Satz 6.6. *$X, Y \in \mathcal{L}^1(\mathbb{P})$, $\mathcal{G} \subset \mathcal{F}$ Teil-σ-Algebra.*

i) $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$ *f.s. für* $a, b \in \mathbb{R}$ *(Linearität)*

ii) $X \leq Y$ *f.s.* $\Rightarrow$ $\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}]$ *f.s.* *(Monotonie)*

iii) $\left|\mathbb{E}[X \mid \mathcal{G}]\right| \leq \mathbb{E}\left[|X| \middle| \mathcal{G}\right]$ *f.s.* *(Dreiecksungleichung)*

iv) Es gelte $\mathbb{E}\left[|XY|\right] < \infty$ *und* Y *sei* $\mathcal{G}$*-messbar, dann ist*

$$\mathbb{E}[XY \mid \mathcal{G}] = Y \cdot \mathbb{E}[X \mid \mathcal{G}] \ \textit{f.s.},$$

insbesondere ist $\mathbb{E}[Y|\mathcal{G}] = Y$ *f.s.* *(„Herausziehen von Bekanntem“)*

v) Sind $\sigma(X)$ *und* $\mathcal{G}$ *unabhängig, so ist* $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ *f.s.* *(Verhalten bei Unabhängigkeit)*

vi) $\mathcal{G}' \subset \mathcal{G}$ *Teil-σ-Algebra, so ist*

$$\mathbb{E}\left[\mathbb{E}[X \mid \mathcal{G}] \middle| \mathcal{G}'\right] = \mathbb{E}[X \mid \mathcal{G}'] \ \textit{f.s.},$$

(„Turmeigenschaft“) insbesondere ist

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$$

Satz 6.6 (Fortsetzung). *$X, Y \in \mathcal{L}^1(\mathbb{P})$, $\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra.*

vii) $0 \leq X_n \nearrow X$ f.s. für $n \to \infty$, so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \underset{n\to\infty}{\nearrow} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

(monotone Konvergenz)

viii) X_n reelle ZVn mit $|X_n| \leq Y \ \forall\, n$ und $X_n \to X$ f.s., so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \xrightarrow[n\to\infty]{} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

(dominierte Konvergenz)

v) $\sigma(X)$ und $\mathscr{G}$ unabhängig $\Rightarrow$ so ist $\mathbb{E}[X \mid \mathscr{G}] = \mathbb{E}[X]$ f.s.

X, Y u.a. ZVn, so gilt $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

denn: • $\mathbb{E}[X]$ ist $\mathscr{G}$-messbar ✓

• sei $A \in \mathscr{G}$, $\mathbb{E}[\mathbb{1}_A \cdot \mathbb{E}[X]] = \mathbb{E}[\mathbb{1}_A] \cdot \mathbb{E}[X]$

$\overset{\text{Unabhängigk.}}{=} \mathbb{E}[\mathbb{1}_A \cdot X]$ ✓

vi) $\mathscr{G}' \subset \mathscr{G}$ Teil-σ-Algebra, so ist $\mathbb{E}\big[\mathbb{E}[X \mid \mathscr{G}] \big| \mathscr{G}'\big] = \mathbb{E}[X \mid \mathscr{G}']$ f.s.

• $\mathbb{E}[X \mid \mathscr{G}']$ ist $\mathscr{G}'$-messbar (n. Def.) ✓

• sei $A \in \mathscr{G}' (\subset \mathscr{G})$

$$\underbrace{\mathbb{E}[\mathbb{1}_A \cdot \mathbb{E}[X \mid \mathscr{G}]]}_{\overset{\shortparallel}{\mathbb{E}[\mathbb{1}_A \cdot X]}} \overset{!}{=} ✓ \underbrace{\mathbb{E}[\mathbb{1}_A \cdot \mathbb{E}[X \mid \mathscr{G}']]}_{\mathbb{E}[\mathbb{1}_A \cdot X]}$$

Spez. für $\mathscr{G}' = \{\emptyset, \Omega\}$ ist $\mathbb{E}[X \mid \mathscr{G}']$ f.s. konst.

