Vorlesung SS 2011

Wahrscheinlichkeitstheorie

Prof. Dr. Franz Merkl

geTEXt von Myriam Schmidt/Robert Herrmann

Inhaltsverzeichnis

1 Bedingte Erwartungen

Aus der Stochastikvorlesung kennen Sie den Erwartungswert

$$E[X] = \int_\Omega X \, dP$$

eine integrierbaren Zufallsvariable X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. Der Begriff der bedingten Erwartung verallgemeinert diesen Erwartungswert-Begriff auf Situationen in denen man Teilinformationen über das zufällige Ergebnis $\omega \in \Omega$ hat.

1.1 Elementare bedingte Erwartungen

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A \in \mathcal{A}$ ein Ereignis mit P(A)>0.

Erinnerung. *Durch $P(\cdot|A) : \mathcal{A} \to [0,1], P(B|A) = \frac{P(A \cap B)}{P(A)}$ wird ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$ definiert. Es heißt bedingtes Wahrscheinlichkeitsmaß zu P gegeben A.*

Erwatungswerte bzgl. dieses Maßes heißen elementare bedingte Erwartungen. Nun besitzt P$(\cdot|A)$ bzgl. P die Dichte $\frac{1_A}{P(A)}$. In der Tat: Für alle $B \in \mathcal{A}$ gilt:

$$P(B|A) = \int_\Omega \frac{1_{B\cap A}}{P(A)} \, dP = \int_B \underbrace{\frac{1_A}{P(A)}}_{Dichte \frac{dP(\cdot|A)}{dP(A)}} \, dP$$

Wir erhalten folgende Formel für elementare bedingte Erwartungen. Für Zufallsvariable X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ oder X $\geq$0 gilt:

$$E_{P(\cdot|A)}[X] = \int_\Omega X \underbrace{\frac{1_A}{P(A)}}_{Dichte} \, dP = \frac{E_P[X \, 1_A]}{P(A)} = \frac{E_P[X \, 1_A]}{E_P[1_A]}$$

Nun betrachten wir statt eines Ereignisses A eine Partition von Ω in n Ereignisse: $\Omega = A_1 \sqcup A_2 \sqcup ... \sqcup A_n$, wobei wir im Moment P$(A_i)$ >0 für i=1,...,n annehmen. Man kann sich die A_i in der Form $A_i = \{Y = y_i\} = Y^{-1}(\{x_i\})$ vorstellen, wobei Y: $\Omega \to \{y_i, ..., y_n\}$ ein Zufallsvariabe sein soll, die nur n-Werte $y_i, ..., y_n$ annehmen kann.
Gegeben X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, kodieren wir die Gesamtheit aller elementar bedingten Erwartungen in einer einzigen Zufallsvariablen Z: $\Omega \to \mathbb{R}$, wobei Z$(\omega) \coloneqq E[X|A_i]$ für $\omega \in A_i$, i=1,...,n.
Z ist also eine „vergröberte Version“von X_i stückweise konstant auf den A_i.

$$Z = \sum_{i=1}^n E[X|A_i] \, 1_{A_i} = \sum_{i=1}^n \frac{E[X \, 1_{A_i}]}{P(A_i)} \, 1_{A_i}$$

Stellt man sich $A_i = \{Y = y_i\}$ vor, so kann man Z in der Form Z=$f(Y) = f \circ Y$ mit $f(y_i) = E[X|A_i] = E[X|Y = y_i]$ i=1,...,n schreiben. Nun bezeichne $\mathcal{F} = \sigma(A_1, ..., A_n) = \{\bigcup_{i \in I} A_i | I \subseteq \{1, ..., n\}\}$=$\sigma$ (Y) ,die von der Partition $A_1, ..., A_n$ erzeugte $\sigma - Algebra$. Die Partition $A_1, ..., A_n$ ist natürlich (bis auf die Reihenfolge) durch $\mathcal{F}$ bestimmt, insofern tragen $\mathcal{F}$ und $A_1, ..., A_n$ die gleiche Information.

Definition 1.1.1.

$$E[X|\mathcal{F}] \coloneqq \sum_{i=1}^n E[X|A_i] 1_{A_i} = \sum_{i=1}^n \frac{E[X \, 1_{A_i}]}{E[1_{A_i})} 1_{A_i}$$

heißt die <u>bedingte Erwartung von X gegeben $\mathcal{F}$</u>. $f : \{y_1, ..., y_n\} \to \mathbb{R}, f(Y) = E[X|Y = y_i]$ heißt <u>faktorisierte bedingte Erwartung von X gegeben Y.</u>

Die bedingte Erwartung E$[X|\mathcal{F}]$ ist also - ebenso wie X - eine Zufallsvariable auf $(\Omega, \mathcal{A}, P)$, die wir uns als „vergröberte Version“, „Schätzung“ oder „Prognose“ von X gegeben $\mathcal{F}$ vorstellen können.

Der Name „faktorisierte bedingte Erwartung“wird durch die Gleichung E$[X|\mathcal{F}] = f \circ Y$, also das kommutative Diagramm $\Omega \overset{E[X|\mathcal{F}]}{\to} \mathbb{R}$, $\Omega \overset{Y}{\to} \{y_1, ..., y_n\} \overset{f}{\to} \mathbb{R}$ gerechtfertigt.

Charakteristische Eigenschaften der bedingten Erwartung:

Lemma 1.1.2. *Unter der Voraussetzung eben gilt:*

1. *$E[X|\mathcal{F}]$ ist $\mathcal{F}$ - $\mathcal{B}(\mathbb{R})$ - messbar (kurz: $\mathcal{F}$ - messbar)*
2. *$\forall A \in \mathcal{F}$ gilt: $E[X\, 1_A] = E[E[X|\mathcal{F}]\, 1_A]$*

Beweis.

1. **Erinnerung.** *Eine Abbildung Z: $\Omega \to \mathbb{R}$ heißt $\mathcal{F}$ - messbar, wenn für alle $B \in \mathcal{B}(\mathbb{R})$ gilt:*

$$Z^{-1}(B) = \{\omega \in \Omega | Z(\omega) \in B\} = \{Z \in B\} \in \mathcal{F}.$$

Nun ist E$[X|\mathcal{F}]$ nach Definiton auf jeder der Mengen $A_1, ..., A_n$ konstant. Für jedes $B \subseteq \mathbb{R}$ ist also $\{$E$[X|\mathcal{F}] \in B\}$ von der Gestalt $\bigcup\limits_{i \in I} A_i$ für geeignetes $I \subseteq \{1, ...n\}$. Es folgt: $\{E[X|\mathcal{F}] \in B\} \in \mathcal{F}$.

2. Es sei $A = \bigcup\limits_{i \in I} A_i \in \mathcal{F}$ für geeignetes I $\subseteq \{1, ..., n\}$. Dann gilt $1_A = \sum\limits_{i \in I} 1_{A_i}$, da die A_i paarweise disjunkt sind. Es folgt:

$$E[X\, 1_A] = E[X \sum_{i \in I} 1_{A_i}] = \sum_{i \in I} E[X\, 1_A] = \sum_{i \in I} \frac{E[X\, 1_A]}{P(A_i)} P(A_i) = \sum_{i \in I} E[X|A_i] E[1_{A_i}] =$$

$$E\left[\sum_{i \in I} E[X|A_i] 1_{A_i}\right] = E\left[\sum_{i=1}^{n} E[X|A_i] \underbrace{1_{A_i} 1_A}_{\substack{=1_{A_i}:\, i \in I \\ =0:\, i \notin I}}\right] = E[E[X|\mathcal{F}] 1_A]$$

□

Verallgemeinerung (Nullmengen in der Partition $A_1, ..., A_n$)**.** *Problem, wenn wir auf die Voraussetzung $P(A_i) > 0\ \forall i \in$ 1,...,n verzichten: Im Fall, dass $P(A_i) = 0$ für manche i tritt in*

$$E[X|\mathcal{F}] = \sum_{i=1}^{n} \frac{E[X\, 1_{A_i}]}{P(A_i)} 1_{A_i}$$

der unbestimmte Ausdruck $\frac{0}{0}$ auf. Allerdings betrifft diese Unbestimmtheit nur den Wert von $E[X|\mathcal{F}]$ auf der Nullmenge unter den A_i. Setzten wir für i=1,...,n:

$$e_i = \begin{cases} E[X|A_i] & \text{falls } P(A_i) > 0 \\ \text{beliebig} & \text{falls } P(A_i) = 0 \end{cases} \quad \text{und}$$

$$E[X|\mathcal{F}] = \sum_{i=1}^{n} e_i 1_{A_i},$$

so ist diese bedingte Erwartung noch eindeutig bis auf Abänderung auf Nullmengen in $\mathcal{F}$.
Sprechweise: $E[X|\mathcal{F}]$ ist p-fast-sicher eindeutig bestimmt.

Die Aussagen 1. und 2. als Lemma bleiben auch mit dieser Verallgemeinerung gültig. Insbesondere kommt es bei der Erwartungswertbildung in 2. auf die Nullmengen unter den A_i nicht an.

1.2 Allgemeine bedingte Erwartungen: Definition und Eindeutigkeit

Im Folgenden sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $\mathcal{F} \subseteq \mathcal{A}$ eine Unter-σ-Algebra und X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. Wir wollen (in Verallgemeinerung des eben betrachteten elementaren Begriffs der bedingten Erwartung) eine bedingte Erwartung von X gegeben $\mathcal{F}$ definieren, die wieder eine „Vergröberung“, „Schätzung“ oder „Prognose“ für X gegeben die Information in $\mathcal{F}$ bedeuten soll. Dazu „drehen wir die Logik gegenüber Abschnitt 1.1 um“:
Die Aussagen des Lemmas in 1.1 werden jetzt zu definierenden Bedingungen:

Definition 1.2.1. *Es sei $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. Eine Zufallsvarible $Z \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ heißt eine bedingte Erwartung von X gegeben einer Unter-σ-Algebra $\mathcal{F} \subseteq \mathcal{A}$, wenn gilt:*

1. *Z ist $\mathcal{F}$-messbar.*
2. *$\forall A \in \mathcal{F}$ gilt: $E[X\, 1_A] = E[Z\, 1_A]$.*

Im Gegensatz zu Abschnitt 1.1, ist diese Definition nicht konstruktiv. Damit haben wir das Problem der Existenz und Eindeutigkeit. Aus Kapitel 1.1 wissen wir schon, dass wir bestenfalls Eindeutigkeit bis auf Abänderung auf einer P-Nullmenge in $\mathcal{F}$ erwarten dürfen.

Satz 1.2.2 (Eindeutigkeit der bedingten Erwartung P- f.s.). *Sind Z und W zwei bedingte Erwartungen der gleichen Zufallsvariablen $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ gegeben $\mathcal{F}$, so gilt Z=W P-f.s.*

Beweis. Nach Voraussetzung 1 sind Z und W $\mathcal{F}$- messbar, also auch Z-W $\mathcal{F}$-messbar. Insbesondere ist $A = \{Z > W\} = \{Z - W > 0\} \in \mathcal{F}$. Es folgt mit Teil 2 in der Definition:

$$E[(Z - W)1_A] = E[Z\, 1_A] - E[W\, 1_A] = E[X\, 1_A] - E[X\, 1_A] = 0$$

Nun ist (Z-W)$1_A \geq 0$, daher folgt (Z-W)$1_A = 0$ P-f.s. Das bedeutet nach Definition von A: P(A)=0 also $Z \leq W$ P-f.s.
Mit vertauschten Rollen von Z und W folgt ebenso W$\leq$Z P-f.s., also zusammen Z=W P-f.s. □

Satz 1.2.3 (von Radon-Nikodym für endliche Maße). *Sind μ und ν endliche Maße auf $(\Omega, \mathcal{A})$, so sind äquivalent:*
1) Jede μ-Nullmenge ist auch eine ν- Nullmenge.

2) ν besitzt eine Dichte bzgl. μ.

Bemerkung. Gelten die beiden äquivalenten Bedingungen des Satzes, so sagt man: ν ist absolut stetig bzgl. μ und schreibt $\nu \ll \mu$.

Beweis. 2)$\Rightarrow$1) Besitzt ν die Dichte f bzgl. μ, so gilt für jede μ-Nullmenge A$\in \mathcal{A}$:

$$\nu(A) = \int\limits_A f d\mu = 0$$

1)$\Rightarrow$2) Wir betrachten die folgende Menge von Funktionen:

$$\mathcal{M} = \{f \colon \Omega \to \mathbb{R} | f \geq 0,\ f\ ist\ \mathcal{A} - messbar, \int\limits_A f\, d\mu \leq \nu(A)\, \forall A \in \mathcal{A}\}$$

Es bezeichne f $\vee$ g das Maximum von f, g: $\Omega \to \mathbb{R}$:

$$f \vee g(\omega) = max\{f(\omega), g(\omega)\}.$$

Es gilt für alle f, g $\in \mathcal{M}$: f $\vee$ g $\in \mathcal{M}$, denn f $\vee$ g ≥ 0, f $\vee$ g ist $\mathcal{A}$ - messbar und für alle A $\in \mathcal{A}$ gilt:

$$\int_A f \vee g \; d\mu = \int_{A \cap \{f \leq g\}} f \vee g \; d\mu + \int_{A \cap \{f > g\}} f \vee g \; d\mu =$$

$$\int_{A \cap \{f \leq g\}} g \; d\mu + \int_{A \cap \{f > g\}} f \; d\mu \overset{f,g \in \mathcal{M}}{\leq} \nu(A \cap \{f \leq g\}) + \nu(A \cap \{f > g\} = \nu(A)$$

Wenn 0$\in \mathcal{M}$ ist $\mathcal{M}$ nicht leer . Wir setzen s:= $\sup\{\int_\Omega f d\mu | f \in \mathcal{M}\}$. Insbesondere ist 0$\leq s \leq \nu(\Omega) < \infty$.
Wir wählen eine Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{M}$ mit $\int_\Omega f_n d\mu \overset{n\to\infty}{\to}$ s.
Wir setzten $g_n = f_1 \vee ... \vee f_n \in \mathcal{M}$. Dann ist $(g_n)_{n\in\mathbb{N}}$ aufsteigend und

$$s \underset{g_n \in \mathcal{M}}{\geq} \int_\Omega g_n d\mu \underset{g_n \geq f_n}{\geq} \int_\Omega f_n d\mu \overset{n\to\infty}{\to} s, \quad \text{also} \int_\Omega g_n d\mu \underset{n\to\infty}{\nearrow} s$$

Setzen wir g $:= \lim\limits_{n\to\infty} g_n$, (punktweise), so folgt (eventuell nach Abänderung auf der μ-Nullmenge $\{g = \infty\}$), so folgt g $\in \mathcal{M}$, da $0 \leq g$ und $\forall A \in \mathcal{M}$ gilt:

$$\int_A g \; d\mu \overset{mon.\ konv.}{=} \lim_{n\to\infty} \int_A g_n \; d\mu \underset{g_n \in \mathcal{M}}{\leq} \nu(A) < \infty. \text{Insbesondere ist} \int_\Omega g \; d\mu = s.$$

Unser Ziel ist es zu zeigen, dass ν die Dichte g bzgl μ besitzt, also $\int_A g \; d\mu = \nu(A) \forall A \in \mathcal{A}$ gilt, wobei wir $\int_A g \; d\mu \leq \nu(A)$ schon wissen.
Sei $\varepsilon > 0$ gegeben. Wir setzen $\mathcal{A}_\varepsilon = \{A \in \mathcal{A} | \nu(A) \leq \int_A g \; d\mu + \varepsilon\mu(A)\}$.
Wir definieren rekursiv zwei Folgen $(A_n)_{n\in\mathbb{N}}$ und $(B_n)_{n\in\mathbb{N}}$ von Ereignissen und eine Folge $(c_n)_{n\in\mathbb{N}}$ von Zahlen $c_n \geq 0$: $A_1 = \Omega$. Für den n-ten Schritt sei A_n schon gegeben. Wir setzten

$$0 \underset{B=\emptyset \text{ betrachten}}{\leq} c_n := \sup\{\mu(B) | B \in \mathcal{A}_\varepsilon, B \subseteq A_n\} \underset{\mu-endl.\ Maß}{<} \infty.$$

Insbesondere können wir ein $B_n \in \mathcal{A}_\varepsilon$ mit $B_n \subseteq A_n$ mit $\mu(B_n) \geq \frac{1}{2}c_n$ wählen: Für $c_n > 0$ ist das offensichtlich wegen $\frac{1}{2}c_n < c_n$ und für $c_n = 0$ können wir einfach $B_n = \emptyset$ setzten.
Schließlich setzen wir $A_{n+1} := A_n \setminus B_n = (\bigcup\limits_{k=1}^{n} B_k)^{\mathrm{C}}$. Nach Konstruktion ist die Folge $(A_n)_{n\in\mathbb{N}}$ absteigend und die $(B_n)_{n\in\mathbb{N}}$ sind paarweise disjunkt, da für $m > n$ gilt:

$$B_m \subseteq A_m \subseteq A_{n+1} = A_n \setminus B_n, \; also B_m \cap B_n = \emptyset.$$

Insbesondere ist $\infty \underset{\mu-endl.}{>} \mu(\bigcup\limits_{n\in\mathbb{N}} B_n) = \sum\limits_{n\in\mathbb{N}} \mu(B_n)$ und daher $\mu(B_n) \overset{n\to\infty}{\to} 0$. Wegen $0 \leq \frac{1}{2}c_n \leq \mu(B_n) \overset{n\to\infty}{\to} 0$ folgt auch $c_n \overset{n\to\infty}{\to} 0$. Setzen wir nun $A_\infty := \bigcap\limits_{n\in\mathbb{N}} A_n = (\bigcup\limits_{n\in\mathbb{N}} B_n)^{\mathrm{C}}$.
Es folgt für alle C $\in \mathcal{A}_\varepsilon$ mit $C \subseteq \mathcal{A}_\infty$, $C \subseteq A_n$, also $\mu(C) \leq c_n$ nach Definition zu c_n $\forall n \in \mathbb{N}$, also $\mu(C) = 0$, da $c_n \overset{n\to\infty}{\to} 0$ und daher $\nu(C) = 0$ nach Voraussetzung 1) des Satzes.

Insbesondere gilt hier $(*) \quad \nu(C) \geq \int_C d\mu + \varepsilon\mu(C)$, denn beide Seiten sind gleich 0. Die Ungleichung $(*)$ gilt jedoch auch für alle C $\in \mathcal{A}$ mit C $\notin \mathcal{A}_\varepsilon$, denn hier gilt sogar $\nu(C) > \int_C d\mu + \varepsilon\mu(C)$ nach

Definition von $\mathcal{A}_\varepsilon$. Damit folgt
$(*)$ $\nu(C) \geq \int_C g d\mu + \varepsilon\mu(C)$ für alle $C \in \mathcal{A}$ mit $C \subseteq A_\infty$. Es folgt für alle A $\in \mathcal{A}$:

$$\nu(A) = \nu(A \setminus A_\infty) + \nu(A \cap A_\infty) \overset{g\in\mathcal{M}}{\geq} \int_{A\setminus A_\infty} g\ d\mu + \int_{A\cap A_\infty} g\ d\mu + \varepsilon\mu(A \cap A_\infty) = \int_A (g + \varepsilon\ 1_{A_\infty}) d\mu$$

Das bedeutet: g+ε $1_{A_\infty} \in \mathcal{M}$, also

$$\int_\Omega g\ d\mu + \varepsilon\mu(A_\infty) = \int_\Omega (g + \varepsilon\ 1_{A_\infty}) d\mu \overset{Def. von\ s}{\leq} s \overset{oben}{=} \int_\Omega g\ d\mu,$$

also $\varepsilon\mu(A_\infty) \leq 0$ und damit $\mu(A_\infty) = 0$ und somit $\nu(A_\infty) = 0$ wieder nach Voraussetzung 1) des Satzes. Es folgt für alle A $\in \mathcal{A}$:

$$(**) \quad \nu(A) = \nu(A \cap A_\infty) + \nu(A \setminus A_\infty) \leq \overbrace{\nu(A_\infty)}^{=0} + \nu(A \setminus A_\infty) = \nu(A \setminus A_\infty) \overset{A_\infty^C = \bigcup_{n\in\mathbb{N}} B_n}{=}$$
$$\nu\left(\bigcup_{n\in\mathbb{N}} (A \cap B_n)\right) \underset{B_n\ pw.\ disj.}{=} \sum_{n\in\mathbb{N}} \nu(A \cap B_n)$$

Nun gilt für den n-ten Summanden:

$$\nu(A \cap B_n) = \nu(B_n) - \underbrace{\nu(B_n \setminus A)}_{\geq \int_{B_n\setminus A} g\ d\mu\ wegen\ g\in\mathcal{M}} \leq \nu(B_n) - \int_{B_n\setminus A} g\ d\mu$$

Wegen $B_n \in \mathcal{A}_\varepsilon$ haben wir $\nu(B_n) \leq \int_{B_n} g\ d\mu + \varepsilon\mu(B_n)$, also eingesetzt:

$$\nu(A \cap B_n) \leq \int_{B_n} g\ d\mu + \varepsilon\mu(B_n) - \int_{B_n\setminus A} g\ d\mu = \int_{B_n\cap A} g\ d\mu + \varepsilon\mu(B_n)$$

In $(**)$ eingesetzt:

$$\nu(A) \leq \sum_{n\in\mathbb{N}} \nu(A\cap B_n) \leq \sum_{n\in\mathbb{N}} \left(\int_{B_n\cap A} g\ d\mu + \varepsilon\mu(B_n) \right) = \sum_{n\in\mathbb{N}} \int_{B_n\cap A} g\ d\mu + \varepsilon \sum_{n\in\mathbb{N}} \mu(B_n) \leq \int_A g\ d\mu + \varepsilon\mu(\Omega),$$

also $\nu(A) \leq \int_A g\ d\mu + \varepsilon\mu(\Omega)$. Dies gilt $\forall\varepsilon > 0$, also folgt $\nu(A) = \int_A g\ d\mu$, was zu zeigen war. □

Bemerkung. Der Satz von Radon-Nikodym git auch für σ-endiche Maße μ ν. Das folgt leicht aus der hier bewiesenen Version (Übung).

Zurück zur Existenz bedingter Erwartungen. Für X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, $X \geq 0$, und eine Unter-σ-Algebra $\mathcal{F} \subseteq \mathcal{A}$ wenden wir den Satz von Radon-Nikodym wie folgt an: μ sei das Maß P eingeschränkt auf $\mathcal{F}$ und ν sei das Maß mit Dichte X bezüglich P, eingeschränkt auf $\mathcal{F}$.

$$\nu : \mathcal{F} \to \mathbb{R},\ \nu(A) = \int_A X dP \text{ für } A \in \mathcal{F}$$

Dann ist jede μ-Nullmenge A $\in \mathcal{F}$ auch eine ν- Nullmenge, denn $\nu(A) = \int\limits_A dP = 0$. Also besitzt ν eine Dichte g bezüglich μ auf $(\Omega, \mathcal{F})$. Insbesondere ist g $\mathcal{F}$-messbar. Es folgt $\forall A \in \mathcal{F}$:

$$E_p \underbrace{[g 1_A]}_{\mathcal{F}} = \int\limits_\Omega g 1_A d\mu = \nu(A) = \int\limits_A X dP = E_p[X 1_A]$$

Also ist g eine bedingte Erwartung von X gegeben $\mathcal{F}$. Für X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ beliebigen Vorzeichens zerlegen wir X in Positiv- und Negativteil: $X = X_+ - X_-$ wobei $X_+ = X \vee 0$ und $X_- = (-X) \vee 0$. Sind nun g_1 und g_2 bedingte Erwartungen von X_+ bzw. X_- gegeben $\mathcal{F}$, so ist g $= g_1 - g_2$ eine bedingte Erwartung von X gegeben $\mathcal{F}$. In der Tat: g $= g_1 - g_2$ ist $\mathcal{F}$-messbar und $\forall A \in \mathcal{F}$ gilt:

$$E[g 1_A] = E[(g_1 - g_2) 1_A] = E[g_1 1_A] - E[g_2 1_A] = E[X_+ 1_A] - E[X_- - 1_A] = E[(X_+ - X_-) 1_A] = E[X 1_A]$$

Damit haben wir gezeigt, dass alle X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ eine bedingte Erwartung gegeben $\mathcal{F}$ besitzen und dass diese bis auf Abänderung auf einer $\mathcal{F}$- Nullmenge eindeutig ist.

E$[X|\mathcal{F}]$ bezeichnet eine beliebige bedingte Erwartung von X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ gegeben $\mathcal{F}$. Damit wird E$[X|\mathcal{F}]$ nur bis auf Abänderung auf P-Nullmengen in $\mathcal{F}$ eindeutig bestimmt.
Gleichungen für $E[X|\mathcal{F}]$ sollten daher immer den Zusatz „p-f.s.“ tragen.
Präziser hätten wir $E[X|\mathcal{F}]$ auch als die Menge aller bedingten Erwartungen X gegeben $\mathcal{F}$ definieren können; dann sollten wir statt „Z=$E[X|\mathcal{F}]$ p-f.s.“besser schreiben „Z $\in E[X|\mathcal{F}]$“. Diese Notation hat sich nicht durchgesetzt.
Man kann sich $E[X|\mathcal{F}]$ als „Prognose“ oder „Schätzung“ für den unbekannten Wert von X vorstellen, wenn man nur Teilinformationen , die durch $\mathcal{F}$ kodiert werden, beobachten kann.
Besonders anschaulich ist der Fall, dass $\mathcal{F}$ durch eine Zufallsvariable Y: $(\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ erzeugt wird. $\mathcal{F} = \sigma(Y) = \{\{y \in A'\} | A \in \mathcal{A}'\}$

Definition 1.2.4. *Eine messbare Abbildung f: $(\Omega', \mathcal{A}') \to \mathbb{R}, \mathcal{B}(\mathbb{R}))$ heißt faktorisierte bedingte Erwartung von X gegeben Y, wenn $f \circ Y$ = f(Y) eine bedingte Erwartung von X gegeben $\sigma(Y)$ ist.*

Diagramm: $\Omega \overset{E[X|\sigma(Y)]}{\to} \mathbb{R}$, $\Omega \overset{Y}{\to} \Omega' \overset{f}{\to} \mathbb{R}$
Faktorisierte bedingte Erwartungen existieren für beliebige Zufallsvariablen Y: $(\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ und X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ (Übung). (Alternativ zur Übung kann man dies auch direkt mit dem Satz von Radon-Nikodym beweisen (hier nicht ausgeführt)).
Man beachte, dass die faktorisierte bedingte Erwartung f nur bis auf Abänderung auf einer $\mathcal{L}_p(Y)$-Nullmenge eindeutig bestimmt ist.

Erinnerung. $\mathcal{L}_P : \mathcal{A} \to [0, 1]$ *bezeichnet die Verteilung von Y:* $\mathcal{L}_P(Y)(A') = P[Y \in A'] = P(Y^{-1}[A'])$ $A' \in \mathcal{A}'$.

Notation. *In Verallgemeinerung des elementaren Falls aus Abschnitt 1.1 schreiben wir* $E[X|Y = y] = f(y)$ $\mathcal{L}_p(Y)$*- f.s. für* $y \in \Omega'$*, wenn f eine faktorisierte bedingte Erwartung von X gegeben Y ist.* $E[X|Y]$ *ist hingegen eine Kurznotation für* $E[X|\sigma(Y)]$*. Allgemeiner steht* $E[X|Y_1, ..., Y_n]$ *für* $E[X|\sigma(Y_1, ..., Y_n)]$.

