integral ist nach dem Transformationssatz

$$\int_{A-y} f_1(x)d\lambda = \int_{T_y(A)} f_1 d\lambda = \int_A f_1 \circ T_y(s)d\lambda(s)$$

und mit dem Satz von Fubini ergibt sich

$$\mu_1 * \mu_2(A) = \int_{\mathbb{R}} \left(\int_A f_1(s-y)d\lambda(s) \right) f_2(y)d\lambda(y) = \int_A \underbrace{\int_{\mathbb{R}} f_1(s-y)f_2(y)d\lambda(y)}_{f_1 * f_2} d\lambda(s)$$

Analog gilt natürlich auch

$$\mu_1 * \mu_2(A) = \int_A \left(\int f_2(s-x) f_1(x) d\lambda(x) \right) d\lambda(s).$$

□

Damit wurde die Faltung im stetigen Fall auf die Faltungsdichte zurückgeführt. Analog werden diskrete Maße gefaltet.

Satz 12 (12.PR)**.** *$\mu_i, i = 1,2$ sind diskrete Maße auf $(\mathbb{R}, \mathcal{B})$ mit Trägermenge D_i (abzählbar). Dann ist $\mu_1 * \mu_2$ diskret verteilt mit der Trägermenge $D = D_1 + D_2 = \{s | s = x + y, x \in D_1, y \in D_2\}$ und für $s \in \mathbb{R}$ gilt*

$$\mu_1 * \mu_2(s) = \sum_{y \in D_2} \mu_1(\{s-y\}) \mu_2(\{y\})$$

Proof.

$$\mu_1 * \mu_2(A) = \int_{D_2} \mu_1(A-y) d\mu_2(y) = \sum_{y \in D_2} \mu_1(A-y) \mu_2(\{y\})$$

Die Trägermenge von $\mu_1 * \mu_2$ ist D, da $D^c - y \subseteq D_1^c$ für $y \in D_2$ und $\mu_1(D_1^c) = 0$ gilt

$$\mu_1 * \mu_2(D^c) = \sum_{y \in D_2} \underbrace{\mu_1(D^c - y)}_{=0} \mu_2(\{y\})$$
$$\implies \mu_1 * \mu_2(D^c) = 0$$

$\mu_1 * \mu_2(A) = \mu_1 * \mu_2(A \cap D)$ und für $s \in D$ gilt

$$\mu_1 * \mu_2(\{s\}) = \sum_{y \in D_2} \mu_1(\{s-y\}) \mu_2(\{y\})$$

oder mit vertauschten $\mu_i, i = 1,2$

$$\mu_1 * \mu_2(\{s\}) = \sum_{x \in D_1} \mu_2(\{s-x\}) \mu_1(\{x\})$$

□

Für stochastischen Größen und die entsprechenden induzierten Wahrscheinlichkeitsmaße wurde diese diskrete Faltung bereits erklärt, o.B.d.A sei X auf $\mathbb{N}$ diskret verteilt, genauso wie Y, dann ist

$$P(X+Y=k) = \sum_{m \geq 0} P(X = k-m) P(Y=m)$$

wenn X,Y unabhängig sind, bzw.

$$P(X+Y=k) = \sum_{m \geq 0} P(X = k-m | Y=m) P(Y=m)$$

wenn X,Y nicht unabhängig sind.

Mit der bedingten Dichte gilt auch im stetigen Fall für die Dichte von $X + Y =: S$ bei abhängigen X und Y

$$f_S(s) = \int f(s-t | Y=t) f_Y(t) d\lambda(t)$$

wobei $f(x|y)$ die Dichte der bedingten Verteilung von $X|Y$ bezeichnet. Auch hier können die marginalen Dichten getauscht werden, d.h. mit $g(y|x)$ als bedingte Dichte $Y|X$

$$f_S(s) = \int g(s-x | X=x) f_X(x) d\lambda(x).$$

Für die Binomialverteilung wurde die Faltung bereits durchgeführt.

Beispiel (Faltung negativer Binomialverteilung). *X ist $NegB_{k,p}$, wenn X die Anzahl der Fehlversuche ($\simeq 0$) bis zum k-ten Erfolg ($\simeq 1$) bei unabhängigen $0-1$ Versuchen mit Erfolgswahrscheinlichkeit p.*

$$P(X=m)=\binom{m+k-1}{k-1}p^k(1-p)^m$$

Die geometrische Verteilung G_p ist ein Spezialfall, $G_p = NegB_{1,p}$ (Variante 2 der geometrischen Verteilung ist hier gemeint).

$$X_i \sim G_p : P(X_i = m) = p(1-p)^m, m \in \mathbb{N}$$

Die Faltung zweier G_p heißt

$$P(X_1+X_2=m)=\sum_{l=0}^{m} p(1-p)^{m-l}p(1-p)^l = (m+1)p^2(1-p)^m = \binom{m+2-1}{2-1}p^2(1-p)^m$$

daher $X_1 + X_2 \sim NegB_{2,p}$.

Genauso ergibt die Faltung $NegB_{k,p} * G_p =: \mu$

$$\mu(\{m\})=\sum_{i=0}^{m}\binom{i+k-1}{k-1}p^k(1-p)^i p(1-p)^{m-i} = p^{k+1}(1-p)^m\sum_{i=0}^{m}\binom{i+k-1}{k-1} = p^{k+1}(1-p)^m\binom{k+m}{k}$$

Mit dem Pascalschen Dreieck $\binom{n+1}{k} = \binom{n}{k-1}\binom{n}{k}$ und Induktion kann gezeigt werden, dass $\sum_{i=0}^{m}\binom{i+k-1}{k-1} = \binom{k+m}{k}$.

Somit ist $NegB_{k,p} * G_p = NegB_{k+1,p}$. Mit Induktion gilt daher sofort $NegB_{k,p} * NegB_{\tilde{k},p} = NegB_{k+\tilde{k},p}$.

Die Summe unabhängiger $NegB$-Verteilungen (mit gleichem p) ist wieder negativ Binomialverteilt.

Bei verschiedenen p ist die Summe nicht nach einer $NegB$ verteilt.

Die Faltung eines (absolut) stetigen Maßes und eines diskreten Maßes ergibt wieder ein Maß mit Dichte.

Satz 13 (13.PR). *μ sei ein Maß mit Dichte f bez. λ. μ daher $\mu \ll \lambda$ und ν sei ein diskretes Maß mit Trägermenge D (abzählbar) auf $(\mathbb{R}, \mathcal{B})$. Dann ist $\mu * \nu \ll \lambda$ und*

$$\mu * \nu(A) = \int_A \underbrace{\sum_{k\in D} f(x-k)\nu(\{k\})\,d\lambda(x)}_{\text{Dichte von } \mu*\nu}.$$

Proof. Da $\mu \ll \lambda$ gilt für eine Nullmenge N, $\lambda(N) = 0$ auch $\mu(N) = 0$. Wegen der Translationsinvarianz $\lambda(N-s) = 0$ für alle $s \in \mathbb{R}$, daher auch

$$\mu(N-s) = 0 \text{ und } \mu * \nu(N) = \int \mu(N-s)\nu(ds) = 0$$
$$\implies \mu * \nu \ll \lambda.$$

Sei $A = (-\infty, t], t \in \mathbb{R}$, dann gilt

$$\mu * \nu(A) = \int \sum_{k\le t-s, k\in D} \nu(\{k\})f(s)d\lambda(s) = \int \sum_{k\in D} \nu(\{k\})1_{(-\infty,t]}(k+s)f(s)d\lambda(s) =$$
$$\sum_{k\in D}\nu(k)\int 1_{(-\infty,t]}(x)f(x-k)d\lambda(x) = \int_{(-\infty,t]}\sum_{k\in D} f(x-k)\nu(k)d\lambda(x)$$

Die Aussage gilt daher für Intervalle. Wegen des Fortsetzungssatzes gilt sie auch für jedes $A \in \mathcal{B}$. □

Beberkung. *Bei mehr als 2 Maßen genügt ein Maß $\mu_i \ll \lambda$.*

Beispiel. *μ sei eine Exponentialverteilung E_{x_1} mit Dichte $f(s) = e^{-s}1_{(0,\inf)}(s)$. ν entspreche einer Geometrischen Verteilung $G_{\frac{1}{2}}$ (Version 1). $\nu(\{k\}) = \frac{1}{2^k}, k \ge 1$.*

Die Lebesgue-Dichte des Faltungsmaßes ist

$$h(x) = \sum_{k\ge 1}\frac{1}{2^k}e^{-(x-k)}1_{(0,\infty)}(x-k) = e^{-(x-1)}\frac{(\frac{e}{2})^{\lfloor x\rfloor}-1}{e-2}$$

was einer stückweisen (auf $[n, n+1)$) gewichteten Exponentialfunktion (Verteiliung) entspricht.

Die Faltung kann auch mit der Verteilungsfunktion durchgeführt werden:

X besitzt F als Verteilungsfunktion Y hat G als VF, X und Y unabhängig. Die Verteilungsfunktion von $X+Y$ ist

$$H(t) = \int F(t-s)dG(s) = \int G(t-s)dF(s)$$

(siehe Übung).

Wie im letzten Beispiel ist $H(.)$ bis auf eine Nullmenge differenzierbar, wenn F und G differenzierbar mit Dichten $F' = f$ und $G' = g$ sind. Die (stückweise) differenzierbar VF H ergibt die Dichte aus Satz 11.PR.

$$\frac{dH}{dx}(s) = \int \frac{dF}{dx}(s-t)\frac{dG}{dx}(t)d\lambda(t).$$

Natürlich wird auch die Verteilung der Differenz $X - Y$ unabhängiger Sgn X, Y mit der Faltung von X und $-Y$ bestimmt.

Wenn Dichten existieren ist die Dichte von $X - Y$ $h(t) = \inf f_X(s) f_Y(s-t) d\lambda(s)$. Die Wahrscheinlichkeit für $X \leq Y$ kann auch mit der FaltungsVF für $X - Y$ berechnet werden.

Lemma 1. *X, Y SGn mit VF F_X, F_Y und Dichten f_X, f_Y. Wenn X u.a. Y, dann ist*

$$P[X \leq Y] = \int_{\mathbb{R}} F_X(t) f_Y(t) d\lambda(t).$$

Proof.

$$P[X \leq Y] = \int_{\{x \leq y\}} f(x,y) d\lambda_2 = \int_{\mathbb{R}} \int_{(-\infty, Y]} f_X(x) f_Y(y) d\lambda_2(x,y) = \int_{\mathbb{R}} F_X(y) f_Y(y) d\lambda(y)$$

□

Beispiel. *X, Y unabhängige Exponentialverteilungen mit Raten λ_1, λ_2, d.h. $X \sim E_{x_{\lambda_1}}, Y \sim E_{x_{\lambda_2}}$.*

$$P(X \leq Y) = \int_{\mathbb{R}+} (1 - e^{-\lambda_1 y}) \lambda_2 e^{-\lambda_2 y} d\lambda(y) = 1 - \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Anteil der Raten von X.

2 Signierte Maße und Zerlegungen

Die Summe von Maßen ist wieder ein Maß, die Differenz von Maßen ist eine zumindest sigma-additive Mengenfunktion.

Definition 5 (signiertes Maß)**.** *Die Mengenfunktion ν ist ein signiertes Maß, wenn auf $(\Omega, \mathcal{A})$ gilt $\nu : \mathcal{A} \to (-\infty, \infty]$ oder $[-\infty, \infty)$ mit $\nu(\emptyset) = 0$ und für disjunkte $A_i \in \mathcal{A}$ ist $\nu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \nu(A_i)$.*

Viele Eigenschaften von Maßen lassen sich auf signierte Maße übertragen, allerdings ein Integral über ein signiertes Maß macht Schwierigkeiten. Es stellt sich aber heraus, dass signierte Maße sich als Differenz zweier Maße darstellen lässt. Das und vieles folgende wäre sofort klar, wenn es für ν eine Dichte $f : \Omega \to \mathbb{R}$ bezüglich eines Maßes μ gibt $\nu(A) = \int_A f d\mu$, was auch $\nu \ll \mu$ zur Folge hat.

$$\nu(A) = \int_A f^+ - f^- d\mu = \underbrace{\int_A f^+ d\mu}_{\mu_1} - \underbrace{\int_A f^- d\mu}_{\mu_2}$$

Für $A \subseteq [f^+ > 0]$ ist eine Menge mit $\nu(A) \geq 0$ und auch jede Teilmenge $B \subseteq A$ erfüllt $\nu(B) \geq 0$.

Definition 6 (positive Menge, negative Menge)**.** *$(\Omega, \mathcal{A}, \nu)$ sei signierter Maßraum. A^+ heißt positive Menge bez. ν, wenn $\forall B \subseteq A^+ : \nu(B) \geq 0$ erfüllt, A^- ist ν-negative Menge, wenn $\nu(B) \leq 0$ für $B \subseteq A^-$.*

Negative Mengen eines signierten Maßraumes bilden einen sigma-Ring (Bezeichnung $\mathcal{A}^-$). $\mathcal{A}^-$ ist $\cap$-stabil und $\triangle$-stabil. Die Monotonie folgt aus der Stetigkeit von signierten Maßen.

Sie Maße sind signierte Maße stetig von unten. Die monoton wachsende Folge $A_n \nearrow A$ wird wieder mit disjunkten Teilmengen dargestellt, $A = \bigcup_{n=1}^{\infty} A_n = A_1 \cup (A_2 \setminus A_1) \dots$.

Bei der Stetigkeit von oben wird noch $|\nu(A_{n_0})| < \infty$ für ein n_0 (damit für $n \geq n_0$) verlangt.

Definition 7 (Hahn-Zerlegung). *Eine Hahn-Zerlegung von Ω besteht aus $\{P, P^c\}$, wobei P eine positive und P^c eine negative Menge ist.*

Satz 14 (1.SZ). *Jeder signierte Maßraum $(\Omega, \mathcal{A}, \nu)$ besitzt eine Hahn Zerlegung von Ω.*

Proof. Sei $\nu : \mathcal{A} \to (-\infty, \infty]$. $\mathcal{A}^-$ das System der negativen Mengen. Sei $c := \inf_{B \in \mathcal{A}'} \nu(B)$ daher gibt es eine Folge B_n mit $\nu(B_n) < c + \frac{1}{n}$ und $B_n \in \mathcal{A}'$.

$N := \bigcup_n B_n$ ist eine negative Menge. Sei $P := N^c$.

Angenommen P enthält eine negative Menge B mit $\nu(B) < 0$, dann wäre $\nu(B \cup N) = \underbrace{\nu(B)}_{<0} + \nu(N) < c$ also c nicht das Infimum.

Demnach ist für jedes $B \subseteq P$ und $\nu(B) < 0$.

$$\mathcal{E}_1 := \sup\{\nu(M) | M \subseteq B\} > 0$$

Es gibt ein $M_1 \subseteq B$ mit $\frac{\mathcal{E}_1}{2} \leq \nu(M_1) \neq \mathcal{E}_1$.

Daher $\nu(B \setminus M_1) = \nu(B) - \nu(M_1) < \nu(B) < 0$. Auch $B \setminus M_1 \notin \mathcal{A}^-$, daher gibt es ein M_2 mit $\nu(B \setminus M_1 \setminus M_2) < 0$.

Mit Induktion gibt es also eine Folge disjunkter Mengen $M_n \subseteq B$ mit $\nu(B \setminus \bigcup_{i=1}^{n} M_i) < 0$.

Betrachte $D = B \setminus \bigcup_{i=1}^{\infty} M_i \implies B = \bigcup M_i \cup D$.

Für M_i gilt $\nu(M_i) > 0$ und $\nu(B) = \sum_i \nu(M_i) + \nu(D)$, daher $0 \leq \sum \nu(M_i) < \infty$ und $\nu(D) < 0$.

Somit muss $\nu(M_i) \to 0$ und für $C \subseteq D \subseteq B \setminus \bigcup_{i=1}^{n} M_i$ jede Teilmenge C erfüllt $\nu(C) \leq \mathcal{E}_n$ mit $\mathcal{E}_n \to 0 \implies \nu(C) \leq 0$.

D ist eine negative Menge mit $\nu(D) < 0$. Da $D \subseteq P = N^c$ wäre $N \cup D$ eine negative Menge mit $\nu(N \cup D) < \nu(N) = 0$. Damit kann $c = \inf_{B \in \mathcal{A}^-} \nu(B)$ nicht stimmen, die Annahme $\nu(B) < 0$ hält nicht, P ist positive Menge. □

Hahn-Zerlegungen sind nicht eindeutig, für $\nu(.) = \inf_{.} f d\lambda$ kann der Trennungspunk gewählt werden. Aber die symmetrische Differenz positiver Mengen ist eine Nullmenge:

Sei $\{P_1, P_1^c\}$ und $\{P_2, P_2^c\}$ jeweils eine Hahn-Zerlegung von Ω. Dann gilt für $A \subseteq P_1 \setminus P_2$, dass $\nu(A) \geq 0$ $(A \subseteq P_1)$ und $\nu(A) \leq 0$ $(A \subseteq P_2^c)$, also $\nu(A) = 0$.

Definition 8 (Jordan-Zerlegung). *Eine Jordan-Zerlegung des signierten Maßes ν bilden Maße μ_1, μ_2 mit*

$$\nu = \mu_1 - \mu_2,$$

wobei μ_1, μ_2 singulär zueinander sind,

$$\mu_1 \perp \mu_2 \text{ d.h. } \exists M \text{ mit } \mu_1(M) = \mu_2(M^c) = 0.$$

Die Singularität führt auf folgende Minimaleigenschaft der Jordan-Zerlegung.