vii) $0 \leq X_n \nearrow X$ f.s. für $n \to \infty$, so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \underset{n\to\infty}{\nearrow} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

$\mathbb{E}[X_n | \mathcal{G}] \leq \mathbb{E}[X_{n+1} | \mathcal{G}]$ (f.s.), insbes. konvergiert $\mathbb{E}[X_n|\mathcal{G}]$

$$\mathbb{E}\Big[\big|\mathbb{E}[X|\mathcal{G}] - \mathbb{E}[X_n|\mathcal{G}]\big|\Big] \leq \mathbb{E}\big[\mathbb{E}[\,|X-X_n|\,|\mathcal{G}]\big] = \mathbb{E}[|X-X_n|] \underset{n\to\infty}{\longrightarrow} 0$$

(Satz v.d. monotonen Konv.)

zudem:
für monotone Folgen von ZVn impliziert $\mathcal{L}^1$-Konv. auch f.s.-Konv.:

Seien $Z_n \nearrow Z$ und $Z_n \xrightarrow[n\to\infty]{\mathcal{L}^1} Z$.

$\Rightarrow \quad Z_n \xrightarrow[n\to\infty]{\mathbb{P}} Z$, d.h. $\forall \varepsilon > 0$: $\mathbb{P}(\underbrace{|Z - Z_n| \leq \varepsilon}_{\overset{\text{Monotonie d. } Z_n}{=} \bigcap_{m=n}^{\infty} \{|Z_m - Z| \leq \varepsilon\}}) \underset{n\to\infty}{\longrightarrow} 1$,

d.h. $\mathbb{P}\left(\bigcup_n \bigcap_{m\geq n} \{|Z - Z_m| \leq \varepsilon\}\right) = 1$.

viii) X_n reelle ZVn mit $|X_n| \leq Y \ \forall\, n$ und $X_n \to X$ f.s., so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \xrightarrow[n\to\infty]{} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

$$Z_n := \sup_{m \geq n} |X_m - X| \leq 2Y, \quad Z_n \searrow_{n\to\infty} 0 \text{ f.s.}$$

$$\mathbb{E}\Big[\big|\underbrace{\mathbb{E}[X|\mathcal{G}] - \mathbb{E}[X_n|\mathcal{G}]}_{=\mathbb{E}[X-X_n|\mathcal{G}]}\big|\Big] \leq \mathbb{E}\big[\mathbb{E}[|X-X_n||\mathcal{G}]\big] = \mathbb{E}[|X-X_n|] \xrightarrow[n\to\infty]{} 0 \text{ (dom. Konv.)}$$

Zur f.s.-Konv.:

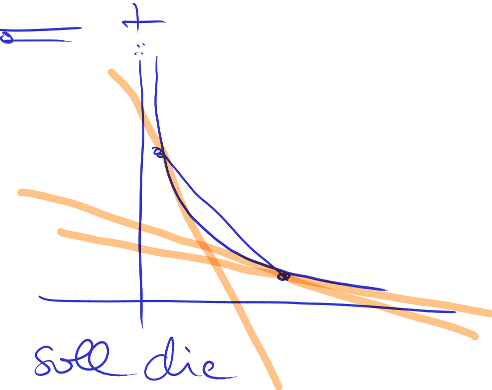
$$\mathbb{E}\Big[\lim_{n\to\infty} \mathbb{E}[Z_n|\mathcal{G}]\Big] \overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \underbrace{\mathbb{E}\big[\mathbb{E}[Z_n|\mathcal{G}]\big]}_{=\mathbb{E}[Z_n]} \underset{\text{mon. Konv.}}{=} 0$$

$$\Rightarrow \mathbb{E}[Z_n|\mathcal{G}] \xrightarrow[n\to\infty]{} 0 \text{ f.s.}$$

$$\big|\mathbb{E}[X|\mathcal{G}] - \mathbb{E}[X_n|\mathcal{G}]\big| \leq \mathbb{E}[\underbrace{|X-X_n|}_{\leq Z_n}|\mathcal{G}] \leq \mathbb{E}[Z_n|\mathcal{G}] \xrightarrow[n\to\infty]{\text{f.s.}} 0$$