1.3 Grundlegende Eigenschaften bedingter Erwartung

Für endliche σ-Algebren $\mathcal{F}$ kennen Sie eine explizite Konstruktion von $E[X|\mathcal{F}]$ aus Abschnitt 1.1. Andererseits ist der Existenzbeweis für$E[X|\mathcal{F}]$ aus Abschnitt 1.2 wegen des nichtkonstruktiven Charakters des Satzes von Radon-Nikodym für Berechnungen nutzlos.
Folgende Berechnungsformel für E$[X|Y]$ ist oft praktisch:

Satz 1.3.1. *Seinen $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und Y: $(\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ Zufallsvariablen. Weiter seinen λ ein σ-endliches Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (oder auf einer messbaren Teilmenge davon, die das Bild von X umfasst) z. B. λ=Lesbesguemaß und μ ein σ- endliches Maß auf $(\Omega', \mathcal{A}')$. Wenn (X,Y): $\Omega^2 \to \mathbb{R} \times \Omega'$ eine Dichte g bzgl. $\lambda \times \mu$ besitzt, so ist f: $\Omega' \to \mathbb{R}$*

$$f(y) = \frac{\int\limits_{\mathbb{R}} x\ g(x,y)\lambda(dx)}{\int\limits_{\mathbb{R}} g(x,y)\lambda(dx)}$$

falls dies wohdefiniert ist, f(y) ist beliebig $\mathcal{A}'$-messbar, sonst eine faktorisierte bedingte Erwartung von X gegeben Y: f(y)=$E[X|Y=y]$ für $\mathcal{L}_P(Y)$- fast alle $y \in \Omega$!

Beweis. Dies folgt unmittelbar aus dem Satz von Fubini: f ist $\mathcal{A}'$-messbar und für alle $A' \in \mathcal{A}'$ gilt:

$$E[X\ 1_{A'}(Y)] = \int\limits_{\mathbb{R}\times\Omega'} x\ 1_{A'} g(x,y)\lambda x \mu(dxdy) \overset{Fubini}{=} \int\limits_{\Omega'} 1_{A'}(y) \int\limits_{\mathbb{R}} x\ g(x,y)\lambda(dx)\mu(dy) =$$

$$\int\limits_{\Omega'} 1_{A'}(y) f(y) \int\limits_{\mathbb{R}} g(x,y)\lambda(dx)\mu(dy) \overset{Fubini}{=} \int\limits_{\mathbb{R}\times\Omega'} f(y) 1_{A'}(y) g(x,y)\lambda x \mu(dxdy) = E[f(y)1_{A'}(y)]$$

Im Fall, dass X: $\Omega \to \{x_1, ..., x_n\}$ und Y: $\Omega \to \{y_1, ..., y_n\}$ direkt sind, führt dies uns zum bekannten aus Abschnitt 1.1 zurück:

$$E[X|Y=y] = \frac{E[X\ 1_{\{Y=y\}}]}{P(Y=y)} = \frac{\sum\limits_x x\ P[X=x, Y=y]}{\sum\limits_x P[X=x, Y=y]}$$

□

Weitere wichtige Eigenschaften der bedingten Erwartung:

Satz 1.3.2. *$E[\cdot|\mathcal{F}]$ ist linear.*
Für $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und $a \in \mathbb{R}$ gilt:

1) $E[X + Y|\mathcal{F}] = E[X|\mathcal{F}] + E[Y|\mathcal{F}]$ P-f.s.

2) $E[a\ X|\mathcal{F}] = aE[X|\mathcal{F}]$ P-f.s

Beweis. Die rechten Seiten in 1) und 2) sind $\mathcal{F}$- messbar und es gilt für $A \in \mathcal{F}$:

1) $E[E[X|\mathcal{F}] + E[Y|\mathcal{F}]1_A] = E[E[X|\mathcal{F}]1_A] + E[E[Y|\mathcal{F}]1_A] = E[X\ 1_A] + E[Y\ 1_A] = E[(X+Y)1_A]$
2) $E[a\ E[X|\mathcal{F}]1_A] = a\ E[E[X|\mathcal{F}]1_A] = a\ E[X1_A] = E[a\ X1_A]$

□

Satz 1.3.3. *Für X$\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ gilt $E[X|\mathcal{F}] \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$.*

Beweis. Es sei A=$\{E[X|\mathcal{F}] \geq 0\}$. Dann gilt A $\in \mathcal{F}$ und es folgt:

$$1)\ E[E[X|\mathcal{F}]_+] = E[E[X|\mathcal{F}]1_A] = E[X\ 1_A] \underset{X\in\mathcal{L}^1}{<} \infty$$
$$2)\ E[E[X|\mathcal{F}]_-] = E[E[X|\mathcal{F}]1_{A^C}] = E[X\ 1_{A^C}] \underset{X\in\mathcal{L}^1}{<} \infty$$

□

Satz 1.3.4. *Für $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ gilt:*

1) $E[X|\{\emptyset, \Omega\}] = E[X]$ *(P-f.s. hier nicht nötig)*

2) $E[X|\mathcal{A}] = X$ *P-f.s.*

3) „Turmeigenschaft“
sind $\mathcal{G} \leq \mathcal{F} \leq \mathcal{A}$ zwei Unter-σ-Algebren von $\mathcal{A}$, so gilt:
$E[E[X|\mathcal{F}]|\mathcal{G}] = E[X|\mathcal{G}]$ *P-f.s*

Beweis. 1) und 2) sind trivial.
Zu 3): E[E[X|$\mathcal{F}$]|$\mathcal{G}$] ist $\mathcal{G}$-messbar und es gilt für alle A $\in \mathcal{G}$:

$$E[[E[X|\mathcal{F}]|\mathcal{G}]1_A] \overset{A \in \mathcal{G}}{=} E[E[X|\mathcal{F}]1_A] \overset{A \in \mathcal{F} \text{ wegen } \mathcal{G} \subseteq \mathcal{F}}{=} E[X1_A].$$

□

Satz 1.3.5 (Monotonie der bedingten Erwartung)**.** *Für $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ mit $X \leq Y$ p-f.s gilt:*

$$E[X|\mathcal{F}] \leq E[Y|\mathcal{F}] \quad P-f.s.$$

Beweis. Für alle A $\in \mathcal{F}$ gilt:

$$E[(E[Y|\mathcal{F}] - E[X|\mathcal{F}])1_A] = E[E[Y - X|\mathcal{F}]1_A] = E[(Y - X)1_A] \geq 0,$$

also E[Y|$\mathcal{F}$] $- E[X|\mathcal{F}] \geq 0$ P-f.s., da diese Differenz $\mathcal{F}$-messbar ist. □

Satz 1.3.6 (Satz von der monotonen Konvergenz für bedingte Erwartungen in $\mathcal{L}^1$)**.** *Es sei $(X_n)_{n \in \mathbb{N}}$ eine monoton steigende Folge in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ mit $X_n \overset{n\to\infty}{\to} X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ P-f.s. Dann gilt:*

$$E[X_n|\mathcal{F}] \overset{n\to\infty}{\to} E[X|\mathcal{F}]$$

P-f.s. und in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ konvergent.

Erinnerung. *Eine Folge $(Z_n)_{n \in \mathbb{N}}$ in $L^1(\Omega, \mathcal{A}, P)$ heißt in $L^1(\Omega, \mathcal{A}, P)$ konvergent gegen eine Zufallsvariable $Z \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, wenn $E[|Z_n - Z|] \overset{n\to\infty}{\to} 0$*

Beweis. Wegen der Monotonie der bedingten Erwartung (Satz 1.3.5) ist auch $(E[X_n|\mathcal{F}])_{n \in \mathbb{N}}$ P-f.s monoton steigend, also P-f.s. konvergent gegen eine $\mathcal{F}$- messbare Zufallsvariable Z. Nun gilt für alle A$\in \mathcal{F}$:

$$E[Z1_A] = \lim_{n\to\infty} E[E[X_n|\mathcal{F}]1_A] \lim_{n\to\infty} E_[X_n1_A] = E[X1_A]$$

Insbesondere ist Z $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, da Z $\overset{P-f.s.}{\geq}$ $X_1 \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$
und E[Z]=$E[Z1_\Omega] = E[X1_\Omega]$=E[X]$< \infty$ Es folgt E[X|$\mathcal{F}$] P-f.s, also $E[X_n|\mathcal{F}] \lim_{n\to\infty} E[X|\mathcal{F}]$ P-f.s. Es folgt $E[X|\mathcal{F}] - E[X_n|\mathcal{F}] \underset{n\to\infty}{\to} 0$ P-f.s. und wir haben die Majorante

$$0 \underset{P-f.s.}{\leq} E[X|\mathcal{F}] - E[X_n|\mathcal{F}] \underset{P-f.s.}{\leq} E[X|\mathcal{F}] - E[X_1|\mathcal{F}] \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

Es folgt mit dem Satz von der dominierten Konvergenz:

$$E[|E[X|\mathcal{F}] - E[X_n|\mathcal{F}]|] = E[E[X|\mathcal{F}] - E[X_1|\mathcal{F}]] \underset{n\to\infty}{\to} 0$$

also E[$X_n|\mathcal{F}$] $\underset{n\to\infty}{\to} E[X|\mathcal{F}]$ in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ □

Satz 1.3.7 (Satz von der dominerten Konvergenz für bedingte Erwartungen). *Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$, die P-f.s. gegen eine Zufallsvariable X konviergiert. Es existiere $Z \in (\Omega, \mathcal{A}, P)$, so dass $|X_n| \leq Z$ P-f.s. gilt. Dan ist $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und es gilt:$E[X_n|\mathcal{F}] \underset{n\to\infty}{\to} E[X|\mathcal{F}]$ P-f.s. und in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$.*

Beweis. X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ ist uns schon bekannt, weil auch $|X| \geq Z$ P-f.s. Wir setzen $U_n := \inf_{m:m\geq n} X_m$ und $V_n := \sup_{m:m\geq n} X_m$ punktweise, dann gilt: $-Z \leq U_n \leq X_n \leq V_n \leq Z$ P- f.s. für alle $n \in \mathbb{N}$ und $U_n \underset{n\to\infty}{\nearrow}$ X, $V_n \underset{n\to\infty}{\searrow}$ X P-f.s.
Aus dem Satz von der monotonen Konvergenz für bedingte Erwartungen (Satz 1.3.6.) folgt:

$$E[U_n|\mathcal{F}] \underset{n\to\infty}{\to} E[X|\mathcal{F}] \text{ und } E[V_n|\mathcal{F}] \underset{n\to\infty}{\to} E[X|\mathcal{F}]\text{P-f.s. in } \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

Aus der Monotonie der bedingten Erwartung folgt P-f.s.:

$$E[U_n|\mathcal{F}] \leq E[X_n|\mathcal{F}] \leq E[V_n|\mathcal{F}], \text{ also } E[X_n|\mathcal{F}] \underset{n\to\infty}{\to} E[X|\mathcal{F}]\text{P-f.s. und in } \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

□

Der folgende Satz verallgemeinert $E[aX|\mathcal{F}] = aE[X|\mathcal{F}]$ $(a \in \mathbb{R})$ wesentlich: Nicht nur konstante a können aus der bedingten Erwartung herangezogen werden, sondern auch $\mathcal{F}$-messbare Funktionen Y. Anschaulich: Gegeben $\mathcal{F}$ ist der Wert von Y bekannt, Y spielt dann eine ähnliche Rolle wie eine Konstante.

Satz 1.3.8. *Es sei $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und Y: $\Omega \to \mathbb{R}$ $\mathcal{F}$- messbar mit $YX \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. Dann gilt:*

$$E[XY|\mathcal{F}] = YE[X|\mathcal{F}] \ P - f.s.$$

Beweis. Wir zeigen das schrittweise

1) für Y=1_A, A $\in \mathcal{F}$
2) für Y=$\sum_{k=1}^{m} \alpha_k 1_{A_k}, \alpha_k \in \mathbb{R},\ A_k \in \mathcal{F}, m \in \mathbb{N}$
3) für Y≥ 0 $\mathcal{F}$- messbar
4) allgemein

zu 1) $1_A[E[X|\mathcal{F}]$ ist $\mathcal{F}$- messbar und integrierbar und es gilt für $B \in \mathcal{F}$:

$$E[1_A E[X|\mathcal{F}]1_B] = E[E[X|\mathcal{F}]1_{A\cap B}] \underset{A\cap B\in\mathcal{F}}{=} E[X1_{A\cap B} = E[1_A X 1_B]$$

Es folgt: $E[1_A X|\mathcal{F}] = 1_A E[X|\mathcal{F}]$ P-f.s.

zu 2)

$$E\left[\sum_{k=1}^{m} \alpha_k 1_{A_k} X|\mathcal{F}\right] = \sum_{k=1}^{m} E[\alpha_k X|\mathcal{F}] \overset{1)}{=} \sum_{k=1}^{m} \alpha_k 1_{A_k} E[X|\mathcal{F}] \ P - f.s.$$

zu 3) Es sei $(Y_n)_{n\in\mathbb{N}}$ eine Folge von $\mathcal{F}$- messbaren Treppenfunktionen wie in 2) mit $0 \leq Y_n \underset{n\to\infty}{\nearrow} Y$ punktweise, z.B. $Y_n = n \wedge 2^{-n}\lfloor 2^n Y \rfloor$.
Wegen $|Y_n X| \leq |YX| \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und $Y_n X \underset{n\to\infty}{\to} YX$ punktweise folgt P-f.s.

$$Y_n E[X|\mathcal{F}] \overset{2)}{=} E[Y_n X|\mathcal{F}] \xrightarrow[n\to\infty]{P-f.s.wg.dom.Konv.} E[YX|\mathcal{F}] \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

Andererseits gilt:

$$Y_n E[X|\mathcal{F}] \xrightarrow{n\to\infty} YE[X|\mathcal{F}] \text{ P-f.s. wegen } Y_n \xrightarrow{n\to\infty} YP-f.s.$$

Zusammen folgt die Behauptung: $E[YX|\mathcal{F}] \xrightarrow{n\to\infty} YE[X|\mathcal{F}]$ P-f.s.

zu 4) Wir zerlegen Y in Positiv- und Negativteil $Y = Y_+ - Y_-$ wobei $Y_+ = Y1_{\{Y\geq 0\}}$, $Y_- = Y1_{\{Y<0\}}$
Weil nach Vorraussetzung YX $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ ist auch $Y_+X = YX1_{\{Y\geq 0\}} \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und $Y_-X = YX1_{\{Y<0\}} \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$.
Weiter sind Y_+ und Y_- wegen $\{Y \geq 0\} \in \mathcal{F}$ $\mathcal{F}$-messbar, da Y $\mathcal{F}$- messbar ist. Es folgt P-f.s.

$$E[YX|\mathcal{F}] = E[Y_+X - Y_-X|\mathcal{F}] = E[Y_+X|\mathcal{F}] - E[Y_-X|\mathcal{F}]] \overset{3)}{=} Y_+E[X|\mathcal{F}] - Y_-E[YX|\mathcal{F}] = YE[X|\mathcal{F}]$$

□

1.4 Bedingte Erwartung für nicht notwendig integrierbare Zufallsvariablen

Bekanntlich ist die Erwartung E[X] nicht nur für X $\in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ definiert, sondern auch für alle Zufallsvariablen X, für die $E[X_+] < \infty$ oder $E[X_-] < \infty$ gilt. E[X] kann dann auch die Werte $\pm\infty$ annehmen.
Auch für die bedingte Erwartung kann man diesen Fall verallgemeinern. wir führen das hier nur für $E[X_-] < \infty$ aus , der andere Fall kann analog behandelt werden.

Lemma 1.4.1. *Es sei X: $\Omega \to \mathbb{R} \cup \{\pm\infty\}$ eine Zufallsvariable und $E[X_-] < \infty$. Es seinen $(Y_n)_{n\in\mathbb{N}}$ und $(Z_n)_{n\in\mathbb{N}}$ zwei aufsteigende Folgen in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ mit $Y_n \underset{n\to\infty}{\nearrow} X$ und $Z_n \underset{n\to\infty}{\nearrow} X$ P-f.s. Dann gilt:*

$$\lim_{n\to\infty} E[Y_n|\mathcal{F}] = \lim_{n\to\infty} E[Z_n|\mathcal{F}] \quad (\lim_{n\to\infty} P-f.s. \text{ gemeint})$$

Bemerkung. so eine Folge $(Y_n)_{n\in\mathbb{N}}$ existiert, z.B. $Y_n = X1_{\{X>-\infty\}} \wedge n$

Beweis. Für $m \in \mathbb{N}$ folgt aus $Z_m \leq X$ und $Y_n \underset{n\to\infty}{\nearrow} X$ P-f.s.:

$$\lim_{n\to\infty} E[Y_n|\mathcal{F}] \underset{Y_n \geq Y_n \cap Z_m}{\geq} \lim_{n\to\infty} E[Y_n \cap Z_m|\mathcal{F}] \underset{mon.konv.}{\overset{Z_m \in \mathcal{L}^1}{=}} E[Z_m|\mathcal{F}] \quad \text{und daher}$$

$$\lim_{n\to\infty} E[Y_n|\mathcal{F}] \geq \lim_{n\to\infty} E[Z_m|\mathcal{F}]$$

Mir vertauschten Rollen von $(Y_n)_{n\in\mathbb{N}}$ und $(Z_m)_{m\in\mathbb{N}}$ folgt ebenso:

$$\lim_{n\to\infty} E[Z_m|\mathcal{F}] \geq \lim_{n\to\infty} E[Y_n|\mathcal{F}] \; P-f.s.$$

also zusammen die Behauptung:

$$\lim_{n\to\infty} E[Y_n|\mathcal{F}] = \lim_{n\to\infty} E[Z_n|\mathcal{F}] \quad P-f.s.$$

□

Das legt folgende Definition nahe:

Definition 1.4.2. *Für Zufallsvariablen X: $\Omega \to \mathbb{R} \cup \{\pm\infty\}$ mit $E[X_-] < \infty$ sei $E[X|\mathcal{F}]$ definiert als eine <u>$\mathcal{F}$-messbare Variante</u> des P-f.s. Limes $\lim_{n\to\infty} E[Y_n|\mathcal{F}]$, wobei $(Y_n)_{n\in\mathbb{N}}$ beliebige aufsteigende Folge in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ mit $Y_n \underset{n\to\infty}{\nearrow} X$ P-f.s. bezeichnet.*

Damit wird $E[X|\mathcal{F}]$ bis auf Abänderung auf einer $\mathcal{F}$-Nullmenge wohldefiniert. Dei Regeln aus Abschnitt 1.3 gelten (mit kleinen Abänderungen) auch für die so verallgemeinerte bedingte Erwartung (hier keine Details). Nur 2 Beispiele ohne Beweis:

Satz 1.4.3. *Es sei X: $\Omega \to \mathbb{R} \cup \{\pm\infty\}$ eine Zufallsvariable mit $E[X_-] < \infty$. Weiter sei Z: $\Omega \to \mathbb{R} \cup \{\pm\infty\}$ eine Zufallsvariable mit folgenden Eigenschaften:*

1) Z ist $\mathcal{F}$-messbar

2) $\forall A \in \mathcal{F}$ existiert $E[Z1_A]$ und es gilt $E[X1_A]$

Dann ist Z=$E[X|\mathcal{F}]$ $\quad P-f.s.$ $\quad$ (ohne Beweis)

Satz 1.4.4 (Satz von der monotonen Konvergenz für bedingte Erwartungen)**.** *Es sei $(X_n)_{n\in\mathbb{N}}$ eine $P-f.s.$ aufsteigende Folge von Zufallsvariablen: $X_n\colon \Omega \to \mathbb{R} \cup \{\pm\infty\}$ mit $E[(X_1)_-] < \infty$ und $X = \lim\limits_{n\to\infty} X_n$ $P-f.s.$ Dann gilt: $E[X_n|\mathcal{F}] \underset{n\to\infty}{\to} E[X|\mathcal{F}] \quad P-f.s.$*

1.5 Die Jensensche Ungleichung

Es sei $I \subset \mathbb{R}$ ein Intervall.

Erinnerung. *Eine Funktion $\varphi : I \longrightarrow \mathbb{R}$ heißt konvex, wenn für alle $x, y \in I$ und alle $\alpha \in [0,1]$ gilt: $\varphi(\alpha x + (1-\alpha)y) \leq \alpha\varphi(x) + (1-\alpha)\varphi(y)$ $(\star)$*

Für konvexe Funktionen $\varphi : I \longrightarrow \mathbb{R}$ gilt:
Für alle $x \in \mathbb{R}$ ist der Differenzquotient $q_x : I\{x\} \to \mathbb{R}, q_x(y) = \dfrac{\varphi(y)-\varphi(x)}{y-x}$ monoton steigend.
Insbesondere existieren für x im Inneren von I die links- und rechtsseitige Ableitung.
$\varphi_l^{'}(x) = \lim\limits_{y\nearrow x} q_x(y)$, $\quad \varphi_v^{'}(x) = \lim q_x(y)$, und es gilt für alle $y < x < z$: $q_x(y) \leq \varphi_l^{'}(x) \leq \varphi_v^{'}(x) \leq q_x(z)$
Insbesondere gilt $\varphi(y) \geq \varphi(x) + \varphi_l^{'}(x)(y-x) \forall y \in I$ Als monoton steigende Funktionen sind $\varphi_l^{'}$ und $\varphi_v^{'}$ messbar.

Satz 1.5.1 (Jensensche Ungleichung)**.** *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, $\varphi : I \to \mathbb{R}$ konvex und $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ mit $X : \Omega \to I$ sowie $\mathcal{F} \subset \mathcal{A}$ eine Unter-σ-Algebra.*
Dann existiert $E[\varphi(x)|\mathcal{F}]$, und es gilt P-f.s. $E[\varphi(x)|\mathcal{F}] \geq \varphi(E[X|\mathcal{F}])$

Bemerkung. Die definierende Bed. $(\star)$ der Konvexität kann als Spezialfall $\mathcal{F} = \{\emptyset, \Omega\}, \mathcal{L}(X) = \alpha\delta_x + (1-\alpha)\delta_y$ der Jensenschen Ungleichung aufgefasst werden.

Beweis. Für alle $x \in I$ wissen wir: $\varphi(X) \geq \varphi(x) + \varphi_l^{'}(x)(X-x)(\star\star)$
Wegen $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ folgt $\varphi(x) + \varphi_l^{'}(x)(X-x) \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ und daher $E[\varphi(X)] \leq E[(\varphi(x) + \varphi_l^{'}(x)(X-x))] < \infty$
Insbesondere existiert $E[\varphi(X)|\mathcal{F}]$
Setzen wir nun $x = E[X|\mathcal{F}]$ in $(\star\star)$ ein, so folgt:
$\varphi(X) \geq \varphi(E[X|\mathcal{F}] + \varpi_l^{'}(E[X|\mathcal{F}])(X - E[X|\mathcal{F}])$
Weil wir nicht wissen, ob hier der 2. Summand in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ liegt, betrachten wir das erst auf den Ereignisse $A_n := \{|\underbrace{\varphi_l^{\prime}(E[X|\mathcal{F}])}_{\mathcal{F}\text{ -messbar}}| \leq n\} \in \mathcal{F}, n \in \mathbb{N}$:

$$1_{A_n}E[\varphi(X)|\mathcal{F}] = E[1_{A_n}\varphi(X)|\mathcal{F}] \geq E[1_{A_N}\varphi(E[X|\mathcal{F}])|\mathcal{F}] + E[\overbrace{\underbrace{1_{A_n}\varphi_l^{\prime}(E[X|\mathcal{F}])}_{\text{beschränkt},\mathcal{F}messbar}\ \underbrace{(X - E[X|\mathcal{F}])}_{\in\mathcal{L}^1}}^{\in\mathcal{L}^1}|\mathcal{F}] =$$
$1_{A_n}\varphi(E[X|\mathcal{F}] + 1_{A_n}\varphi_l^{'}(E[X|\mathcal{F}])E[X - E[X|\mathcal{F}]|\mathcal{F}] = 1_{A_n}\varphi(E[X|\mathcal{F}])$ P-f.s.

Weil das für alle $n \in \mathbb{N}$ gilt. folgt:
$E[\varphi(X)|\mathcal{F}] \geq \varphi(E[X|\mathcal{F}])$ P-f.s.

□

Beispiel 1.5.2. Für $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ gilt:
$E[X^2|\mathcal{F}] \geq E[X|\mathcal{F}]^2$
Das verallgemeinert die bekannte Ungleichung $E[X^2] \geq E[X]^2$

Als Anwendung besprechen wir die Beschränktheit der bedingten Erwartungswertbildung in den Räumen $\mathcal{L}^P(\Omega, \mathcal{A}, P), 1 \leq p < \infty$
Für $X : \Omega \to \mathbb{R}$ messbar sei $\|X\|_P = (\int\limits_\Omega |X|^P dP)^{\frac{1}{P}}$
$\mathcal{L}^P(\Omega, \mathcal{A}, P) := \{X : \Omega \to \mathbb{R} | X \text{ ist } \mathcal{A}\text{-messbar}, \|X\|_P < \infty\}$
Es sei $\mathcal{L}^P(\Omega, \mathcal{A}, P) = \mathcal{L}^P(\Omega, \mathcal{A}, P)/\sim$, wobei $\sim$ die Äquivalenzrelation $X \sim Y \Leftrightarrow X = Y$ P-f.s. bezeichnet.
Es bezeichne:
$[X]_\sim = \{y \in \mathcal{L}^P(\Omega, \mathcal{A}, P) | Y \sim X\}$ die Äquivalenzklasse von $X \in \mathcal{L}^P(\Omega, \mathcal{A}, P)$.

Bemerkung. In der Maßtheorie bzw. Funktionalanalysis lernen Sie:
$\|\cdot\|_P$ ist eine Halbnorm auf $\mathcal{L}^P(\Omega, \mathcal{A}, P)$ mit $\|X\|_P = 0 \Leftrightarrow X \sim 0$, induziert also eine (wieder mit $\|\cdot\|_P$ bezeichnet) Norm auf $\mathcal{L}^P(\Omega, \mathcal{A}, P)$. Mit dieser Norm wird $\mathcal{L}^P(\Omega, \mathcal{A}, P)$ eine Banachraum, also ein vollständiger normierter Raum.