Ist ν^+, ν^- eine Jordanzerlegung von $\nu = \nu^+ - \nu^-$ und $\nu = \mu_1 - \mu_2$ für beliebige Maße μ_1, μ_2, dann gilt $\nu^+(A) \leq \mu_1(A)$ und $\nu^-(A) \leq \mu_2(A)$ für alle $A \in \mathcal{A}$.

Mit obiger Trennungsmenge M gilt

$$\begin{aligned}
\nu^+(A) = \underbrace{\nu^+(A \cap M)}_{=0} + \nu^+(A \cap M^c) = \nu^+(A \cap M^c) = \nu(A \cap M^c) \\
\nu^-(A) = \nu^-(A \cap M) + \underbrace{\nu^-(A \cap M^c)}_{=0} = \nu^-(A \cap M) = -\nu(A \cap M) \\
\implies \nu^+(A) = \mu_1(A \cap M^c) - \mu_2(A \cap M^c) \leq \mu_1(A) \\
\nu^-(A) = \mu_2(A \cap M) - \mu_1(A \cap M) \leq \mu_2(A)
\end{aligned}$$

Mit der Hahn-Zerlegung findet man auch eine Jordan-Zerlegung.

Satz 15 (2.SZ Jordan'scher Zerlegungssatz). *Jedes signierte Maß ν besitzt eine eindeutige Jordan-Zerlegung $\nu = \mu_1 - \mu_2$.*

Proof. Aus der Hahn-Zerlegung $\Omega = P \cup P^c$ erhält man mit

$$\mu_1(A) = \nu(A \cap P) \text{ und } \mu_2(A) = -\nu(A \cap P^c)$$

zwei singuläre Maße. Die Eindeutigkeit folgt aus der Minimaleigenschaft.

Sei $\nu = \nu^+ - \nu^-$ mit singulären Maßen ν^+, ν^-, dann gilt $\nu^+ \leq \mu_1 \wedge \nu^- \leq \mu_2$ und $\mu_1 \leq \nu^+ \wedge \mu_2 \leq \nu^-$. □

Def (Totalvariation)**.** *Durch $\mu_1 + \mu_2$ aus der Jordan-Zerlegung entsteht ein Maß, der Variation bzw. Totalvariation mit der Bezeichnung $|\nu| = \mu_1 + \mu_2$ genannt wird. Die Rechtfertigung der Bezeichnung gilt.*

Lemma 2. *Für das signierte Maß ν gilt*

$$\forall A \in \mathcal{A} : |\nu(A)| \leq |\nu|(A).$$

Proof. Die Aussage folgt sofort aus

$$|\nu(A)| = |\mu_1(A) - \mu_2(A)| \leq |\mu_1(A)| + |\mu_2(A)| = |\nu|(A).$$

□

Die Zerlegungssätze für signierte Maße werden für sigma-endliche Maße konkretisiert.

Definition 9 (Lebesgue-Zerlegung)**.** *ν und μ seinen sigma-endliche Maße auf $(\Omega, \mathcal{A})$. Eine Lebesgue-Zerlegung von ν besteht aus Maßen ν_c, ν_s, sodass*

$$\nu = \nu_c + \nu_s \qquad \text{und} \qquad \nu_c \ll \mu \qquad \text{und} \qquad \nu_s \perp \mu$$

also ν_c absolut stetig bezüglich μ und ν_s singulär zu μ ist.

ν ist in Bezug auf μ in einen stetigen und einen singulären Anteil zerlegt.

Die zentrale Aussage ist, dass eine Lebesgue-Zerlegung für sigma-endliche Maße immer existiert und eindeutig ist. Für den Beweis können verschiedene Wege genommen werden. Eine Möglichkeit ist die Verwendung des Satzes von Riesz.

Satz 16 (Darstellungssatz von Riesz-Frechet)**.** *$(V, < .,. >)$ sei ein Hilbertraum und $H : V \to \mathbb{R}$ eine reellwertige Abbildung. Dann sind folgende Aussagen äquivalent:*

- *H ist stetig und linear*
- *$\exists a \in V$ mit $H(x) =< x, a >$ und a ist eindeutig bestimmt.*

Der Hilbertraum hier ist $\mathcal{L}_2(\mu)$ (also die quadratisch integrierbaren Funktionen) mit $< f, g >= \int fg d\mu$.

Die Abbildung $H : \mathcal{L}^2(\mu) \to \mathbb{R}$ ist also genau dann stetig und linear, wenn es ein $f \in \mathcal{L}_2(\mu)$ gibt mit $H(g) = \int gf d\mu \forall g \in \mathcal{L}_2$.

Satz 17 (3.SZ Zerlegungssatz von Lebesgue)**.** *Ein sigma-endliches Maß ν auf dem sigma-endlichen Maßraum $(\Omega, \mathcal{A}, \mu)$ hat eine eindeutige Lebesgue Zerlegung.*

Satz 18 (4.SZ Satz von Radon-Nikodym)**.** *Für sigma-endliche Maße μ, ν gilt $\nu \ll \mu \iff \nu$ ist Maß mit Dichte, d.h.*

$$\nu(A) = \int_A f d\mu$$

mit f.ü. eindeutiger Dichte $f \geq 0$; $f = \frac{d\nu}{d\mu}$.

Satz 3 und 4. Da beide Maße als sigma-endlich vorausgesetzt werden, kann (wie üblich) sich auf den endlichen Fall beschränken. Sei für das Maß $\gamma := \mu + \nu$.

$\gamma(\Omega) < \infty$. Für $f, g \in \mathcal{L}^2(\gamma)$ gilt

$$\int |f - g| d\nu \leq \int |f - g| d\mu \leq (1 + \gamma(\Omega)) ||f - g||_2 \text{ (Norm auf } \mathcal{L}^2(\gamma))$$

Das bedeutet $h \to \int h d\nu$ ist eine stetige Abbildung $\mathcal{L}_2(\gamma) \mapsto \mathbb{R}$ und natürlich auch eine lineare Abbildung.

Nach dem Satz von Riesz-Frechet existiert ein $g \in \mathcal{L}_2(\gamma)$ mit

$$(1) \int h d\nu = \int hg d\gamma \forall h \in \mathcal{L}_2(\gamma) \text{ oder}$$
$$(2) \int \tilde{h}(1-g) d\gamma = \int \tilde{h} d\mu \forall \tilde{h} \in \mathcal{L}_2(\gamma).$$

Betrachtet man $h = 1_{[g<0]}$ so gilt

$$0 \le \underbrace{\nu([g<0])}_{=0} = \int h d\nu = \int 1_{[g<0]} g d\gamma \le 0$$

$$\text{da, } \underbrace{\int 1_{g<0} g\nu}_{=0} + \underbrace{\int 1_{g<0} g\mu}_{=0}$$

Das erste Integral ist über eine Nullmenge und das Zweite weil: Wäre $\mu[g<0] > 0$, dann wäre das Integral < 0.

Betrachtet man $\tilde{h} = 1_[g > 1]$ ist genauso

$$\mu([g>1]) = \int 1_{[g>1]}(1-g) d\gamma \le 0$$

und $g \le 1$ γ-f.ü.

Es gilt γ-f.ü. $0 \le g \le 1$ und $\tilde{\gamma} := (1-g)\gamma$ ist ein Maß mit Dichte $(1-g)$, $\tilde{\gamma}(A) = \int_A (1-g) d\gamma$.

Für jedes $h \ge 0$ existiert eine Folge $h_n \in \mathcal{L}_2(\tilde{\gamma})$ mit $h_n \nearrow h$ (beispielsweise sind Treppenfunktionen aus $\mathcal{L}^2$). Nach dem Satz von der monotonen Konvergenz gilt

$$\int h d\tilde{\gamma} = \int \lim h_n d\tilde{\gamma} = \lim \int h_n d\tilde{\gamma} \overbrace{=}^{\text{wegen (2)}} \lim \int h_n d\mu = \int h d\mu.$$

Also gilt für alle $h \ge 0$ (1) und (2),

$$\int h(1-g) d\gamma = \int h d\mu$$

Sei $E = [g=1]$ und $h = 1_E$, dann folgt $\int h d\mu = 0 \implies \mu([g=1]) = 0$.

Seien die (Spur-)Maße

$$\nu_c(A) = \nu(A \setminus E) \qquad \text{und} \qquad \nu_s(A) = \nu(A \cap E)$$

Offensichtlich ist $\nu = \nu_c + \nu_s$ und die Maße sind singulär aufeinander, und $\nu_s \perp \mu$ da $\mu(E) = 0$ und $\nu_s(E^c) = 0$.

Sei A eine μ-Nullmenge $\mu(A) = 0$, dann genügt $A \cap E = \emptyset$ anzunehmen da $\mu(e) = 0$, dh. $\int_A (1-g) d(\mu+\nu) = 0$ (nach 2), da $1-g > 0$ auf A ist $\mu(A) + \nu(A) = 0 \implies \nu(A) = 0$ und auch $\nu_c(A) = 0$.

Daher ist $\nu_c \ll \mu$.

Es bleibt noch die Dichte zu finden, sei

$$f := \frac{g}{1-g} 1_{E^c}.$$
$$\int_A f d\mu \overbrace{=}^{\text{wegen (2)}} \int_A 1_{E^c} g d(\mu+\nu) \overbrace{=}^{wegen (1)} \int 1_A 1_{E^c} d\nu = \nu(A \setminus E) = \nu_c(A)$$

und die Dichte ist f.

□

Der Satz von Radon-Nikodym ist damit auch gezeigt, da wenn $\nu \ll \mu$ wird $\nu_s \equiv 0$ in der Lebesgue-Zerlegung gesetzt.

Beispiel. $\mu_i, i = 1, 2$ *seinen Poissonverteilungen mit Raten* $\lambda_i > 0$

$$\mu_1(\{k\}) = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, k \in \mathbb{N}$$

$\mu_1 \ll \mu_2$ *und* $\mu_2 \ll \mu_1$*, da* $\mu_i(\{k\}) > 0 \forall k \in \mathbb{N}$.

$$\mu_2(\{k\}) = \left(\frac{\lambda_2}{\lambda_1}\right)^k e^{\lambda_1 - \lambda_2} \mu_1(\{k\})$$

$$\textit{also } \frac{d\mu_2}{d\mu_1} = \left(\frac{\lambda_2}{\lambda_1}\right)^k e^{\lambda_1 - \lambda_2} \textit{ und}$$

$$\mu_2(A) = \sum_{k \in A} \left(\frac{\lambda_2}{\lambda_1}\right)^k e^{\lambda_1 - \lambda_2} \mu_1(\{k\})$$

Definition 10 (Äquivalenz)**.** *Die Maße* μ_1, μ_2 *heißen äquivalent (*$\mu_1 \approx \mu_2$*), wenn* $\mu_1 \ll \mu_2$ *und* $\mu_2 \ll \mu_1$*.*

Lemma 3. *Die sigma-endlichen Maße* μ_1 *und* μ_2 *seinen äquivalent* $\mu_1 \approx \mu_2$*. Dann gilt für die RN-Dichten*

$$\frac{d\mu_1}{d\mu_2} = \left(\frac{d\mu_2}{d\mu_1}\right)^{-1} \mu_1\text{-f.ü.}$$

Proof. Da $\mu_1 \ll \mu_2 \ll \mu_1$ ergibt die Anwendung der Kettenregel mit $\mu_2 = f_2 \cdot \mu_1$

$$\mu_1(A) = \int_A f_1 d\mu_2 = \int_A \underbrace{f_1}_{\frac{d\mu_1}{d\mu_2}} \underbrace{f_2}_{\frac{d\mu_2}{d\mu_1}} d\mu_1$$

Da $\mu_1(A) = \int_A 1 d\mu_1$ und die RN-Dichte f.ü. eindeutig ist, gilt μ_1-f.ü.

$$\frac{d\mu_1}{d\mu_2}\frac{d\mu_2}{d\mu_1} = 1 \implies \frac{d\mu_1}{d\mu_2} = \left(\frac{d\mu_2}{d\mu_1}\right)^{-1}$$

("Kehrwert und Ableitung sind vertauschbar.") □

Ein einfaches Beispiel zeigt, dass die (sigma-)Endlichkeit der Maße unverzichtbar ist.

Beispiel. $\mathcal{A} = \{\Omega, \emptyset\}$*,* $\mu_1(\emptyset) = 0 = \mu_2(\emptyset)$ *und* $\mu_1(\Omega) = 1, \mu_2(\Omega) = \infty$*.* $\mu_1 \approx \mu_2$

$$\mu_1(\Omega) = \int_\Omega f d\mu_2$$

ist für keine messbare Funktion $f \geq 0$ *möglich.*

Im obigem Beispiel (Poissonverteilung) waren $\mu_i \ll \zeta$ (Zählmaß) und die Dichten wurden über den "Umweg" über ζ bestimmt.

3 Dichten und absolute Stetigkeit

Lebesgue-Stieltjes Maße auf $(\mathbb{R}, \mathcal{B})$ besitzen eine Verteilungsfunktion F von μ. Angenommen F ist überall stetig differenzierbar, dann gilt

$$\frac{dF}{dx} = \frac{d}{dx}\mu((-\infty, x]) = f \text{ und}$$

$$\mu((-\infty, x]) = \int_{-\infty}^{\infty} f(t) dt.$$

Nach dem Hauptsatz der Differenzial- und Integralrechnung gilt

$$f = \frac{dF}{dx}$$

und die Dichte entspricht der Ableitung der Verteilungsfunktion.

Im Gegensatz zu den vorigen (Existenz-)Sätzen zu Dichten, besteht jetzt eine einfache Möglichkeit eine Dichte (bezüglich λ) zu finden.

Für eine Verallgemeinerung des obigen Prinzips, wird die stetige Differenzierbarkeit gegen schwächere Eigenschaften der Dichten gewechselt.

Definition 11 (Beschränkte Variation). *$f : [a, b] \to \mathbb{R}$ ist von beschränkter Variation mit Schranke M, wenn für jede endliche Partition des Intervalls, $a = x_0 < x_1 < ... < x_n = b$*

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \leq M$$

Definition 12 (Totalvariation). *Die Totalvariation von f ist*

$$V_a^b f := \sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|, n \in \mathbb{N} \right\}$$

wobei das Supremum über alle solche Partitionen gebildet wird. Die Menge $BV(a, b)$ beherbergt alle f von beschränkter Variation auf $[a, b]$.

Beispiel. *$f(x) = x \sin(1/x)$ mit $f(0) = 0$ ist auf $[0, 1]$ stetig, aber nicht von beschränkter Variation. Sei die Partition definiert durch*

$$x_i = \left(\frac{\pi}{2}(1 + 2i)\right)^{-1}, i = 1, 2, ...$$

womit

$$|f(x_i) - f(x_{i-1})| \geq \frac{1}{\pi(i + \frac{3}{2})}.$$

Die rechte Seite ist nicht summierbar, für jedes M existiert eine Anzahl n mit $\sum^n |f(x_i) - f(x_{i-1})| > M$.

Lemma 4. *Wenn $f \in BV(a, b)$ und $a < c < b$ gilt*

$$V_a^c f + V_c^b f = V_a^b f$$

Proof. Für eine beliebige Partition mit n Parametern des Intervalls (a, b) sei $s(a, b) := \sum |f(x_i) - f(x_{i-1})|$ mit $x_0 = a$, $x_n = b$.

Ist c ein Punkt von $a = x_0 < ... < x_n = b$ gilt

$$s(a, c) + s(c, b) \leq V_a^b f$$

da die Partitionen von (a, c) und (c, b) beliebig waren, gilt auch

$$V_a^c f + V_c^b f \leq V_a^b f$$

Ist c kein Punkt und $x_i < c < x_i + 1$

$$s(a, b) \leq s(a, x_j) + |f(x_{j+1}) - f(c)| + |f(c) - f(x_j)| + s(x_{j+1}, b).$$

Auch das gilt für beliebige Partitionen, somit $V_a^b f \leq V_a^c f + V_c^b f$ □

Lemma 5. *$f \in BV(a, b)$ dann sind auf (a, b)*

$$x \to V_a^x f \text{ und } x \to V_a^x f - f(x)$$

monoton wachsende Funktionen.

Proof. $V(x) := V_a^x f$. Für $y > x$ gilt nach vorigem Lemma $V_a^y f = V_a^x f + V_x^y f$, d.h. $V(.)$ ist wachsend.

$$\begin{aligned} f(y) - f(x) \leq |f(y) - f(x)| \leq V_x^y f &= V_a^y f - V_a^x f \\ \implies V_a^x f - f(x) &\leq V_a^y f - f(y) \end{aligned}$$

□

Die Darstellung von Funktionen mit beschränkter Variation durch monotone Funktionen ist auch eine Charakterisierung.

Satz 19 (1.DS)**.** *$f \in BV(a,b)$ genau dann, wenn $f = v - w$ mit monoton wachsenden Funktionen v, w.*

Proof. Es ist nur mehr $\Longleftarrow$ zu zeigen.

Sei $f = v - w$

$$\sum |f(x_i) - f(x_{i-1})| = \sum |v(x_i) - v(x_{i-1}) + w(x_{i-1}) - w(x_i)| \leq \sum |v(x_i) - v(x_{i-1})| + \sum |w(x_i) - w(x_{i-1})|.$$

Da jede monotone Funktion von beschränkter Variation ist, gilt

$$\sum |f(x_i) - f(x_{i-1})| \leq V_a^b v + V_a^b w < \infty.$$

□

Eine monotone Funktion hat nur höchstens abzählbar viele Unstetigkeitsstellen. $f : [a,b] \mapsto \mathbb{R}$ monoton steigend, d.h. $f_-(x) \leq f(x) \leq f_+(x)$ (also jeweils der linksseitige bzw. rechtsseitige Grenzwert).

Jede Unstetigkeitsstelle x_0 erzeugt ein Intervall $(f_-(x_0), f_+(x_0))$, diese Intervalle sind eindeutig und disjunkt, wovon es nur höchstens abzählbar viele geben kann.

Eine Funktion beschränkter Variation, $f \in BV(a,b)$ ist stetig λ-f.ü.

Wenn $f \in BV(a,b)$ mit $f = F - G$ und F, G monoton wachsend, dann kann F, G an der Unstetigkeitsstellen durch den rechtsseitigen Grenzwert ersetzt werden.

F_+, G_+ sind dann Verteilungsfunktionen von Lebesgue-Stieltjes Maßen.

$$f_+ := F_+ - G_+ = f \lambda\text{-f.ü.}$$

f_+, f ist als Verteilungsfunktionen eines signierten LS-Maßes zu verstehen.

Eine Verteilungsfunktion F ist zerlegbar in $F = F_s + F_d$, wobei F_s stetig ist und F_d eine Verteilungsfunktion eines diskreten Maßes ist. (Abzählbare Trägermenge $\{x_i\}_{i \in I}$ mit $\mu(\{x_i\}) > 0$).

Der stetige Anteil (F_s) lässt sich weiter zerlegen. Dazu betrachten wir absolut stetige Funktionen.

Definition 13 (absolut stetig)**.** *Die Funktion $f : [a,b] \to \mathbb{R}$ heißt absolut stetig, wenn zu beliebigem $\epsilon > 0$ ein $\delta > 0$ existiert, sodass aus $\sum_{i=1}^n (b_i - a_i) < \delta$ folgt $\sum_{i=1}^n |f(b_i) - f(a_i)| < \epsilon$. Dabei sind (a_i, b_i) disjunkte Teilintervalle von $[a,b]$, die Anzahl der Intervalle ist n (beliebig).*

Solche absolut stetigen Funktionen liegen "zwischen" stetig differenzierbaren und gleichmäßig stetigen Funktionen.

Wenn $f \in C^1[a,b]$, dann ist f absolut stetig, da

$$|f(b_i) - f(a_i)| \leq \sum_{x \in (a,b)} |f'(x)|(b_i - a_i), a_i < b_i$$

Lemma 6. *f sei auf $[a,b]$ absolut stetig. Dann ist f gleichmäßig stetig und $f \in BV[a,b]$.*

Proof. Nach Voraussetzung existiert zu $\epsilon > 0$ ein $\delta > 0$ sodass aus $(b-a) < \delta \implies |f(b) - f(a)| < \epsilon$ egal was a, b ist. Damit ist f gleichmäßig stetig.

Für eine Partition $a = x_0 < x_1 < ... < x_n = b$ mit $(x_i - x_{i-1}) < \delta \forall i$ gilt für jede endliche Partition (a_i, b_i) des Intervalls (x_{i-1}, x_i)

$$\sum_{i=1}^m |f(b_i) - f(a_i)| \leq V_{x_{i-1}}^{x_i} f < \epsilon$$

nach Voraussetzung der absoluten Stetigkeit.

$V_a^b f$ ist additiv, also

$$V_a^b f = \sum_{i=1}^n V_{x_{i-1}}^{x_i} < n\epsilon$$

daher ist $V_a^b f$ endlich. □

Absolute Stetigkeit ist stärker als gleichmäßige Stetigkeit, aber glm. Stetigkeit ist nicht hinreichend. Lipschitz Stetigkeit wiederum ist hinreichend.

Wenn $f \in C^1[a,b]$, also stetig differenzierbar ist, dann ist f absolut stetig.

Für $a_i, b_i \in [a,b]$ und $\sup_{x\in[a,b]} |f'(x)| < \infty$ ist

$$|f(b_i) - f(a_i)| \leq \sup_{x\in[a,b]} |f'(x)||b_i - a_i|$$

und daher ist f absolut stetig.

Im folgenden soll geklärt werden, wie weit die komfortable Situation einer stetig differenzierbaren Verteilungsfunktion mit daraus folgender Maßdichte auf absolut stetige Verteilungsfunktionen übertragen werden kann. Damit könnte dann auch die Dichte eines Maßes μ mit $\mu \ll \lambda$ bestimmt werden.

Für eine absolut stetige Funktion diene eine Verteilungsfunktion eines signierten Maßes als Vorlage. Das rechtfertigt folgender Satz.

Satz 20 (2.DS)**.** *$f : [a,b] \to \mathbb{R}$ ist absolut stetig. Dann ist $f = F - G$, wobei beide Funktionen F, G monoton wachsend und absolut stetig sind.*

Proof. Zunächst wird gezeigt, dass $x \to V_a^x f$ eine absolut stetige Funktion ist.

Für $\epsilon > 0$ existiert ein δ, sodass eine Partition $P = \{(a_i, b_i)|i = 1, ..., n\}$ disjunkter Intervalle mit $\sum(b_i - a_i) < \delta$ existiert, sodass $\sum_{i=1}^n |f(b_i) - f(a_i)| < \epsilon$.

Wird jedes einzelne Intervall (a_i, b_i) wieder zerlegt, $a_i \leq a_{i_0} < ... < a_{i_{m_i}} = b_i$, ist

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}(a_{i_j} - a_{i_{j-1}}) = \sum_{i=1}^{n} b_i - a_i < \delta$$

und nach Voraussetzung

$$\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{m_i}|f(a_{i_j}) - f(a_{i_{j-1}})}_{=V_{a_{i-1}}^{a_i}} < \epsilon$$

Da die Zerlegung von (a_i, b_i) beliebig war, ist auch

$$\sum_{i=1}^{n} \sup\left\{\sum_{j=1}^{m_i} |f(a_{i_j}) - f(a_{i_{j-1}})| \middle| a_i = a_{i_0} < ... < a_{i_{m_i}} = b_i\right\} = \sum_{i=1}^{n} V_{a_i}^{b_i} f = \sum_{i=1}^{n} \underbrace{|V_a^{b_i} f - V_a^{a_i} f|}_{=V_{a_i}^{b_i} f} < \epsilon$$

also $v(x) = V_a^x f$ ist absolut stetig. Dann ist auch $v(x) - f(x)$ absolut stetig (als Summe von absolut stetigen Funktionen) Nach vorigem Lemma sind $v(.)$ und $v(.) - f(.)$ beide monoton wachsend. □

Die Bezeichnung "absolut stetig" für Maße $\mu \ll \lambda$ kommt von einer der gleichmäßigen Stetigkeit entsprechenden Eigenschaft für Funktionen.

Satz 21 (3.DS)**.** *μ sei ein LS-Maß auf $(\mathbb{R}, \mathcal{B})$ mit Verteilungsfunktion F. $[a,b] \subseteq \mathbb{R}$.*
$\mu \ll \lambda$ ist äquivalent mit:

1. *ϵ, δ-Kriterium: Zu $\epsilon > 0$ existiert ein $\delta > 0$, sodass aus $\lambda(A) < \delta \implies \mu(A) < \epsilon$, wenn μ endlich ist, $A \in \mathcal{B}$.*

2. *$F|_{[a,b]}$ ist Verteilungsfunktion von μ auf $([a,b], \mathcal{B}|_{[a,b]})$ ist absolut stetig.*

Proof. 1. Wenn das ϵ, δ Kriterium gilt und $\lambda(A) = 0$ aber $\mu(A) > 0$, dann ist für $\epsilon = \frac{\mu(A)}{2}$, zwar $\lambda(A) < \delta$ für alle $\delta > 0$ aber es gilt nicht $\mu(A) < \epsilon$.

Es sei $\mu \ll \lambda$. Da μ endlich ist (und daher natürlich sigma-endlich) existiert eine RN-Dichte f mit $\mu(A) = \int_A f d\lambda$.

Da μ endlich $\lambda([f = \infty]) = 0$ $\lambda([f > m]) \to 0$ für $m \to \infty$. Mit $\Omega = [a,b]$ folgt

$$\int_{\Omega} f 1_{[f>m]} d\lambda \to 0$$

wegen der Stetigkeit von oben des Maßes μ. Für $\epsilon > 0$ wähle m mit

$$\int_\Omega f 1_{[f>m]} d\lambda < \frac{\epsilon}{2}$$

und $\delta < \frac{\epsilon}{2m}$, dann gilt

$$\mu(A) = \int_A f 1_{[f>m]} d\lambda + \int_A f 1_{[f\leq m]} d\lambda \leq \frac{\epsilon}{2} + m\lambda(A), \text{ wenn}$$
$$\lambda(A) < \delta \implies \mu(A) < \epsilon.$$

Bemerkung: Teil 1) gilt für beliebige (endliche) Maße μ, ν mit $\mu \ll \nu$.

2. Wenn $\mu \ll \lambda$ gilt nach dem ϵ, δ Kriterium für $A = \bigcup_{i=1}^n (a_i, b_i]$ mit $\lambda(A) = \sum_{i=1}^n (b_i - a_i) < \delta$

$$\mu(A) = \sum_{i=1}^n F(b_i) - F(a_i) = \sum |F(b_i) - F(a_i)| < \epsilon$$

und F ist absolut stetig.

Sei F absolut stetig und $A \in \mathcal{B} \cap [a, b]$.

λ ist (auch) das äußere Maß, somit gilt

$$\lambda(A) = \inf \left\{ \sum_{i=1}^\infty (b_i - a_i) \middle| A \subseteq \bigcup_{i=1}^\infty (a_i, b_i] \right\}$$

Sei $A = N$ eine λ-Nullmenge, $\lambda(N) = 0$, dann existieren disjunkte Intervalle $(a_i, b_i]$ mit $N \subseteq \bigcup_i (a_i, b_i]$ und $\sum_i (b_i - a_i) < \delta$

Für jede endliche Auswahl gilt $\sum_{i=1}^n (b_i - a_i) < \delta$ und nach Voraussetzung

$$\sum_{i=1}^n F(b_i) - F(a_i) < \epsilon$$

für jedes $\epsilon > 0$. Das bedeutet

$$\mu(N) \leq \mu\left(\bigcup_{i=1}^\infty (a_i, b_i]\right) \leq \sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon$$

N ist auch eine μ-Nullmenge.

□

Satz 22 (4.DS)**.** *Eine Funktion $F : [a, b] \to \mathbb{R}$ besitzt genau dann eine Integraldarstellung*

$$F(x) = \int_{(a,x]} f d\lambda + c, x \in [a, b]$$

mit $f \in \mathcal{L}_1(\lambda|_{[a,b]})$, wenn F absolut stetig ist.

Proof. F als Verteilungsfunktion eines signierten Maßes aufgefasst erklärt den Zusammenhang:

Ist F absolut stetig, so existieren monoton wachsende Funktionen G, H mit $F = G - H$ und beide G, H sind absolut stetig.

Die zugehörigen LS-Maße seien μ_G, μ_H und beide $\mu_G \ll \lambda, \mu_H \ll \lambda$ mit RN-Dichten

$$g := \frac{d\mu_G}{d\lambda} \text{ und } h := \frac{d\mu_H}{d\lambda}, \text{ somit ist}$$
$$F(x) - F(a) = \mu_G((a, x]) - \mu_H((a, x]) - F(a)$$
$$\implies F(x) = F(a) + \int_{(a,x]} g - h d\lambda$$

Besitzt F die Integraldarstellung mit f

$$F(x) = F(a) + \int_{(a,x]} f d\lambda$$

dann ist f^+ die Dichte des Maßes μ_G, f^- die von H und $\mu_G \ll \lambda, \mu_G \ll \lambda$. Die Verteilungsfunktionen beider Maße G, H sind absolut stetig.

Für $x \in (a, b]$ ist

$$F(x) = G(x) - G(a) - (H(x) - H(a))$$

und F ist die Differenz zweier absolut stetigen Funktionen und auch absolut stetig. □

Für die Klärung, in wie weit die Integraldarstellung mit der Ableitung $f = F'$ gelingt wird zuerst die Lebesgue-Zerlegung von μ bezüglich λ untersucht

$$\mu = \mu_c + \mu_s, \mu_s \perp \lambda, \mu_c \ll \lambda$$

Für folgenden Satz, der den Zusammenhang zwischen Nullmengen und Ableitung angibt, wird ein Hilfssatz benötigt.

Lemma 7. *$A \in \mathcal{B}|_{[a,b]}$ für ein endliches Intervall $[a, b]$ und μ sei ein endliches Maß auf $\mathcal{B}|_{[a,b]}$. Wenn $\mu(A) = 0$, dann ist für $n \in \mathbb{N}$*

$$\lambda \left(x \in A \middle| \overline{\lim_{h \searrow 0}} \frac{\mu((x-h, x+h])}{2h} > \frac{1}{n} \right) = 0$$

Proof. ohne Beweis. Der Beweis ist eine Folgerung des Überdeckungslemmas von Vitali, Deteils findet man in Elstrodt, "Maß- und Integrationstheorie". □

Satz 23 (5.DS). *μ sei ein endliches Maß auf $\mathcal{B}|_{[a,b]}$ mit Verteilungsfunktion $F(x) = \mu([a, x])$. Wenn $\mu(A) = 0$, dann ist $F' = 0$ λ-f.ü. auf A.*

Proof. F ist wachsend, $F' = 0$ bedeutet

$$\frac{F(x+h) - F(x-h)}{2h} = \frac{\mu((x-h, x+h])}{2h} \to 0$$

wenn $h \to 0$ für λ-f.ü. Punkte $x \in A$.

Wir zeigen

$$\overline{\lim_{h \to 0}} \frac{\mu(x-h, x+h]}{2h} = \overline{\lim_{k \to \infty}} \frac{\mu(x - \frac{1}{k}, x + \frac{1}{k})}{\frac{2}{k}} \tag{1}$$

sonst könnte die Menge

$$A_n := \left\{ x \in A \middle| \overline{\lim_{h \searrow 0}} \frac{\mu(x-h, x+h]}{2h} > \frac{1}{n} \right\}$$

nicht unbedingt messbar sein, $A_n \in \mathcal{B}$.

(1) gilt, da für $\frac{1}{k+1} < h \leq \frac{1}{k}$

$$\frac{\mu(x - \frac{1}{k+1}, x + \frac{1}{k+1}]}{\frac{1}{k+1}} \frac{\frac{1}{k+1}}{\frac{1}{k}} \leq \frac{\mu(x-h, x+h]}{h} \leq \frac{\mu(x - \frac{1}{k}, x + \frac{1}{k}]}{\frac{1}{k}} \frac{\frac{1}{k}}{\frac{1}{k+1}}$$

Für $k \to \infty$ gilt mit $\frac{k+1}{k} \to 1$ auch (1) ($\overline{\lim}$ von abzählbaren Funktionen ist messbar).

Nach vorigem Lemma ist $\lambda(A_n) = 0$ und daher auch der Grenzwert $n \to \infty$, also

$$\lambda \left(\bigcup_{n=1}^{\infty} A_n \right) = 0$$

und $F' = 0$ λ-f.ü. □

Satz 24 (6.DS). *μ sei ein endliches Maß auf $([a,b], \mathcal{B}|_{[a,b]})$ und singulär zu λ, $\mu \perp \lambda$. Die Verteilungsfunktion F_μ von μ erfüllt $F' = 0$ λ-f.ü. auf $[a,b]$.*

Proof. Wenn $\mu \perp \lambda$ existiert ein $A \in \mathcal{B}|_{[a,b]}$ mit $\mu(A) = 0$ und $\lambda(A^c) = 0$. Nach Satz 5.DS gilt $F' = 0$ λ-f.ü. auf A. Da $\lambda(A^c) = 0$ gilt das auch auf $[a,b]$ λ-f.ü. □

Beispiel. *Wir haben die Cantor-Menge*

$$C = \{x \in [0,1] | x = \sum_{i=1}^{\infty} \frac{x_i}{3^i}, x_i = 0 \vee 2\}$$

bereits behandelt, mit $\lambda(C) = 0$ und die Funktion für $x \in C$

$$F_C(x) = \sum_{i=1}^{\infty} \frac{\frac{x_i}{2}}{2^i}$$

bzw. die Fortsetzung auf $[0,1]$

$$\tilde{F_C}(x) = \begin{cases} F_C(x) & x \in C \\ \sup\{F_C(y) | y \in C, y \leq x\} & x \notin C \end{cases}$$

diskutiert. $\tilde{F_C}$ ist monoton und stetig. Daher gehört $\tilde{F_C}$ als Verteilungsfunktion zu einem Maß (sogar einem Wahrscheinlichkeitsmaß).

Die Verteilung heißt Cantor-Verteilung.

$\tilde{F_C}$ ist eine stetige Verteilungsfunktion eines zu λ singulären Maßes (Der Vollständigkeit wegen sei $\tilde{F_C}(x) = 0, x \leq 0$ und $\tilde{F_C}(x) = 1, x > 1$).