Satz 1.5.3. *Für alle $p \in [1, \infty]$ und alle $X \in \mathcal{L}^P(\Omega, \mathcal{A}, P)$ gilt:*

$$\|E[X|\mathcal{F}]\|_P \leq \|X\|_P$$

Bemerkung. Für $1 \leq p < q < \infty$ gilt:

$$\mathcal{L}^q(\Omega, \mathcal{A}, P) \subset \mathcal{L}^P(\Omega, \mathcal{A}, P)$$

da $E[|X|^P] \leq E[1 + |X|^q] < \infty$ für $X \in \mathcal{L}^q(\Omega, \mathcal{A}, P)$

Beweis. Die Abb. $\varphi : \mathbb{R} \to \mathbb{R}, \varphi(t) = |t|^P$ ist konvex. Die Jensensche Ungleichung liefert:
$|E[X|\mathcal{F}]|^P \leq E[|X|^P|\mathcal{F}]$ und daher $\|E[X|\mathcal{F}]\|_P^P = E[|E[X|\mathcal{F}]|^P] < E[E[|X|^P|\mathcal{F}]] = E[|X|^P] = \|X\|_P^P$ also die Behauptung. □

Insbesondere liefert bedingte Erwartungsbildung eine beschränkte lineare Abbildung

$$E[\cdot|\mathcal{F}] : \mathcal{L}^P(\Omega, \mathcal{A}, P) \to \mathcal{L}^P(\Omega, \mathcal{A}, P), [X]_\sim \mapsto [E[X|\mathcal{F}]]_\sim$$

1.6 $\mathcal{L}^2$ -Theorie der bedingten Erwartung

Besonders schön wird die Theorie bed. Erwartungen im Raum $\mathcal{L}^2(\Omega, \mathcal{A}, P)$. Dieser Raum bildet bekanntlich mit dem Skalarprodukt $\langle [X]_\sim, [Y]_\sim \rangle := E[XY]$ sogar einen Hilbertraum, also einen Banachraum, dessen Norm durch das Skalarprodukt induziert wird: $\|x\|_2 = \sqrt{\langle X, X \rangle}, X \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$. Ist nun $\mathcal{F} \subset \mathcal{A}$ eine Unter-σ-Algebra, so ist $\mathcal{L}^2(\Omega, \mathcal{F}, P) \subset \mathcal{L}^2(\Omega, \mathcal{A}, P)$
Damit können wir $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ als Teilraum von $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ auffassen.
Weil $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ vollständig ist, ist dieser Teilraum sogar abgeschlossen in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$

Satz 1.6.1. *Für alle $X \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ und alle $Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P), \mathcal{F}$ Unter-σ-Algebra sind Y und $X - E[X|\mathcal{F}]$ orthogonal zueinander. Weil $E[\cdot|\mathcal{F}] : \mathcal{L}^2(\Omega, \mathcal{A}, P) \to \mathcal{L}^2(\Omega, \mathcal{F}, P)$ linear ist und $E[X|\mathcal{F}] = X$ für alle $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ gilt, können wir das auch so formulieren:*
$E[\cdots|\mathcal{F}] : \mathcal{L}^2(\Omega, \mathcal{A}, P) \to \mathcal{L}^2(\Omega, \mathcal{F}, P)$ ist die orthogonale Projektion auf $\mathcal{L}^2(\Omega, \mathcal{F}, P)$

Beweis. $E[\underbrace{\overbrace{\underbrace{Y}_{\in\mathcal{L}^2}\cdot\underbrace{(X-E[X|\mathcal{F}]}_{\in\mathcal{L}^2}}^{\in\mathcal{L}^1}}] = E[E[\underbrace{Y}_{\mathcal{F}-messbar}\cdot(X-E[X|\mathcal{F}])|\mathcal{F}]] = E[Y\cdot E[X-E[X|\mathcal{F}]|\mathcal{F}]] = 0$ □

Bemerkung. Das Lemma motiviert eine alternative Einführung der bedingten Erwartungen für $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$. In der Funktionalanalysis beweist man:
Ist V ein abgeschlossener Teilraum eines Hilbertraums H , so gibt es eine orthogonale Projektion π von H auf V mit folgenden Eigenschaften:

1) Für alle $x \in V$ ist $\pi(x) = x$
2) Für alle $x \in H$ und $y \in V$ ist $y \perp x - \pi(x)$

Damit hätten wir unabhängig von Satz von Radon-Nikodyn $E[\cdot|\mathcal{F}] : \mathcal{L}^2(\Omega,\mathcal{A},P) \to \mathcal{L}^2(\Omega,\mathcal{F},P)$ auch als orthogonale Projektion definieren können. Die charackteristischen Eigenschaften erhält man dann so:

1) Für alle $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ ist $E[X|\mathcal{F}] \in \mathcal{L}^2(\Omega,\mathcal{F},P)$ eine Klasse von $\mathcal{F}$-messbaren Zufallsvariablen
2) Für $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ und $Y \in \mathcal{L}^2(\Omega,\mathcal{F},P)$ ist $0 = E[Y(X-E[X|\mathcal{F}])] = E[YX] - E[YE[X|\mathcal{F}]]$. Insbesondere erhalten wir für $Y = 1_A, A \in \mathcal{F} : E[1_A X] = E[1_A E[X|\mathcal{F}]]$

Erinnerung (Satz des Pythagoras)**.**

Lemma 1.6.2. *Sind $X, Y \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ orthogonal zueinander, so gilt:*

$$E[(X+Y)^2] = E[X^2] + E[Y^2]$$

Allgemeiner: Sind $X_1,\dots,X_n \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ paarweise orthogonal zueinander, so gilt:

$$E\left[\left(\sum_{k=1}^n X_k\right)^2\right] = \sum_{k=1}^n E[X_k^2]$$

Beweis. Es gilt:
$E[(\sum_{k=1}^n X_k)^2] = \sum_{k=1}^n \sum_{l=1}^n \underbrace{E[X_k X_l]}_{=0 \text{ für } k\neq l} = \sum_{k=1}^n E[X_k^2]$ □

Wir betrachten folgenden Spezialfall:
Für $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ sind $X - E[X|\mathcal{F}]$ und $E[X|\mathcal{F}] - E[X] \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ orthogonal zueinander.
Es folgt:
$\mathrm{Var}(X) = E[(X-E[X])^2] = E[\underbrace{((X-E[X|\mathcal{F}]) + (E[X|\mathcal{F}]-E[X]))}_{\text{Summanden orthogonal zueinander}}^2] \underset{\text{Pythagoras}}{=} E[(X-E[X|\mathcal{F}])^2] + E[(E[X|\mathcal{F}]-E[X])^2] = E[E[(X-E[X|\mathcal{F}])]] + \mathrm{Var}(E[X|\mathcal{F}])$

Definition 1.6.3. *Für $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ definieren wir die bedingte Varianz von X gegeben $\mathcal{F}$ durch:*

$$\mathrm{Var}(X|\mathcal{F}) := E[(X-E[X|\mathcal{F}])^2|\mathcal{F}]$$

Damit haben wir gezeigt:

Lemma 1.6.4. *Für $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ gilt:*

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|\mathcal{F})] + \mathrm{Var}(E[X|\mathcal{F}])$$

Beispiel 1.6.5. Ein fairer Spielwürfel wir so lange geworfen, bis zum ersten Mal eine "6 " erscheint. Es bezeichnet S die Summe der Augenzahlen. Wir berechnen $E[S]$ und $\mathrm{Var}(S)$

Zu $E[S]$ Es sei N die Anzahl der Würfe bis zum ersten Auftreten von "6 ". Dann ist N geometrisch verteilt:
$P[N=n]=\left(\frac{5}{6}\right)^{n-1}\cdot\frac{1}{6}$ für $n\in\mathbb{N}$
$P[N=\infty]=0$, und bedingt auf $N=n$ sind die ersten $N-1$ Augenzahlen $X_1,\dots,X_{N-1}$ iid gleichverteilt auf $[1,2,3,4,5]$ und $X_N=6$
$E[S|N=n]=\sum\limits_{k=1}^{n}E[X_k|N=n]=3(n-1)+6=3n+3$, also $E[S|N]=3N+3$
Also: $E[S]=E[E[S|N]]=E[3N+3]=3E[N]+3\underset{E[N]=6}{=}3\cdot 6+3=21$
$E[N]=6$ sieht man einfach so:
$E[N]=1\cdot P[N=1]+\underbrace{E[N|N\geq 2]}_{1+E[N]\text{ wg. Gedächtnislosigkeit der geom. Vert.}}\cdot P[N\geq 2]=\frac{1}{6}+(1+E[N])\cdot\frac{5}{6}$
wegen $E[N]<\infty$ folgt hieraus $E[N]=6$

Zu $\mathrm{Var}(S)$ Es gilt:

$$\mathrm{Var}(S)=E[\mathrm{Var}(S|N)]+\mathrm{Var}(E[S|N])$$

Nun gilt:

$$\mathrm{Var}(S|N)=\mathrm{Var}(\sum_{k=1}^{N-1}X_k+6|N)\underbrace{=}_{\substack{\uparrow\\ \text{Unabhängigkeit der }X_1,\dots,X_{N-1} ggn.\ N}}\sum_{k=1}^{N-1}\mathrm{Var}(X_k|N)\overbrace{=}^{\substack{\text{Gegeben }N=n>k\text{ ist }X_k\text{ gleicherverteilt auf }\{1,\dots,5\}\\ \downarrow}}(N-1)\cdot 2$$

denn die Varianz der Gleichverteilung auf $\{1,\dots,5\}$ ist 2.
Es folgt:

$$E[\mathrm{Var}(S|N)]=E[2(N-1)]=2\cdot\underbrace{E[N]}_{=6}-2=10$$

Weiter:
$\mathrm{Var}(E[S|N])=\mathrm{Var}(3N+3)=9\,\mathrm{Var}(N)$
Statt $\mathrm{Var}(N)$ direkt (was auch möglich ist) zu berechnen, bedingen wir auf die von der Partition $\{N=1\},\{N\geq 2\}$ erzeugte σ-Algebra $\mathcal{F}$.

$$\mathrm{Var}(N)=E[\mathrm{Var}(N|\mathcal{F})]+\mathrm{Var}(E[N|\mathcal{F}])$$

Nun ist

$$\mathrm{Var}(N|\mathcal{F})=\underbrace{\mathrm{Var}(N|N=1)1_{\{N=1\}}}_{=0}+\underbrace{\mathrm{Var}(N|N\geq 2)}_{\mathrm{Var}(N)}1_{\{N\geq 2\}}\overset{\substack{\text{Gedächtnislosigkeit der geom. Verteilung}\\ \uparrow}}{=}\mathrm{Var}(N)1_{\{N\geq 2\}}$$

$$E[\mathrm{Var}(N|\mathcal{F})]=\mathrm{Var}(N)\cdot P[N\geq 2]=\frac{5}{6}\mathrm{Var}(N)$$

Weiter:
$E[N|\mathcal{F}]=1\cdot 1_{\{N=1\}}+\underbrace{E[N|N\geq 2]}_{1+E[N]=7}\cdot 1_{\{N\geq 2\}}=1_{\{N=1\}}+7\cdot 1_{\{N\geq 2\}}$
Also:
$\mathrm{Var}(E[N|\mathcal{F}])=\mathrm{Var}(1_{\{N=1\}}+7\cdot 1_{\{N\geq 2\}})=\mathrm{Var}(6\cdot 1_{\{N\geq 2\}})=36\cdot\mathrm{Var}(1_{\{N\geq 2\}})=$
$36\cdot(P[N\geq 2]-P[N\geq 2]^2)=36\cdot(\frac{5}{6}-(\frac{5}{6})^2)=5$

$\operatorname{Var}(N) = \frac{5}{6}\operatorname{Var}(N) + 5$ Nun ist $\operatorname{Var}(N) < \infty$, da $P[N \geq k]$ für $k \to \infty$ exponentiell schnell fällt.
Es folgt:
$\operatorname{Var}(N) = 30$ und damit $E[\operatorname{Var}(N|\mathcal{F})] = \frac{5}{6}\operatorname{Var}(N) = 25$
Oben eingesetzt:

$$\operatorname{Var}(S) = E[\operatorname{Var}(S|N)] + \operatorname{Var}(E[S|N]) = 10 + 25 = 35$$

2 Martingale

2.1 Motivation: Approximation bed. Erwartungen

Wir besprechen als Einführung eine Konstruktion bed. Ewartungen durch Approximation durch elementare bed. Erwartungen. Hierzu sei $X \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ eine Zufallsvariable und $\mathcal{F} \subseteq \mathcal{A}$ eine Unter-σ-Algebra. Wir nehmen an, dass $\mathcal{F}$ eine abzählbares Erzeugendensystem besitzt: $\mathcal{F} = \sigma(A_n : n \in \mathbb{N})$ mit $A_n \in \mathcal{F}$ Ereignisse $\forall\ n \in \mathbb{N}$.
Zum Beispiel ist das stets der Fall, wenn $\mathcal{F} = \sigma(Y)$ mit einer Zufallsvariable $Y : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $n \in \mathbb{N}$, denn hier ist

$$\left\{\{Y \in B(x,r)\} \,|\, x \in \mathbb{Q}^n, r \in \mathbb{Q}^+\right\}$$

ein abzählbares Erzeugendensystem ($B(x,r)$ offene Kugel, x Mittelpunkt, r Radius).
Wir sehen für $k \in \mathbb{N}$: $\mathcal{F}_n = \sigma(A_1, \ldots, A_n)$.
Insbesondere ist $\mathcal{F}_n$ endlich und es gilt:

$$\mathcal{F} = \mathcal{F}_\infty := \sigma(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n)$$

Offensichtlich gilt: $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \ldots \subseteq \mathcal{F}_\infty$
Solche aufsteigenden Familien von Unter-σ-Algebren bekommen eine eigene Bezeichnung:

Definition 2.1.1. *Es sein $(\Omega, \mathcal{A})$ ein Ereignisraum. Eine aufsteigende Folge $(\mathcal{F}_n)_{n\in\mathbb{N}}$ von Unter-σ-Algebren von $\mathcal{A}$ heißt Filtration von $\mathcal{A}$.*
Sei allgemeiner $(J, \leq)$ eine linear geordnete Menge und $(\mathcal{F}_t)_{t\in J}$ eine Familie von Unter-σ-Algebren von $\mathcal{A}$. Wenn $\forall\ s,t \in J$ mit $s \leq t$ gilt: $\mathcal{F}_s \subseteq \mathcal{F}_t$, so heißt $(\mathcal{F}_t)_{t\in J}$ eine Filtration in $\mathcal{A}$ über $(J, \leq)$.

Oben haben wir eine Filtration $(\mathcal{F}_n)_{n\in N\cup\{\infty\}}$ in $\mathcal{A}$ über $(N \cup \{\infty\}, \leq)$.
Die bedingte Erwartung $X_n := E[X|\mathcal{F}_n]$, $n \in \mathbb{N}$ sind wie in Abschnitt 1.1 als Elementare bedingte Erwartungen explizit gegeben. Wir wollen nun $X_\infty := E[X|\mathcal{F}_\infty]$ als Limes der $X_n, n \in \mathbb{N}$ darstellen.
Aus der Turmeigenschaft der bedingten Erwartung wissen wir für $n, m \in \mathbb{N} \cup \{\infty\}$ mit $n \leq m$:

$$X_n = E[X|\mathcal{F}_n] \underbrace{=}_{\mathcal{F}_n \subseteq \mathcal{F}_m} E[E[X|\mathcal{F}_m]|\mathcal{F}_n] = E[X_m|\mathcal{F}_n]$$

Bezeichnung dafür:

Definition 2.1.2. *Es sei $(\mathcal{F}_t)_{t\in J}$ eine Filtration in $\mathcal{A}$ über $(J, \leq)$. Eine Familie $(X_t)_{t\in J}$ von reellen Zufallsvariablen heißt adaptiert an diese Filtration, wenn X_t bzgl. $\mathcal{F}_t$ messbar ist $\forall\ t \in J$. Gilt $X_t \in \mathcal{L}^1(\Omega, \mathcal{F}_t, P)$ und $X_s = \overline{E[X_t|\mathcal{F}_t]}$ P-f.s. $\forall\ s,t \in J$ mit $s \leq t$, so heißt $(X_t)_{t\in J}$ eine Martingal bzgl. der Filtration $(\mathcal{F}_t)_{t\in J}$*

Definition 2.1.3. *Es sei $(\mathcal{F}_t)_{t\in I}$ eine Filtration in $\mathcal{A}$ über $(I, \subseteq)$. Eine Familie $(X_t)_{t\in I}$ von ZV heißt adaptiert, wenn für alle $t \in I X_t$ bzgl. $\mathcal{F}_t$ messbar ist.*
Gilt $X_t \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ l.o.t?? und $E[X_t|\mathcal{F}_s] = X_s P$-f.s. für alle $s,t \in I$ mit $s \leq t$, so heißt $(X_t)_{t\in I}$ ein Martingal bzgl. $(\mathcal{F}_t)_{t\in I}$.
Anschauung: Martingale beschreiben "faire"Glücksspiele.

Martingale in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ besitzen folgende schöne Eigenschaft:

Satz 2.1.4 (Orthogonalität der Zuwächse eines $\mathcal{L}^2$-Martingals)**.** *Ist $(X_t)_{t\in I}$ eine Martingal bzgl. $(\mathcal{F}_t)_{t\in I}$ mit $X_t \in \mathcal{L}^2(\Omega, \mathcal{F}_t, P)$ für alle $t \in I$, so gilt für alle r, s, t mit $r \leq s \leq t$:*
$X_r, X_s - X_r$ und $X_t - X_s$ sind paarweise orthogonal zueinander, d.h.:

1. $E[X_r(X_s - X_r)] = 0$
2. $E[X_r(X_t - X_s)] = 0$
3. $E[(X_s - X_r)(X_t - X_s)] = 0$

Stärker gilt sogar:

i. $E[X_r(X_s - X_r)|\mathcal{F}_r] = 0 P$*-f.s.*

ii. $E[X_r(X_t - X_s)|\mathcal{F}_s] = 0 P$*-f.s.*

iii. $E[(X_s - X_r)(X_t - X_s)|\mathcal{F}_s] = 0 P$*-f.s.*

Beweis. 1.,2.,3. erhält man aus i.,ii.,iii., indem man die (unbedingte) Erwartung bildet.
i. erhält man aus ii. indem man (r, s, t) durch (r, r, s) ersetzt.

Zu ii.: $E[X_r(X_t - X_s)|\mathcal{F}_s] = X_r(\underbrace{E[X_t|\mathcal{F}_s]}_{=X_s(\text{ Martingaleigenschaft})} - \overbrace{E[X_s|\mathcal{F}_s]}^{=X_s,\text{da}(X_k)_{k\in\mathbb{Z}}\text{adaptiert}}) = 0P$-f.s.

Indem wir in ii. (r, s, t) durch (s, s, t) ersetzen, erhalten wir

iv. $E[X_s(X_t - X_s)|\mathcal{F}_s] = 0P-$f.s.

und damit iii. als Differenz von iv. und ii. ◻

$[X_n = E[\underset{\in\mathcal{L}^2}{\underset{\uparrow}{X}} \,|\, \underbrace{\sigma(A_1, ..., A_n)}_{=\mathcal{F}_n}]$

Für unser MArtingal $(X_n)_{n\in\mathbb{N}}$ von oben (und allgemeiner für jedes Martingal in $\mathcal{L}^2$ bedeutet das:
Die Zuwächse $X_{n+1} - X_n, n \in \mathbb{N}$, sind paarweise orthogonal.
Aus dem Satz von Pythagoras folgt für $m > n$:

$$E[(X_m - X_n)^2] = E\left[(\sum_{k=n}^{m-1}(X_{k+1} - X_k))^2\right] = \sum_{k=n}^{m-1} E[(X_{k+1} - X_k)^2]$$

Nun ist die Folge $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{L}^2$ beschränkt. d.h. $\sup\limits_{n\in\mathbb{N}} E[X_n^2] < \infty$
In der Tat: Aus $X_n = E[X|\mathcal{F}_n]$ und $X \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ folgt $E[X^2] \underbrace{\subseteq}_{\text{Jensen oder } X_n \perp X - X_n} E[X^2] < \infty$ für alle $n \in \mathbb{N}$.

Nun gilt:

Satz 2.1.5. *Jedes in $\mathcal{L}^2$ beschränktes Martingal $(X_n)_{n\in\mathbb{N}}$ bzgl. einer Filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ ist in $\mathcal{L}^2$ konvergent. Es gibt also ein $Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ mit $E[(X_n - Y)^2] \underset{n\to\infty}{\longrightarrow} 0$*
Genauer gesagt kann man $Y \in \mathcal{L}^2(\Omega, \mathcal{F}_\infty, P)$ wählen, wobei $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n)$

Beweis. Wir zeigen, dass $(X_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{L}^2(\Omega, \mathcal{F}_\infty, P)$ bildet, d.h.

$$\lim_{k\to\infty} \sup_{m,n:m>n\geq k} E[(X_m - X_n)^2] = 0$$

Es sei $c := \sup\limits_{n\in\mathbb{N}} E[X_n^2] \underset{\substack{\uparrow \\ \text{Vorraussetzung}}}{<} \infty$

Es gilt:

$$\sum_{k=1}^{n-1} E[(X_{k+1}-X_k)^2] \underset{\substack{\uparrow \\ \text{orthogon. d. Zuwächse}}}{=} E[(X_n-X_1)^2] \le E[X_1^2] + E[(X_n-X_1)^2] \underset{\substack{\uparrow \\ X_1 \perp X_n - X_1}}{=} E[X_n^2] \le c < \infty \quad \forall n \in \mathbb{N}$$

Insbesondere ist $\sum\limits_{k=1}^{\infty} E[(X_{k+1}-X_k)^2]$ in $\mathbb{R}$ konvergent, also

$$\sup_{m,n:m>n\ge k} E[(X_m-X_n)^2] = \sup_{m,n:m>n\ge k} \sum_{l=n}^{m-1} E[(X_{l+1}-X_l)^2] = \sum_{l=k}^{\infty} E[(X_{l+1}-X_l)^2] \underset{k\to\infty}{\longrightarrow} 0$$

Damit ist gezeigt, dass $(X_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$ bildet.
Aus der Funktionalanalysis/Maßtheorie verwenden wir, dass $\mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$ ein vollständiger Raum ist, d.h. jedes Cauchyfolge in $\mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$ konvergiert in $\mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$. Es folgt die Behauptung. $\square$

Wir zeigen jetzt, dass für den $\mathcal{L}^2$- Grenzwert Y von $(X_n)_{n\in\mathbb{N}}$ sogar gilt $Y = E[X|\mathcal{F}_\infty]$:

Satz 2.1.6. *Es sei $X \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ und $(\mathcal{F}_n)_{n\in\mathbb{N}\cup\{\infty\}}$ eine Filtration über $(\Omega,\mathcal{A})$ mit $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n)$. Dann gilt:*

$$E[X|\mathcal{F}_n] \underset{n\to\infty}{\longrightarrow} E[X|\mathcal{F}_\infty] \text{ in } \mathcal{L}^2(\Omega,\mathcal{A},P)$$

Beweis. Wir wissen $E[X|\mathcal{F}_n] \underset{n\to\infty}{\longrightarrow} Y$ in $\mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$ für ein $Y \in \mathcal{L}^2(\Omega,\mathcal{F}_\infty,P)$. Wir zeigen:
$Y = E[X|\mathcal{F}_\infty] P$-f.s. Weil Y bzgl. $\mathcal{F}_\infty$ messbar ist, müssen wir nur noch $E[Y1_A] = E[X1_A]$ für alle $A \in \mathcal{F}_\infty$ zeigen.
Im Fall $A \in \mathcal{F}_n, n \in \mathbb{N}$ folgt dann

$$E[(X_m-Y)1_A] \underbrace{\le}_{\text{Cauchy Schwarz}} E[(X_m-Y)^2]^{\frac{1}{2}} \underbrace{E[1_A^2]^{\frac{1}{2}}}_{\le 1} \underset{m\to\infty}{\longrightarrow} 0$$

wegen $X_m \underset{m\to\infty}{\longrightarrow} Y$ in $\mathcal{L}^2$
also $E[X_m 1_A] \underset{m\to\infty}{\longrightarrow} E[Y1_A]$
Für $m \ge n$ folgt wegen $A \in \mathcal{F}_n \subseteq \mathcal{F}_m$ und $X_m = E[X|\mathcal{F}_m]$:
$E[X_m 1_A] = E[X1_A]$ Zusammen folgt die behauptung:
$E[Y1_A] = E[X1_A]$
Nun setzten wir: $\mathcal{D} = \{A \in \mathcal{F}_\infty | E[Y1_A] = E[X1_A]\}$
Nach dem oben Gezeigten wissen wir:
$\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n \subseteq \mathcal{D}$
$\mathcal{D}$ ist ein Dynkin-System:

1. $\emptyset \in \mathcal{D}$ ist klar

2. Aus $A \in \mathcal{D}$ folgt $A^c \in \mathcal{D}$ wegen

$$E[Y1_{A^c}] = E[Y1_\Omega] - E[Y1_A] \overset{A\in\mathcal{D},\Omega\in\mathcal{F}_1\subseteq\mathcal{D}}{=} E[X1_\Omega] - E[X1_A] = E[X1_{A^c}]$$

3. Sind $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkte Elemente von $\mathcal{D}$ und setzen wir $A = \bigcup\limits_{n\in\mathbb{N}} A_n$, so folgt $A \in \mathcal{D}$.
   Wegen

$$E[Y1_A] \overset{\text{dom.Konv.mit Maj.}|Y|\in\mathcal{L}^1}{\overset{\downarrow}{=}} \lim_{n\to\infty} E[Y \underbrace{\sum_{k=1}^{n} 1_{A_k}}_{\underset{n\to\infty}{\longrightarrow} 1_A \text{ punktweise}}] = \lim_{n\to\infty} \sum_{k=1}^{n} E[Y1_{A_k}] \underset{A_k\in\mathcal{D}}{=} \lim_{n\to\infty} \sum_{k=1}^{n} E[X1_{A_k}] =$$

$$\lim_{n\to\infty} E[X \sum_{k=1}^{n} 1_{A_k}] \overset{\text{Maj.}|X|}{\overset{\downarrow}{=}} E[X \sum_{k=1}^{\infty}] \underset{A=\bigcup\limits_{k\in\mathbb{N}} A_k}{=} E[X1_A], \text{also } A \in \mathcal{D}$$

Da $\mathcal{F}$ $\cap$-stabil sind und $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq ...$ gilt, ist auch $\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n$ $\cap$-stabil. Es folgt aus dem Dynkin-Lemma: $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n) \subseteq \mathcal{D}$, also die Behauptung:

$$E[Y1_A] = E[X1_A] \ \forall A \in \mathcal{F}_\infty.$$

□

Zusammenfassung: Damit haben wir unsere gewünsche Approximation bedingter Erwartungen $E[X|\sigma(A_n : n \in \mathbb{N})]$ für $X \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ gefunden.
Für $n \to \infty$ konvergieren die elementaren bed. Erwartungen
$E[X|\sigma(A_1, \ldots, A_n)]$ in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ gegen $E[X|\sigma(A_n : n \in \mathbb{N})]$

2.2 Sub- und Supermartingale, Doob-Zerlegung

Man kann sich Martingale $(X_n)_{n\in\mathbb{N}}$ bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}}$ als faire Glücksspiele vorstellen: Wenn X_n das Kapital des Spielers zur Zeit n bedeutet und $\mathcal{F}_n$ die bis zur Zeit n beobachtbare Information, bedeutet die Martingalbedingung $E[X_m|\mathcal{F}_n] = X_n$ für $m \geq n$ folgendes:
Gegeben die beobachtbare Information „bis heute“(n), ist die Prognose für den zukünftigen Wert X_m gleich dem heutigen Wert X_n.
Variante :„für den Spieler günstige bzw. ungünstige Spiele“.