Die bisherigen Sätze sind für Verteilungsfunktionen auf endlichen Intervallen formuliert. Durch Erweiterung auf $\mathbb{R}$ (stückweise Verknüpfung) bleiben die Aussagen für VF bzw. LS-Maße auf $(\mathbb{R}, \mathcal{B})$ erhalten.

So ist beispielsweise die VF F eines LS-Maßes μ auf jeder μ-Nullmenge differenzierbar (λ-f.ü.) und $F' = 0$ λ-f.ü.

Zur Klärung, ob F' die Dichte des LS-Maßes ist, betrachten wir die einseitigen Ableitungen.

$$\text{rechtsseitig} \begin{cases} \delta^r f(x) := \lim\limits_{n\to\infty} \sup_{x<y<x+1/n} \frac{f(y)-f(x)}{y-x} \\ \delta_r f(x) := \lim\limits_{n\to\infty} \inf_{x<y<x+1/n} \frac{f(y)-f(x)}{y-x} \end{cases}$$

$$\text{linksseitig} \begin{cases} \delta^l f(x) := \lim\limits_{n\to\infty} \inf_{x-1/n<y<x} \frac{f(y)-f(x)}{y-x} \\ \delta_l f(x) := \lim\limits_{n\to\infty} \sup_{x-1/n<y<x} \frac{f(y)-f(x)}{y-x} \end{cases}$$

Satz 25 (7.DS). *Für das LS-Maß μ mit $\mu \ll \lambda$ gilt $F' = \frac{d\mu}{d\lambda}$ λ-f.ü, wobei F die VF von μ ist.*

Proof. Es sei $\bar{\delta F}$ die "obere" Ableitung $\bar{\delta F} = \max(\delta^r F, \delta^l F)$.

Die Menge $[\bar{\delta F} > \frac{d\mu}{d\lambda}]$ soll eine λ-Nullmenge sein, d.h. für alle $q \in \mathbb{Q}$ soll

$$\lambda[\bar{\delta F} > q > \frac{d\mu}{d\lambda}] = 0$$

sein.

Dazu betrachten wir

$$\nu(A) := \int_{A \cap [\frac{d\mu}{d\lambda} \geq q]} \left(\frac{d\mu}{d\lambda} - q \right) d\lambda$$

ν ist ein LS-Maß und $\nu(\frac{d\mu}{d\lambda} < q) = 0$. Nach Satz 5.DS gilt für die VF F_ν von ν

$$F'_\nu = 0 \text{ auf } [\frac{d\mu}{d\lambda} < q] \text{ } \lambda\text{-f.ü.}$$

Das signierte Maß γ erfüllt

$$\gamma(A) := \mu(A) - \lambda(A) = \int_A \left(\frac{d\mu}{d\lambda} - q \right) d\lambda \leq \int_{A \cap [\frac{d\mu}{d\lambda} \geq q]} (\frac{d\mu}{d\lambda} - q) d\lambda = \nu(A)$$

Daher folgt für $x, y \in \mathbb{R}$

$$
\begin{aligned}
\gamma((\min(x,y), \max(x,y)]) &= \mu((\min, \max]) - \lambda(\min, \max)q \\
\implies \frac{F(y)-F(x)}{y-x} - q &\leq \frac{F_\nu(y) - F_\nu(x)}{y-x} \\
\implies \bar{\delta F} - q &\leq F'_\nu = 0 \text{ auf } [\frac{d\mu}{d\lambda} < q] \\
\text{d.h. } \lambda\left(\frac{d\mu}{d\lambda} < q < \bar{\delta F}\right) &= 0
\end{aligned}
$$

Dieselbe Vorgangsweise für die "untere" Ableitung angewandt auf $\delta(\bar{-F})$ liefert

$$
\begin{aligned}
&\lambda\left(\underline{\delta F} < \frac{d\mu}{d\lambda}\right) = 0 \\
\implies &F' = \frac{d\mu}{d\lambda} \lambda\text{-f.ü.}
\end{aligned}
$$

□

Die VF F eines beliebigen LS-Maßes μ auf $(\mathbb{R}, \mathcal{B})$ wurde bereits in eine stetige VF F_0 und eine diskrete VF F_d zerlegt. Nach der Lebesgue-Zerlegung des Maßes μ_0 (von F_0 erzeugt) $\mu_0 = \mu_c + \mu_s$ $\mu_c \ll \lambda$ und $\mu_s \perp \lambda$ mit VFen $F = F_c + F_s + F_d$.

Die vorigen Ergebnisse zusammengefasst bedeutet, F ist λ-f.ü. differenzierbar

$$
F' = \underbrace{F'_c}_{=\frac{d\mu}{d\lambda}} + \underbrace{F'_s}_{=0} + (F'_d)\lambda\text{-f.ü.}
$$

Ein LS-Maß μ mit F für das $F' = 0$ λ-f.ü. gilt, ist zu λ singulär. Das folgt daraus, dass der absolut stetige Teil μ_c die Dichte $F' = \frac{d\mu_c}{d\lambda}$ λ-f.ü. besitzt.

$$
\mu_c(A) = \int_A \underbrace{\frac{d\mu_c}{d\lambda}}_{=0\lambda\text{-f.ü.}} d\lambda = 0
$$

Aus der Zerlegung ergibt sich, dass für jede monoton wachsende Funktion

$$
\underbrace{F(b) - F(a)}_{\mu(a,b]} \geq \underbrace{\int_{[a,b]} F' d\lambda}_{\mu_c([a,b])}
$$

gilt und F ist λ-f.ü. differenzierbar.

Wenn F keine VF ist (also nicht rechtsstetig), wird F an den Unstetigkeitsstellen (sind höchstens abzählbar viele) geändert, ohne dass sich bezüglich λ etwas ändert, (F' λ-f.ü. oder $\int F' d\lambda$ etc).

Damit sind aber auch Funktionen von beschränkter Variation λ-f.ü. differenzierbar, da sie als Differenz monoton wachsender Funktionen darstellbar ist.

Die Aussagen zusammen ergeben

Satz 26 (8.DS (Satz von Lebesgue für Differenzierbarkeit))**.** $F : [a,b] \to \mathbb{R}$ *sei von beschränkter Variation. Dann ist* F λ*-f.ü. differenzierbar.*
Ist F *monoton steigend, dann ist* F λ*-f.ü. differenzierbar und*

$$
F(b) - F(a) \geq \int_{[a,b]} F' d\lambda
$$

Die einzelnen Ergebnisse dieses Abschnitts münden in der Formulierung des Hauptsatzes der Differential- und Integralrechung. Die gewohnte Form des Hauptsatzes für das Riemann-Integral wird die absolute Stetigkeit nicht benötigt, dagegen gilt der Hauptsatz für $\mathcal{L}_1$-Funktionen in folgender Form.

Satz 27 (9.DS (Hauptsatz der DI-Rechnung für das Lebesgue-Integral))**.** *$f : [a, b] \to \mathbb{R}$ sei $\in \mathcal{L}_1$, also λ-integrierbar. Dann ist*

$$F(x) := \int_{[a,x]} f d\lambda$$

absolut stetig und λ-f.ü. differenzierbar mit $F' = f$ λ-f.ü.
Ist $F : [a, b] \to \mathbb{R}$ absolut stetig, dann existiert λ-f.ü. F' mit

$$F(x) - F(a) = \int_{[a,x]} F' d\lambda, x \in [a, b]$$

Proof. F kann zerlegt werden in

$$F(x) = F_1(x) - F_2(x) = \int_{[a,x]} f^+ d\lambda - \int_{[a,x]} f^- d\lambda$$

F_1, F_2 sind monoton und $\mu^+(A) = \int_A f^+ d\lambda$ ist absolut stetig bezüglich λ und $F_1' = \frac{d\mu^+}{d\lambda} = f^+$ genauso mit f^-.

Ist F absolut stetig, dann ist das (signierte) Maß absolut stetig mit $F' = \frac{d\mu}{d\lambda}$ und $\int_A F' d\lambda = \int_A \frac{d\mu}{d\lambda} d\lambda = \mu(A)$ mit $A = [a, x]$ ergibt den Hauptsatz. □

3.1 Weitere Eigenschaften der Dichten

Im folgenden seien die behandelten Maße als sigma-endlich angenommen.

- Wenn $\nu_i \ll \mu, i = 1, ..., n$, dann ist $\tau := \sum_{i=1}^n \nu_i$ ein sigma-endliches Maß, $\tau \ll \mu$ und die Dichte ist auch die Summe

$$\frac{d\tau}{d\mu} = \sum_{i=1}^{n} \frac{d\nu_i}{d\mu} \mu\text{-f.ü.}$$

- T sei $\Omega \to \Omega'$ und $\mathcal{A}/\mathcal{A}'$ messbar. Wenn $\nu \ll \mu$, dann sind die induzierten Maße ν^T und μ^T auf $(\Omega', \mathcal{A}')$ auch $\nu^T \ll \mu^T$ absolut stetig.

  Da für $B \in \mathcal{A}'$ mit $\mu^T(B) = 0$

$$\mu^T(B) = \mu(T^{-1}(B)) = 0 \implies \nu(T^{-1}(B)) = 0.$$

  μ^T, ν^T sind auch endlich, da für $B_n \nearrow \Omega'$ und μ, ν endlich sind.

  Es kann aber die Dichte $\frac{d\nu^T}{d\mu^T}$ keinerlei Verwandtschaft zu $\frac{d\nu}{d\mu}$ haben. Das hängt davon ab, in wieweit T Nullmengen in Nullmengen abbildet.

- μ auf $(\mathbb{R}, \mathcal{B})$ sei $\mu \ll \lambda$ und T eine messbare (sogar monotone) Abbildung. Dann ist μ^T i.a. nicht mehr $\mu^T \ll \lambda$ absolut stetig.

  Betrachtet man beispielsweise die Cantor-Funktion $F_C : C \to [0, 1]$ bzw. $F_C^{-1} : [0, 1] \to C$

$$\lambda^{F_C^{-1}}(C) = \lambda([0, 1]) = 1 \text{ aber } \lambda(C) = 0.$$

- Allgemein werden Nullmengen nicht von bijektiven Transformationen erhalten.

  Beispiel. $\Omega = \{1, 2\}, \mathcal{A} = 2^\Omega, T(x) = 3 - x, T : \Omega \to \Omega$ *bijektiv, sei* $\mu(1) = 1, \mu(2) = 0$*, dann ist* $A = \{2\}$ *eine Nullmenge, aber* $\mu^T(A) = \mu(T^{-1}(\{2\})) = \mu(1) = 1$

  Für bijektive Abbildungen $T : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ kann die Dichte dargestellt werden.

  Genauer muss T nur eine "Links-Inverse" haben. Es existiere ein $\tilde{T} : (\Omega', \mathcal{A}') \to (\Omega, \mathcal{A})$ mit $\tilde{T} \circ T = id_\Omega$, also $id_\Omega(\omega) = \omega, \omega \in \Omega$. Dann kann mit dem allgemeinen Transformationssatz die Dichte von ν^T bzw. μ^T bestimmt werden.

Satz 28 (10.DS). *ν, μ sind sigma-endlich und $\nu \ll \mu$. $T : \Omega \to \Omega'$ besitzt ein $\tilde{T}$ mit $\tilde{T} \circ T = id_\Omega$ dann gilt $\nu^T \ll \mu^T$ mit der Dichte*

$$\frac{d\nu^T}{d\mu^T} = \frac{d\nu}{d\mu} \circ \tilde{T} \mu^T\text{-f.ü.}$$

Proof. $\nu^T \ll \mu^T$ ist klar. T ist injektiv, wenn $T(x) = T(y) \implies x = \tilde{T} \circ T(x) = \tilde{T} \circ T(y) = y$.
$\frac{d\nu}{d\mu} = f$, dann gilt für $B' \in \mathcal{A}'$

$$\nu^T(B') = \nu(T^{-1}(B')) = \int_{T^{-1}(B')} f d\mu = \int 1_{T^{-1}(B')}(\omega) f(\omega) d\mu = \int 1_{B'}(T(\omega)) f \circ \tilde{T} \circ T(\omega) d\mu(\omega)$$

nach dem Transformationssatz

$$= \int 1_{B'}(\omega') f \circ \tilde{T}(\omega') d\mu^T$$

Also ist $\frac{d\nu^T}{d\mu^T} = f \circ \tilde{T}$ □

Betrachtet man Maße $\mu \ll \lambda$, dann ist auch $\mu^T \ll \lambda$, wenn T λ-Nullmengen immer in λ-Nullmengen abbildet.

Sei $T : \mathbb{R} \to \mathbb{R}$ bijektiv und differenzierbar (λ-f.ü.) und auch T^{-1} ist λ-f.ü. differenzierbar $T' \neq 0$ λ-f.ü. und $T^{-1} \neq 0$. T sei wachsend

$$\lambda^T((a, b]) = \lambda((T^{-1}(a), T^{-1}(b)]) = \int_{[T^{-1}(a), T^{-1}(b)]} d\lambda$$

Auf $(a, b]$ sei $T' \neq 0, T^{-1'} \neq 0$, dann ist

$$\lambda^T(a, b] = \int 1_{(T^{-1}(a), T^{-1}(b)]} d\lambda(s = T(\omega)) = \int 1_{(a,b]}(s) |T^{-1'}(s)| d\lambda(s)$$

d.h. die Dichte von λ^T ist $|T^{-1'}(s)| = \frac{d\lambda^T}{d\lambda}$.

Die Bestimmung der Dichte einer Transformation einer stochastischen Größe $Y = T(X)$, wobei die Verteilung von X, $P^X \ll \lambda$ mit F_X als Verteilungsfunktion, erfolgt über die Verteilungsfunktion von X. T sei wieder differenzierbar mit $T' \neq 0$ (auf $[a, b]$) und wachsend.

$$F_Y(y) = P[Y \in (-\infty, y]] = P[T(X) \in (-\infty, y]] = P[X \in (-\infty, T^{-1}(y))] = F_X(T^{-1}(y))$$

und die Dichte von Y ist λ-f.ü.

$$f_Y(y) = f_X(T^{-1}(y)) |T^{-1'}(y)|$$

Diese beiden Beispiele für die Bestimmung von Dichten können auf $\mathbb{R}^k$ verallgemeinert werden.

$T : \mathbb{R}^k \to \mathbb{R}^k$ sei ein Diffeomorphismus, also T ist bijektiv und die Jakobi-Determinanten

$$\left| det\left(\frac{\delta T_i}{\delta x_j}(x)\right)\right| \neq 0, x \in \mathbb{R}^k \text{ und}$$
$$\left| det\left(\frac{\delta T_i'}{\delta y_j}(y)\right)\right| \neq 0, y \in \mathbb{R}^k$$

dann ist $\lambda^T \ll \lambda$ (hier ist $\lambda = \lambda_k$ das k-dimensionale Lebesgue-Maß) und die Dichte ist

$$\frac{d\lambda_k^T}{d\lambda_k} = \left| det\left(\frac{\delta T_i^{-1}}{\delta y_j}\right)(y)\right|.$$

Wenn $X \in \mathbb{R}^k$ eine SG mit $P^X \ll \lambda_k$ und Dichte $\frac{dP^X}{d\lambda_k} = f_X$, T sei bijektiv, diffbar und $|det T^{-1}| \neq 0$, dann gilt für $Y = T \circ X$ $P^Y \ll \lambda_k$ mit der Dichte

$$f_Y(y) = f_X(T^{-1}(y)) \left| det\left(\frac{\delta T_i^{-1}}{\delta y_j}\right)(y)\right|.$$

3.1.1 Stückweise Anwendung

$\mathbb{R}^k = \bigcup_{i=1}^n B_i$ und $\underbrace{T|_{B_i}}_{T_i}$ hat obige Eigenschaften.

$$|J_i^{-1}| := \left| det \frac{\delta T_i^{-1}}{\delta x_j} \right| \neq 0 \text{ auf } B_i \text{ für}$$

$$T := \sum_{i=1}^n T_i 1_{B_i} \text{ ist die Dichte von } Y = T \circ X$$

$$f_Y(y) = \sum_{i=1}^n f(T_i^{-1})|J_i^{-1}|$$

3.2 Illustrative Beispiele

1.

$$\mu = \xi_{\mathbb{N}}, X \sim B_{N,p} = \nu, \nu(k) = P[X = k] = \binom{N}{k} p^k (1-p)^{N-k}, k \leq N$$

$$T : x \mapsto N - x, x \leq N, (Tx = x, x \geq N)$$

$$Y = T \circ X = N - X$$

$$P[Y = y] = P[T(X) = y] = P[N - X = y] = P[X = N - y] = \nu(N - y) = \underbrace{\binom{N}{N-y}}_{\binom{N}{y}} P^{N-y}(1-p)^y$$

$$Y \sim B_{N,1-p} \text{ (Anzahl der ”0”)}$$

2.

$$X \sim \mathcal{N}_k(\mu, \Sigma), f_X(x) = \frac{1}{(2\pi)^{k/2}\sqrt{det\Sigma}} exp\{-1/2(x-\mu)^T\Sigma^{-1}(x-\mu)\}$$

$$A \text{ reguläre Matrix}, Y = AX, f_Y(y) = f_X(A^{-1}y)|A^{-1}|$$

$$f_Y(y) = \frac{1}{(2\pi)^{k/2}\sqrt{det\Sigma}} exp\{-1/2 \quad \underbrace{(A^{-1}y - A^{-1}A\mu)^T\Sigma^{-1}()}_{(y-A\mu)^T \underbrace{(A^{-1}\overset{-1}{\Sigma^{-1}} A^{-1})}_{(A\Sigma A^T)^{-1} = \tilde{\Sigma}} (y-A\mu)} \quad \}|A^{-1}|$$

$$\text{Wegen } \frac{|A^{-1}|}{\sqrt{det\Sigma}} = \frac{1}{\sqrt{(detA)^2 det\Sigma}} = \frac{1}{\sqrt{det\tilde{\Sigma}}}$$

$$Y \sim \mathcal{N}(A\mu, A\Sigma A^T)$$

3. $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y = e^X$ log-Normalverteilung, $X = \log(Y)$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{logY-\mu}{\sigma^2})} \frac{1}{|y|}, y \geq 0$$

4. (X_1, X_2) u.a. exponential

$$f(x_1, x_2) = e^{-x_1-x_2}, x_1, x_2 \geq 0$$

$$Y = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 \end{pmatrix} = T \circ X, y_1 = x_1^2 + x_2^2, y_2 = x_1, x_1 = y_2, x_2 = \sqrt{y_1 - y_2^2}, y_1 \geq y_2^2$$

$$|J| = \begin{vmatrix} 2x_1 & 2x_2 \\ 1 & 0 \end{vmatrix} = |-2x_2| = 2|x_2|$$

$$|J^{-1}| = \frac{1}{2x_2} = \frac{1}{2\sqrt{y_1 - y_2^2}}$$

$$\text{Dichte von } Y : f_Y(y) = e^{-y_2 - \sqrt{y_1 - y_2^2}} \frac{1}{2\sqrt{y_1 - y_2^2}} \text{ für } y_1 > y_2^2 \geq 0$$

5. $X \sim \mathcal{N}(0,1)$, $Y = X^2$ stückweise invertierbar auf $\mathbb{R}^+$

$$T_1^{-1}(y) = \sqrt{y}, y \geq 0, B_1 = (0, \infty)$$
$$T_2^{-1}(y) = -\sqrt{y}, y < 0, B_2 = (-\infty, 0]$$
$$\text{allgemein gilt dann} f_Y(y) = f_X(\sqrt{y})\left|\frac{1}{2\sqrt{y}}\right| + f_X(-\sqrt{y})\left|-\frac{1}{2\sqrt{y}}\right|$$
$$\text{bei symmetrischer Dichte von } X, \text{ also } f_X(\sqrt{y}) = f_X(-\sqrt{y})$$
$$f_Y(y) = \frac{1}{\sqrt{y}} f_X(\sqrt{y})$$
$$\mathcal{N}(0,1) \text{ für } f_X : f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} \frac{1}{\sqrt{y}} \sim \gamma(\frac{1}{2}, \frac{1}{2})$$
$$\chi^2\text{-Verteilung mit Freiheitsgrad 1}$$

4 L_p-Räume

Der Raum der integrierbaren Funktionen bilden einen Vektorraum $\mathcal{L}(\mu)$ (genauer $(\mathcal{L}(\mu), +, \mathbb{R})$) mit $||f|| = \int |f| d\mu$ sogar einen normierten Raum, also einen metrischen Raum, wenn man $f = g[\mu]$ als gleich ansieht.

Def (Pseudonorm, Pseudometrik)**.** $||.||$ *bildet eine Pseudonorm (=Halbnorm), es gilt also*

$$||\alpha f|| = |\alpha| ||f||$$
$$||f + g|| \leq ||f|| + ||g|| \triangle\text{-}\textit{Ungleichung (damit bereits } ||f|| \geq 0 \textit{ und } ||0|| = 0)$$

Norm: $||f|| = 0 \implies f = 0$ *Das ist vorläufig nicht erfüllt.*
$||.||$ *führt zu einer Pseudometrik, also* $d(x,y) := ||y - x||$ *mit*

$$d(x,y) = d(y,x); d(x,x) = 0; \text{ und } d(x,y) \leq d(x,z) + d(z,y)$$

Metrik: $d(x,y) = 0 \iff x = y$

Die Funktionenräume der p-fach integrierbaren Funktionen bilden wesentliche, abstrakte Funktionenräume deren Analyse wichtig ist (besonders L_2).

Definition 14 (p-fach integrierbar)**.** f *heißt p-fach integrierbar für* $1 \leq p < \infty$ *wenn*

$$||f||_p := \left(\int |f|^p d\mu\right)^{1/p} < \infty$$

Raum: $\mathcal{L}^p(\mu)$
Wenn $f = g[\mu]$ als gleich angesehen wird, bildet dieser Raum einen normierten Vektorraum.
Wenn eine Funktion beschränkt ist, dann ist auch

$$\int |f|^p d\mu < \infty \forall p \geq 1$$

Es sei daher $||f||_\infty := \inf\{K \geq 0 | \mu(|f| > k) = 0\} = ess \sup |f|$ essentielles Supremum.
$||f||_\infty$ bildet genauso eine Pseudonorm.
Eine Norm (bzw. Metrik) entsteht durch Übergang zum Quotientenraum. Es sei

$$\mathcal{N} := \{f \text{ messbar und } f = 0[\mu] \text{ f.ü.}\}$$

Definition 15 (Lp-Raum)**.**

$$L^p(\Omega, \mathcal{A}, \mu) := \mathcal{L}^p(\mu)/\mathcal{N} = \{\bar{f} = f + g; f \in \mathcal{L}^p, g \in \mathcal{N}\}$$

L^p enthält die Repräsentanten und alle Elemente aus einer Klasse haben die gleiche Norm: $||f||_p = ||g||_p$, wenn $f = g[\mu]$.

Um zu zeigen, dass $L^p(\mu)$ wirklich ein normierter Vektorraum ist, benötigen wir noch die $\triangle$-Ungleichung, die gleich folgt.

Die Norm (bzw. Metrik) ergibt eine Konvergenz Definition im L^p.

Definition 16 (Konvergenz im p-ten Mittel). *f_i sei eine Folge in $\mathcal{L}^p(\mu)$.*

$$f_n \to^{L^p} f \text{ wenn } ||f_n - f||_p \to 0, n \to \infty.$$

Für SG X in einem W-Raum $(\Omega, \mathcal{A}, P)$ heißt

$$\int X dp = \mathbb{E}X \text{ Erwartung}$$

$$\text{und } \int X^k dP = \mathbb{E}X^k \text{ k-tes Moment}$$

$$\text{sowie falls } \mathbb{E}X < \infty\text{: } \mathbb{E}|X - \mathbb{E}X|^k \text{ k-tes zentrales Moment}$$

Besondere Rolle spielt das 2. zentrale Moment:

$$\text{Varianz } \mathbb{E}(X - \mathbb{E}X)^2 \text{ und } ||X - \mathbb{E}X||_2 \text{ Streuung}$$

Beispiel. *$X \sim \mathcal{N}(\mu, \sigma^2)$ Normalverteilung besitzt alle Momente, da*

$$\forall p : \mathbb{E}X^p = \int_{-\infty}^{\infty} x^p \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2}} dx < \infty$$

Def. *Die Folge von SG X_i konvergiert gegen X im p-ten Mittel, wenn $\mathbb{E}|X_n - X|^p \to 0$.*
Schreibweise: $X_n \to^{L^p} X$.

Aus der Konvergenz der Integrale folgt natürlich i.a. nicht, dass die Funktionen konvergieren $[\mu]$.

Beispiel. *X_n ... SG mit Wahrscheinlichkeitsmaß $P[X_n = 0] = 1 - \frac{1}{n^{p+1}} = 1 - P[X = n]$ auf $\{0, n\}$. Es sei $X_0 \sim \delta_0$. Dann gilt*

$$\mathbb{E}|X_n - X_0|^p = n^p \frac{1}{n^{p+1}} = \frac{1}{n} \to 0 \implies X_n \to^{L^p} X_0$$

Aber limsup$X_n = \infty$ also $X_n \nrightarrow X_0$ P-f.s. $\lim X_n$ existiert nicht.

Auch, wenn die Funktionen beschränkt sind, muss nicht $f_n \to^{L^p} f$ folgen.

Beispiel.

1. *$f_n := 1_{[}a_n, b_n]$ mit $a_n = \frac{n - \lfloor\sqrt{n}\rfloor^2}{2\lfloor\sqrt{n}\rfloor + 1}$ und $b_n = \frac{n+1-\lfloor\sqrt{n}\rfloor^2}{2\lfloor\sqrt{n}\rfloor+1}$ bildet eine Folge auf $([0,1], \mathcal{B}|_{[0,1]}, \lambda|_{[0,1]})$ die nicht punktweise (λ-f.ü.) konvergiert.*

   Da $b_n - a_n = \frac{1}{2\lfloor\sqrt{n}\rfloor+1} \to 0$ gilt $||f_n - 0||_p \to 0 \forall p \geq 1$.

   Dass die Umkehrung zwischen punktweiser und L_p-Konvergenz auch nicht uneingeschränkt gilt, scheint überraschender zu sein.

2. *$f_n := \frac{1}{n^{1/p}} 1_{[0,n]}$ konvergiert glm. $f_n \to 0$ aber*

$$\int f_n^p d\lambda = \frac{1}{n} \int_{[0,n]} d\lambda = 1$$

$$||f_n||_p \nrightarrow ||f_0||_p, f_0 = 0$$

In dem letzten Beispiel war $||f_n||_p \to ||f_0||$ nicht erfüllt.

Fügt man weitere Bedingungen ein, dann kann Äquivalenz zwischen $||f_n - f||_p \to 0$ und $f_n \to f \mu$-f.ü. erreicht werden.

In Beispiel 2) ist die Folge konvergent auch konvergent im Maß

$$\lambda(|f_n - f_0| > \epsilon) = 0$$

wenn für $\epsilon > 0 n$ gewählt wird, sodass $\frac{1}{n^{1/p}} < \epsilon \implies n > (\frac{1}{\epsilon})^p$.

Konvergenz im Maß garantiert nicht L_p-Konvergenz. Die Umkehrung gilt.

Satz 29 (1.LP). *Für $1 \leq p \leq \infty$ gelte $f_n \to^{L_p} f; ||f_n - f||_p \to 0$, dann gilt $\forall \epsilon > 0 : \lim_{n\to\infty} \mu(|f_n - f| > \epsilon) = 0, f_n \to^{\mu} f$*

Proof. Folgt unmittelbar aus der Markov-Ungleichung (Satz 1.UG) mit $\phi(x) = |x|^p$

$$\mu(|f_n - f| > \epsilon) \leq \frac{\int |f_n - f|^p d\mu}{\epsilon^p} = \frac{||f_n - f||_p^p}{\epsilon^p}$$

□

Def (vollständig)**.** *Ein metrischer Raum M heißt vollständig, wenn jede Cauchy-Folge konvergiert, d.h. Cauchy-Folge: $x_n \in M : \forall \epsilon > 0 \exists N_\epsilon \forall m, n \geq N_\epsilon : d(x_m, x_n) < \epsilon$ hat Grenzwert $x_0 \in M$.*

L^p-Raum für $1 \leq p \leq \infty$ ist vollständig (folgt später).

4.1 Wichtige Ungleichungen für p-integrierbare Funktionen

Lemma 8. *$p, q \in (1, \infty)$ und $\frac{1}{p} + \frac{1}{q} = 1$ (zueinander konjugiert) für $x, y \geq 0$ gilt $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$.*

Proof. $f(x) := \frac{x^p}{p} + \frac{y^q}{q} - xy$; $f' = x^{p-1} - y$; $f'' = (p-1)x^{p-2} \geq 0$ konvex mit Minimierung in $x_0 = y^{\frac{1}{p-1}}$ da $\frac{p}{p-1} = q$ gilt $x_0^p = y^q$ und $f(x_0) = 0 \implies f(x) \geq 0$ □

Satz 30 (Allgemeinerer Satz von Young (Young'sche Ungleichung))**.** *$f : \mathbb{R}^+ \to \mathbb{R}^+$, $f(0) = 0$, strikt monoton, $f(x) \to \infty, x \to \infty$, f stetig diffbar, also Umkehrfunktion $f^{-1} = g$. Mit den Stammfunktionen $F(x) = \int_0^x f(t)dt$, $G(y) = \int_0^y g(t)dt$ gilt $xy \leq F(x) + G(y)$.*

Siehe Skizze. Spezialfall $p = q = 2 : xy \leq \frac{x^2}{2} + \frac{y^2}{2} \iff (x-y)^2 \geq 0$

Satz 31 (Hölder'sche Ungleichung)**.** *Für $\frac{1}{p} + \frac{1}{q} = 1$ für $p \in [1, \infty]$ (wenn $p = 1$ ist $q = \infty$) gilt für $f \in \mathcal{L}^p, g \in \mathcal{L}^q$ (oder rechte Seite ∞ möglich)*

$$||f \cdot g||_1 \leq ||f||_p \cdot ||g||_q$$

und daher $f \cdot g \in \mathcal{L}^1$.

Proof. $p = 1$ bzw. $p = \infty$ unmittelbar (UE)

$p \in (1, \infty)$ Definiere $\tilde{f} := \frac{|f|}{||f||_p}, \tilde{g} := \frac{|g|}{||g||_q}$. Nach Young $\tilde{f}\tilde{g} \leq \frac{\tilde{f}^p}{p} + \frac{\tilde{g}^q}{q}$ eingesetzt ins Integral und Monotonie:

$$\int \tilde{f}\tilde{g} d\mu \leq \frac{1}{p} \frac{\int |f|^p d\mu}{||f||_p^p} + \frac{1}{q} \frac{\int |g|^q d\mu}{||g||_q^q} = 1$$

$$\implies \int |fg| d\mu \leq ||f||_p ||g||_q$$

□

Spezialfall: $f, g \in \mathcal{L}^2 \implies \int |fg| d\mu \leq ||f||_2 ||g||_2$. Für SG X, Y bezüglich $(\Omega, \mathcal{A}, P)$ mit 2. Moment: $\mathbb{E}XY \leq \mathbb{E}|X||Y| \leq \sqrt{\mathbb{E}X^2 \mathbb{E}Y^2}$.

Satz 32 (Minkowski-Ungleichung)**.** *$f, g \in \mathcal{L}^p(\mu), 1 \leq p \leq \infty$ Dann gilt:*

$$||f + g||_p \leq ||f||_p + ||g||_p$$

Proof. $p = 1$ oder $p = \infty$ Übung

$\infty > p > 1 :$

$$(f+g)^p = (f+g)(f+g)^{p-1} = f(f+g)^{p-1} + g(f+g)^{p-1}$$

$$|f+g|^p \leq |f| \cdot |f+g|^{p-1} + |g| \cdot |f+g|^{p-1} \text{ also durch integrieren}$$

$$||f+g||_p^p \leq \int |f| \cdot |f+g|^{p-1} + \int |g| \cdot |f+g|^{p-1} \text{mit Hölder}$$

$$\leq ||f||_p \cdot ||(f+g)^{p-1}||_q + ||g||_p \cdot ||(f+g)^{p-1}||_q + \int |g| \cdot |f+g|^{p-1} \text{ mit } q = \frac{p}{p-1}$$

$$= (||f||_p + ||g||_p) \cdot |||f+g|^{p-1}||_q + \int |g| \cdot |f+g|^{p-1} = (||f||_p + ||g||_p) \cdot ||f+g||_p^{p-1} + \int |g| \cdot |f+g|^{p-1}$$

$$|||f+g|^{p-1}||_q = \left(\int |f+g|^{(p-1)q}\right)^{\frac{p-1}{p}} = \left(\int |f+g|^p\right)^{\frac{p-1}{p}} = ||f+g||_p^{p-1}$$

Dividiert durch den letzten Wert ist

$$||f+g||_p = \frac{||f+g||_p^p}{||f+g||_p^{p-1}} \leq ||f||_p + ||g||_p$$

Die linke Seite ist definiert (also endlich) weil aus $f, g \in \mathcal{L}^p$ folgt

$$|f+g|^p \leq (2 \cdot \max(|f|, |g|))^p \leq 2^p \cdot \underbrace{\max(|f|^p, |g|^p)}_{\leq |f|^p + |g|^p}.$$

□

Mit der Minkowski-Ungleichung ist jetzt L^p ein Vektorraum ($f, g \in \mathcal{L}^p \implies f+g \in \mathcal{L}^p$) und $||f||_p$ eine Norm mit der Dreiecksungleichung.

Lemma 9 (Streuungsungleichung). *SG* $X, Y \in \mathcal{L}^p$. $SD(X+Y) \leq SD(X) + SD(Y)$

Die Bezeichnung L_∞ kann man aus folgendem Zusammenhang rechtfertigen:

Satz 33 (2.LP). *μ ... endlich, dann gilt*

$$\lim_{p\to\infty} ||f||_p = ||f||_\infty$$

(wenn f nicht p-fach integrierbar ist, gilt ∞ auf beiden Seiten).