Definition 2.2.1. *Es sei $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum, $(\mathcal{F}_s)_{s\in I}$ eine Filtration darauf über$(I, \leq)$ und $(X_s)_{s\in I}$ eine adaptierte Familie von Zufallsvariblen in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$. Diese Familie heißt ein Submartingal, wenn $X_s \leq E[X_t|\mathcal{F}_s]$ $P - f.s.$ $\forall s, t \in I$ mit $s \leq t$ und Supermartingal, wenn $X_s \geq E[X_t|\mathcal{F}_s]$ $P - f.s.$ $\forall s, t \in I$ mit $s \leq t$.*

Merke. Submartingal: heutiger Wert unter künftiger Erwartung, tendenziel ansteigend, günstig für den Spieler.
Supermartingale: heutiger Wert üer künftiger Erwartung, tendenziell fallend, ungünstig für den Spieler „Supermartingale sind nicht super “.

Submartingale können als Summe eines Martingals und einer „vorhersehbaren“ansteigenden Folge geschrieben werden.

Definition 2.2.2. *Es sei $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ eine Fitration. Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen heißt vorhersehbar, wenn $(X_{n+1})_{n\in\mathbb{N}_0}$ adaptiert ist.*

Anschaulich gesagt: Der Wurf von X_{n+1} ist uns schon zur Zeit n bekannt.

Lemma 2.2.3 (Doob-Zerlegung)**.** *Es sei $(F_n)_{n\in\mathbb{N}_0}$ eine Filtration in $(\Omega, \mathcal{A}, P)$ und $(X_n)_{n\in\mathbb{N}_0}$ eine adaptierte Folge in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$. Dann gibt es eine vorhersehbare Folge in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ $(A_n)_{n\in\mathbb{N}_0}$ und ein Martingal $(M_n)_{n\in\mathbb{N}_0}$ bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, mit folgenden Eigenschaften:*

1) $A_0 = 0$

2) $X_n = A_n + M_n \ \forall n \in \mathbb{N}_0$

Die A_n und M_n, $n \in \mathbb{N}_0$, sind $P - f.s.$ eindeutig.

Beweis. Wir konstruieren A_n, M_n, $n \in \mathbb{N}_0$ rekursiv: Rekursionsanfang: $A_0 = 0$ und damit $M_0 = X_0 - A_0 = X_0$ sind durch 1) und 2) vorgegeben. Rekursionsschritt: Es sei $A_0, \dots, A_n$ und $M_0, \dots, M_n$ in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ bereits konstruiert, vorhersehbar bzw. Martingal. Die gesuchte Zerlegung $X_{n+1} = A_{n+1} + M_{n+1}$ muss folgende Bedingung erfüllen:

$$E[X_{n+1}|\mathcal{F}_n] = \underbrace{E[A_{n+1}|\mathcal{F}_n]}_{\text{soll } A_{n+1} \text{ sein, vorhersehbar}} + \underbrace{E[M_{n+1}|\mathcal{F}_n]}_{\text{soll } M_n \text{ sein, Martingaleig.}} = A_{n+1} + M_n \ P - f.s.$$

$E[M_{n+1}|\mathcal{F}_n] = M_n$ soll $P - f.s.$ gelten. Uns bleibt $P - f.s.$ nur folgende Möglichkeit der Definition:

1)
$$A_{n+1} \coloneqq E[X_{n+1}|\mathcal{F}_n] - M_n \in \mathcal{L}^1(\Omega, \mathcal{A}, P) \ \textit{und}$$

2)
$$M_{n+1} \coloneqq X_{n+1} - A_{n+1} \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

□

Offensichtlich ist A_{n+1} $\mathcal{F}_n$ -messbar und M_{n+1} $\mathcal{F}_{n+1}$-messbar und es gilt:

$$E[M_{n+1}|\mathcal{F}_n] = E[X_{n+1}|\mathcal{F}_n] - E[A_{n+1}|\mathcal{F}_n] = E[X_{n+1}|\mathcal{F}_n] - A_{n+1} = M_n \ P - f.s.$$

Insbesondere ist $(M_n)_{n\in\mathbb{N}_0}$ ein Martingal, denn es gilt :

Bemerkung. Eine adaptierte Folge $(M_n)_{n\in\mathbb{N}_0}$ in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ ist ein Martingal (bzw. Submartingal bzw. Supermatingal) bzgl. $(F_n)_n$ genau dann, wenn $\forall n \in \mathbb{N}_0$ gilt:

$$E[M_{n+1}|\mathcal{F}_n] = M_n \ P - f.s. \quad \text{(bzw. „}\geq\text{"bzw. „}\leq\text{"statt „}=\text{")}$$

Beweis. „⇒"ist trivial.
„⇐": Wir zeigen induktiv $\forall k \in \mathbb{N}_0$:

$$E[M_{n+k}|\mathcal{F}_n] \underset{bzw. \leq,\geq}{=} M_n \ P - f.s.$$

Für k=0 folgt dies sofort aus der Adaptiertheit. Für den Induktionsschritt k→ k+1 rechnen wir:

$$E[M_{n+k+1}|\mathcal{F}_n] \underset{Turmeigenschaft:\ \mathcal{F}_{n+k} \geq \mathcal{F}_n}{=} E[E[M_{n+k+1}|\mathcal{F}_{n+k}]|\mathcal{F}_n] = E[M_{n+k}|\mathcal{F}_n] = M_n \ P - f.s.$$

□

Lemma 2.2.4 (Doob-Zerlegung von Sub- und Supermartingalen)**.** *Es sei $X_n = A_n + M_n, n \in \mathbb{N}_0$, die Doob- Zerlegung einer Adaptierten Folge $(X_n)_{n\in\mathbb{N}_0}$ in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$. Dann ist $(X_n)_{n\in\mathbb{N}_0}$ ein Submartingal (bzw. Supermartingal) genau dann, wenn $(A_n)_{n\in\mathbb{N}_0}$ $P - f.s.$ monoton steigend bzw. monoton fallend ist.*

Beweis. Es gilt:

$$A_{n+1} - A_n = E[X_{n+1}|\mathcal{F}_n] - M_n - A_n = E[X_{n+1}|\mathcal{F}_n] - X_n$$

Das ist ≥ 0 $P - f.s.$ für alle n genau dann, wenn $(X_n)_n$ ein Submartingal bilden und ≤ 0 $P - f.s.$ für alle n genau dann, wenn $(X_n)_n$ ein Supermartingal bilden. □

2.3 Stoppzeiten und optional sampling

Bei Glücksspielen tritt es oft auf, dass so lange gespielt wird bis ein Ereignis eintritt z.B.: „Würfeln bis zum 1. Mal eine 6 erscheint“ oder „wir spielen solange bis das Kapital eine Grenze M unterschreitet oder erreicht“. Der begriff der Stoppzeit abstrahiert diesen Begriff der „zufälligen Spieldauer“:

Definition 2.3.1. *Es sei $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ eine Filtration über $(\Omega, \mathcal{A})$. Eine Zufallsvariable $T : \Omega \to \mathbb{N}_0 \cup \{+\infty\}$ heißt Stoppzeit, wenn für alle $n \in \mathbb{N}_0$ gilt: $\{T \leq n\} \in \mathcal{F}_n$.*
Ist allgemeiner $(\mathcal{F}_t)_{t\in I}$ eine Fitration bzgl. einer linearen geordneten Menge $(I, \leq)$, so heißt $T : \Omega \to I$ eine Stoppzeit , wenn für alle $t \in I$ gilt: $\{T \leq t\} \in \mathcal{F}_t$.

Bemerkung. Äquivalent hätten wir statt $\{T \leq n\} \in \mathcal{F}$ für alle n auch $\{T = n\} \in \mathcal{F}_n$ für alle $n \in \mathbb{N}_0$ fordern können.

Ist nämlich $\{T \leq n\} \in \mathcal{F}_n$ für alle $n \in \mathbb{N}_0$, so folgt für alle $n \in \mathbb{N}$:

$$\{T = n\} = \underbrace{\{T \leq n\}}_{\in\mathcal{F}_n} \setminus \underbrace{\{T \leq n-1\}}_{\in\mathcal{F}_{n-1}\subseteq\mathcal{F}_n} \in \mathcal{F}_n,\ sowie \{T = 0\} = \{T \leq 0\}$$

Ist umgekehrt $\{T = n\} \in \mathcal{F}_n$ für alle $n \in \mathbb{N}_0$, so folgt:

$$\{T \leq n\} = \bigcup_{k=0}^{n} \underbrace{\{T = k\}}_{\in\mathcal{F}_k\subseteq\mathcal{F}_n} \in \mathcal{F}_n.$$

Beispiel 2.3.2 (Würfeln bis zur ersten 6). Sind $(X_n)_{n\in\mathbb{N}_0}$ Zufallsvariablen in {1,2,3,4,5,6} und setzen wir $T = \inf\{n \in \mathbb{N}_0 | X_n = 6\}$, so ist T eine Stoppzeit bzgl, der von den $(X_n)_n$ erzeugten Filtration $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, $\mathcal{F}_n = \sigma(X_0, ..., X_n)$.

Allgemeiner: „Eintrittszeiten sind Stopzeiten “. Ist $(X_n)_{n\in\mathbb{N}_0}$ eine bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ adaptierte Folge von Zufallsvariablen $X_n : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$, so ist für alle $A \in \mathcal{A}'$ die zufällige Zeit $T_A : \Omega \to \mathbb{N}_0 \cup \{\infty\}$, $T_A = \inf\{n \in \mathbb{N}_0 | X_n \in A\}$ eine Stoppzeit. In der Tat: $\forall n \in \mathbb{N}_0$ gilt:

$$\{T_A \leq n\} = \bigcup_{k=0}^{n} \underbrace{\{X_k \in A\}}_{\in\mathcal{F}_k,\ da (X_n)_n adaptiert} \in \mathcal{F}_n.$$

Lemma 2.3.3. *Sind $S, T : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ zwei Stoppzeiten bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}}$, so auch ihr Maximum $S \vee T$ und ihr Minimum $S \wedge T$.*

Beweis. $\forall n \in \mathbb{N}_0$ ist:

$$\{S \wedge T \leq n\} = \underbrace{\{S \leq n\}}_{\in\mathcal{F}_n} \cup \underbrace{\{T \leq n\}}_{\in\mathcal{F}_n} \in \mathcal{F}_n \ und$$
$$\{S \vee T \leq n\} = \underbrace{\{S \leq n\}}_{\in\mathcal{F}_n} \cap \underbrace{\{T \leq n\}}_{\in\mathcal{F}_n} \in \mathcal{F}_n$$

□

Definition 2.3.4. *Es sei $T : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ eine Stoppzeit bzgl. $(F_n)_{n\in\mathbb{N}_0}$ auf $(\Omega, \mathcal{A})$ Wir definieren:*

$$\mathcal{F}_T = \{A \in \mathcal{A} | A \cap \{T \leq n\} \in \mathcal{F}_n \forall n \in \mathbb{N}_0\}$$

$\mathcal{F}_T$ heißt die Menge der bis zur Stoppzeit beobachtbaren Ereignisse.

Bemerkung. Offensichtlich ist $\mathcal{F}_T$ wieder eine σ-Algebra. Weiter gilt: $\mathcal{F}_T = \{A \in \mathcal{A} | A \cap \{T = n\} \in \mathcal{F}_n \ \forall n \in \mathbb{N}_0\}$ (ohne Beweis).

Definition 2.3.5. *Für eine Familie $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen und eine zufällige Zeit $X_{T(\omega)}(\omega)$ für $T(\omega) < \infty$, $\omega \in \Omega$ und $X_T(\omega) :=$ beliebig aber fest.*

Beispiel 2.3.6. $X_T(\omega) = 0$ für $T(\omega) = \infty$

Stoppzeit: $T : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ mit $\{T \le n\} \in \mathcal{F}_n$ $(n \in \mathbb{N}_0)$ Zur Stoppzeit T beobachtbare Ereignisse:

$$\mathcal{F}_T = \{A \in \mathcal{A} | A \cap \{T \le n\} \in \mathcal{F}_n \quad (n \in \mathbb{N}_0)\}$$

Lemma 2.3.7. *Sind $(X_n)_{n\in\mathbb{N}_0}$ adaptiert und T eine Stoppzeit bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ so ist X_T bzgl. $\mathcal{F}_T$ messbar*

Beweis. Für alle messbaren Mengen A im Zielraum Ω' der X_n gilt:

$$\{X_T \in A\} = \bigcup_{n\in\mathbb{N}_0} \{T = n, X_n \in A\} \cup \{T = \infty, 0 \in A\} \in \mathcal{A}$$

und für alle $n \in \mathcal{N}_0$:

$$\{X_T \in A\} \cap \{T = n\} = \underbrace{\{X_n \in A\}}_{\in\mathcal{F}_n, da (X_n)_n adaptiert} \cap \overbrace{\{T = n\}}^{\in\mathcal{F}_n,\text{ da T Stoppzeit}} \in \mathcal{F}_n$$

Es folgt: $\{X_T \in A\} \in \mathcal{F}_T$ $\square$

Lemma 2.3.8. *Es sei $(X_n)_{n\in\mathbb{N}}$ ein Martingal und T eine Stoppzeit bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ mit $T \le m$ für ein $m \in \mathbb{N}_0$. Dann $E[X_m|\mathcal{F}_T] = X_T \quad P - f.s.$*

Beweis. Nach Lemma 2.3.6 ist X_T $\mathcal{F}_T$ messbar. Damit bleibt für alle $A \in \mathcal{F}_T$ zu zeigen:

$$E[X_m 1_A] = E[X_T 1_A]$$

Nun gilt für alle $k = 0, 1, ..., m$:

$$E[X_T 1_A 1_{\{T=k\}}] = E[X_k 1_{\underbrace{A \cap \{T = k\}}_{\in\mathcal{F}_k}}] = E[E[X_m|\mathcal{F}_k] 1_{\underbrace{A \cap \{T = k\}}_{\mathcal{F}_k - messbar}}] =$$
$$E[E[X_m 1_{A\cap\{T=k\}}|\mathcal{F}_k]] = E[X_m 1_A 1_{\{T=k\}}]$$

und daher

$$E[X_T 1_A] = E[X_T 1_A \sum_{k=0}^m 1_{\{T=k\}}] = \sum_{k=0}^m E[X_T 1_A 1_{\{T=k\}}] \overset{s.o.}{=} \sum_{k=0}^m E[X_m 1_A 1_{\{T=k\}}] =$$
$$E[X_m 1_A \sum_{k=0}^m 1_{\{T=k\}}] = E[X_m 1_A]$$

$\square$

Korollar. Unter der Vorraussetzung von Lemma 2.3.7 gilt:

$$E[X_T] = E[X_0]$$

Beweis.

$$E[X_T] = E[E[X_m|\mathcal{F}_T]] = E[X_m] = E[E[X_m|\mathcal{F}_0]] \underset{\substack{\uparrow \\ Martingaleigenschaft}}{=} E[X_0]$$

□

Anschaulich interpretiert: Aus einem fairen Spiel kann durch Stoppen mit einer beschränkten Stoppzeit „ohne prophetische Fähigkeiten“kein unfaires Spiel gewonnen werden.
Verzichtet man auf die Beschränktheit $T \leq m$, so wird die Aussage des Korollars im Allgemeinen falsch. Gegenbeispiel dazu:

Beispiel 2.3.9 (Verdopplungsstrategie). Es sei $Z_k,\ k \in \mathbb{N}$ i.i.d. auf $\{\pm 1\}$ gleichverteilte Zufallsvariablen: $\mathcal{L}(Z_k) = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$; sowie $X_n = \sum\limits_{k=1}^{n} 2^{k-1} Z_k$ und $\mathcal{F}_n = \sigma(Z_1, ...Z_n)$ für $n \in \mathbb{N}$, insbesondere $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Dann ist $(X_n)_{n\in\mathbb{N}_0}$ ein Martingal bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, denn für alle $n \in \mathbb{N}_0$ gilt:

$$E[X_{n+1}|\mathcal{F}_n] = E[\underbrace{X_n}_{\mathcal{F}_n-messbar} + 2^n Z_{n+1}|\mathcal{F}_n] = X_n + 2^n E[\underbrace{Z_{n+1}|\mathcal{F}_n}_{\text{unabhängig}}] = X_n + 2^n \underbrace{E[Z_{n+1}]}_{0} = X_n \quad P-f.s.$$

Nun sei $T = \inf\{n \in \mathbb{N} | Z_n = 1\}$. T ist eine Stoppzeit. Es gilt: $\{T = n\} = \{Z_n = 1, Z_k = -1 \text{ für } k < n\} \in \mathcal{F}_n$ und $\{T = \infty\} = \{Z_k = -1 \text{ für alle } k \in \mathbb{N}\}$. Insbesondere folgt $T < \infty\ P-f.s.$
Nun gilt für $n \in \mathbb{N}$:

$$X_T 1_{\{T=n\}} = X_n 1_{\{Z_n=1, Z_k=-1 \text{ für } k<n\}} = \underbrace{\left(2^{n-1} - \sum_{k=1}^{n-1} 2^{k-1}\right)}_{=1} 1_{\{T=n\}} = 1_{\{T=n\}}$$

also $X_T 1_{\{T<\infty\}} = 1_{\{T<\infty\}}$ und daher $X_T = 1\ P-f.s.$ Wegen $X_0 = 0$ folgt: $E[X_T] = 1 \neq 0 = E[X_0]$.

Lemma 2.3.10. *Sind $S \leq T$ zwei Stoppzeiten, so gilt: $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

Beweis. Es sei $A \in \mathcal{F}_S$ und $n \in \mathbb{N}_0$. Dann gilt:

$$A \cap \{T = n\} \underset{S\leq T}{=} \left(\bigcup_{k=0}^{n} \underbrace{A \cap \{S = k\}}_{\in\mathcal{F}_k\subseteq\mathcal{F}_n, da\ A\in\mathcal{F}_S} \right) \cap \underbrace{\{T = n\}}_{\in\mathcal{F}_n, \text{ da T Stoppzeit}} \in \mathcal{F}_n, \text{ also } A \in \mathcal{F}_T.$$

□

2.4 Optional sampling und gleichgradige Integrierbareit

Satz 2.4.1 (Optional sampling für Martintegrale und beschränkte Stoppzeiten)**.** *Es seinen $S \leq T$ zwei beschränkte Stoppzeiten, $T \leq m \in \mathbb{N}_0$ und $(X_n)_{n\in\mathbb{N}_0}$ ein Martintegral bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$. Dann gilt:*

$$E[X_T|\mathcal{F}_S] = X_S \quad P-f.s.$$

Beweis.

$$E[X_T|\mathcal{F}_S] \overset{\substack{Lemma\ 2.3.7 \\ \downarrow}}{=} E[E[X_m|\mathcal{F}_T]|\mathcal{F}_S] \underset{\substack{\uparrow \\ \mathcal{F}_S\subseteq\mathcal{F}_T\ Turmeigenschaft}}{\overset{\substack{Lemma\ 2.3.10 \\ \downarrow}}{=}} E[X_m|\mathcal{F}_S] \overset{\substack{Lemma\ 2.3.7 \\ \downarrow}}{=} X_S \quad P-f.s.$$

□

Beispiel 2.4.2. 1) Ist $(X_n)_{n\in\mathbb{N}}$ ein Martingal und T eine Stoppeit bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, so sind auch alle $T\wedge n$, $n \in \mathbb{N}_0$, beschränkte Stoppzeiten mit $T\wedge 0 \le T\wedge 1 \le ...$ Es folgt aus dem Satz:

$$E[X_{T\wedge(n+1)}|\mathcal{F}_{T\wedge n}] = X_{T\wedge n} \quad P-f.s. \text{ für alle } n \in \mathbb{N}_0.$$

Anders gesagt: Das „gestoppte Martingal“$(X_{T\wedge n})_{n\in\mathbb{N}_0}$ ist wieder ein Martingal bzgl. $(\mathcal{F}_{T\wedge n})_{n\in\mathbb{N}_0}$. $(X_{T\wedge n})_{n\in\mathbb{N}_0}$ ist auch ein Martingal bzgl. $(\mathcal{F}_{T\wedge n})_{n\in\mathbb{N}_0}$. In der Tat gilt für $n \in \mathbb{N}_0$:

$$\begin{aligned} E[X_{T\wedge(n+1)|\mathcal{F}_n}] &= E[X_{n+1}\underbrace{1_{\{T>n\}}}_{\in\mathcal{F}_n} + [X_{T\wedge n}\underbrace{1_{\{T\le n\}}}_{\in\mathcal{F}_n}|\mathcal{F}_n] = \\ &\underbrace{E[X_{n+1}|\mathcal{F}_n]}_{=X_n \ Martingal} 1_{\{T>n\}} + E[\underbrace{X_{T\wedge n}}_{messbar\ bzgl.\ \mathcal{F}_{T\wedge n}\subseteq\mathcal{F}_n}|\mathcal{F}_n]1_{\{T\le n\}} \\ &= X_n 1_{\{T>n\}} + \underbrace{X_{T\wedge n}}_{X_T} 1_{\{T\le n\}} = X_{T\wedge n} \quad P-f.s. \end{aligned}$$

2) Wir betrachten die „einfache Irrfahrt“:
$(Z_n)_{n\in\mathbb{N}}$ seinen i.i.d. Zufallsvariablen mit Werten $\{\pm 1\}$, gleichverteilt in $\{\pm 1\}$:
$P[Z_n = 1] = P[Z_n = -1] = \frac{1}{2}$ für alle $n \in \mathbb{N}$. Wir setzen $S_n := \sum\limits_{k=1}^{n} Z_k$ und $\mathcal{F}_n = \sigma(Z_1, ..., Z_n)$, $\mathcal{F}_0 = \{\emptyset, \Omega\} \quad (n \in \mathbb{N}_0)$.
Die einfache Irrfahrt $(S_n)_{n\in\mathbb{N}_0}$ ist ein Martingal bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, denn für alle $n \in \mathbb{N}_0$ gilt:

$$E[S_{n+1}|\mathcal{F}_n] = E[\underbrace{S_n}_{\mathcal{F}_n-messbar} + Z_{n+1}|\mathcal{F}_n] = S_n + E[Z_{n+1}|\mathcal{F}_n] = S_n + E[Z_{n+1}] = S_n \quad P-f.s.$$

Auch $(S_n^2 - n)_{(n\in\mathbb{N})}$ ist ein Martingal:

$$\begin{aligned} E[S_{n+1}^2 - (n+1)|\mathcal{F}_n] &= E[S_n + Z_{n+1})^2 - (n+1)|\mathcal{F}_n] = \\ E[\ \underbrace{S_n^2 - n}_{\mathcal{F}_n-messbar} &+ 2\underbrace{S_n}_{\mathcal{F}_n-messbar} Z_{n+1} + \underbrace{Z_{n+1}^2 - 1}_{=0}|\mathcal{F}_n] = \\ S_n^2 - n + 2S_n E[Z_{n+1}|\mathcal{F}_n] = S_n^2 - n + 2S_n &\underbrace{E[Z_{n+1}]}_{=0} = S_n^2 - n \quad P-f.s. \end{aligned}$$

Es folgt für jede Stoppzeit T und alle $n \in \mathbb{N}_0$:

$$0 = E[S_0] = E[S_{T\wedge n}]\ und\ 0 = E[S_0^2 - 0] = E[S_{T\wedge n}^2 - (T\wedge n)],\ also\ E[T\wedge n] = E[S_{T\wedge n}^2]$$

„Ruinproblem“beim Glücksspiel: 2 Spieler A und B wetten immer wieder 1 Euro auf den Ausgang eines fairen Münzwurfs. S_n bedeutet die Bilanz von A zur Zeit n, wenn Z_k den Gewinn von A in der k-ten Runde bedeutet. Die beiden Spieler stoppen das Spiel, wenn Spieler A a Euro verloren hat (Ruin von A) oder Spieler B b Euro verloren hat (Ruin von B).
$a, b \in \mathbb{N}$. Stoppzeit: $T = \inf\{n \in \mathbb{N}_0 : S_n \in \{-a, b\}\}$. Wegen $-a \le S_{T\wedge n} \le b$ folgt:
$S_{T\wedge n}^2 \le a^2 \vee b^2$ und daher

$$E[T] \xleftarrow[mon.\ konv.]{\infty\leftarrow n} E[T\wedge n] = E[S_{T\wedge n}^2] \le a^2 \vee b^2 < \infty$$

Insbesondere folgt $E[T] < \infty$, also $T < \infty$ $P-f.s.$ und daher $S_{T\wedge n} \xrightarrow{n\to\infty} S_T$ $P-f.s.$
<u>Aus der nächsten Vorlesung:</u>

$E[S_{T\wedge n}] = E[S_0] = 0 \xrightarrow[n\to\infty]{\text{dom.Konv.}} = E[S_T] = -aP[S_T = -a] + bP[S_T = b]$
Es folgt: $-aP[S_T = -a] + bP[S_T = b] = 0$, anders gesagt:

$$\frac{P[\text{„Ruin des Spielers A"}]}{P[\text{„Ruin des Spielers B"}]} = \frac{P[S_T = -a]}{P[S_T = b]} = \frac{b}{a}$$

Zudem folgt $P[S_T = -a] + P[S_T = b] = 1$ wegen $T < \infty$ P-f.s.
Es folgt: $P[S_T = -a] = \dfrac{b}{a+b}$, $P[S_T = b] = \dfrac{a}{a+b}$

Satz 2.4.3 (Optional Sampling für beschränkte Stoppzeiten und Sub- und Supermartingale)**.** *Sind $S \leq T$ zwei beschränkte Stoppzeiten und $(X_n)_{n\in\mathbb{N}_0}$ ein Sub-(bzw. Super-),martingal, dann gilt:*

$$E[X_T|\mathcal{F}_S] \underset{\text{bzw. } \leq}{\geq} X_S P\text{-f.s.}$$

Beweis. Es sei $X_n = \underbrace{A_n}_{\text{vorhersehbar, steigend (bzw. fallend)}} + \overbrace{M_n}^{\text{Martingal}}, n \in \mathbb{N}_0$ die Doob-Zerlegung von $(X_n)_{n\in\mathbb{N}}$
Dan gilt $A_T \geq A_S$ (bzw. $A_T \leq A_S$) P-f.s. wegen $S \leq T$ und $E[M_T|\mathcal{F}_S] = M_S$ P-f.s. nach Satz 2.4.1
Es folgt:

$$E[X_T|\mathcal{F}_S] = E[A_T|\mathcal{F}_S] + E[M_T|\mathcal{F}_S] \underset{\text{bzw. } \leq wg. A_T \geq A_S}{\overset{\uparrow}{\geq}} \overbrace{E[A_S|\mathcal{F}_S]}^{=A_S,\text{ da } (A_n)_{n\in\mathbb{N}} \text{ adaptiert}} + E[M_T|\mathcal{F}_S] = A_S + M_S = X_S \text{ P-f.s.}$$

□

Das Beispiel der Verdoppelungsstrategie zeigt, dass man im optional sampling Theorem i.A. nicht auf die Beschränktheit der Stoppzeit verzichten kann. Der folgende Begriff dient „als Ersatz"für die Beschränktheit.