Proof.

$$\int |f|^p d\mu \leq K^p \int d\mu = K^p \mu(\Omega) \qquad K := ess\sup_c\{c|\mu(|f|>c)=0\}(K=\infty \text{ möglich}) \qquad |f| \leq K[\mu]$$

$$\implies ||f||_p^p \leq ||f||_\infty^p \mu(\Omega) \qquad \implies ||f||_p \leq ||f||_\infty (\mu(\Omega)^{\frac{1}{p}})$$

$$\limsup_{p\to\infty} ||f||_p \leq ||f||_\infty \qquad \text{da } \limsup_{p\to\infty}(\mu(\Omega)^{\frac{1}{p}}) = 1$$

Die andere Ungleichung folgt aus der Markoff-Ungleichung $\phi(x) = x^p$

$$\int |f|^p d\mu \geq C^p \mu(|f| > c) \forall C > 0 \qquad \liminf_{p\to\infty} ||f||_p \geq C \cdot \liminf_{p\to\infty} \underbrace{(\mu(|f|>C))^{\frac{1}{p}}}_{>0 \text{ wenn } C<||f||_\infty = K}$$

$$\implies \liminf ||f||_p \geq C \forall 0 < C < ||f||_\infty \implies \liminf ||f||_p \geq ||f||_\infty$$

$$\text{Aus } ||f||_\infty \leq \liminf_p ||f||_p \limsup_p ||f||_p \leq ||f||_\infty \implies \lim_p ||f||_p = ||f||_\infty$$

□

Für endliche Maße gelten spezielle Aussagen, wie etwa:

Satz 34 (3.LP, Schachtelung der L^p-Räume). *μ sei endlich.* $1 \leq p \leq \tilde{p} \leq \infty$.

$$\mathcal{L}^{\tilde{p}}(\mu) \subset \mathcal{L}^p(\mu)$$

Also auf $||f||_{\tilde{p}} < \infty$ folgt $||f||_p < \infty$.

Proof. Es sei $\tilde{p} = \infty$, da $||f||_\infty = \int\{K > 0|\mu(|f| > K) = 0\}$ also $|f| \leq ||f||_\infty[\mu]$-f.s

$$\int |f|^p d\mu \leq \int ||f||_\infty^p d\mu = |f||_\infty^p \mu(\mathbb{R}) < \infty$$

Sonst $\tilde{p}$ bel. $p < \tilde{p}$: Da $|f|^p = |f|^p \cdot 1_{(|f|\leq 1)} + |f|^p 1_{(|f|>1)} \leq 1 + |f|^{\tilde{p}}$ und Minkowski-Ungl.

$$\int |f|^p d\mu \leq \int 1 d\mu + \int |f|^{\tilde{p}} d\mu = \mu(\Omega) + ||f||_{\tilde{p}}^{\tilde{p}} < \infty$$

□

Für sigma-endliche (aber nicht endliche) Maße gilt diese Schachtelung i.a. nicht. Es gibt sogar sigma-endliche Maße μ, wo die umgedrehte Schachtelung gilt $\mathcal{L}^p(\mu) \subseteq \mathcal{L}^{\tilde{p}}(\mu)$.

Weder die Schachtelung noch die Normkonvergenz gilt i.A. für nicht endliche Maße. Etwa $f = 1$ ist für λ: $||f||_p = 1 \forall p \geq 1, ||f||_\infty = \infty$.

Nach Satz 8.MS ist L_∞ ein Banachraum. Die Vollständigkeit wird für alle L_p-Räume nachgewiesen.

Satz 35 (4.LP, Satz von Riesz-Fischer)**.** *Eine L_p konvergente Folge $f_n \in \mathcal{L}_p, 1 \leq p \leq \infty$ mit Grenzfunktion f, $||f_n - f||_p \to 0$, führt auf $f \in \mathcal{L}_p$. $L_p(\Omega, \mathcal{A}, \mu)$ sind für jedes $p \geq 1$ Banachräume.*

Proof. $p = \infty$ ist Satz 8.MS. Für endliches $p \geq 1$ gehen wir den Umweg über die Konvergenz im Maß. f_n sei eine Cauchyfolge aus $\mathcal{L}_p$. $\implies$ f_n ist Cauchyfolge im Maß.

Nach Satz 10.MS existiert eine μ-fast glm. konvergente Teilfolge $f_{n_k} \to f$ μ-fast glm.

Daher gilt auch $f_{n_k} \to f$ μ-f.ü. Diese Teilfolge ist nach Voraussetzung auch L_p-konvergent.

Für $\epsilon > 0 \exists n_\epsilon : ||f_n - f_m||_p^p < \epsilon; n, m \geq n_\epsilon$ und für festes $n \geq n_\epsilon ||f_n - f_{n_k}||_p^p < \epsilon; n_k \geq n_\epsilon$, dh. $|f_n - f|^p = \liminf_{k \to \infty} |f_n - f_{n_k}|^p$

$$\int |f_n - f|^p d\mu = \int \liminf_k |f_n - f_{n_k}|^p d\mu \leq \liminf_k \int |f_n - f_{n_k}|^p d\mu < \epsilon \text{ mit Lemma von Fatou}$$

Somit ist f der Grenzwert der Folge f_n. Aus der Dreiecksungleichung folgt

$$||f||_p = ||f - f_n + f_n||_p \leq ||f_n||_p + \underbrace{||f - f_n||_p}_{<\epsilon^{\frac{1}{p}}} \leq \inf_n ||f_n||_p + \epsilon^{\frac{1}{p}} < \infty$$

und $f \in \mathcal{L}_p$. □

L_p-Räume sind Banachräume und L_2 ist sogar ein Hilbertraum mit $< f, g >:= \int f \cdot g d\mu$.

Es folgt sofort, dass Grenzwerte in $\mathcal{L}_p$ μ-f.ü. eindeutig sein müssen,

$$f_n \to^{L_p} f, f_n \to^{L_p} g : ||f - g||_p \leq ||f - f_n|| + ||f_n - g|| < \epsilon, n > n_\epsilon$$
$$\implies ||f - g||_p = 0, \int |f - g|^p d\mu = 0 \implies |f - g| = 0 \implies f = g\mu\text{-f.ü.}$$

Die Existenz einer Teilfolge, die konvergiert, wird sich oft als sehr nützlich erweisen.

Das ist speziell für separable Räume (abzählbar dichte Teilmenge) der Fall.

Sei $([0, 1], \mathcal{A}, \lambda)$: L_p separabel $p \geq 1$, L_∞ nicht!

4.1.1 L^p für $0 < p < 1$

TODO

4.2 Dualraum von L_p

Prinzipiell existiert zu einem normierten linearen Raum $(V, ||.||)$ der Vektorraum der linearen Funktionale $T : V \to \mathbb{R}, T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$, die auch stetig sind. Um die Stetigkeit definieren zu können, muss V ein Banachraum sein.

Def (Dualraum, beschränkt)**.** *Der Dualraum V' zu V ist $V' = \{F | F : V \to \mathbb{R},$ linear und stetig$\}$. Auch V' wird durch $||T|| := \sup\{|T(x)| : ||x|| \leq 1\}$ zu einem Banachraum. T heißt beschränkt, wenn $||T|| \leq C < \infty$.*

T ist genau dann linear und stetig, wenn T linear und beschränkt ist. (UE)

Wir haben den Satz von Riesz-Frechet für die Darstellung der Elemente von V' bereits verwendet. Wenn $(V, < ., . >)$ ein Hilbertraum ist, dann hat $F \in V'$ eine eindeutige Darstellung durch ein $f \in V$ mit $F(v) =< v, f >$.

Für Hilberträume können V und V' in dem Sinn identifiziert werden, dass eine Isometrie zwischen V und V' existiert.

Für den Banachraum $(L^p, ||.||_p)$ kann eine solche Isometrie gefunden werden.

Lemma 10. $1 \leq p \leq \infty$ *Zu $f \in \mathcal{L}^p$ existiert eine Folge $t_n \in \mathcal{L}^p$ von Treppenfunktionen mit $||t_n||_p \leq ||f||_p \forall n$ und $\lim_{n \to \infty} ||f - t_n||_p = 0$.*

Proof. Für die Konstruktion des Integrals wurde beispielsweise die Folge von Treppenfunktionen

$$t_n(w) := \begin{cases} n & f(w) \geq n \\ \frac{k-1}{2^n} & \frac{k-1}{2^n} \leq f(w) \leq \frac{k}{2^n} \end{cases}$$

verwendet. $|t_n| \nearrow |f|$ mit $|t_n| \leq |f|$, daher $||t_n||_p \leq ||f||_p$ und $|f - t_n|^p \leq 2^p |f|^p \in L^p$
Da $\lim\limits_{n\to\infty} |f - t_n|^p = 0$ gilt nach dem Satz der dominierten Konvergenz

$$\lim_{n\to\infty} ||f - t_n||_p^p = \lim_{n\to\infty} \int |f - t_n|^p = \int \underbrace{\lim_{n\to\infty} |f - t_n|^p}_{=0} = 0$$

für $p < \infty$.
Bei $p = \infty$ konvergieren obige t_n sogar gleichmäßig gegen $f \in L_\infty$ und $||f - t_n||_\infty \to 0$. □

Zwischen L^q und $(L^p)'$ sei die kanonische Abbildung

$$K : L^q \to (L^p)' \text{ mit } K(f)(g) := \int f \cdot g d\mu, f \in L^q, g \in L^p$$

K ist eine lineare Abbildung.

Satz 36 (8.LP)**.** *Für* $\frac{1}{p} + \frac{1}{q} = 1$ *ist die kanonische Abbildung* K *eine Isometrie, also* $||K(f)||_p' = ||f||_q$

Proof. $K(f)$ ist aus dem Dualraum von L^p mit der Norm (des Dualraums)

$$||K(f)||_p' := \sup\{ \underbrace{|K(f)(g)|}_{|\int fg| \leq \int |fg| \leq ||f||_q ||g||_p = ||f||_q} : ||g||_p = 1, g \in L^p\}$$

mit $f \in L^q$.
Daher gilt $||K(f)||_p' \leq ||f||_q$
Für die umgekehrte Ungleichung betrachte

$$||K(f)||_p' \geq |\int f \cdot \frac{g}{||g||_p} d\mu| \forall g \in L^p$$

$$\implies ||K(f)||_p' ||g||_p \geq |\int fg d\mu|$$

Ersetze g durch $\tilde{g} := |g| \cdot sign(f)$, dann ist $|\tilde{g}| \leq |g|$ (wo $f = 0$, ist $\tilde{g} = 0$) und auch $||\tilde{g}||_p \leq ||g||_p$

$$\implies ||K(f)||_p' ||g||_p \geq ||K(f)||_p' \cdot ||\tilde{g}||_p \geq |\underbrace{\int f \cdot \tilde{g}}_{\geq 0}| = \int |fg| d\mu = ||fg||_1$$

Fallunterscheidung

1. $q = 1$, d.h. $f \in L^1$, wähle $g = 1 \in \mathcal{L}_\infty$. $\implies ||K(f)||_p' \geq ||f||_1$
2. $q \in (1, \infty)$ wähle $g = |f|^{q-1} = |f|^{\frac{q}{p}}$ da $\int g^p = \int |f|^q < \infty$, $f \in L^q \implies g \in L^p$

$$||K(f)||_p' ||g||_p \geq ||fg||_1 = |||f|^q||_1 = ||f||_q^q = ||f||_q \cdot \underbrace{||f||_q^{q-1}}_{(\int |f|^q)^{\frac{q-1}{q}}} = ||f||_q \cdot (\int g^p)^{\frac{1}{p}} = ||f||_q ||g||_p$$

$$\implies ||K(f)||_p' \geq ||f||_q$$

3. $q = \infty$ also $||f||_\infty < \infty$
   Für jedes $\epsilon > 0$ ist $\{|f| > (1-\epsilon)||f||_\infty\} =: A_\epsilon$ keine Nullmenge. $0 < \mu(A_\epsilon) < \infty$.
   Wähle $g = \frac{1}{\mu(A_\epsilon)} 1_{A_\epsilon}$ ist aus L_1, $||g||_1 = 1$

$$\implies ||K(f)||_p' \geq ||fg||_1 = \int |fg| d\mu = \int |f||g| d\mu = \int \underbrace{|f| 1_{A_\epsilon}}_{> (1-\epsilon)||f||_\infty \int 1_{A_\epsilon} \frac{1}{\mu(A_\epsilon)} d\mu} \frac{1}{\mu(A_\epsilon)} d\mu$$

   für alle $\epsilon > 0$, d.h. $||K(f)||_p' \geq ||f||_\infty$

In allen Fällen folgt mit oben $||K(f)||'_p = ||f||_q$.

$\square$

Die kanonische Abbildung liefert eine Isometrie und ist daher injektiv.
Angenommen $f, g \in L^q$ und $f \neq g$ in L^q, d.h. $||f-g||_q \neq 0$.

$$\begin{aligned} ||K(f)-K(g)||'_p &= ||K(f-g)||'_p = ||f-g||_q \neq 0 \\ \implies &||K(f)-K(g)||'_p \neq 0 \text{ und } K(f) \neq K(g) \end{aligned}$$

$K(.)$ ist auch surjektiv, daher gilt:

Satz 37 (9.LP)**.** $1 \leq p < \infty$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$.
Dann ist L^q *isomorph zu* $(L^p)'$*, dem Dualraum von* L^p*. Man kann dann* $(L^p)' \simeq L^q$ *also* $(L^p)'$ *mit* L^q *identifizieren.*

ohne Beweis. Für den Nachweis der Surjektivität sei auf die Literatur (Buch der VO) verwiesen.

Beispiel. *Wenn* $\Omega = \mathbb{N}$ *und* $\mathcal{A} = 2^{\mathbb{N}}, \mu = \zeta$ *(Zählmaß), dann ist* L^p *ein Folgenraum mit der Bezeichnung* l^p*.* $(a_n) \in l^p$*, wenn* $\sum_{n=1}^{\infty} |a_n|^p < \infty$.
Die kanonische Abbildung für $a \in L^q, b \in L^p, a = (a_n)_{n\in\mathbb{N}}, b = (b_n)_{n\in\mathbb{N}}$ *ist*

$$K(a)(b) = \sum_{i=1}^{\infty} a_i - b_i$$

und entsprechend dem letzten Satz gilt auch $(l^p)' \simeq l^q$.
Für endliches Ω*, etwa* $\Omega = \{1, ..., m\}$ *und* $a = (a_1, ..., a_m)$ *und* $b = (b_1, ..., b_m)$ *ist auch* $(L^\infty)' \simeq L^1$.
Wenn $a \in l^\infty$ *(endliche Folge ist immer beschränkt) und* $b \in l^1$ *(endliche Folge immer summierbar) existiert für jedes* $b \in l^1$ *auch ein* $a \in l^\infty$ *mit* $K(a) \simeq b$.
Im Allgemeinen gilt Satz 9.LP für $p = \infty$ *nicht.*

Beispiel. $(\mathbb{N}, 2^{\mathbb{N}}, \zeta)$ *und* l^p *sei der Folgenraum. Die Teilmenge* $(\hat{l}_\infty)$ *der beschränkten Folgen* l_∞ *(Spur) sei die Menge der konvergenten Folgen. Definiere für* $(a_n) \in \hat{l}_\infty$

$$F((a_n)) = \lim_{n\to\infty} a_n$$

F ist linear. F ist auch stetig: $a_{n_k} \to b_n, k \to \infty$*, d.h.* $\limsup_{k\to\infty} |a_{n_k} - b_n| = 0$*.* (b_n) *konvergent gegen* b_0.

$$\begin{aligned} |a_{n_k} - b_0| &\leq |a_{n_k} - b_n| + |b_n - b_0| \\ \limsup_k |a_{n_k} - b_0| &\leq \limsup_k \underbrace{|a_{n_k} - b_n|}_{\to 0} + \lim \underbrace{|b_n - b_0|}_{\to 0} \\ &\lim_k F((a_{n_k})) = F((b_n)) \end{aligned}$$

F ist nur auf der Teilmenge definiert, kann bei Beibehaltung der Eigenschaften (linear, stetig) auf l_∞ *erweitert werden. Ein* $(b_n) \in l_1$ *kann es nicht geben, sodass* $\sum a_n b_n = \lim a_n$ *für* $a_n \in l_1 \cap l_\infty$.

4.3 Gleichmäßige Integrierbarkeit

Für die Konvergenz im L_p werden zur punktweisen Konvergenz noch weitere strukturelle Eigenschaften der Familie von Funktionen aus L_p benötigt. Ein Fall, wo aus $f_n \to f$ μ-f.ü. doch L_p-Konvergenz folgt, ist gegeben, falls die Integrale $||f_n||_p$ konvergieren.

Satz 38 (10.LP)**.** *Für* $1 \leq p < \infty$ *und die Folge* $f_n \in L_p$ *gelte* $f_n \to f$ *μ-f.ü. und* $||f_n||_p \to ||f||_p$ *dann ist* f_n *L_p-konvergent,* $\lim_{n\to\infty} ||f_n - f||_p = 0$

Proof. Da $|f + f_n|^p \leq (2 \cdot \max(|f|, |f_n|))^p \leq 2^p(|f_n|^p + |f|^p)$ gilt

$$2^p(\int |f|^p + \int |f|^p) = \int \liminf ((2^p(|f|^p + |f_n|^p)) - |f - f_n|^p) \, d\mu \text{ da } \lim |f - f_n|^p = 0$$

wegen Fatou ist weiter

$$\leq \liminf \left(2^p \int |f|^p + 2^p \int |f_n|^p - \int |f - f_n|^p\right) = 2^p \int |f|^p + 2^p \lim \int |f_n|^p + \liminf(- \int |f - f_n|^p)$$

Insgesamt ist

$$2^p(\int |f|^p + \int |f|^p) \leq 2^p(\int |f|^p + \int |f|^p) - \limsup \int |f - f_n|^p$$

das ist nur möglich, wenn $\limsup \int |f - f_n|^p = 0$ also $||f_n - f||_p \to 0$. □

Für $p = \infty$ gilt obige Aussage auch für sigma-endliche Maße i.A. nicht.