Definition 2.4.4. *Eine Menge $\mathcal{M}$ von Zufallsvariablen heißt gleichgradig integrierbar, wenn*

$$\sup_{X\in\mathcal{M}} E[|X|1_{\{|X|\geq a\}}] \xrightarrow{a\to\infty} 0$$

Für $p \geq 1$ heißt sie p-fach gleichgradig integrierbar, wenn

$$\sup_{X\in\mathcal{M}} E[|X|^p 1_{\{|X|\geq a\}}] \xrightarrow{a\to\infty} 0$$

Einfach gleichgradige Integrierbarkeit ist also gleichgradige Integrierbarkeit.

Lemma 2.4.5. *Folgende Aussagen sind äquivalent:*

1) *$\mathcal{M}$ ist p-fach gleichgradig integrierbar*
2) $\sup_{X\in\mathcal{M}} E[(|X|^p - a)_+] \xrightarrow{a\to\infty} 0$

Beweis.

1)⇒2) Wegen $(|X|^p - a)_+ \leq |X|^p \mathbb{1}_{\{|X|\geq a^{\frac{1}{p}}\}}$ $(X \in \mathcal{M})$ folgt:

$$0 \leq \sup_{X\in\mathcal{M}} E[(|X|^p - a)_+] \underset{\xrightarrow{wg.1)}[a\to\infty]0}{\leq} \sup_{X\in\mathcal{M}} E[|X|^p \mathbb{1}_{\{|X|\geq a^{\frac{1}{p}}\}}]$$

und damit auch

$$\sup_{X\in\mathcal{M}} E[(|X|^p - a)_+] \xrightarrow[a\to\infty]{0}$$

2)⇒1) Für $|X| \geq a \geq 0$ gilt:

$$|X|^p \leq 2|X|^p - a^p = 2(|X|^P - \frac{a^p}{2})$$

Es folgt für $a \geq 0$:

$$|X|^p 1_{\{|X|\geq a\}} \leq 2(|X|^p - \frac{a^p}{2})_+$$

und daher

$$0 \leq \sup_{X\in\mathcal{M}} E[|X|^p 1_{\{|X|\geq a\}}] \leq 2 \sup_{X\in\mathcal{M}} E[(|X|^p - \frac{a^p}{2})_+] \xrightarrow[wg.2)]{a\to\infty} 0 \Rightarrow \sup_{X\in\mathcal{M}} E[|X|^p 1_{\{|X|\geq a\}}] \xrightarrow{a\to\infty} 0.$$

□

Praktisches hinreichendes Kriterium zum Nachweis der p-fachen gleichgradigen Integrierbarkeit:

Lemma 2.4.6. *Ist $\mathcal{M}$ eine Menge von Zufallsvariablen über $(\Omega, \mathcal{A}, P)$ und $1 \leq p < q$. Ist $\mathcal{M}$ in $\mathcal{L}^q(\Omega, \mathcal{A}, P)$ beschränkt, d.h. $\sup_{X\in\mathcal{M}} E[|X|^q] < \infty$, so ist $\mathcal{M}$ p-fach gleichgradig integrierbar.*

Beispiel 2.4.7. Haben die Zufallsvariablen in $\mathcal{M}$ beschränkte Erwartungswerte und beschränkte Varianzen, so ist $\mathcal{M}$ gleichgradig integrierbar.

Beweis. Für $a > 0, X \in \mathcal{M}$ gilt:

$$|X|^p 1_{\{|X|\geq a\}} \underset{\substack{\uparrow \\ p-q<0}}{\leq} a^{p-1}|X|^q$$

Es folgt:

$$\sup_{X\in\mathcal{M}} E[|X|^p 1_{\{|X|\geq a\}}] \leq a^{p-q} \underbrace{\sup_{X\in\mathcal{M}} E[|X|^q]}_{<\infty} \xrightarrow{n\to\infty} 0$$

□

Gleichgradige Integrierbarkeit vererbt sich unter Stoppen von Martingalen:

Satz 2.4.8. *Ist $(X_n)_{n\in\mathbb{N}_0}$ eine p-fach gleichgradig integrierbares Martingal oder nichtnegatives Submartingal:*

$$\mathcal{M} = \{X_T | T \textit{ ist eine P-f.s. endliche Stoppzeit}\}$$

ist p-fach gleichgradig integrierbar. ($p \geq 1$)

Beweis. $(|X_n|)_{n\in\mathbb{N}}$ ist ein nichtnegatives Submartingal nach der Jensenschen Ungleichung, da $|\cdot|$ konvex ist. Wegen der Konvexität von $t \mapsto (t^p - a)_+$ $(t \geq 0, a \geq 0)$ folgt für alle Stoppzeiten T und alle $n \in \mathbb{N}_0, a \geq 0$:

$$E[(|X_{T\wedge n}|^p - a)_+] \overset{(\star)}{\leq} E[(E[|X_n||\mathcal{F}_{T\wedge n}]^p - a)_+] \underset{\substack{\uparrow \\ \text{Jensen, } t\mapsto(|t|^p-a)_+, t\in\mathbb{R} \text{ konvex}}}{\leq} E[(|X_n|^p - a)_+]$$

Bemerkung ($\star$). $|X_{T\wedge n}| \leq E[|X_n||\mathcal{F}_{T\wedge n}]$ wg. optional sampling für beschränkte Stoppzeiten $t \mapsto (t^p - a)_+, t \geq 0$ ist monoton steigend

Ist nun T eine P-f.s. endliche Stoppzeit, so folgt:

$$E[(|X_T|^p - a)_+] \overset{\text{mon. Konv.}}{=} \lim_{n\to\infty} E[(|X_T|^p - a)_+ 1_{\{T\leq n\}}] = \lim_{n\to\infty} E[(|X_{T\wedge n}|^p - a)_+ 1_{\{T\leq n\}}]$$

$$\leq \sup_{n\in\mathbb{N}} E[(|X_{T\wedge n}|^p - a)_+] \overset{\text{Bem. }(\star)}{\leq} \sup_{n\in\mathbb{N}} E[(|X_n|^p - a)_+]$$

Und daher:

$$\sup\{E[(|X_T|^p - a)_+] | T \text{ ist P-f.s. endliche Stoppzeit}\} \leq \sup_{n\in\mathbb{N}} E[(|X_n|^p - a)_+] \xrightarrow{(\star)} [a\to\infty]$$

$(\star)$ Voraussetzung: $(X_n)_{n\in\mathbb{N}} p-$ fach gleichg. integrierbar $\square$

Bemerkung. Ist $\mathcal{M}$ p-fach gleichgradig integrierbar, $(p \geq)1$. so ist $\mathcal{M}$ auch in $\mathcal{L}^2$ beschränkt: $\sup\{E[|X|^p] | X \in \mathcal{M}\} < \infty$.

Beweis. Für $a \geq 0$ gilt: $|X|^p \leq |X|^p 1_{\{|X|\geq a\}} + a^p 1_{\{|X|\geq a\}}$
also $\sup\limits_{X\in\mathcal{M}} E[|X|^p] \leq \underbrace{\sup_{X\in\mathcal{M}} E[|X|^p 1_{\{|X|\geq a\}}]}_{\xrightarrow{a\to\infty} 0} + \underbrace{a^p}_{<\infty}$ für geeignetes a. $\square$

Lemma 2.4.9. *Ist $(X_n)_{n\in\mathbb{N}_0}$ p-fach gleichgradig integrierbar ($p \geq 1$), und gilt $X_n \xrightarrow{n\to\infty} X$ P-f.s., so folgt $X_n \xrightarrow{n\to\infty} X$ in $\mathcal{L}^P(\Omega, \mathcal{A}, P)$, d.h. $E[|X_n - X|^p] \xrightarrow{n\to\infty} 0$.*

Beweis. Erinnern Sie sich an das Lemma von Faton aus der Maßtheorie: Sind $Y_n \geq 0$ ($n \in \mathbb{N}_0$) Zufallsvariablen, so gilt:

$$E[\liminf_{n\to\infty} Y_n] \leq \liminf_{n\to\infty} E[Y_n]$$

(gilt analog auch für bel. Maße statt W'maße)
Insbesondere gilt:

$$E[|X|^P] \underset{\text{Faton}}{\leq} \liminf_{n\to\infty} E[|X_n|^P] \overset{\text{gleichgr. Integr.bar}\Rightarrow\text{Beschränktheit (obige Bemerkung)}}{<} \infty \text{ also } X \in \mathcal{L}^P(\Omega, \mathcal{A}, P)$$

Weiter gilt: $|X|^P \mathbb{1}_{\{|X|\geq a\}} \xrightarrow{a\to\infty} 0$ P-f.s. , also $E[|X|^P \mathbb{1}_{\{|X|\geq a\}}] \xrightarrow{a\to\infty} 0$ P-f.s. wegen der dom. Konv. mit der Majorante $|X|^P \in \mathcal{L}^1$.
Gegeben $\varepsilon > 0$ können wir wegen der p-fachen gleichgradigen Integrierbarkeit der $(X_n)_{n\in\mathbb{N}}$ ein $a > 0$ finden, so dass $\sup\limits_{n\in\mathbb{N}} E[|X_n|^P \mathbb{1}_{\{|X_n|\geq a\}}] < \varepsilon$ und $E[|X|^P \mathbb{1}_{\{|X|\geq a\}}] < \varepsilon$.
Wir schätzen ab:

$$|X_n - X|^P \leq \overbrace{|X_n - X|^P \mathbb{1}_{\{|X|<a, |X_n|<a\}}}^{(I)_n} + \overbrace{[2(|X| \vee |X_n|)]^P \mathbb{1}_{\{|X_n|\vee|X|\geq a\}}}^{(II)_n}$$

Es gilt: $(I)_n \xrightarrow{n\to\infty} 0$ P-f.s. und $|(I)_n| \leq (2a)^P < \infty$ also $E[(I)_n] \xrightarrow{n\to\infty} 0$ P-f.s. nach dem Satz für die dom. Konvergenz.
Weiter: $(II)_n \leq 2^P |X_n|^P \mathbb{1}_{\{|X_n|\geq a\}} + 2^P |X|^P \mathbb{1}_{\{|X|\geq a\}}$
und damit für das gewählte $a > 0$:

$$E[|X_n - X|^P] \leq \underbrace{E[(I)_n]}_{\xrightarrow{n\to\infty} 0} + 2^P \underbrace{E[|X_n|^P \mathbb{1}_{\{|X_n|\geq a\}}]}_{<\varepsilon} + 2^P \underbrace{E[|X|^P \mathbb{1}_{\{|X|\geq a\}}]}_{<\varepsilon}$$

Das bedeutet:

$$\limsup_{n\to\infty} E[|X_n - X|^P] \leq 2 \cdot 2^P \varepsilon$$

Weil $\varepsilon > 0$ beliebig war, folgt:

$$E[|X_n - X|^P] \xrightarrow{n\to\infty} 0$$

□

Korollar. Ist $(X_n)_{n\in\mathbb{N}}$ ein p-fach gleichgradig integrierbares Martingal oder nicht negatives Submartingal und ist T eine P-f.s. endliche Stoppzeit, so gilt:

$$X_{T\wedge n} \xrightarrow{n\to\infty} X_T \text{ in } L^P$$

Beweis. Es gilt $X_{T\wedge n} \xrightarrow{n\to\infty} X_T$ P-f.s. und $(X_{T\wedge n})_{n\in\mathbb{N}}$ ist nach Satz 2.4.7 p-fach gleichgradig integrierbar. Die Behauptung folgt nun aus Lemma 2.4.9. □

Damit erhalten wir:

Satz 2.4.10 (Optional Sampling für gleichgradig integrierbare Martingale)**.** *Sind $S \leq T$ zwei P-f.s. endliche Stoppzeiten und ist $(X_n)_{n\in\mathbb{N}_0}$ ein gleichgradig integrierbares Martingal, so gilt:*

$$E[X_T|\mathcal{F}_S] = X_S \text{ P-f.s.}$$

Beweis. Aus Korollar 2.4.10 wissen wir: $X_{T\wedge n} \xrightarrow{n\to\infty} X_T$ in $\mathcal{L}^1$, also auch $\forall m \in \mathbb{N}_0$:

$$X_{S\wedge m} \overset{\text{für } n\geq m}{\underset{\text{wg. Satz (opt. sampling für beschränkte Stoppzeiten)}}{=}} E[X_{T\wedge n}|\mathcal{F}_{S\wedge m}] \xrightarrow{n\to\infty} E[X_T|\mathcal{F}_{S\wedge m}] \text{ in } L^1$$

Das bedeutet: $X_{S\wedge m} = E[X_T|\mathcal{F}_{S\wedge m}]$ P-f.s.
Wir zeigen jetzt:

$$E[X_T|\mathcal{F}_S]\mathbb{1}_{\underbrace{\{S<m\}}_{\in\mathcal{F}_S}} = \underbrace{E[X_T|\mathcal{F}_{S\wedge m}]}_{\text{mb. bzgl. } \mathcal{F}_{S\wedge m}\subseteq\mathcal{F}_S} \mathbb{1}_{\underbrace{\{S<m\}}_{\in\mathcal{F}_S}} \text{ P-f.s.}$$

In der Tat: Beide Seiten sind $\mathcal{F}_S$-messbar und es gilt für alle $A \in \mathcal{F}_S$

$$E[E[X_T|\mathcal{F}_S] \underbrace{\mathbb{1}_{\{S<m\}}\mathbb{1}_A}_{\mathcal{F}_S\text{-messbar}}] = E[X_T \underbrace{\mathbb{1}_{\{S<m\}\cap A}}_{\in\mathcal{F}_{S\wedge m}(\star)}] = E[E[X_T|\mathcal{F}_{S\wedge m}]\mathbb{1}_{\{S<m\}}\mathbb{1}_A]$$

<u>Zu $(\star)$:</u> Für $k \in \mathbb{N}$: $\{S < m\} \cap A \cap \{S\wedge m \leq k\} = \{S \wedge m \leq k \wedge (m-1)\} \cap A \in \mathcal{F}_{k\wedge(m-1)} \subseteq \mathcal{F}_{k_S}$.
Zusammen folgt die Behauptung:

$$E[X_T|\mathcal{F}_S]\mathbb{1}_{\{S<m\}} = E[X_T|\mathcal{F}_{S\wedge m}]\mathbb{1}_{\{S<m\}} \text{ P-f.s.}$$

Wir erhalten:

$$X_S\mathbb{1}_{\{S<m\}} = X_{S\wedge m}\mathbb{1}_{\{S\wedge m\}} = E[X_T|\mathcal{F}_S]\mathbb{1}_{\{S<m\}} \text{ P-f.s.}$$

also

$$X_S = E[X_T|\mathcal{F}_S] \text{ P-f.s., da } m \text{ beliebig war.}$$

□

Bevor wir das Theorem auf Sub-und Supermartingale erweitern, eine technische Vorbereitung:

Lemma 2.4.11. *Sind $(X_i)_{i\in I}$ und $(Y_i)_{i\in I}$ p-fach gleichgradig integrierbar, so auch $(X_i + Y_i)_{i\in I}$ und $(bX_i)_{i\in I}$ $\;b \in \mathbb{R}$*

Beweis. Es gilt für $a > 0$:
$$|X_i + Y_i|^P \mathbb{1}_{\{|X_i+Y_i|\geq a\}} \leq (2|X_i|)^P \mathbb{1}_{\{|X_i|\geq \frac{a}{2}\}} + (2|Y_i|)^P \mathbb{1}_{\{|Y_i|\geq \frac{a}{2}\}}$$
Also
$$\sup_{i\in I} E[|X_i + Y_i|^P \mathbb{1}_{\{|X_i+Y_i|\geq a\}}] \leq 2^P \underbrace{\sup_{i\in I} E[|X_i|^P \mathbb{1}_{\{|X_i|\geq \frac{a}{2}\}}]}_{\xrightarrow{a\to\infty} 0} + 2^P \underbrace{\sup_{i\in I} E[|Y_i|^P \mathbb{1}_{\{|Y_i|\geq \frac{a}{2}\}}]}_{\xrightarrow{a\to\infty} 0}$$
und für $b \neq 0$:
$$\sup_{i\in I} E[|bX_i|^P \mathbb{1}_{\{|bX_i|\geq a\}}] = |b|^P \sup_{i\in I} E[|X_i|^P \mathbb{1}_{\{|X_i|\geq \frac{a}{|b|}\}}] \xrightarrow{a\to\infty} 0$$
Für $b = 0$ ist die Aussage trivial. □

Satz 2.4.12 (Optional Sampling für gleichgradig integrierbare Sub- und Supermartingale)**.** *Sin $S \leq T$ zwei P-f.s. endliche Stoppzeiten und ist $(X_n)_{n\in\mathbb{N}}$ ein gleichgradig integrierbares Sub- (bzw. Super-)martingal, so gilt:*
$$E[X_T|\mathcal{F}_S] \underset{\substack{\uparrow \\ \textit{bzw.} \leq}}{\geq} X_S \text{ P-f.s.}$$

Beweis. Wir zeigen das nur für Submartingale. Für Supermartingale wechsle man das Vorzeichen.
Es sei $X_n = A_n + M_n$, $n \in \mathbb{N}_0$ die Doob-Zerlegung in eine vorhersehbare, steigende Folge $(A_n)_{n\in\mathbb{N}}$ mit $A_0 = 0$ und ein Martingal $(M_n)_{n\in\mathbb{N}_0}$.
Insbesondere gilt: $0 \leq A_n \underset{n\to\infty}{\nearrow} A_\infty$ für eine geeignete Zufallsvariable A_∞ mit Werten in $[0, \infty)$. Nun gilt für alle $n \in \mathbb{N}_0$:
$$E[A_n] = E[A_n] + \underbrace{E[M_n - M_0]}_{=0} = E[\underbrace{A_n - M_n}_{=X_n} - \underbrace{M_0}_{=X_0}] = E[X_n] - E[X_0]$$
Weil $(X_n)_{n\in\mathbb{N}_0}$ in L^1 beschränkt ist (Bemerkung 2.4.8), folgt:
$$\sup_{n\in\mathbb{N}_0} E[A_n] = \sup_{n\in\mathbb{N}_0} E[X_n] - E[X_0] < \infty$$
Also
$$E[A_\infty] \overset{\substack{\text{mon. Konv.} \\ \downarrow}}{=} \lim_{n\to\infty} E[A_n] < \infty$$
Folglich ist $(A_n)_{n\in\mathbb{N}_0}$ gleichgradig integrierbar, denn es gilt für alle $a > 0$:
$$\sup_{n\in\mathbb{N}_0} E[|A_n|\mathbb{1}_{\{|A_n|\geq a\}}] = \sup_{n\in\mathbb{N}_0} E[A_n \mathbb{1}_{\{A_n\geq a\}}] \leq E[\underbrace{A_\infty \mathbb{1}_{\{A_\infty\geq a\}}}_{\xrightarrow{a\to\infty} 0 \text{ punktweise, mit Maj.} A_\infty}] \xrightarrow[\text{dom. Konv.}]{a\to\infty} 0$$
Mit Lemma 2.4.12 ist auch $(M_n = X_n - A_n)_{n\in\mathbb{N}_0}$ gleichgradig integrierbar, da $(X_n)_{n\in\mathbb{N}}$ und $(A_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar sind.
Mit Satz 2.4.11 (Opt. Sampling für gleichgradig integrierbare Martingale) folgt:
$$E[M_T|\mathcal{F}_S] = M_S \text{ P-f.s.}$$
Andererseits gilt: $A_T > A_S$, da $T \geq S$ und $(A_n)_{n\in\mathbb{N}}$ monoton steigend ist. Es folgt:
$$E[A_T|\mathcal{F}_S] \geq E[A_S|\mathcal{F}_S] \underset{\substack{\uparrow \\ \mathcal{F}_S\text{-messbar, da } (A_n)_{n\in\mathbb{N}} \text{ vorhersehbar, also erst recht adaptiert ist}}}{=} A_S$$
Und daher:
$$E[X_T|\mathcal{F}_S] = E[M_T|\mathcal{F}_S] + E[A_T|\mathcal{F}_S] \geq M_S + A_S = X_S \text{ P-f.s.}$$
□

2.5 Doobsche Ungleichung

Die Doobschen Ungleichungen erlauben die Kontrolle der Verteilung des Maximums:

$$\overline{X}_n := \max_{k=0,\dots,n} X_k$$

eines (Sub-)Martingals $(X_n)_{n\in\mathbb{N}}$ mit Hilfe der Verteilung des Endwertes X_n

Satz 2.5.1 (1. Doob-Ungleichung). *Ist $(X_n)_{n\in\mathbb{N}}$ ein Submartingal, so gilt $\forall\ a > 0$:*

$$a \cdot P[\overline{X}_n \geq a] \leq E[X_n \mathbb{1}_{\{\overline{X}_n \geq a\}}]$$

Bemerkung. Mit der Abschätzung $E[X_n \mathbb{1}_{\{\overline{X}_n \geq a\}}] \leq E[(X_n)_+]$ folgt auch $aP[\overline{X}_n \geq a] \leq E[(X_n)_+]$.

Beweis. Es sei $T := \inf\{k \in \mathbb{N}_0 | X_k \geq a\}$ die erste Zeit zu der das Niveau a erreicht oder überschritten wird. Folglich ist T eine Stoppzeit. Weiter gilt:

$$\{T \leq n\} = \{\overline{X}_n \geq a\} = \{X_{T\wedge n} \geq a\}$$

Damit folgt

$$aP[\overline{X}_n \geq a] = aP[X_{T\wedge n} \geq a] = E[a\mathbb{1}_{\{X_{T\wedge n}\geq a\}}]$$

Nun gilt: $X_{T\wedge n} \leq E[X_n|\mathcal{F}_{T\wedge n}]\ P - f.s.$ nach dem Optional Sampling Satz, also

$$E[X_{T\wedge n}\mathbb{1}_{\{X_{T\wedge n}\geq a\}}] \leq E[E[X_n|\mathcal{F}_{T\wedge n}]\mathbb{1}\underbrace{\{X_{T\wedge n} \geq a\}}_{\mathcal{F}_{T\wedge n}-messbar}] = E[X_n\mathbb{1}_{\{X_{T\wedge n}\geq a\}}] = E[X_n\mathbb{1}_{\{X_n\geq a\}}]$$

Zusammen folgt die Behauptung. □

Korollar (1.Doob-Ungleichung-L^P-Version). Ist $(X_n)_{n\in\mathbb{N}_0}$ ein nichtnegatives Submartingal, so gilt $\forall a > 0$ und $p \geq 1$:

$$a^p P[\overline{X}_n \geq a] \leq E[X_n^p]$$

Bemerkung. Insbesondere gilt das, wenn $X_n = |M_n|$, $n \in \mathbb{N}$, mit einem Martingal $(M_n)_{n\in\mathbb{N}}$ gilt, denn nach der Jensen'schen Ungleichung ist $(X_n)_{n\in\mathbb{N}}$ ein Submartingal.

Beweis.

$$a^p P[\overline{X}_n \geq a] = a^p P[\overline{X}_n^p \geq a^p] \overset{\text{2.5.1, da } (X_n)_{n\in\mathbb{N}} \text{ nach Jensen Submartingal}}{\leq} E[X_n^p \mathbb{1}_{\{\overline{X}_n^p \geq a^p\}}] \leq E[X_n^p]$$

□

Satz 2.5.2 (2. Doobsche L^P-Ungleichung). *Ist $(X_n)_{n\in\mathbb{N}}$ ein nichtnegatives Submartingal und $p > 1$, so gilt:*

$$E[\overline{X}_n^p] \leq (\frac{p}{p-1})^p E[X_n^p]$$

Bemerkung. Genau wie oben gilt das insbesondere im Fall $X_n = |M_n|$ mit einem Martingal $(M_n)_{n\in\mathbb{N}}$.

Beweis. Aus Satz 2.5.1 wissen wir für $a > 0$, $n \in \mathbb{N}_0$:

$$aP[\overline{X}_n \geq a] \leq E[X_n\mathbb{1}_{\{\overline{X}_n\geq a\}}].$$

Wir multiplizieren das mit a^{p-2} und integrieren von 0 bis $t > 0$:

$$\int_0^t a^{p-1} \underbrace{P[\overline{X}_n \geq a]}_{E[\mathbb{1}_{\{\overline{X}_n\geq a\}}]} da \leq \int_0^t a^{p-2} E[X_n\mathbb{1}_{\{\overline{X}_n\geq a\}}] da$$

Wegen $X_n \geq 0$ können wir auf beiden Seiten die Integrationsreihenfolge nach Fubini vertauschen:

$$E[\underbrace{\int_0^t a^{p-1}\mathbb{1}_{\{\overline{X}_n \geq a\}}da}_{=\int_0^{t\wedge\overline{X}_n} a^{p-1}da=\frac{1}{p}(t\wedge\overline{X}_n)^p}] \leq E[X_n \underbrace{\int_0^t a^{p-2}\mathbb{1}_{\{\overline{X}_n \geq a\}}da}_{=\int_0^{t\wedge\overline{X}_n} a^{p-2}da=\frac{1}{p-1}(t\wedge\overline{X}_n)^{p-1}}]$$

also:

$$E[\frac{1}{p}(t\wedge\overline{X}_n)^p] \leq E[X_n\frac{1}{p-1}(t\wedge\overline{X}_n)^{p-1}] \quad (*)$$

Erinnern Sie sich an die Höldersche Ungleichung aus der Maßtheorie: Für Zufallsvariablen $X \geq 0$ und $Y \geq 0$ und $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$, d.h. $q = \frac{p}{p-1}$ gilt:

$$E[XY] \leq E[X^p]^{\frac{1}{p}} E[Y^q]^{\frac{1}{q}}$$

In dieser Situation erhalten wir:

$$E[X_n(t\wedge\overline{X}_n)^{p-1}] \leq E[X_n^p]^{\frac{1}{p}} E[(t\wedge\overline{X}_n)^{\overbrace{(p-1)q}^{=p}}]^{\frac{1}{q}}$$

In $(*)$ eingesetzt:

$$\frac{1}{p}E[(t\wedge\overline{X}_n)^p] \leq \frac{1}{p-1}E[X_n^p]^{\frac{1}{p}} E[(t\wedge\overline{X}_n)^p]^{\overbrace{\frac{1}{q}}^{=1-\frac{1}{p}}}$$

Im Fall $E[(t\wedge\overline{X}_n)^p] =: u > 0$ können wir durch $u^{1-\frac{1}{p}}$ dividieren und erhalten in jedem Fall ($u = 0, u > 0$)

$$\frac{1}{p}E[\underbrace{(t\wedge\overline{X}_n)^p}_{\geq 0 \xrightarrow{t\to\infty} \overline{X}_n^p \ pktw.}] \leq \frac{1}{p-1}E[X_n^p]^{\frac{1}{p}}$$

für $t \to \infty$ erhalten wir mit der monotonen Konvergenz:

$$\frac{1}{p}E[\overline{X}_n^p] \leq \frac{1}{p-1}E[X_n^p]^{\frac{1}{p}}$$

Also

$$E[\overline{X}_n^p] \leq (\frac{p}{p-1})^p E[X_n^p]$$

□

Bemerkung. Weil trivialerweise $E[X_n^p] \leq E[\overline{X}_n^p]$ wegen $X_n \leq \overline{X}_n$, bedeutet das: $||\overline{X}_n||_p$ und $||X_n||_p$ können durch Konstanten wechselseitig abgeschätzt werden.