Beispiel. $(\mathbb{N}, 2^{\mathbb{N}}, \zeta)$ *mit dem (auf* $\mathbb{N}$*) sigma-endlichen Zählmaß* ζ*. Sei* $f_n = 1_{\{1,\ldots,n\}}$*, dann gilt* $||f_n||_\infty = 1$ *und* $f_n \to f = 1$ *auf* $\mathbb{N}$ *mit* $||f||_\infty = 1$*, aber* $|f_n - f| = 1_{\{n+1,\ldots\}}$ *mit* $||f_n - f||_\infty = 1 \nrightarrow 0$.

Dass für $p \in [1, \infty)$ die Umkehrung nicht ohne weiteres gilt, wurde im Beispiel A (LP 4) demonstriert. Dort war $||f_n - 0||_p \to 0$ für jedes p, aber f_n konvergiert nicht punktweise.

Die Umkehrung gelingt, wenn punktweise Konvergenz durch die (schwächere) Konvergenz im Maß ersetzt wird. Im Beispiel A (LP 4) gilt

$$\lambda(|f_n - f| > \epsilon) = \lambda([a_n, b_n]) = \frac{1}{2\lfloor n \rfloor - 1} \to 0.$$

Satz 39 (11.LP)**.** *Für* $1 \leq p < \infty$ *konvergiert die Folge* f_n *im p-ten Mittel,* $||f_n - f||_p \to 0$ *genau dann, wenn* $f_n \to^\mu f$ *im Maß mit* $f \in L_p$ *und* $||f_n||_p \to ||f||_p$

Proof. $\implies$ Aus $||f_n - f||_p$ folgt Konvergenz im Maß (Satz 1.LP) und wegen der Vollständigkeit ist $f \in L_p$. Die Konvergenz der Integrale ergibt sich aus der (umgekehrten) Dreiecksungleichung. $||f||_p \leq ||f - f_n||_p + ||f_n||_p$ und $||f_n||_p \leq ||f - f_n||_p + ||f||_p$

$$|||f||_p - ||f_n||_p| \leq ||f - f_n||_p \to 0 \implies ||f_n||_p \to ||f||_p$$

Wenn $f_n \to f$ im Maß, dann betrachtet man wieder eine Teilfolge f_{n_k} mit $f_{n_k} \to f\mu$-f.ü. (sogar μ-fast gleichmäßig). Nach dem vorigen Satz folgt daraus auch $||f_{n_k} - f||_p \to 0$.

Angenommen f_n konvergiert nicht in L_p, dann existiert eine Teilfolge f_m, die $||f - f_{m_j}||_p \geq \epsilon \forall j$ aber $f_{m_j} \to f$ im Maß, dann gibt es wieder eine Teilfolge $f_{m_{j_k}}$ mit $||f - f_{m_{j_k}}||_p \to 0, k \to \infty$, wie oben. Das ist ein Widerspruch, und $||f_n - f||_p \to 0$. □

Aus der Konvergenz im Maß wird eine Konvergenzart von Maßen abgeleitet.

Wenn für alle $A \in \mathcal{A}\mu_n(A) \to \mu(A)$ für Maße μ und μ_n auf $(\Omega, \mathcal{A})$ konvergiert μ_n schwach gegen μ (bzw. in Totalvariation), d.h.

$$\lim_{n\to\infty} \sup |\mu_n(A) - \mu(A)| = 0.$$

Bei LS-Maßen mit Verteilungsfunktionen F_n, F bedeutet das für $A = (-\infty, x]$

$$\lim_{n\to\infty} F_n(x) = F(x)$$

wenn die Maße endlich sind. Es gibt gute Gründe hier die Konvergenz nur für die Stetigkeitspunkte C_F von $F(.)$ zu fordern, $x \in C_F : \lim\limits_{n\to\infty} F_n(x) = F(x)$.

Allgemein kann die Konvergenz auch auf die Dichten (falls vorhanden) zurückgeführt werden.

Satz 40 (12.LP, Satz von Scheffe)**.** ν_n, ν *sind endliche Maße mit* $\nu_n(\Omega) = \nu(\Omega)$*.* μ *sei sigma-endlich und* $\nu_n \ll \mu, \nu \ll \mu$*.*
Dann gilt

$$\lim_{n\to\infty} \frac{d\nu_n}{d\mu} = \frac{d\nu}{d\mu} \implies \sup_{A\in\mathcal{A}} |\nu_n(A) - \nu(A)| \to 0$$

Proof. Die punktweise Konvergenz der RN-Dichten und Konvergenz der Integrale bringt nach Satz 10.LP

$$\left|\int_A \frac{d\nu_n}{d\mu} - \int_A \frac{d\nu}{d\mu}\right| \leq \int_A \left|\frac{d\nu_n}{d\mu} - \frac{d\nu}{d\mu}\right| d\mu \leq \int \left|\frac{d\nu_n}{d\mu} - \frac{d\nu}{d\mu}\right| d\mu = ||\frac{d\nu_n}{d\mu} - \frac{d\nu}{d\mu}||$$

Da $||\frac{d\nu_n}{d\mu}|| = \nu_n(\Omega) \to \nu(\Omega) = ||\frac{d\nu}{d\mu}||$ nach Voraussetzung liefert der vorige Satz die Aussage. □

Der Satz gilt auch (wie im Beweis), wenn $\nu(\Omega) \to \nu(\Omega)$. Meist wird aber die Verteilungskonvergenz von W-Maßen mit Dichten untersucht.

Diskrete Verteilung: $(\mathbb{N}, 2^{\mathbb{N}})$ mit Zählmaß als dominierendes Maß. Es genügt die Konvergenz der Punktwahrscheinlichkeiten für Verteilungskonvergenz $P_n \to P$.

Beispiel. *$P_n \sim H_{N,M,k}$ Hypergeometrische Verteilung wobei für $n, N_n, M_n, \in \mathbb{N}$ Folgen mit $N_n \to \infty$ und $\frac{M_n}{N_n} \to p, n \to \infty$ für $0 < p < 1$.*

Somit $P_n(\{i\}) = \frac{\binom{M_n}{i}\binom{N_n - M_n}{k-i}}{\binom{N_n}{k}}, i \in \mathbb{N}$

Asymptotisch gilt für den Binomialkoeffizienten $\binom{n}{x} \sim \frac{n^x}{x!}, n \to \infty$ da $\binom{n}{x} = \frac{n^x}{x!}\frac{n(n-1)\cdots(n-x+1)}{n^x} \to \frac{n^x}{x!}, n \to \infty$.

Daher

$$P_n(\{i\}) \sim \frac{\frac{M_n^i}{i!}\frac{(N_n-M_n)^{k-i}}{(k-i)!}}{\frac{N_n^k}{k!}} = \binom{k}{i}\left(\frac{M_n}{N_n}\right)^i \left(1 - \frac{M_n}{N_n}\right)^{k-i}$$

$$\lim_{n\to\infty} P_n(\{i\}) = \binom{k}{i} p^i (1-p)^{k-i}$$

Binomial $B_{k,p}$. Die Hypergeometrische Verteilung wird durch die Binomialverteilung approximiert, $N \to \infty H_{N,M,k} \sim B_{k,p=M/N}$. (Ziehung mit (Binomial) und ohne (Hypergeom.) Zurücklegen.) Erwartungswert $k\frac{M}{N} \to kp$

Beispiel. *$P_n \sim B_{n,p_n}$ Binomialv.; es gelte $n \to \infty$ und $p_n \to 0$ mit $np_n \to \theta$.*

$$P_n(\{k\}) = \binom{n}{k} p_n^k (1-p_n)^{n-k}$$

$$\Longrightarrow P_n(\{k\}) \sim \frac{n^k}{k!} \underbrace{p_n^k}_{\to\theta^k} (1-p_n)^n \underbrace{(1-p_n)^{-k}}_{\to 1}$$

Für jede Nullfolge $p_n \to 0$ gilt $\lim_{p_n \nearrow 0} (1-p_n)^{\frac{1}{p_n}} \to e^{-1}$ und wegen $np_n \to \theta$ folgt

$$P_n(\{k\}) \sim \frac{n^k}{k!} e^{-\theta} \text{ Poisson mit Rate } \theta \qquad B_{n,p_n} \sim Poi_\theta$$

Die Erwartungswerte für $X_n \sim B_{n,p_n}$, $X \sim Poi_\theta$ konvergieren $\mathbb{E}X_n = np_n \to \theta = \mathbb{E}X$.

Auch bei stetigen Verteilungen werden Verteilungskonvergenzen über die Konvergenz der Dichten untersucht.

Beispiel. *$(\mathbb{R}, \mathcal{B}, \lambda)$, $X_n \sim \gamma(n, 1)$ mit Dichte*

$$\frac{dP^{X_n}}{d\lambda}(x) = \frac{x^{n-1}}{(n-1)!} e^{-x}, x \geq 0$$

Gamma-Verteilung mit $\mathbb{E}X_n = n$

Sei $Y_n := \frac{X_n - n}{\sqrt{n}} = \frac{X_n}{\sqrt{n}} - \sqrt{n}$ mit der Dichte nach dem Transformationssatz für Dichten

$$f_n := \frac{dP^{Y_n}}{d\lambda}(y) = \frac{(\sqrt{n}y + n)^{n-1} e^{-\sqrt{n}y} e^{-n}}{(n-1)!} \sqrt{n}$$

Die Sirling'sche Formel

$$n! \sim \sqrt{2\pi}\sqrt{n} n^n e^{-n}$$

ergibt

$$f_n(y) \sim \frac{1}{\sqrt{2\pi}} \frac{(\sqrt{n}y+n)^{n-1}}{n^{n-1}} e^{-\sqrt{n}y}$$

Wegen

$$n \log(1+\frac{x}{\sqrt{n}}) - x\sqrt{n} \to -\frac{x^2}{2}$$

für $n \to \infty$, gilt

$$f_n(y) \sim \frac{1}{\sqrt{2\pi}}(1+\frac{y}{\sqrt{n}})^n e^{-\sqrt{n}y} \text{ und}$$
$$\lim_{n\to\infty} f_n(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

Die Grenzverteilung ist die Standardnormalverteilung $P^{Y_n} \to N(0,1)$ und $\mathbb{E}Y_n = \frac{n}{\sqrt{n}} - \sqrt{n} = 0$.

Die Verteilung im letzten Beispiel entsteht als standardisierte Summe unabhängiger Exponentialverteilungen $EX_1, Z_i \sim EX_1$ u.a. $Y_n = \frac{\sum Z_i - \sum \mathbb{E}Z_i}{\sqrt{\sum Var(Z_i)}} \to N(0,1)$ diese Summen konvergieren allgemein.

Ein großer Vorteil für eine Familie von Funktionen ist gegeben, wenn die Integrale der Funktionen "gleichgradig" sind, also sich einheitlich verhalten. Das wäre der Fall, wenn eine integrierbare Funktion $g \in \mathcal{L}_1$ existiert, sodass $|f_i| \leq g$.

Eine Funktion f ist genau dann integrierbar, wenn ein $g \in \mathcal{L}_1$ existiert, sodass

$$\int_{|f|>g} |f| d\mu = 0$$

Eine andere Konsequenz der Integrierbarkeit gibt das folgende Lemma.

Lemma 11. *$f \in \mathcal{L}_1$ auf $(\Omega, \mathcal{A}, \mu)$, dann existiert für jedes $\epsilon > 0$ ein $c_\epsilon > 0$ und ein A_ϵ mit $\mu(A_\epsilon) < \infty$ mit*

$$\int_{|f|>c_\epsilon} |f| d\mu < \epsilon \text{ und } \int_{A_\epsilon^c} |f| d\mu < \epsilon$$

Proof. Da $f \in L_1$ muss $\mu([f = \infty]) = 0$ sein. Für die Folge $f_n := |f| 1_{[|f|>n]}$ gilt f_n fallend, $f_n \to |f| 1_{[|f|=\infty]}$ und $f_n \to 0 \mu$-f.ü.

$f_n \leq |f|$ und mit dominierter Konvergenz

$$\lim_n \int f_n d\mu = \lim \int_{[|f|>n} |f| d\mu = 0$$

Genauso für $A_n := [|f| \geq \frac{1}{n}]$ und $A_n^c = [|f| < \frac{1}{n}]$, $A_n^c \searrow [|f| = 0]$

$$\lim_{n\to\infty} \int_{A_n^c} |f| d\mu = \int \lim_{n\to\infty} |f| 1_{A_n^c} = 0$$

Da $\int |f| \geq \int |f| 1_{[|f|>\frac{1}{n}]} \geq \frac{1}{n} \int 1_{[|f|>\frac{1}{n}]}$ ist $\mu(A_n) < \infty$ und $\mu(A_n^c) \to 0$. □

Die Erweiterung obiger Eigenschaften auf eine Familie integrierbarer Funktionen ist folgende Definition.

Definition 17 (gleichmäßig integrierbar)**.** *Die Familie $f_i, i \in I$ auf $(\Omega, \mathcal{A}, \mu)$ heißt gleichmäßig integrierbar oder gleichgradig integrierbar wenn für jedes $\epsilon > 0$ ein $c_\epsilon > 0$ und ein $A_\epsilon \in \mathcal{A}$ mit $\mu(A_\epsilon) < \infty$ existieren, sodass*

$$\sup_{i\in I} \int_{|f_i|>c_\epsilon} |f_i| d\mu < \epsilon \text{ und } \sup_{i\in I} \int_{A_\epsilon^c} |f_i| d\mu < \epsilon$$

Satz 41 (13.LP, Charakterisierungen der gmI). *Für eine Familie $\mathcal{F} = \{f_i, i \in I\}$ messbarer Funktionen auf $(\Omega, \mathcal{A}, \mu)$ sind folgende Bedingungen äquivalent:*

1. *$\mathcal{F}$ ist gleichmäßig integrierbar*
2. *$\forall \epsilon > 0 \exists g_\epsilon \geq 0, g_\epsilon \in L_1$ mit $\sup_{i \in I} \int_{|f_i| > g_\epsilon} |f_i| d\mu < \epsilon$*
3. *$\forall \epsilon > 0 \exists g_\epsilon \geq 0, g_\epsilon \in L_1$ mit $\sup_{i \in I} \int (|f_i| - g_\epsilon)^+ d\mu < \epsilon$*
4. *$\sup \int |f_i| d\mu < \infty$ und $\forall \epsilon > 0 \exists g_\epsilon \geq 0, g_\epsilon \in L_1$ und $\delta > 0$ mit $\int_A g_\epsilon d\mu < \delta \implies \sup_{i \in I} \int_A |f_i| d\mu < \epsilon$*

Proof. i) $\implies$ ii): Nach Voraussetzung existiert c_ϵ und A_ϵ^c wie in der Definition, so ist $g_\epsilon := c_\epsilon 1_{A_\epsilon}$ ein Kandidat für ii)

$$\int g_\epsilon d\mu = c_\epsilon \mu(A_\epsilon) < \infty \implies g_\epsilon \in \mathcal{L}_1$$

$$\int_{[|f_i| > g_\epsilon]} |f_i| d\mu = \int_{[] \cap A_\epsilon} |f_i| + \int_{[] \cap A_\epsilon^c} |f_i| \leq \int_{\underbrace{[|f_i| > c_\epsilon] \cap A_\epsilon}_{[|f_i| > c_\epsilon 1_{A_\epsilon}]}} |f_i| + \int_{A_\epsilon^c} |f_i| \leq \int_{|f_i| > c_\epsilon} |f_i| + \epsilon \leq 2\epsilon$$

ii) $\implies$ iii): $|f_i| 1_{[|f_i| > g_\epsilon]} \geq (|f_i| - g_\epsilon)^+$ also gilt auch

$$\int_{|f_i| > g_\epsilon} |f_i| d\mu \geq \int (|f_i| - g_\epsilon)^+ d\mu$$

und iii) gilt, nach Anwendung von $\sup_{i \in I}$.
TODO □

Es ergeben sich unmittelbar viele Situationen (Familien), wo gmI gilt.

- Ist die Familie $\mathcal{F}$ durch $g \in \mathcal{L}_1$ dominiert, d.h. $|f_i| \leq g \forall i$, dann ist $\mathcal{F}$ gmI, wegen 3.

$$0 = \sup \int (|f_i| - g)^+ < \epsilon$$

- Eine endliche Familie integrierbarer Funktionen $\mathcal{F}$ ist immer gmI.
- Wird die Familie $\mathcal{F}$ durch die Familie $\mathcal{G}$ dominiert, also zu jedem $f \in \mathcal{F}$ existiert ein $g \in \mathcal{G}$ mit $|f| \leq |g|$, dann ist auch $\mathcal{F}$ gmI, vorausgesetzt $\mathcal{G}$ ist gmI.
- Die Summe $\mathcal{F} + \mathcal{G} = \{f + g | f \in \mathcal{F}, g \in \mathcal{G}\}$ ist auch gmI, wenn $\mathcal{F}$ und $\mathcal{G}$ gmI sind.
- Die Integrale sind gleichmäßig über $\mathcal{F}$ (gmI) beschränkt, da $C = \sup_{f \in \mathcal{F}} \int |f| d\mu < \infty$ eine Folgerung aus gmI ist.