2.6 Überquerung eines Intervalls

Ist $(X_n)_{n\in\mathbb{N}}$ eine Folge von (adaptierten) Zufallsvariablen und ist $a < b$, so definieren wir zufällige Zeiten (bzw. Stoppzeiten) $(T_k)_{k\in\mathbb{N}_0}$ und $(S_k)_{k\in\mathbb{N}_0}$ rekursiv ($k \in \mathbb{N}$)

$$\begin{aligned} T_0 &:= 0 \\ S_k &= \inf\{n \geq T_{k-1|X_n\leq a}\} \\ T_k &= \inf\{n \geq S_{k-1|X_n\geq b}\} \end{aligned}$$

Das Zeitintervall von S_k bis T_k ist also das Intervall der k-ten Überquerung von $[a,b]$ von unten nach oben, kurz der „k-ten Aufkreuzung“von $[a,b]$. Man beachte $0 = T_k \leq S_1 < T_1 < S_2 < T_2 < ...$
Die S_k und T_k sind Stoppzeiten, falls $(X_n)_{n\in\mathbb{N}}$ adaptiert sind. Weiter bezeichne für $n \in \mathbb{N}$ $U_n := \sum_{k=1}^{\infty} \mathbb{1}_{\{T_k \leq n\}}$ die Anzahl der Aufkreuzungen von $[a,b]$ bis zur Zeit n.
Zur Vereinfachung der Notation betrachten wir zunächst den Spezialfall $[a,b] = [0,1]$:

Lemma 2.6.1 (Aufkreuzungslemma). *Es sei $(X_n)_{n\in\mathbb{N}}$ ein Supermartingal. Dann gilt für die Anzahl U_n der Aufkreuzungen von $[0,1]$ bis zur Zeit $n \in \mathbb{N}_0$:*

$$E[U_n] \leq E[(X_n)_-]$$

Bemerkung. Schärfer gilt $E[U_n] \leq E[(X_n)_-] - E[(X_0)_-]$. Wir zeigen hier nur die einfache Version, weil sie für unsere Zwecke genügt.

Beweis. Für $k \in \mathbb{N}$, $n \in \mathbb{N}_0$ gilt:
$X_{T_k \wedge n} \geq 1$ für $T_k \leq n$ und
$X_{T_k \wedge n} = X_n \geq -(X_n)_-$ für $T_k > n$
Das bedeutet:

$$X_{T_k \wedge n} \geq \mathbb{1}_{\{T_k \leq n\}} - (X_n)_- \mathbb{1}_{\{T_k > n\}}$$

Bilden wir hier die Erwartung auf dem Ereignis $\{S_k \leq n\} \geq \{T_k \leq n\}$:

$$E[X_{T_k \wedge n} \mathbb{1}_{\{S_k \leq n\}}] \geq E[\mathbb{1}_{\{T_k \leq n\}}] - E[(X_n)_- \mathbb{1}_{\{S_k \leq n < T_k\}}]$$

Nun gilt:

$$E[\underbrace{(X_{T_k \wedge n} - X_{S_k \wedge n})}_{=0 \text{ für } S_k = n} \mathbb{1}_{\{S_k \leq n\}}] = E[(X_{T_k \wedge n} - X_{S_k \wedge n}) \mathbb{1}_{\underbrace{\{S_k < n\}}_{\in \mathcal{F}_{S_k \wedge n}}}] \overset{T_k \wedge n \geq S_k \wedge n}{\overset{\downarrow}{\leq}} 0$$

Also

$$0 \geq E[\underbrace{(X_{S_k \wedge n} \mathbb{1}_{\{S_k \leq n\}}}_{\text{nach Def. von } S_k \leq 0}] \geq E[X_{T_k \wedge n} \mathbb{1}_{\{S_k \leq n\}}] \overset{s.o.}{\geq} E[\mathbb{1}_{\{T_k \leq n\}}] - E[(X_n)_- \mathbb{1}_{\{S_k \leq n < T_k\}}]$$

Also

$$E[\mathbb{1}_{\{T_k \leq n\}}] \leq E[\underbrace{(X_n)_-}_{\geq 0} \mathbb{1}_{\underbrace{\{S_k \leq n < T_k\}}_{\leq \{T_{k-1} \leq n < T_k\}}}] \leq E[(X_n)_- \mathbb{1}_{\{T_{k-1} \leq n < T_k\}}]$$

Summieren wir über k:

$$E[U_n] = E[\sum_{k\in\mathbb{N}} \mathbb{1}_{\{T_k \leq n\}}] \overset{mon.\ Konv.}{\overset{\downarrow}{=}} \sum_{k\in\mathbb{N}} E[\mathbb{1}_{\{T_k \leq n\}}] \leq \sum_{k\in\mathbb{N}} E[\underbrace{(X_n)_-}_{\geq 0} \mathbb{1}_{\{T_{k-1} \leq n < T_k\}}] \overset{mon.\ Konv.}{\overset{\downarrow}{=}}$$
$$E[(X_n)_- \sum_{k\in\mathbb{N}} \mathbb{1}\underbrace{\{T_{k-1} \leq n < T_k\}}_{\text{bilden in K Partition von } \Omega}] = E[(X_n)_-].$$

□

Für ein beliebiges Intervall $[a,b]$ statt $[0,1]$ erhalten wir:

Korollar. Es sei $(X_n)_{n\in\mathbb{N}}$ ein Supermartingal und $a < b$. Dann gilt für die Auswahl U_n der Aufkreuzungen des Intervalls $[a,b]$ bis zur Zeit $n \in \mathbb{N}$

$$E[U_n] \leq \frac{E[(X_n)_-]}{b-a}$$

Beweis. Wir skalieren $Y_n := \frac{X_n-a}{b-a}$, $n \in \mathbb{N}_0$ Aufkreuzungen von $[a,b]$ durch X sind das Gleiche wie Aufkreuzungen von $[0,1]$ durch Y. Es gilt: $\frac{X_n-a}{b-a} = (Y_n)_-$. Zudem ist Y ein Supermartingal. Die Behauptung folgt damit aus Lemma 2.6.1. □

Korollar. Ist $(X_n)_{n\in\mathbb{N}}$ ein Supermartingal mit $\sup\limits_{n\in\mathbb{N}_0} E[(X_n)_-] < \infty$, so wird $\forall a < b$ das Intervall $[a,b]$ $P-f.s.$ nur endlich oft überquert.

Beweis. Aus $\sup\limits_{n\in\mathbb{N}_0} E[(X_n)_-] < \infty$ folgt auch $\sup\limits_{n\in\mathbb{N}_0} E[(X_n-a)_-] < \infty$ wegen $(X_n-a)_- \leq (X_n)_- + a_+$. Es sei $U_\infty := \lim\limits_{n\to\infty} U_n$ die Anzahl aller Aufkreuzungen von $[a,b]$. Dann gilt:

$$E[U_\infty] \overset{\substack{mon.\ Konv.\\ \downarrow}}{=} \lim_{n\to\infty} E[U_n] \overset{\substack{Kor\ 2.6.2\\ \downarrow}}{\leq} \sup_{n\in\mathbb{N}_0} \frac{E[(X_n)-a)_-]}{b-a} < \infty$$

Es folgt $U_\infty < \infty$ $P-f.s.$ □

2.7 Martingalkonvergenzsätze

Wir beobachten:

Lemma 2.7.1. *Ist $(x_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{R}$, die jedes rationale Intervall $[a,b]$, $a<b$, $a,b\in\mathbb{Q}$, nur endlich oft überquert, so konvergiert x_n für jedes $n\to\infty$ gegen ein $x\in\mathbb{R}\cup\{\pm\infty\}$*

Beweis. Angenommen , das ist nicht der Fall, also:

$$\liminf_{n\to\infty} x_n < \limsup_{n\to\infty} x_n$$

Dann gibt es $a<b$, $a,b\in\mathbb{Q}$ mit $\liminf\limits_{n\to\infty} x_n < a < b < \limsup\limits_{n\to\infty} x_n$. Das Intervall $[a,b]$ wird dann jedoch ∞ oft überquert im Gegensatz zur Voraussetzung. □

Satz 2.7.2 (Martingalkonvergenzsatz)**.** *Ist $(X_n)_{n\in\mathbb{N}_0}$ ein Supermartingal bzgl. $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ mit $\sup\limits_{n\in\mathbb{N}_0} E[(X_N)_-] < \infty$, so konvergiert $(X_n)_{n\in\mathbb{N}}$ für $n\to\infty$ $P-f.s.$ in $\mathbb{R}$.*
Eventuell nach Abänderung auf einer Nullmenge ist $X_\infty := \lim\limits_{n\to\infty} X_n$ $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n)$-messbar und es gilt $X_\infty \in L^1(\Omega, \mathcal{F}_\infty, P)$

Beweis. Nach Korollar 2.6.3 und da es nur abzählbar viele rationale Intervalle $[a,b]$ gibt, wird $P-f.s.$ für alle rationalen $a<b$ das Intervall $[a,b]$ nur endlich oft überquert. Nach Lemma 2.7.1 konvergiert $(X_n)_{n\in\mathbb{N}}$ $P-f.s.$ in $\mathbb{R}\cup(\pm\infty)$.
Setzen wir $Y := \lim\limits_{n\to\infty} X_n$, so folgt:

$$E[Y_+] \underset{\substack{\uparrow\\ Fatou\ Lemma}}{\leq} \liminf_{n\to\infty} \underbrace{E[(X_n)_+]}_{\leq E[(X_0)_+]} \leq E[(X_0)_+] < \infty \quad und$$

$$E[Y_-] \underset{\substack{\uparrow\\ \lim\limits_{n\to\infty} X_n \text{ existiert } P-f.s.}}{=} E[\liminf_{n\to\infty}(X_n)_-] \underset{\substack{\uparrow\\ Fatou}}{\leq} \liminf_{n\to\infty} E[(X_n)_-] \underset{\substack{\uparrow\\ Voraussetzung}}{<} \infty$$

Also: $E[|Y|] = E[Y_+] + E[Y_-] < \infty$. Insbesondere gilt $|Y| < \infty$ $P-f.s.$
Setzen wir $X_\infty := \mathbb{1}_{\{|Y|<\infty\}} Y$, so folgt $X_\infty = Y$ $P-f.s.$, also $X_n \xrightarrow{n\to\infty} X_\infty$ $P-f.s.$, $E[|X_\infty|] = E[|Y|] < \infty$ und X_∞ ist $\mathcal{F}_\infty$-messbar, da alle X_n und damit auch alle Y $\mathcal{F}_\infty$-messbar sind. □

Bemerkung. Das Beispiel der Verdopplungsstrategie zeigt, dass man im Allgemeinen nicht $X_n \xrightarrow{n\to\infty} X_\infty$ in $L^1(\Omega, \mathcal{F}_\infty, P)$ erwarten kann.

Unter der Voraussetzung gleichgradiger Integrierbarkeit erhalten wir:

Satz 2.7.3 (Martingalkonvergenz bei gleichgradiger Integrierbarkeit)**.** *Es sei $(X_n)_{n\in\mathbb{N}_0}$ ein gleichgradig integrierbares Martingal oder Sub- oder Supermartingal. Dann konvergiert X_n für $n \to \infty$ $P-f.s.$ und in $L^1(\Omega, \mathcal{F}_\infty, P)$ gegen ein $X_\infty \in L^1(\Omega, \mathcal{F}_\infty, P)$ und es gilt $P-f.s.$:*

$$\begin{aligned} E[X_\infty|\mathcal{F}_n] &= X_n \quad \textit{für Martingale} \\ E[X_\infty|\mathcal{F}_n] &\leq X_n \quad \textit{für Supermartingale} \\ E[X_\infty|\mathcal{F}_n] &\geq X_n \quad \textit{für Supmartingale} \quad (n \in \mathbb{N}_0) \end{aligned}$$

Beweis. Zunächst sei $(X_n)_n$ ein Supermartingal. Aus der gleichgradigen Integrierbarkeit folgt:

$$\sup_{n\in\mathbb{N}_0} E[(X_n)_-] < \infty, \text{ also ist der Martingalkonvergenzsatz 2.7.2 anwendbar.}$$

Insbesondere ist $X_\infty = \lim_{n\to\infty} X_n \in L^1(\Omega, \mathcal{F}_\infty, P)$ wohldefiniert. Aus Lemma 2.4.9 folgt $X_n \xrightarrow{n\to\infty} X_\infty$ in $L^1(\Omega, \mathcal{F}_\infty, P)$ und daher $\forall m \in \mathbb{N}_0$

$$X_m \underset{\substack{\uparrow \\ \text{für } n\geq m\ P-f.s., \text{ da X Supermartingal}}}{\geq} E[X_n|\mathcal{F}_m] \xrightarrow{n\to\infty} E[x_\infty|\mathcal{F}_m] \text{ in } L^1(\Omega, \mathcal{F}_\infty, P)$$

Also $X_m \geq E[X_\infty|\mathcal{F}_m]$ $P-f.s.$, also die Behauptung für Supermartingale. Für Submartingale folgt die Behauptung hieraus indem wir $(-X_n)_n$ statt $(X_n)_n$ betrachten.
Nimmt man die Aussagen für Super- und Submartingale zusammen, folgt die Behauptung für Martingale. □

Satz 2.7.4 (L^P-Martingalkonvergenzsatz)**.** *Ist $(X_n)_{n\in\mathbb{N}_0}$ ein in L^P beschränktes Martingal , wobei $P > 1$, so konvergiert $(X_n)_{n\in\mathbb{N}_0}$ $P-f.s.$ und in L^P.*

Beweis. Weil $(X_n)_n$ in L^P beschränkt ist, ist es gleichgradig integrierbar, da $p > 1$ (Lemma 2.4.6). Wir zeigen jetzt, dass $(X_n)_{n\in\mathbb{N}_0}$ auch p- fach gleichgradig integrierbar ist. $(X_n)_{+n\in\mathbb{N}_0}$ ist ein nichtnegatives Submartingal. $(\overline{X_n})_+ := \max_{k=0,\dots,n} (X_n)_+$, so folgt nach der Doobschen L^P- Ungleichung:

$$E[(X_n)_+^P] \leq c(p)E[(X_n)_+^P] \quad \textit{mit } c(p) = (\frac{p}{p-1})^P$$

und daher

$$E[\sup_{n\in\mathbb{N}_0} (X_n)_+^P] \underset{\substack{\uparrow \\ mon.\ Konv}}{=} \lim_{n\to\infty} E[(\overline{X_n})_+^P] \leq \sup_{n\in\mathbb{N}_0} E[(X_n)_+^P] \underset{\substack{\uparrow \\ (X_n)_n \text{ in } L^P \text{ beschränkt}}}{<} \infty$$

Es folgt für $a > 0$ mit der Abkürzung $Y := \sup_{n\in\mathbb{N}_0} (X_n)_+$:

$$\sup_{n\in\mathbb{N}_0} E[(X_n)_+^P 1_{\{X_n\geq a\}}] \leq E[\underbrace{Y^P 1_{\{Y\geq a\}}}_{\xrightarrow[n\to\infty]{} 0 \textit{ punktweise mit Majorante } Y^P\in\mathcal{L}^1}] \xrightarrow[dom.\ kon.]{n\to\infty} 0$$

Also ist $((X_n)_+)_{n\in\mathbb{N}_0}$ p-fach gleichgradig integrierbar. Das gleiche Argument, angewandt auf $(X_n)_- = (-X_n)_+$, zeigt, dass auch $((X_n)_-)_{n\in\mathbb{N}_0}$ p-fach gleichgradig integrierbar ist und daher auch $(X_n) = (X_n)_+ - (X_n)_-$, $n \in \mathbb{N}_0$ (z.B wg. Lemma 2.4.12) Aus dem Martingalkonvergenzsatz 2.7.3 folgt bei gleichgradiger Integrierbarkeit $P-f.s.$ Konvergenz und Konvergenz in L^1 von $(X_n)_{n\in\mathbb{N}_0}$ und aus Lemma 2.4.9 dann auch Konvergenz in L^P. □

Beispiel 2.7.5 (Approximation in L^P durch Bisektion). Es sei $(\Omega, \mathcal{A}, P)$=$([0,1[, \mathcal{B}([0,1[), \ unif[0,1[)$. Weiter sei $f \in \mathcal{L}^P(\Omega, \mathcal{A}, P)$, d.h. $\int\limits_0^1 |f(x)|^P dx < \infty$ für ein $p \geq 1$. Für $n \in \mathbb{N}$ setzen wir

$$f_n(x) = \sum_{k=0}^{2^n-1} \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(x) \underbrace{2^n \int\limits_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} f(x)dx}_{Mittelwert\ von\ f\ auf [\frac{k}{2^n}, \frac{k+1}{2^n})}$$

Dann gilt: $f_n \xrightarrow{n\to\infty} f_1 \ P-f.s.$ und in L^P $\int\limits_0^1 |f(x)-f_n(x)|^P dx \xrightarrow{n\to\infty} 0$ und $f_n(x) \xrightarrow{n\to\infty} f(x)$ für unif [0,1)-fast alle x. In der Tat:

$$\mathcal{F}_n := \sigma\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) : k = 0, ..., 2^n - 1\right) \quad n \in \mathbb{N}_0$$

ist eine Filtration mit $\mathcal{F}_\infty := \sigma(\bigcup\limits_{n\in\mathbb{N}_0} \mathcal{F}_n) = \mathcal{A}$. Nun ist $f_n = E[f|\mathcal{F}_n]$, $n \in \mathbb{N}_0$, ein p-fach integrierbares Martingal, da $f \in \mathcal{L}^P$.
Aus den Martingalkonvergenzsätzen 2.7.3 und 2.7.4 folgt Konvergenz $P-f.s.$ und in L^P gegen ein $g \in \mathcal{L}^P$. Die noch fehlende Aussage $f = g \ P-f.s.$ folgt aus folgender Verschärfung von Lemma 2.1.5.

Lemma 2.7.6. *Es sei* $X \in \mathcal{L}^P(\Omega, \mathcal{A}, P)$ *und* $(\mathcal{F}_n)_{n\in\mathbb{N}\cup\{\infty\}}$ *eine Filtration über* $(\Omega, \mathcal{A})$ *mit* $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n)$, $p \geq 1$. *Dann gilt :* $E[X|\mathcal{F}_n] \xrightarrow{n\to\infty} E[X|\mathcal{F}_\infty] \ P-f.s.$ *und in* $\mathcal{L}^P(\Omega, \mathcal{A}, P)$.

Beweis. $(E[X|\mathcal{F}_n])_{n\in\mathbb{N}}$ ist p-fach gleichgradig integrierbar, da

$$\sup_{n\in\mathbb{N}} E[(|E[X|\mathcal{F}_n]|^P - a)_+] \underset{Jensen}{\leq} E[(|X|^P - a)_+] \xrightarrow[\text{dom. konv. mit Majorante } |X|^P \in L^1]{a\to\infty} 0$$

Nach den Martingalkonvergenzsätzen konvergiert $E[X|\mathcal{F}_n]$ für $n \to \infty \ P-f.s.$ und in L^P gegen ein $X_\infty \in \mathcal{L}^1(\Omega, \mathcal{F}_\infty, P)$. Wir zeigen $X_\infty = E[X|\mathcal{F}_\infty]$, indem wir $E[X_\infty 1_A] = E[E[X|\mathcal{F}_\infty]1_A]$ $(*)$ für alle $A \in \mathcal{F}_\infty$ zeigen:
Die Menge $\mathcal{D}$ aller $A \in \mathcal{F}_\infty$ mit der Eigenschaft $(*)$ ist offensichtlich ein Dynkin-System (Beweis wie in Lemma 2.1.5).
$\mathcal{D}$ enthält das $\cap$-stabile System $\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n$, denn $\forall n \in \mathbb{N}$ und $A \in \mathcal{F}_n$ gilt:

$$E[X_\infty \mathbb{1}_A] \overset{E[X|\mathcal{F}_n] \xrightarrow{n\to\infty} X_\infty \text{ in } L^1}{=} \lim_{m\to\infty} E[E[X|\mathcal{F}_m]\mathbb{1}_A] \overset{\text{für } m\geq n \text{ ist } A\in\mathcal{F}_m}{=} \lim_{m\to\infty} E[X\mathbb{1}_A] \overset{A\in\mathcal{F}_\infty}{=} E[E[X|\mathcal{F}_\infty]\mathbb{1}_A]$$

Aus dem Dynkin-Lemma folgt: $\mathcal{F}_\infty = \sigma(\bigcup\limits_{n\in\mathbb{N}} \mathcal{F}_n) \subseteq \mathcal{D}$, also die Behauptung. $\square$

Der folgende Satz zeigt, dass Martingale mit beschränkten Zuwächsen entweder konvergieren oder „wild oszillieren “

Satz 2.7.7 (Alternative zur Konvergenz bei Martingalen)**.** *Es sei* $(X_n)_{n\in\mathbb{N}_0}$ *ein Martingal bzgl.* $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ *mit gleichmäßig beschränkten Zuwächsen:*
$\exists M \in \mathbb{N} \ \forall n \in \mathbb{N}_0$*:* $|X_{n+1} - X_n| \leq M$. *Dann gilt* $P-f.s.$*:*
entweder: $\lim\limits_{n\to\infty} X_n$ *existiert in* $\mathbb{R}$
oder: $\inf\limits_{n\in\mathbb{N}_0} X_n = -\infty$ *und* $\sup\limits_{n\in\mathbb{N}_0} X_n = +\infty$

Beweis. O.B.d.A. dürfen wir $X_0 = 0$ annehmen, sonst betrachten wir das Martingal $X'_n = X_n - X_0$, $n \in \mathbb{N}_0$, statt X_n. Es sei $M \in \mathbb{N}$ eine Schranke für die Zuwächse $|X_{n+1} - X_k|$, $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$:
Wir setzen

$$T_k = \inf\{n \in \mathbb{N}_0 | X_n \geq k\} \overset{X_0=0}{\geq} 1.$$

Dann gilt $\forall n \in \mathbb{N}$: $X_{T_k \wedge n} \leq k + M$. In der Tat:

1. Fall: $T_k \leq n$: Hier gilt: $X_{T_k \wedge n} = X_{T_k} = \underbrace{X_{T_k - 1}}_{<k} + \underbrace{X_{T k} - X_{T_k - 1}}_{\leq M} < k + M$

2. Fall: $T_k > n$: Hier gilt: $X_{T_k \wedge n} = X_n < k \leq k + M$

Nach dem Beispiel 2.4.2 (a) gilt: $(X_{T_k \wedge n})_{n \in \mathbb{N}_0}$ ist wieder Martingal. Weiter gilt: $\sup\limits_{n \in \mathbb{N}_0} E[(X_{T_k \wedge n})_+] \leq k + M < \infty$
Nach dem Martingalkonvergenzsatz 2.7.2, angewandt auf das Martingal $(-X_{T_k \wedge n})_{k \in \mathbb{N}}$, folgt:
$X_{T_k \wedge n}$ konvergiert für $n \to \infty$ $P - f.s.$ in $\mathbb{R}$, genauer gesagt sogar in $(-\infty, k + M]$. Das bedeutet

$P - f.s.$:
- für $T_k = \infty$: X_n konvergiert in $\mathbb{R}$
- für $T_k < \infty$: $\sup\limits_{n \in \mathbb{N}} X_n \geq k$ (nach Def. von T_k)

Weil dies $\forall k \in \mathbb{N}$ gilt, folgt $P - f.s.$:
entweder: $\lim\limits_{n \to \infty} X_n$ existiert in $\mathbb{R}$
oder: $\sup\limits_{n \in \mathbb{N}_0} X_n = +\infty$

Das gleiche Argument, angewandt auf $(-X_n)_{n \in \mathbb{N}_0}$ statt $(X_n)_n$ liefert $P - f.s.$:
entweder: $\lim\limits_{n \to \infty} X_n$ existiert in $\mathbb{R}$
oder: $\inf\limits_{n \in \mathbb{N}_0} X_n = -\infty$

Zusammen folgt die Behauptung. □

Nun betrachten wir Martingale $(X_k)_{k \in \mathbb{Z}_0^-}$ bzgl. einer Filtration $(\mathcal{F}_k)_{k \in \mathbb{Z}_0^-}$, wobei $\mathbb{Z}_0^- := \{-n | n \in \mathbb{N}_0\}$
Solche Martingale nennt man auch Rückwärtsmartingale. Wir setzen $\mathcal{F}_{-\infty} := \bigcap\limits_{k \in \mathbb{Z}_0^-} \mathcal{F}_k$. $\mathcal{F}_{-\infty}$ ist als Durschnitt von σ-Algebren wieder eine σ-Algebra.

Satz 2.7.8 (Rückwärtsmartingal-Konvergenzsatz)**.** *Jedes Rückwärtsmartingal $(X_n)_{n \in \mathbb{Z}^-}$ konvergiert für $n \to -\infty$ in L^1 und $P - f.s.$ gegen $E[X_0 | \mathcal{F}_{-\infty}]$. Ist $X_0 \in L^P(\Omega, \mathcal{F}_0, P)$, so konvergiert $(X_n)_n$ sogar in L^P $(p \geq 1)$.*

Fortsetzung. Schon gezeigt: $\forall a < b$ ist die Anzahl U_∞ der Aufkreuzungen von $[a, b]$ $P - f.s.$ endlich. Weil es nur abzählbar viele rationale Intervalle gibt, folgt aus Lemma 2.7.1: $(X_k)_{k \in \mathbb{Z}_0^-}$ konvergiert für $k \to -\infty$ $P - f.s.$ gegen eine Zufallsvariable $X_{-\infty}$ mit Werten in $\mathbb{R} \cup \{\pm\infty\}$ Ist nun $X_0 \in L^P(\Omega, \mathcal{F}_0, P)$ für $p \geq 1$ (für $p = 1$ gilt das stets), so ist $(X_n)_{n \in \mathbb{Z}_0^-}$ p-fach gleichgradig integrierbar, denn $\forall k \in \mathbb{Z}_{\not{0}}^-$ gilt:

$$E[(|X_k|^P - a)_+] = E[(E[X_0 | \mathcal{F}_k]|^P - a)_+] \underset{Jensen}{\leq} E[(|X_0|^P - a)_+]$$

und daher

$$\sup_{n \in \mathbb{Z}_{\not{0}}^-} E[(|X_k|^P - a)_+] \leq E[\underbrace{(|X_0|^P - a)_+}_{\xrightarrow{n \to \infty} 0 \text{ punktw.}}] \xrightarrow[\text{dom. kon. mit Majorante } |X_0|^P \in L^1]{n \to \infty} 0$$

Weiter ist $E[(|X_{-\infty}|^P] < \infty$, denn

$$E[(|X_{-\infty}|^P] \underset{Fatou}{\leq} \liminf_{k \to -\infty} E[(|X_k|^P] \underset{Jensen, X_k = E[X_0 | \mathcal{F}_k] P - f.s.}{\leq} E[(|X_0|^P] < \infty$$

Aus der p-fachen gleichgradigen Integrierbarkeit folgt mit Lemma 2.4.9 $X_k \xrightarrow{k\to-\infty} X_{-\infty}$ in L^P.
Wir zeigen jetzt $X_{-\infty} = E[X_0|\mathcal{F}_{-\infty}]$ $P-f.s.$ Zunächst betrachten wir $X_{-\infty} = \liminf_{k\to-\infty} X_k$ $P-f.s.$ und dieser $\liminf$ ist $\forall m \in \mathbb{Z}_{\not 0}^-$ $\mathcal{F}_m$-messbar, denn $\forall k \le m$ ist X_k $\mathcal{F}_m$-messbar. Das bedeutet $A \in \mathcal{B}(\mathbb{R}\cup\{\pm\infty\}$ $\{\liminf_{k\to-\infty} X_k \in A\} \in \bigcap_{m\in\mathbb{Z}_0^-} \mathcal{F}_m = \mathcal{F}_\infty$, d.h. $\liminf_{k\to-\infty} X_k$ ist $\mathcal{F}_{-\infty}$- messbar.
Weiter gilt für $B \in \mathcal{F}_{-\infty}$:

$$E[X_{-\infty}\mathbb{1}_B] \underset{X_k \xrightarrow{k\to-\infty} X_{-\infty}\in L^1}{=} \lim_{k\to-\infty} E[X_k\mathbb{1}_B] = \lim_{k\to-\infty} E\left[E[X_0|\mathcal{F}_k]\mathbb{1}_{\underset{\in\mathcal{F}_{-\infty}\subseteq\mathcal{F}_k}{B}}\right] = E[X_0\mathbb{1}_B]$$

Zusammen folgt $E[X_0|\mathcal{F}_{-\infty}] = X_{-\infty}$ $P-f.s.$ □

Als eine Anwendung erhalten wir:

Satz 2.7.9 (Teilaussage des starken Gesetzes der großen Zahlen)**.** *Ist $(Z_n)_{n\in\mathbb{N}}$ eine i.i.d. Folge von Zufallsvariablen in $L^P(\Omega,\mathcal{A},P)$, $p \ge 1$, so konvergiert $\frac{1}{n}\sum_{k=1}^{n} Z_k$ für $n\to\infty$ $P-f.s.$ und in $L^P(\Omega,\mathcal{A},P)$.*

Beweis. Für $X_{-n} := \frac{1}{n}\sum_{k=1}^{n} Z_k$ und $\mathcal{F}_{-n}, (Z_k)_{k>n})$ MAn beachte $\mathcal{F}_{-n-1} = \frac{1}{n+1}(nX_{-n} + Z_{n+1})$ ist $\mathcal{F}_{-n}$-messbar. Nun gilt $\forall k = 1,...,n$ gilt: $E[Z_k|\mathcal{F}_{-n}] = E[Z_1|\mathcal{F}_{-n}]$ $P-f.s.$
In der Tat: Da $(Z_k, X_{-n}, (Z_l)_{l>n})$ und $(Z_1, X_{-n}, (Z_l)_{l>n})$ die gleiche gemeinsame Verteilung haben, gilt $\forall A \in \mathcal{F}_{-n}$:

$$E[Z_k\mathbb{1}_A] = E[Z_1\mathbb{1}_A] = E[E[Z_1|\mathcal{F}_{-n}]\mathbb{1}_A]$$

Es folgt $\forall n \in \mathbb{N}$

$$nE[Z_n|\mathcal{F}_{-n}] = \sum_{k=1}^{n} E[Z_k|\mathcal{F}_{-n}] = E[\underbrace{\sum_{k=1}^{n} Z_k}_{=nX_{-n}}|\mathcal{F}_{-n}] = E[nX_{-n}|\mathcal{F}_{-n}] = nX_{-n}\ P-f.s.$$

also $E[Z_n|\mathcal{F}_{-n}] = X_{-n}$ $P-f.s.$
Das bedeutet $(X_{-n})_{-n\in\mathbb{Z}^-}$ ist ein Rückwärtsmartingal bzgl. $(\mathcal{F}_{-n})_{-n\in\mathbb{Z}^-}$. Dei Behauptung folgt nun aus dem Rückwärtsmartingalkonvergenzsatz 2.7.8 □

Bemerkung. Später werden wir sehen, dass der Limes $P-f.s.$ konst gleich $E[Z_n]ist$

2.8 Lemmata von Borel-Cantelli

Für eine Folge $(A_n)_{n\in\mathbb{N}}$ von Ereignissen sei $\limsup_{n\to\infty} := \bigcap_{n\in\mathbb{N}}\bigcup_{m:m\ge n} A_m$ das Ereignis , dass A_n für ∞ viele $n\in\mathbb{N}$ eintritt. Anders gesagt:

$$\mathbb{1}_{\limsup_{n\to\infty} A_n} = \limsup_{n\to\infty} \mathbb{1}_{A_n}.$$

Erinnerung.

Satz 2.8.1 (1. Lemma von Borel-Cantelli)**.** *Wenn $\sum_{n\in\mathbb{N}} P(A_n) < \infty$,so gilt $P(\limsup_{n\to\infty} A_n) = 0$*

Beweis. Wegen $\mathbb{1}_{\limsup_{n\to\infty} A_n} \le \sum_{n=k}^{\infty} \mathbb{1}_{A_k} \forall k\in\mathbb{N}$ gilt:

$$P\left(\limsup_{n\to\infty} A_n\right) = E\left[\mathbb{1}_{\limsup_{n\to\infty} A_n}\right] \le E\left[\sum_{n=k}^{\infty} \mathbb{1}_{A_n}\right] \underset{mon.\ kon.}{=} \sum_{n=k}^{\infty} P(A_n) \xrightarrow[wg.\ \sum_{n=1}^{\infty} P(A_n)<\infty]{n\to\infty} 0$$

also $P\left(\limsup_{n\to\infty} A_n\right) = 0.$ □

In diesem Abschnitt beweisen wir mehrere Versionen von „Umkehrungen “dieses Lemmas unter Zusatzvoraussetzungen. Wir beginnen mit einer Unabhängigkeitsannahme und schwächen diese später ab.

Satz 2.8.2 (2. Lemma von Borel- Cantelli)**.** *Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger Ereignisse mit $\sum_{n\in\mathbb{N}} P(A_n) = \infty$, so gilt: $P(\limsup_{n\to\infty} A_n) = 1$*
Für unabhängige A_n, $n \in \mathbb{N}$, gilt also ein „0-1-Gesetz“:

$$P(\limsup_{n\to\infty} A_n) = \begin{cases} 0, \text{ falls } \sum_{n\in\mathbb{N}} P(A_n) < \infty \\ 1, \text{ falls } \sum_{n\in\mathbb{N}} P(A_n) = \infty \end{cases}$$

Beweis. Für $n \in \mathbb{N}$ gilt:

$$P(\bigcup_{m:m\geq n} A_m = 1 - P((\bigcup_{m:m\geq n} A_m)^C) = 1 - P(\bigcup_{m:m\geq n} A_m^C) \quad (*)$$

Nun gilt wegen der σ-Stetigkeit von p:

$$P(\bigcup_{m:m\geq n} A_m^C) = \lim_{k\to\infty} P(\bigcup_{m=n}^{k} \underbrace{A_m^C}_{\text{unabhängig}} = \lim_{k\to\infty} \prod_{m=n}^{k} P(A_m^C) \quad (**)$$

Wir schätzen:

$$\prod_{m=n}^{k} P(A_m^C) = \prod_{m=n}^{k} \underbrace{(1 - P(A_m)}_{\geq 0} \underset{\forall x\in\mathbb{R}: 1-x\leq e-x}{\leq} \prod_{m=n}^{k} e^{-P(A_m)} = e^{-\sum_{m=n}^{k} P(A_m)} \xrightarrow[\sum_{m=n}^{\infty} P(A_m)=\infty, da \sum_{n\in\mathbb{N}} P(A_m)=\infty]{n\to\infty} 0$$

also auch $\sum_{m=n}^{\infty} P(A_m) = \infty$, da $\sum_{n\in\mathbb{N}} P(A_m) = \infty$
In $(**)$ und $(*)$ eingesetzt erhalten wir: $P(\bigcup_{m\geq n} A_m) = 1$ und daher auch $P(\bigcap_{n\in\mathbb{N}} \bigcup_{m:m\geq n} A_m) = 1$, also die Behauptung. □

Beispiel 2.8.3. Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen auf $]0,1[$ gleichverteilten Zufallsvariablen und $(a_n)_{n\in\mathbb{N}}$ eine Folge positiver Zahlen mit $a_n \xrightarrow{n\to\infty} \infty$. Dann gilt $P - f.s.$:

$$\liminf_{n\to\infty} a_n X_n = \begin{cases} 0, & \text{falls } \sum_{n\in\mathbb{N}} \frac{1}{a_n} = \infty \\ \infty, & \text{falls } \sum_{n\in\mathbb{N}} \frac{1}{a_n} < \infty \end{cases}$$

Beweis. Zunächst betrachten wir den Fall $\sum_{n\in\mathbb{N}} \frac{1}{a_n} = \infty 0$. Es sei $\varepsilon > 0$. Für alle genügend großen n, sagen wir für $n \geq n_0(\varepsilon)$, gilt $\frac{\varepsilon}{a_n} \leq 1$.
Für diese n schließen wir:

$$P[a_n X_n < \varepsilon] = P\left[X_n < \frac{\varepsilon}{a_n}\right], \text{ also}$$

$$\sum_{n=n_0(\varepsilon)}^{\infty} P[a_n X_n \varepsilon] = \sum_{n=n_0(\varepsilon)}^{\infty} \frac{\varepsilon}{a_n} = \infty$$

Wegen der Unabhängigkeit der $\{a_nX_n < \varepsilon\}$, $n \in \mathbb{N}$, folgt aus dem 2. Borel-Cantelli-Lemma:

$$P[a_nX_n < \varepsilon \text{ für } \infty \text{ viele } n] = 1, \text{ also } P-f.s. \liminf_{n\to\infty} a_nX_n \leq \varepsilon.$$

Weil das für alle rationale $\varepsilon > 0$ gilt, folgt:
$\liminf\limits_{n\to\infty} a_nX_n = 0 \ P-f.s.$
Nun betrachten wir den Fall $\sum\limits_n \frac{1}{a_n} < \infty$ Es sei $M > 0$, $M \in \mathbb{R}$. Für $n \in \mathbb{N}$ schließen wir:
$P[a_nX_n < M] = P[X_n \leq \frac{M}{a_n}] \leq \frac{M}{a_n}$ und daher $\sum\limits_{n\in\mathbb{N}} P[a_nX_n < M] \leq \sum\limits_{n\in\mathbb{N}} \frac{M}{a_n} < \infty$
Es folgt aus dem 1. Borel-Cantelli-Lemma:
$P[a_nX_n < M$ für ∞ viele n$] = 0$, also $P[\liminf\limits_{n\to\infty} a_nX_n < M] = 0$.
Weil das für alle rationalen $M > 0$ gilt, folgt $\liminf\limits_{n\to\infty} a_nX_n = \infty \ P-f.s.$ □

Verallgemeinerung der beiden Borel-Cantelli-Lemmata zu einer bedingten Version: Notation: Für ein Ereignis A und eine σ-Algebra $\mathcal{F} \subseteq \mathcal{A}$ schreiben wir $P[A|\mathcal{F}] := E[\mathbb{1}_A|\mathcal{F}]$. $P[A|\mathcal{F}]$ wird bedingte Wahrscheinlichkeit von A gegeben $\mathcal{F}$ genannt.

Satz 2.8.4 (bedingtes Borel-Cantelli-Lemma)**.** *Es sei $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ eine Filtration und $(A_n)_{n\in\mathbb{N}}$ eine Folge von Ereignissen mit $A_n \in \mathcal{F}_n$ $\forall n \in \mathbb{N}$. Dann gilt $P-f.s.$:*

$$\sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} = \infty \textit{ genau dann, wenn } \sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}] = \infty$$

Beweis. Es sei $X_n = \sum\limits_{k=1}^n \mathbb{1}_{A_k}$ und $Y_n = \sum\limits_{k=1}^n E[\mathbb{1}_{A_k}|\mathcal{F}_{k-1}]$ für $n \in \mathbb{N}_0$. Offensichtlich ist $(X_n)_{n\in\mathbb{N}_0}$ adaptiert und $(Y_n)_{n\in\mathbb{N}}$ sogar vorhersehbar bezüglich $(\mathcal{F}_n)_{n\in\mathbb{N}}$ und es gilt $Y_0 = 0$.
Weiter ist $M_n = X_n - Y_n$, $n \in \mathbb{N}_0$, ein Martingal, denn es gilt:

$$E[M_n|\mathcal{F}_{n-1}] = E[\underbrace{M_{n-1}}_{\mathcal{F}_{n-1}-messbar} + \mathbb{1}_{A_n} - \underbrace{E[\mathbb{1}_{A_n}|\mathcal{F}_{n-1}]}_{\mathcal{F}_{n-1}-messbar}|\mathcal{F}_{n-1}] = M_{n-1} + E[\mathbb{1}_{A_n}|\mathcal{F}_{n-1}] - E[\mathbb{1}_{A_n}|\mathcal{F}_{n-1}] = M_{n-1}.$$

Anders gesagt: $X_n = M_n + Y_n$, $n \in \mathbb{N}_0$, ist die Doob-Zerlegung von $(X_n)_{n\in\mathbb{N}_0}$. Nun hat $(M_n)_{n\in\mathbb{N}}$ $P-f.s.$ beschränkte Zuwächse, da $|M_n - M_{n-1}| = |\mathbb{1}_{A_n} - P[A_n|\mathcal{F}_{n-1}]| \leq 1$ Aus Satz 2.7.7 (alternative zur Konvergenz von Martingalen) folgt folgende Alternative: $P-f.s.$ gilt:

<u>entweder:</u>	$\lim\limits_{n\to\infty} M_n$ existiert in $\mathbb{R}$ (1.Fall)
<u>oder:</u>	$\liminf\limits_{n\to\infty} M_n = -\infty$ und $\limsup\limits_{n\to\infty} M_n = +\infty$ (2.Fall)

Im 1.Fall gilt $X_n \xrightarrow{n\to\infty} \infty$ genau dann, wenn $Y_n = X_n - M_n \xrightarrow{n\to\infty} \infty$, d.h. die Behauptung des Satzes gilt im 1.Fall.
Im 2.Fall schließen wir $P-f.s.$:

$$\sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} = \lim_{n\to\infty} X_n = \lim_{n\to\infty} (\underbrace{Y_n}_{\geq 0,\ steigend\ in\ n} + M_n) = \underbrace{\lim_{n\to\infty} Y_n}_{\geq 0} + \underbrace{\limsup_{n\to\infty} M_n}_{=+\infty} = +\infty \textit{ und}$$

$$\sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}] = \lim_{n\to\infty} Y_n = \lim_{n\to\infty} (\underbrace{X_n}_{\geq 0\ steigend} - M_n) = \underbrace{\lim_{n\to\infty} X_n}_{\geq 0} - \underbrace{\liminf_{n\to\infty} M_n}_{=-\infty} = +\infty$$

□

Bemerkung. Die beiden „unbedingten" Borel-Cantelli-Lemmata 2.8.1, 2.8.2 sind in der bedingten Version 2.8.4 wie folgt enthalten:

1) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von Ereignissen mit $\sum\limits_{n\in\mathbb{N}} P(A_n) < \infty$ und $(\mathcal{F}_n)_{n\in\mathbb{N}}$ irgendeine Filtration mit $A_n \in \mathcal{F}_n$ $(n \in \mathbb{N})$, so folgt $P-f.s.$:

$$\sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}] < \infty \ wegen \ E[\sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}]] \underset{mon.\ konv.}{=} \sum_{n\in\mathbb{N}} E[P[A_n|\mathcal{F}_{n-1}]] = \sum_{n\in\mathbb{N}} P(A_n) < \infty,$$

also $\sum\limits_{m} \mathbb{1}_{A_n} < \infty$, d.h. das Ereignis $\limsup\limits_{n\to\infty} A_n$ tritt nicht ein nach der bedingten Version des Borel-Cantelli-Lemmas 2.8.4.

2) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Ereignissen mit $\sum\limits_{n\in\mathbb{N}} P(A_n) = \infty$ und $\mathcal{F}_n = \sigma(A_1, ..., A_n)$ für $n \in \mathbb{N}_0$, so ist

$$P[A_n|\mathcal{F}_{n-1}] = E[\mathbb{1}_{A_n}] = P[A_n] \ P-f.s. \ (n \in \mathbb{N}), \ also$$

$$\sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}] = \sum_{n\in\mathbb{N}} P[A_n] = \infty \ P-f.s.$$

und damit nach der bedingten Version 2.8.4:
$\sum\limits_{n\in\mathbb{N}} \mathbb{1}_{A_n} = \infty$, d.h. $\limsup\limits_{n\to\infty} A_n$ tritt $P-f.s.$ auf.

Beispiel 2.8.5 (Polya-Urne). Eine Urne enthält zunächst eine rote und eine blaue Kugel. Wir ziehen immerwieder zufällig eine Kugel und legen sie zurück zusammen mit einer neuen Kugel der gleichen Farbe. R_n und B_n bezeichne dei Anzahl roter bzw. blauer Kugeln nach dem n-ten Zug in der Urne, $B_0 = R_0 = 1$. Wir zeigen, dass $P-f.s.$ unendlich oft eine blaue Kugel gezogen wird.
Hierzu bezeichne $\mathcal{F}_n = \sigma(R_k, B_k : \ k = 1, ..., n)$ $(n \in \mathbb{N}_0$ die σ-Algebra der bis zum n-ten Zug beobachtbaren Ereignisse und $A_n = \{B_n = B_{n-1} + 1\}$, $n \in \mathbb{N}$, das Ereignis, das im n-ten Zug eine blaue Kugel gez. wird. Es gilt also $A_n = \{B_n = B_{n-1}, R_n = R_{n-1} + 1\}$.
Man beachte: Die A_n $n \in \mathbb{N}$ sind abhängig!
Nach der Modellbeschreibung gilt:

$$P[A_n|\mathcal{F}_{n-1}] = \frac{B_{n-1}}{B_{n-1} + R_{n-1}} \underset{fehlt}{=} \frac{B_{n-1}}{n+1} \underset{fehlt}{\geq} \frac{1}{n+1}$$

und daher

$$\sum_{n\in\mathbb{N}} P[A_n|\mathcal{F}_{n-1}] \geq \sum_{n\in\mathbb{N}} \frac{1}{n+1} = \infty$$

Aus der bedingten Version des Borel-Cantelli-Lemmas folgt $\sum\limits_{n\in\mathbb{N}} \mathbb{1}_{A_n} = \infty \ P-f.s.$, also die Behauptung.

2.9 Verzweigungsprozesse

Verzweigungsprozesse (Synonym: Galton-Watson Prozesse) sind in der mathematischen Biologie Modelle zur Beschreibung von Ppulationsdynamiken. Wir analysieren hier nur das einfachste Modell in dieser Klasse.
Eine Population besteht aus N_t Individuen zur Zeit t, $t \in \mathbb{N}_0$. Anfangs ist nur 1 Individuum vorhanden: $N_0 = 1$. Zu jeder Zeit t spaltet sich jedes vorhandene Individuum in eine zufällige, nach einer Verteilung μ auf $\mathbb{N}_0$ verteilten Anzahl von Nachkommen auf, unabhängig von den anderen Individuen und unabhängig von der Vorgeschichte. Wir modellieren wie folgt:
$(X_{t,i})_{t\in\mathbb{N}, i\in\mathbb{N}}$ seien i.i.d. μ-verteilte Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Dabei soll $Z_{t,i}$ die Nachkommenanzahl des i-ten Individuums zur Zeit t-1 bedeuten, falls es dieses Individuum gibt, $t \in \mathbb{N}$. Wir setzen $\mathcal{F}_t := \sigma(Z_{s,i} : s = 1, ..., t, i \in \mathbb{N})$, $t \in \mathbb{N}_0$ und definieren rekursiv:
$N_0 := 1,$ Insbesondere ist $(N_t)_{t\in\mathbb{N}_0}$ bezüglich $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ adaptiert,

$N_t := \sum\limits_{i=1}^{N_{t-1}} Z_{t,i}$ und $N_1 = Z_{1,1}$ ist μ-verteilt.
Wir definieren für $\varphi \in [0,1]$:

$$f(\varphi) := E\left[\varphi^{N_i} = \sum_{n=0}^{\infty} \varphi^n \mu(\{n\})\right], \textit{ wobei } 0^0 := 1$$

Lemma 2.9.1. *Es gilt:* $E[\rho^{N_t}|\mathcal{F}_{t-1}] = f(\rho)^{N_{t-1}}$ *für* $t \in \mathbb{N}, \rho \in [0,1]$. *Für* $\rho \in [0,1]$ *ist* $(\rho^{N_t})_{t\in\mathbb{N}_0}$ *ein Martingal für* $f(\rho) = \rho$, *ein Submartingal für* $f(\rho) \geq \rho$ *und ein Supermartingal für* $f(\rho) \leq \rho$ *bzgl.* $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$

Beweis. Für $t \in \mathbb{N}, n \in \mathbb{N}_0$ gilt P-f.s.:

$$E[\rho^{N_t}|\mathcal{F}_{t-1}]\mathbb{1}_{\underbrace{\{N_{t-1} = n\}}_{\in\mathcal{F}_{t-1}}} = E[\rho^{N_t}\mathbb{1}_{\{N_{t-1}=n\}}|\mathcal{F}_{t-1}] = E[\rho^{\sum_{i=1}^n Z_{t,i}}\mathbb{1}_{\{N_{t-1}=n\}}|\mathcal{F}_{t-1}]$$

$$= E[\prod_{i=1}^n \rho^{Z_{t,i}}|\mathcal{F}_{t-1}] \underset{\text{Fakt. alle unabh. \& unabh. von } \mathcal{F}_{t-1}}{\uparrow}= \prod_{i=1}^n \underbrace{E[\rho^{\overbrace{Z_{t,i}}^{\text{gleiche Vert. wie} N_1}}]}_{=f(\rho)}\mathbb{1}_{\{N_{t-1}=n\}} = f(\rho)^n\mathbb{1}_{\{N_{t-1}=n\}} = f(\rho)^{N_{t-1}}\mathbb{1}_{\{N_{t-1}=n\}}$$

und daher P-f.s.:

$$E[\rho^{N_t}|\mathcal{F}_{t-1}] = \sum_{n\in\mathbb{N}_0} E[\rho^{N_t}|\mathcal{F}_{t-1}]\mathbb{1}_{\{N_{t-1}=n\}} = \sum_{n\in\mathbb{N}_0} f(\rho)^{N_{t-1}}\mathbb{1}_{\{N_{t-1}=n\}} = f(\rho)^{N_{t-1}}$$

Insbesonder ist im Fall $f(\rho) = \rho$ die Folge (ρ^{N_t}) ein Martingal.
Nun ist für $f(\rho) \underset{(\geq)}{\leq} \rho : f(\rho)^{N_{t-1}} \underset{(\geq)}{\leq} \rho^{N_{t-1}}$. Hieraus folgt die Behauptung. ☐

Wir analysieren nun die erz. Fkt. f, insbesondere im Hinblick auf Fixpunkte.

Lemma 2.9.2. *Die Funktion* $f : [0,1] \to [0,1], f(\rho) = E[\rho^{N_1}]$ *ist monoton steigend und konvex. Sie ist strikt konvex für* $P[N_1 > 1] > 0$, *d.h.*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y) \textit{ für } \lambda \in (0,1)\ 0 \leq x < y \leq 1$$

Es gilt $f(1) = 0, f(0) = E[0^{N_1}] = E[\mathbb{1}_{\{N_1=0\}}] = P[N_1 = 0] = \mu(\{0\})$.
Weiter existiert die linksseitige Ableitung in 1: $f_l'(1) = \lim\limits_{\rho\nearrow 1} \dfrac{f(\rho)-1}{\rho-1}$ *in* $[0,\infty)$, *und es gilt* $f_l'(1) = E[N_1]$

Beweis. Da alle FUnktionen $p_n : \rho \to \rho^n,\ \rho \in [0,1],\ n \in \mathbb{N}$, monoton steigend und konvex sind, ist auch

$$E[\rho^{N_1}] = \sum_{n\in\mathbb{N}_0} p_n(\rho)\mu(\{n\}) \text{ monoton steigend und konvex.}$$

Nun ist auch p_n für $n > 1$ strikt konvex, also ist auch f strikt konvex, sobald ein p_n mit $n > 1$ positives Gewicht hat, d.h. sobald $P[N_1 > 1] > 0$.
$f(1) = 1$ und $f(0) = \mu(\{0\})$ ist klar. Für $\rho < 1$ erhalten wir:

$$\frac{f(\rho)-f(1)}{\rho-1} = \frac{E[\rho^{N_1}]-1}{\rho-1} = E\underbrace{\left[\frac{\rho^{N_1}-1}{\rho-1}\right]}_{\text{geom. Reihe}} = E\underbrace{\left[\sum_{k=0}^{N_1-1}\rho^k\right]}_{\nearrow N_1 \text{ für } \rho\nearrow 1} \xrightarrow[\text{mon. Konv.}]{\rho\nearrow 1} E[N_1]$$

Also $f_l'(1) = E[N_1]$.

☐

Für die Fixpunktstruktur von f erhalten wir:

Lemma 2.9.3. *(1) "Trivialfall: keine zeitliche Änderung": Falls $N_1 = 1$ P-f.s. ist $f(\rho) = \rho$, $0 \leq \rho \leq 1$*

(2) ßubkritischer Fall: weniger als 1 Nachkommen pro Individuum erwartet": Falls $E[N_1] < 1$ ist $f(\rho) > \rho$ für $0 \leq \rho < 1$ und $f(1) = 1$. Hier gibt es also nur der Fixpunkt 1 von f.

(3) "kritischer Fall: 1 Nachkommen pro Individuum erwartet": Falls $E[N_1] = 1$, aber nicht P-f.s. $N_1 = 1$, so gilt ebenfalls $f(\rho) = \rho\ \forall 0 \leq \rho < 1$ und $f(1) = 1$. Auch hier gibt es also nur den Fixpunkt 1 von f.

(4) SSuperkritischer Fall: mehr als 1 Nachkommen pro Individuum erwartet"Falls $E[N_1] > 1$, so existiert genau ein $\rho^\star \in [0,1)$ mit $f(\rho^\star) = \rho^\star$. Es gilt $f(\rho) > \rho$ für $0 \leq \rho < \rho^\star, f(\rho) < \rho$ für $\rho^\star < \rho < 1$ und $f(1) = 1$. In diesem Fall hat f genau 2 Fixpunkte: $\rho^\star$ und 1.

Beweis. (1) ist klar.

(2) Wegen Konvexität von f gilt für $0 \leq \rho < 1$:

$$f(\rho) \geq \underbrace{f(1)}_{=1} + \underbrace{(\rho - 1)}_{>0} \underbrace{f_l'(1)}_{<1} > 1 + (\rho - 1) = \rho$$

(3) Wegen der strikten Konvexität von f gilt für $0 \leq \rho < 1$:

$$f(\rho) > \underbrace{f(1)}_{=1} + (\rho - 1) \underbrace{f_l'(1)}_{=1} = \rho$$

(4) Wegen $f_l'(1) > 1$ ist für alle $\rho < 1$ genügend nahe bei 1. $\dfrac{f(\rho) - f(1)}{\rho - 1} > 1$, also $f(\rho) - \underbrace{f(1)}_{=1} < \rho - 1$, d.h. $f(\rho) < \rho$.
Weil $f(0) \geq \mu(\{0\})$ und f stetig ist, gibt es nach dem ZWS ein $\rho^\star \in [0,1)$ mit $f(\rho^\star) = \rho^\star$. f hat mindestens zwei Fixpunkte $\rho^\star$ und 1. Es kann nicht mehr als zwei Fixpunkte haben, da es strikt konvex ist.