Damit ist etwa $\mathcal{F} := \{f(x) = \frac{1}{x^{1+\epsilon}} 1_{(1,\infty)(x)} | \epsilon > 0\}$ auf $(\mathbb{R}, \mathcal{B}, \lambda)$ keine gmI Familie, obwohl jedes $f \in \mathcal{F}$ integrierbar ist und $\sup_{f \in \mathcal{F}} ||f||_\infty = 1 < \infty$.
Das wäre anders, wenn statt λ ein endliches Maß μ vorliegen würde.
Ist μ endlich und $\sup_{f \in \mathcal{F}} ||f||_\infty < K$ ist wegen 3. sofort auch $\mathcal{F}$ gmI mit $g_\epsilon = K$.

- Jede Familie $\mathcal{F}$, die gmI ist wegen $\sup_{f \in \mathcal{F}} \int |f| < \infty$ Teilmenge von $\mathcal{L}_1$, $\mathcal{F} \subseteq \mathcal{L}_1$ für $(\Omega, \mathcal{A}, \mu)$.

Für endliche Maße (also alle W-Maße) wird der Nachweis der gmI wesentlich einfacher.

Satz 42 (14.LP). *Die Familie $\mathcal{F}$ messbarer Funktionen auf dem endlichen Maßraum $(\Omega, \mathcal{A}, \mu)$ ist genau dann gmI, wenn*

$$C = \sup_{i \in I} \int |f_i| < \infty \text{ oder } \lim_{c \to \infty} \sup_i \mu(|f_i| \geq c) = 0$$

$$\text{und } \forall \epsilon > 0 \exists \delta > 0 \text{ mit } \mu(A) < \delta \implies \sup_i \int_A |f_i| < \epsilon$$

Proof. $\Longrightarrow$: $C = \sup \int |f_i| < \infty$ folgt aus gmI, dann gilt auch $\mu(|f_i| > c) \le \frac{1}{c}\int |f_i| \le \frac{C}{c} \to 0$ für $c \to \infty$ somit $\lim\limits_{c\to\infty} \sup_i \mu(|f_i| \ge c) = 0$.

Das ϵ-δ-Kriterium b) gilt ebenfalls, zu $\epsilon > 0$ existiert wegen gmI ein $c_\epsilon > 0$ mit $\sup_i \int_{|f_i|>c_\epsilon} |f_i| < \frac{\epsilon}{2}$ und daher gilt wenn $\mu(A) \le \delta := \frac{\epsilon}{2c_\epsilon}$:

$$\int_A |f_i| = \int_{A\cap[|f_i|\le c_\epsilon]} |f_i| + \int_{A\cap[|f_i|>c_\epsilon]} |f_i| \le c_\epsilon \mu(A) + \frac{\epsilon}{2} < \epsilon$$

$\Longleftarrow$: Es gelte $\sup_i \mu(|f_i| \ge c) \to 0$, es existiert ein c_δ mit $\mu(|f_i| \ge c_\delta) < \delta$, da auch das ϵ-δ-Kriterium gilt, folgt für $A_{\epsilon,i} := [|f_i| > c_\delta]$

$$\sup_j \int_{A_{\epsilon,i}} |f_j| d\mu < \epsilon$$

daher auch $\sup_i \sup_j \int_{A_{\epsilon,i}} |f_j| d\mu < \epsilon$ und $\int_{|f_i|>c_\delta} |f_i| d\mu < \epsilon$, die zweite Bedingung für gmI. □

Wenn die Funktionen f_i als Dichten verstanden werden von $\nu_i(A) := \int_A |f_i| d\mu$.

Diese (endliche) Maße ν_i sind daher gleichmäßig absolut stetig bezüglich dem endlichen Maß μ.

Für den Nachweis der gmI einer Familie $\mathcal{F}$ messbarer Funktionen auf dem endlichen Maßraum $(\Omega, \mathcal{A}, \mu)$ genügt es, eine (stärker als $x \to x$ wachsende) Funktion zu finden, die gleichmäßig über $\mathcal{F}$ integrierbar ist. Dieses Kriterium ist oft einfacher zu prüfen.

Lemma 12. *$\mathcal{F} \subseteq \mathcal{L}_1$ auf dem endlichen Maßraum $(\Omega, \mathcal{A}, \mu)$ ist genau dann gmI, wenn eine Funktion $H : [0,\infty) \to [0,\infty)$ mit*

$$\lim_{x\to\infty} \frac{H(x)}{x} = \infty \text{ und } \sup_{f\in\mathcal{F}} \int H(|f|) d\mu < \infty$$

existiert.
Die Funktion H ist als monoton wachsend und konvex wählbar.

Für SG X bzw. eine Familie SGn $\mathcal{F}$ reduziert sich die Bedingung für gmI auf

$$\sup_{X\in\mathcal{F}} \mathbb{E}[|X| 1_{[|X|>K]}] < \epsilon$$

wobei zu $\epsilon > 0$ ein $K \ge 0$ mit obiger Bedingung existiert.

Lemma 13. *Folge X_n SGn mit $\mathbb{E}|X_n| < \infty$ ist gmI, genau wenn*

$$\lim_{x\to\infty} (\sup_{n\ge1} \mathbb{E}[|X_n| 1_{[|X_n|>x]}]) = 0$$

Proof. $\Longrightarrow$: X_n sei gmI, da $\sup_n \mathbb{E}|X_n| < \infty$ gilt mit der Markoff-Ungleichung für jedes $x > 0$

$$xP(|X_n| > x) \le \mathbb{E}|X_n| \forall n$$

Für $\delta > 0$ und ausreichend großes x gilt

$$\sup_n P(\underbrace{|X_n| > x}_{A_{\epsilon,n}}) \le \frac{1}{x} \sup_{n\ge1} \mathbb{E}|X_n| < \delta$$

gilt $\mathbb{E}|X_n| 1_{A_{\epsilon,n}} \le \epsilon$ und daher auch

$$\lim_{x\to\infty} \sup \sup_n \mathbb{E}[|X_n| 1_{[|X_n|>x]}] < \epsilon.$$

$\Longleftarrow$: Es gelte $\lim\limits_{x\to\infty} (\sup \mathbb{E}[|X_n| 1_{|X_n|>x}]) = 0$.

Für x groß genug gilt

$$\sup_n \mathbb{E}|X_n| \le \sup_n \{\mathbb{E}|X_n| 1_{|X_n|>x} + x\} < \infty$$

Für die zweite Bedingung der gmI wird für $\epsilon > 0$ ein x gewählt, sodass $\sup_n \mathbb{E}[|X_n| 1_{|X_n|>x}] < \frac{\epsilon}{2}$.

Für $\delta < \frac{\epsilon}{2x}$ und $A \in \mathcal{A}$ mit $P(A) < \delta$ folgt

$$\mathbb{E}|X_n|1_A = \mathbb{E}|X_n|1_{A\cap[|X_n|>x]} + \mathbb{E}|X_n|1_{A\cap[|X_n|\leq x]} \leq \underbrace{\mathbb{E}|X_n|1_{|X_n|>x}}_{<\frac{\epsilon}{2}} + \mathbb{E}x\underbrace{1_{A\cap[|X_n|\leq x]}}_{\leq 1_A} \leq \epsilon.$$

$(\delta x < \frac{\epsilon}{2})$ □

Beispiel. *Die Familie SGn $\mathcal{F} = \{X_i | i \in I\}$ enthalte quadratisch integrierbare Funktionen, $\mathcal{F} \supseteq \mathcal{L}_2$ auf $(\Omega, \mathcal{A}, P)$. Wenn die 2.Momente beschränkt sind, etwa $\sup_i Var(X_i) < \infty$ dann ist $\mathcal{F}$ gmI. Aus $(\mathbb{E}|X_i|)^2 \leq \mathbb{E}X_i^2$ (wegen der Jensen Ungl.) ist auch $\sup \mathbb{E}|X_i| < \infty$.*

Weiters gilt

$$\sup_i \int |X_i|^2 d\mu = \sup_i \left((\mathbb{E}|X_i|)^2 + Var(X_i)\right) < \infty$$

Mit $H(x) = x^2$ sind die Bedingungen der letzten Proposition erfüllt. Eigentlich würde die Aussage auch für $H(x) = x^{1+a} \forall a > 0$ gelten.

Wie es aus den vorigen Aussagen vermuten lässt, ist die Äquivalenz von L_p-Konvergenz und Konvergenz im Maß, wenn die Familie (Folge) auch gmI ist.

Als Vorbereitung dienen die folgenden Hilfssätze.

Lemma 14. *f_n ist eine gmI-Folge auf $(\Omega, \mathcal{A}, \mu)$ dann gilt das Lemma von Fatou in der Form*

$$\int \liminf_n f_n \leq \liminf_n \int f_n \leq \limsup_n \int f_n \leq \int \limsup_n f_n$$

Proof. Für die Anwendung des Lemmas von Fatou müssen integrierbare Funktionen mit $f_n \geq g_1$ bzw. $f_n \leq g_2$ existieren.

Da die f_n gmI sind, existiert zu $\epsilon > 0$ ein $g \geq 0$, $g \in \mathcal{L}_1$ und $\sup_n \int_{|f_n|>g} |f_n| d\mu < \epsilon$ und

$$\int f_n = \int_{f_n > -g} f_n + \int_{f_n \leq -g} f_n \geq \int_{f_n > -g} f_n - \epsilon$$

$$\text{Da} \int_{f_n \leq -g} f_n = -\int_{f_n \leq -g} |f_n| \geq -\int_{|f_n|>g} |f_n| \geq -\epsilon$$

Betrachte $\tilde{f}_n := f_n 1_{[f_n > -g]} \geq -g$. $\tilde{f}_n$ erfüllt die Voraussetzungen des Lemmas von Fatou.

$$\int \liminf_n \tilde{f}_n \leq \liminf_n \int \tilde{f}_n = \liminf_n \underbrace{\int f_n 1_{f_n > -g}}_{\leq \sup \int |f_n| < \infty}$$

Da $f_n \leq \tilde{f}_n \implies \liminf f_n \leq \liminf \tilde{f}_n$ somit $\int \liminf f_n \leq \int \liminf \tilde{f}_n \leq \liminf \int \tilde{f}_n$.

Nach oben gilt $\int \tilde{f}_n \leq \int f_n + \epsilon$ also ist der letzte Term $\leq \liminf \int f_n + \epsilon$.

Die Ungleichung für $\limsup$ wird analog wegen $\limsup f_n = -\liminf(-f_n)$ geführt. □

Lemma 15. *f_n sei gmI Folge. Aus der Konvergenz im Maß $f_n \xrightarrow{\mu} f$ folgt $f \in \mathcal{L}_1$, $\lim \int f_n = f$ und $\lim_n ||f_n - f|| = 0$.*
Die Folgerungen treffen auch für $f_n \to f\mu$-f.ü. zu.

Proof. Es gelte $f_n \to f\mu$-f.ü. und daher auch $|f_n| \to |f|$. Nach dem vorigen Lemma (von Fatou) ist

$$\int \lim |f_n| \leq \underbrace{\liminf \int |f_n|}_{=\lim \int |f_n| \leq C = \sup \int |f_n|} \leq \limsup \int |f_n| \leq \int \lim |f_n|$$

Also ist $f \in \mathcal{L}_1$. Da $||f_n||_1 \to ||f||_1$ und $f_n \to f\mu$-f.ü. gilt nach Satz 10.LP auch $||f_n - f||_1 \to 0$.

Gilt $f_n \to f$ im Maß, dann existiert wieder eine Teilfolge $f_{n,k}, f_{n,k} \to f\mu$-f.ü. für diese Teilfolge gilt $||f_{n,k} - f||_1 \to 0$ nach vorigem Argument.

Angenommen $||f_n - f|| \nrightarrow 0$ müsste es eine andere Teilfolge geben, die $||f_{m,j} - f||_1 > \epsilon$ erfüllt für alle $j \geq 1$. $f_{m,j} \to f$ im Maß gilt trotzdem. So müsste $f_{m,j}$ wieder eine Teilfolge $f_{m,j,l}$ besitzen, die wiederum $||f_{m,j,l} - f||_1 \to 0$ erfüllt. Das ist ein Widerspruch. □

Die letzten Aussagen über gmI Folgen werden zum wesentlichen Kriterium für L_p-Konvergenz zusammengefügt.

Satz 43 (15.LP, Vitali-Kriterium). *Die Folge $f_n \in \mathcal{L}_p, 1 \leq p < \infty$ konvergiert in L_p $||f_n - f||_p \to 0 \iff f_n \xrightarrow{\mu} f$ und $|f_n|^p$ ist gmI.*

Proof. $\implies$: Aus $||f_n - f||_p \to 0$ folgt nach Satz 11.LP $f_n \xrightarrow{\mu} f$ und $||f_n||_p \to ||f||_p$ mit $f \in \mathcal{L}_p$.

Daher ist $C = \sup \int |f_n|^p < \infty$.

Für die zweite Bedingung der gmI sei $\epsilon > 0$ und $\delta := \frac{\epsilon^{1/p}}{2}$ sowie $g := |f|^p + \sup_{i=1}^{n_\epsilon} |f_i|^p$ wobei n_ϵ so bestimmt wird, dass $\underbrace{||f_n - f||_p}_{\to \int |f_n - f|^p < \epsilon/2} < \delta \forall n \geq n_\epsilon$.

Damit gilt für ein $A \in \mathcal{A}$

$$\int_A g d\mu < \delta \implies \int_A |f_n|^p d\mu < \epsilon$$

für $n \leq n_\epsilon$ nach Definition von g.

Wenn $n > n_\epsilon$ betrachte

$$||f_n 1_A||_p \leq \underbrace{||(f_n - f)1_A||_p}_{<\frac{\epsilon}{2}} + \underbrace{||f 1_A||_p}_{<\frac{\epsilon}{2} \text{ wenn } \int_A g < \delta}$$

Auch in diesem Fall folgt aus $\int_A g d\mu < \delta \implies \int_A |f_n|^p d\mu < \epsilon$

$\impliedby$: $|f_n|^p$ sei gmI und Maß-konvergent. Es gibt wieder eine Teilfolge $f_{n,k}$ mit $f_{n,k} \to f \mu$-f.ü. $\implies |f_{n,k}|^p \to |f|^p \mu$-f.ü. $|f_{n,k}|^p$ ist gmI, nach vorigem Lemma muss $||f_{n,k}||_p \to ||f||_p$ und $||f||_p < \infty$.

Nach Satz 10.LP ist dann auch $||f_{n,k} - f||_p \to 0$. Würde $||f_n - f||_p \nrightarrow 0$ müsste eine Teilfolge $f_{m,j}$ mit $||f_{m,j} - f||_p > \epsilon$ die aber wieder $f_{m,j} \to f$ im Maß erfüllt. Davon müsste eine nächste Teilfolge $f_{m,j,k}$ mit $||f_{m,j,k} - f||_p \to 0$ geben, was ein Widerspruch ist. □

Die Sätze 11.LP und 15.LP zusammengefasst

$$||f_n - f||_p \to 0 \iff f_n \xrightarrow{\mu} f \land f \in L_p \land ||f_n||_p \to ||f||_p \iff f_n \xrightarrow{\mu} f \land |f_n|^p \text{ ist gmI}$$

kann als Kriterium für gmI verwendet werden:

Für eine im Maß konvergente Folge $f_n \xrightarrow{\mu} f$ sind $|f_n|^p$ gmI $\iff f \in L_p \land ||f_n||_p \to ||f||_p, ||f||_p < \infty$.

Wie die L_p-Konvergenz von einer Metrik (sogar einer Norm), lässt sich auch die Konvergenz im Maß (für sigma-endliches Maß) aus einer Metrik erzeugen.

Lemma 16. *μ sei ein sigma-endliches Maß mit $A_n \nearrow \Omega, \mu(A_n) < \infty$. Zu messbaren Funktionen $f, g : \Omega \to \mathbb{R}$ sei*

$$d(f,g) = \sum_{n=1}^{\infty} \frac{2^{-n}}{1 + \mu(A_n)} \int_{A_n} \min(1, |f - g|) d\mu$$

Dann ist d eine Metrik, sodass für die Folge f_n, f gilt $f_n \xrightarrow{\mu} f \iff d(f_n, f) \to 0$.

Eine weitere Bedingung für f.ü. Konvergenz aus L_p-Konvergenz abzuleiten ist die schnelle Konvergenz.

Satz 44 (15.LP, Schnelle Konvergenz). *f_n und f aus L_p für ein $1 \leq p < \infty$. Dann folgt aus*

1. *$\sum_{n \geq 1} ||f_n - f||_p < \infty$ oder*
2. *Für $A \in \mathcal{A}$ mit $\mu(A) < \infty$ und jedes $\epsilon > 0$ gilt $\sum_{n \geq 1} \mu(A \cap [|f_n - f| > \epsilon]) < \infty$*

dass $f_n \to f \mu$-f.ü.

Proof. Aus 1. folgt wegen

$$\mu(A \cap [|f_n - f| > \epsilon]) \leq \mu([|f_n - f| > \epsilon]) \leq \frac{1}{\epsilon^p} ||f_n - f||_p^p$$

sofort 2. Für 2. wird ein messbares $A \in \mathcal{A}$ mit $\mu(A) < \infty$ angenommen, o.B.d.A sei $A = \Omega$. Mit $D_n := [|f_n - f| > \epsilon]$ und $D_\epsilon := \limsup_n D_{n\to\infty}$ folgt nach dem Lemma von Borel-Cantelli

$$\mu(D_\epsilon) = \mu(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} [|f_k - f| > \epsilon]) = 0$$

Sei $N = \bigcup_{m=1}^{\infty} D_{1/m}$, daher $\mu(N) = 0$ und $f_n \to f$ auf N^c. $\square$