□

Satz 2.9.4. *(1) Im Trivialfall $N_1 = 1$ P-f.s. ist $N_t = 1\ \forall\ t \in \mathbb{N}_0$ P-f.s.*

(2) und (3) Im Subkritischen und kritischen Fall: $E[N_1] \leq 1$ und $P[N_1 = 1] < 1$ gilt P-f.s.: $N_t = 0$ für $t \to \infty$ schließlich, d.h. die Population stirbt mit Wahrscheinlichkeit 1 aus.

(4) Im superkritischen Fall $E[N_1] > 1$ gilt: $P[N_t \geq 1\ \forall\ t \in \mathbb{N}_0] > 0$, d.h. die Population stirbt mit positiver Wahrscheinlichkeit nicht aus.

Beweis. (1) ist klar.

Wir definieren rekursiv: $f^{(0)} = id, f^{(n+1)} = f \otimes f^{(n)}\ n \in \mathbb{N}_0$, also $f^{(n)}(\rho) = \underbrace{f(f(\ldots f}_{n\text{-mal}}(\rho)\ldots))$.
Durch Iteration erhalten wir aus Lemma 2.9.1:

$$E[\rho^{N_{t+n}}]|\mathcal{F}_t = f^{(n)}(\rho)^{N_t} \text{ für } t, n \in \mathbb{N}_0, \rho \in [0,1]$$

Insbesondere

$$E[\rho^{N_t}] = f^{(t)}(\rho) \quad (t \in \mathbb{N}_0)$$

(2) und (3) Im subkritischen und kritischen Fall gilt $f(\rho) \geq \rho \ \forall \ \rho \in [0,1]$. Also ist $(f^{(t)}(0))_{t\in\mathbb{N}_0}$ monoton steigend und damit konvergent. Wegen der Stetigkeit von f folgt für $s := \lim\limits_{t\to\infty} f^{(t)}(0) : f(s) = \lim\limits_{t\to\infty} f^{(t+1)}(0) = s$, also $s = 1$, da f nur den Fixpunkt 1 besitzt.

Interpretation:

$$P[N_t = 0] = E[\mathbb{1}_{\{N_t=0\}}] = E[O^{N_t}] = f^{(t)}(0) \xrightarrow{t\to\infty} 1$$

Die Behauptung folgt nun, da aus $N_t = 0$ auch $N_s = 0 \ \forall \ s \geq t$ folgt.

$$P[N_t = 0 \text{ für } t \to \infty \text{ schließlich}] = P[\exists t : N_t = 0] \geq \sup_{t\in\mathbb{N}_0} P[N_t = 0] = 1$$

(4) Im superkritischen Fall gilt $\forall \ \rho \in [0,\rho^\star] : 0 \leq f(\rho) \underset{\substack{\uparrow \\ \text{Monotonie von } f}}{\leq} f(\rho^\star) = \rho^\star$, also $f(\rho) \in [0,\rho^\star]$.

Es folgt durch Iteration: $f^{(t)}(0) \in [0,\rho^\star] \ \forall \ t \in \mathbb{N}_0$.
Wir schließen:
$P[N_t = 0] = f^{(t)}(0) \leq \rho^\star \ \forall \ t \in \mathbb{N}_0$, also $P[\exists t : N_t = 0] = P[N_t = 0 \text{ für } t \to \infty \text{ schließlich}] = \lim\limits_{t\to\infty} P[N_t = 0] \leq \rho^\star < 1$

□

Bemerkung. Man sieht leicht, dass $f(t)(0) \xrightarrow{t\to\infty} \rho^\star$, also ist $\rho^\star$ die Aussterbbarkeit der Population. (Übung)

3 0-1-Gesetze und Gesetze der großen Zahlen

3.1 Das 0-1-Gesetz von Kolmogoroff

In der Martingaltheorie haben wir schon die $\sigma-$Algebra $\mathcal{F}_\infty = \bigcup\limits_{t\in\mathbb{Z}_0^-} \mathcal{F}_t$ der Ereignisse der ∞ fernen Vergangenheit kennengelernt.
Die folgende Definition verallgemeinert das:

Definition 3.1.1. *Es sei $(\mathcal{F}_i)_{i\in I}$ eine Familie von σ-Algebren über einen Ergebnisraum Ω, indiziert mit einer unendlichen Familie I. Die terminale σ-Algebra, engl. „tail field“dazu wird durch*

$$\tau := \bigcap_{\substack{E\subset I \\ E \text{ endlich}}} \sigma\left(\bigcup_{i\in I\setminus E} \mathcal{F}_i\right)$$

Die Elemente von τ heißen terminale Ereignisse bzgl. $(\mathcal{F}_i)_{i\in I}$. (engl.:„tailevents“). Ist $(X_i)_{i\in I}$ eine Familie von Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega,\mathcal{A},P)$, so heißt die terminale σ-Algebra τ zu $(\sigma(X_i))_{i\in I}$ auf terminale σ-Algebra zu $(X_i)_{i\in I}$. Es gilt also:

$$\tau := \bigcap_{\substack{E\subset I \\ E \text{ endlich}}} \sigma(X_i : i \in I \setminus E)$$

Bemerkung. 1. τ ist als Durchschnitt von σ-Algebren wieder eine σ-Algebra

2. Anschaulich gesprochen besteht die terminale σ-Algebra aus allen Ereignissen, die man mit Hilfe der X_i entscheiden kann, wenn man endlich viele der X_i ignoriert.

3. Im Fall $I = \mathcal{Z}_0^-$, $\mathcal{F}_t = \sigma(X_s : s \leq t)$, $t \in I$ stimmt die terminale σ-Algebra zu den $(X_t)_{t\in I}$ mit der σ-Algebra $\mathcal{F}_{-\infty} = \bigcap\limits_{t\in\mathbb{Z}_{\not{0}}^-} \mathcal{F}_t$ der „Ereignisse in der ∞ fernen Vergangenheit“überein.

4. *Sind*$(X_n)_{n\in\mathbb{N}}$ reellwertige Zufallsvariablen, so sind die Zufallsvariablen

$$\liminf_{n\to\infty} X_n, \quad \limsup_{n\to\infty} X_n, \quad \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k, \quad \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k$$

messbar bzgl. der terminalen σ-Algebra $\tau = \bigcap\limits_{n\in\mathbb{N}} \sigma(X_n)_{k\geq n}$ (Übung)

Für die letzten beiden Zufallsvariablen folgt das daraus, dass es bei der Mittelbildung $\frac{1}{n}\sum\limits_{k=1}^{n} X_k$ im Limes $n\to\infty$ auf ein festes Anfangsstück $X_1, ..., X_m$ nicht ankommt.

5. Manchmal sollte man sich I besser räumlich als zeitlich vorstellen:
Ist $I = \mathbb{Z}$ und stellt man sich die $i \in \mathbb{Z}^d$ als „Atome “vor , so kamm man sich terminale Ereignisse als „makroskopische Ereignisse “vorstellen, bei denen es nicht auf endlich viele Atome ankommt.

6. Ist I eine unbeschränkte Teilmenge von $\mathcal{R}^n$, so wird eine Variante der terminalen σ-Algebra durch $\mathcal{F} = \bigcap\limits_{B\subseteq I \text{ beschränkt}} \sigma(X_i : i \in I\setminus B)$ definiert. Man kann sich ihre Elemente „als Ereignisse im ∞ Fernen “vorstellen.

Satz 3.1.2 (0-1-Gesetz von Kolmogoroff)**.** *Ist* $(\mathcal{F}_i)_{i\in I}$ *eine unendliche Familie von unabhängigen* σ*-Algebren über einem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, P)$*, so hat jedes Ereignis A in der zugehörigen terminalen* σ*-Algebra* τ *entweder die Wahrscheinlichkeit* $P(A) = 0$ *oder* $P(A) = 1$

Beweis. Wir kürzen für $J \subseteq I$ ab:

$$\mathcal{F}_J := \sigma\left(\bigcup_{i\in J} \mathcal{F}_i\right). \text{ Es gilt also} \tau = \bigcap_{\substack{E\subseteq I\\ E \text{ endlich}}} \mathcal{F}_{I\setminus E}$$

Erinnerung (Stochastikvorlesung)**.** *Sind* $\mathcal{G}_1, \mathcal{G}_2$ *unabhängige* $\cap$*-stabile Ereignissysteme, so sind* $\sigma(\mathcal{G}_1)$*,* $\sigma(\mathcal{G}_2)$ *ebenfalls unabhängig.* $(*)$

Wir zeigen hiermit jetzt: Sind $J, K \subseteq I$ disjunkt , so sind $\mathcal{F}_J$ und $\mathcal{F}_k$ unabhängig. In der Tat: Setzen wir $G_L = \{\bigcap\limits_{i\in E} A_i | E \subseteq L \text{ } endlich, A_i \in \mathcal{F}_i \forall i \in E\}$ für $L \subseteq I$, so ist G_L ein $\cap$- stabiles Erzeugendensystem von $\mathcal{F}_L$. Weil die $(\mathcal{F}_i)_{i\in I}$ unabhängig ist und $J \cap K = \emptyset$, sind G_L und G_K unabhängig voneinander. Also sind auch $\mathcal{F}_J = \sigma(G_J)$ und $\mathcal{F}_K = \sigma(G_K)$ voneinander unabhängig, wegen $(*)$. □

Insbesondere ist für alle endlichen $E \subset J$ die σ-Algebra $\mathcal{F}_E$ unabhängig von $\mathcal{F}_{J\setminus E}$, also wegen $\mathcal{F} \subseteq \mathcal{F}_{I\setminus E}$ erst recht unabhängig von $\mathcal{F}$.
Nun ist $\mathcal{H} = \bigcup\limits_{\substack{E\subset I\\ E \text{ endlich}}} \mathcal{F}_E$ ein $\cap$-stabiles Erzeugendensystem von $\mathcal{F}_I$ und nach dem eben Gezeigten ist es unbahängig von τ. Nochmal mit $(*)$ folgt: $\mathcal{F}_I = \sigma(\mathcal{H})$ ist unabhängig von $\tau = \sigma(\tau)$.
Wegen $\tau \subseteq \mathcal{F}_i$ folgt hieraus: τ ist unabhängig von sich selbst. Es gilt also $\forall A \in \tau$:
$P(A) = P(A \cap A) = P(A)P(A)$, also $P(A) \in [0, 1]$

Definition 3.1.3. *Es sei* $(\Omega, \mathcal{A}, P)$ *ein Wahrscheinlichkeitsraum. Wir nennen eine Unter-*σ*-Algebra* $\mathcal{F} \subseteq \mathcal{A}$ $P - f.s.$ *trivial, wenn* $\forall A \in \mathcal{F}$ *gilt* $P(A) = 0$ *oder* $P(A) = 1$.

Das 0-1-Gesetz von Kolmogoroff besagt also, dass die terminale σ-Algebra einer unendlichen Familie von unabhängigen σ-Algebren $P - f.s.$ trivial ist.

Lemma 3.1.4. *Jede bezüglich einer* $P{-}f.s.$ *trivialen* σ*-Algebra* $\mathcal{F}$ *messbare reellwertige Zufallsvariable X ist* $P - f.s.$ *konstant, d.h. es gibt ein* $a \in \mathbb{R}$ *mit* $P[X = a] = 1$.

Beweis. Es sei $F : \mathbb{R} \to [0,1]$, $F(t) = P(X \leq t)$ die Verteilungsfunktion von X. Weil $\{X \leq t\} \in \mathcal{F}$ für $t \in \mathbb{R}$, nimmt nach Voraussetzung F nur die Werte 0 und 1 an.
Nun wissen wir: $F(t) \to 0$ für $t \to -\infty$ und $F(t) \to 1$ für $t \to +\infty$. Es folgt : $F(t) = 0$ für alle genügend kleinen t, und $F(t) = 1$ für alle genügend großen t. Weil F monoton steigt, besitzt F genau eine Sprungstelle $a = \sup\{t \in \mathbb{R} | F(t) = 0\} = \inf\{t \in \mathbb{R} | F(t) = 1\}$ bei der der Wert von F von 0 auf 1 springt. Es folgt:

$$P[X = a] = \lim_{t \searrow a} F(t) - \lim_{s \nearrow a} F(s) = 1$$

□

Bemerkung. Die Aussage des Lemmas gilt auch wenn X Werte in $\mathbb{R} \cup \{\pm\infty\}$ annimmt.

Beispiel 3.1.5. Sind $(X_n)_{n\in\mathbb{N}}$ unabhängige Zufallsvariablen, so sind folgende Zufallsvariablen $P-f.s.$ konstant:

1) $\limsup\limits_{n\to\infty} X_n, \quad \liminf\limits_{n\to\infty} X_n$

2) $\limsup\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{n} X_k, \quad \liminf\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{n} X_k$

All diese Zufallsvariablen sind nämlich bzgl. der terminalen σ-Algebra messbar. Außerdem haben folgende Ereignisse entweder die Wahrscheinlichkeit 0 oder 1:

$$\{\lim_{n\to\infty} X_n \ existiert\}, \quad \{\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k \ existiert\}, \quad \{\lim_{n\to\infty} \sum_{k=1}^{n} \ existiert\}.$$

All diese Ereignisse sind nämlich terminal.

Satz 3.1.6 (Starkes Gesetz der großen Zahlen)**.** *Es seinen $(X_n)_{n\in\mathbb{N}}$ i.i.d. Zufallsvariablen in $\mathcal{L}^P(\Omega, \mathcal{A}, P)$, $p \geq 1$. Dann gilt $P-f.s.$ und in L^P:*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{n\to\infty} E[X_1].$$

Beweis. Nach dem eben Gezeigten und Satz 2.7.9 wissen wir: $Y := \lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{n} X_k$ existiert $P-f.s.$ und in L^P und ist $P-f.s.$ konstant. Es folgt:

$$E[X_1] \underset{(X_n)_{n\in\mathbb{N}} \ i.i.d}{\underset{\uparrow}{=}} E\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] \xrightarrow{n\to\infty} E[Y] = Y \ P-f.s.$$

also $Y = E[X_1] \ P-f.s.$ □

Beispiel 3.1.7 (Perkulationscluster). Wir stellen uns die Punkte $i \in \mathbb{Z}^2$ als Städte vor, die zufällig durch Telefonleitungen verbunden werden: $I = \{i, j \in \mathbb{Z}^2 | ||i-j||_2 = 1\}$ sei die Menge der möglichen Leitungen. Jede Leitung sei - unabhängig von den anderen - mit Wahrscheinlichkeit p durchlässig und $(1-p)$ unterbrochen. D bezeichne die Menge der durchlässigen Leitungen.
<u>Formal:</u> $X_e = \mathbb{1}_{\{e\in I\}}$ seien i.i.d $p\delta_1 + (1-p)\delta_0$- verteilte Zufallsvariablen. Zwei Städte $i, j \in \mathbb{Z}^2$ sind verbunden, in Zeichen $i \overset{D}{\longleftrightarrow} j$, wenn es einen Pfad $i = i_0, i_1, ..., i_n = j$ mit $\{i_{m-1}, i_m\} \in D$, für $m = 1, ..., n$ gilt. Offensichtlich ist $i \overset{D}{\longleftrightarrow} j$ eine Äquivalenzrelation. Ihre Äquivalenzklassen heißen <u>Perkulationscluster</u>. Es sei$C(i)$ der Perkulationscluster einer Stadt $i \in \mathbb{Z}^2$.

Frage. *Mit welcher Wahrscheinlichkeit gibt es einen unendlichen Perkulationscluster?*

Satz 3.1.8.
$$P\left[\exists i \in \mathbb{Z}^2 : |C(i)| = \infty\right] \in \{0,1\}$$

Beweisskizze. Wir zeigen $A = \{\exists i \in \mathbb{Z}^2 : |C(i)| = \infty\}$ ein terminales Ereignis bzgl. $(X_e)_{e\in I}$ ist. Die Behauptung folgt danach aus den 0-1-Gesetz von Kolmogoroff.
Es sei $E \subseteq I$ endlich. Es ist klar (ohne das hier formal zu beweisen), dass es einen unendlichen Cluster genau dann gibt, wenn es einen unendlichen Pfad $e_1, e_2, ...$ in D gibt. Pfade sollen hier stets aus verschiedenen Kanten bestehen. Nun gibt es einen unendlichen Pfad in D genau dann, wenn es einen unendlichen Pfad in $D \setminus E$ gibt. Dies ist wiederum äquivalent dazu, dass es ein $i \in \mathbb{Z}^2$ gibt, so dass es für jedes $n \in \mathbb{N}$ einen Pfad in $D \setminus E$ der Länge n mit Start in i gibt. Es gilt also:
$$A = \bigcup_{i\in\mathbb{Z}^2} \bigcap_{n\in\mathbb{N}} \bigcup_{\substack{i=i_1,...,i_n \\ Pfad\ in\ I\setminus E}} \bigcap_{m=1}^{n} \{X_{\{i_m,i_{m-1}\}} = 1\}. \quad A \in \sigma(X_e : e \in I \setminus E)$$
Weil für alle endlichen $E \subseteq I$ gilt, ist A ein terminales Ereignis. □

Bemerkung. Ein berühmter Satz von Kesten besagt $P[A] = 0$ für $p \leq \frac{1}{2}$ und $P[A] = 1$ für $p > \frac{1}{n}$.
Eine analoge Situation tritt oft auf:
Wenn ein 0-1-Gesetz besagt, dass eine Wahrscheinlichkeit 0 oder 1 ist, bleibt es im Allgemeinen eine schwierige Frage, zu entscheiden, welche der beiden Fälle eintritt.

3.2 Austauschbarkeit

Zunächst sei $\Omega = \mathbb{R}^{\mathbb{N}}$, versehen mit der σ-Algebra $\mathcal{A} = \sigma(Z_n : n \in \mathbb{N})$, die von den kanonischen Projektionen $Z_n : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, $(\omega_k)_{k\in\mathbb{N}} \mapsto \omega_n$ erzeugt wird.

Definition 3.2.1. *Ein Ereignis $A \in \mathcal{A}$ heißt <u>austauschbar</u>, wenn für jede Bijektion $\tau : \mathbb{N} \to \mathbb{N}$, die höchstens endlich viele Zahlen nicht fest lässt, gilt:*
$$\forall \omega : (\omega_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : (\omega_n)_{n\in\mathbb{N}} \in A \Leftrightarrow (\omega_{\tau(n)})_{n\in\mathbb{N}} \in A \tag{1}$$
Die Menge aller austauschbaren Ereignisse bildet offensichtlich eine σ-Algebra Z.

Beispiel 3.2.2. Das Ereignis $\{\sum_{k=1}^{n} Z_k \xrightarrow{n\to\infty} 0\}$ ist austauschbar, weil es bei der unendlichen Reihe nicht auf die Permutation eines Anfangsstücks ankommt.
Das Ereignis ist nicht terminal, weil der Wert der unendlichen Reihe sehr wohl von jedem Anfangsstück abhängt. Umgekehrt ist das terminale Ereignis austauschbar, denn wenn es auf jedes Anfangsstück nicht ankommt, kommt es auch nicht darauf an, ob ein Anfangsstück permutiert ist.

Satz 3.2.3 (0-1-Gesetz von Hewitt und Savage)**.** *Es sei P ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, sodass $(Z_n)_{n\in\mathbb{N}}$ i.i.d sind. Dann gilt für alle austauschbaren Ereignisse A: $P[A] = 0 \vee P[A] = 1$.*

Beweis. Es sei A ein austauschbares Ereignis. Es sei $\mathcal{F}_n := \sigma(Z_n : k = 1, ..., n)$, $n \in \mathbb{N}$, die von den Z_n erzeugte Filtration. Dann ist $X_n = P[A|\mathcal{F}_n]$, $n \in \mathbb{N}$ ein Martingal mit Werten $P - f.s.$ in $[0,1]$. Weil $A \in \sigma(\bigcup_{n\in\mathbb{N}} A_n) =: \mathcal{F}_\infty$ ist $P[A|\mathcal{F}_\infty] = \mathbb{1}_A\ P - f.s.$
Nach Lemma 2.7.6 konvergiert X_n für $n \to \infty$ $P - f.s.$ und in L^1 gegen $P[A|\mathcal{F}_\infty] = \mathbb{1}_A$. Gegeben $n \in \mathbb{N}$, betrachten wir die Bijektion
$$\tau_n : \mathbb{N} \to \mathbb{N}, \tau_n(k) = \begin{cases} k+n & \text{für } k = 1, ..., n \\ k-n & \text{für } k = n+1, ..., z_n \\ k & sonst\ (k > z_n) \end{cases}$$

τ_n vertauscht also $\{1, ..., n\}$ mit $\{n+1, ..., z_n\}$.
Wir setzen $\tau_n : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$, $\tau_n(X_m)_{m\in\mathbb{N}} = (X_{\tau(m)})_{m\in\mathbb{N}}$.
Weil die $(Z_n)_{n\in\mathbb{N}}$ i.i.d sind, stimmt das Bildmaß $\pi_n P$ von P mit P überein, denn es stimmt auf dem $\cap$-stabilen Erzeugendensystem von $\mathcal{A}$, der von den Ereignissen $\{Z_k \in \mathcal{B}_k$ für $k = 1, ..., m\}$ mit $m \in \mathbb{N}$, $\mathcal{B}_1, ..., \mathcal{B}_m \in \mathcal{B}(\mathbb{R})$, gebildet und mit P überein. Weil A ein austauschbares Ereignis ist, folgt $\mathbb{1}_A = \mathbb{1}_A \circ \pi_n$. Es folgt mit der Abkürzung $Y_n = X_n \circ \pi_n$:

$$||Y_n - \mathbb{1}_A||_1 = ||X_n \circ \pi_n - \mathbb{1}_A \circ - \mathbb{1}_A \pi_n||_1 \overset{\substack{\pi_n P = P\\ \downarrow}}{=} ||X_n - \mathbb{1}_A||_1.$$

Damit wissen wir sowohl $X_n \xrightarrow{n\to\infty} \mathbb{1}_A$ in L^1 als auch $Y_n \xrightarrow{n\to\infty} \mathbb{1}_A$ in L^1. Weil $X_n, Y_n \in [0,1]$ $P - f.s.$, schließen wir hieraus:

$$X_n Y_n \xrightarrow{n\to\infty} \underbrace{\mathbb{1}_A \mathbb{1}_A}_{=\mathbb{1}_A} \ in\ L^1.$$

In der Tat:

$$||X_nY_n - \mathbb{1}_A||_1 = ||X_nY_n - \mathbb{1}_A\mathbb{1}_A||_1 \le ||(X_n - \mathbb{1}_A)Y_n||_1 + ||\mathbb{1}_A(Y_n - \mathbb{1}_A||_1 \le ||X_n - \mathbb{1}_A||_1 + ||Y_n - \mathbb{1}_A||_1 \xrightarrow{n\to\infty} 0.$$

Nun ist X_n bzgl. $\mathcal{F}_n = \sigma(Z_1, ..., Z_n)$ messbar und $Y_n = X_n \circ \pi_n$ bzgl. $\sigma(Z_{n+1}, ..., Z_{2n})$ messbar. Insbesondere sind X_n, Y_n undabhängig (da die erzeugenden σ-Algebren unabhängig sind). Es folgt:

$$P[A]^2 = E[\mathbb{1}_A]^2 = \lim_{n\to\infty} E[X_n]E[Y_n] \overset{\substack{\text{Unabhängigkeit}\\ \downarrow}}{=} \lim_{n\to\infty} E[X_nY_n] = E[\mathbb{1}_A] = P[A],\ also P[A] \in \{0,1\}$$

□

In der Statistik hat man bei der Modellbildung oft die Situation, dass man nichts Genaues über Abhängigkeiten zwischen den Elementen einer Stichprobe $X_1, ..., X_n$ weiß. Insbesondere wäre eine Unabhängigkeitsannahme oft nicht gut gerechtfertigt. Die Modellannahme, dass alle $X_1, ..., X_n$ „gleichberechtigt“, also „austauschbar“ sind, erscheint oft plausibler:

Definition 3.2.4. *Zufallsvariablen $X_1, ..., X_n$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ heißen austausch- bar, wenn für alle Permutationen σ von $\{1, ..., n\}$ der Zufallsvektor $(X_{\sigma(1)}, ..., X_{\sigma(n)})$ die gleiche gemeinsame Verteilung wie $(X_1, ..., X_n)$ hat. Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$ heißt austauschbar, wenn $X_1, ...X_n$ für alle $n \in \mathbb{N}$ austauschbar sind.*

Bei austauschbaren Zufallsvariablen steckt die gesamte „relevante Information“ bereits in der „empirischen Verteilung“, die wie folgt definiert ist:

Definition 3.2.5. *Es seien $(X_1, ..., X_n)$ Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$ mit Werten in $(\Omega', \mathcal{A}')$. Die empirische Verteilung ist das zufällige Wahrscheinlichkeitsmaß $L_n : \Omega \to \mathcal{M}_1(\Omega', \mathcal{A}') = \{Q : \mathcal{A}' \to [0,1] |\ Q$ ist Wahrscheinlichkeitsmaß auf $(\Omega', \mathcal{A}')\}$, $L_n(\omega) = \frac{1}{n}\sum_{k=1}^{n} \delta_{X_k(\omega)}$ Für jedes $A \in \mathcal{A}'$ ist also $L_n(\omega)(A) = \frac{1}{n}\sum_{k=1}^{n} \delta_{X_k(\omega)}(A)$ die relative Häufigkeit von A unter $X_1(\omega), ..., X_n(\omega)$. Ist allgemeiner $m \le n$ und bezeichne $I_{n,m}$ die Menge aller injektiven Abbildungen von $\{1, ..., m\}$ nach $\{1, ..., n\}$, so nennen wir das zufällige Maß $L_{n,m} : \mathcal{M}_1((\Omega')^m, (\mathcal{A})^{\otimes m})$, $L_{n,m}(\omega) = \frac{1}{|I_{n,m}|}\sum_{b\in I_{n,m}} \delta_{X_{\sigma(1)}(\omega), ..., X_{\sigma(m)}(\omega)})$ die m-fache (gemeinsame) empirische Verteilung von $X_1, ..., X_m$.*

Die $L_{m,n}$ enthalten - ähnlich wie die Ordungsstatistik - alle Informationen über $X_1, ..., X_m$, wenn man die Reihenfolge der $X_1, ..., X_m$ ignoriert. Wir setzen für eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen:

$$\mathcal{F}_{-n} := \sigma(L_{k,m} : k \ge n, m \le k) = \sigma(L_{k,m}(\cdot)(A) : k \ge n, m \le k, A \in \mathcal{A}'^{\otimes m})\ und\ \mathcal{F}_{-\infty} := \bigcap_{n\in\mathbb{N}} \mathcal{F}_{-n}.$$

$(\mathcal{F}_{-n})_{-n\in-\mathbb{N}\cup\{-\infty\}}$ ist eine Rückwärtsfiltration; $\mathcal{F}_{-n}$ enthält die gesamte Information über die $(X_m)_{m\in\mathbb{N}}$, die verbleibt, wenn man die Reihenfolge von $X_1, ..., X_n$ ignoriert.