Satz 6.7 (Jensen'sche Ungleichung für die bedingte Erwartung)**.** $X \in \mathcal{L}^1(\mathbb{P})$, $k : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *konvex, dann gilt*

$$(\ast) \qquad \mathbb{E}\big[k(X)\,\big|\,\mathscr{G}\big] \geq k\big(\mathbb{E}[X \mid \mathscr{G}]\big) \quad \textit{f.s.}$$



beachte: $(k(X))^- \in \mathcal{L}^1(\mathbb{P})$, $(k(X))^+$ muss nicht in $\mathcal{L}^1(\mathbb{P})$ liegen, dann soll die Ungl. als $\infty \geq \ldots$ gelesen werden

(Ungl. ist f.s. wohldefiniert)

Erinnerung: Zu $a \in \mathbb{R}$ mit $k(a) < \infty$ gibt $m (= m(a))$ mit

$$k(z) \geq k(a) + m(z-a) \quad \forall z \in \mathbb{R}$$

Falls k affin-linear:
$(\ast)$ gilt mit „=" ✓

Es gilt $\quad k(x) = \sup\{ax+b : (a,b) \in S \cap \mathbb{Q}^2\}$

mit $\quad S = \{(a,b) \in \mathbb{R}^2 : \; ax+b \leq k(x) \;\; \forall x \in \mathbb{R}$

Sei $(a_n, b_n)_{n \in \mathbb{N}}$ eine Aufzählung von $S \cap \mathbb{Q}^2$:

$$\mathbb{E}[\underbrace{k(X)}_{\geq a_n X + b_n} \mid \mathscr{G}] \underset{(f.s.)}{\geq} \sup_{n \in \mathbb{N}}(a_n \mathbb{E}[X \mid \mathscr{G}] + b_n) = k(\mathbb{E}[X \mid \mathscr{G}])$$

$\lrcorner$

Bemerkung 6.8. X, Y reelle ZVn mit gemeinsamer Dichte $f_{X,Y}$, d.h.

$$\mathbb{P}\big((X,Y)\in A\big) = \int_A f_{X,Y}(x,y)\,\lambda^{\otimes 2}\big(d(x,y)\big) \quad \text{für } A\in\mathcal{B}(\mathbb{R}^2),$$

$Y\in\mathcal{L}^1(\mathbb{P})$.

Sei für $x\in\mathbb{R}$

$$f_X(x) := \int_{\mathbb{R}} f_{X,Y}(x,y)\,\lambda(dy) \quad \text{die Marginaldichte von } X,$$

$$\varphi(x) := \int_{\mathbb{R}} y\frac{f_{X,Y}(x,y)}{f_X(x)}\,\lambda(dy)\mathbf{1}_{f_X(x)>0}.$$

$\mathbb{E}[\mathbb{1}_B(X)\cdot Y]$

$\shortparallel$

$\int_{\mathbb{R}^2} \mathbb{1}_B(x)\cdot y\cdot f_{X,Y}(x,y)\,\lambda^{\otimes 2}(d(x,y))$

$\shortparallel$

Dann gilt

oft auch knapp notiert als $\mathbb{E}[Y|X]$ $\longrightarrow$

$$\mathbb{E}[Y|\sigma(X)] = \varphi(X) \ \text{ f.s.}$$

$= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_B(x)\,y\,f_{X,Y}(x,y)\,\lambda(dy)\,\lambda(dx)$

- $\varphi(X)$ ist $\sigma(X)$-m.b. ✓

- $B\in\mathcal{B}(\mathbb{R})$. $\mathbb{E}[\mathbb{1}_{\{X\in B\}}\varphi(X)] = \int_{\mathbb{R}} \mathbb{1}_B(x)\varphi(x) f_X(x)\,\lambda(dx)$

$= \int_{\mathbb{R}} \mathbb{1}_B(x)\int_{\mathbb{R}} y\,\frac{f_{X,Y}(x,y)}{\cancel{f_X(x)}}\,\lambda(dy)\,\mathbb{1}_{f_X(x)>0}\cdot \cancel{f_X(x)}\,\lambda(dx)$

Bericht 6.9. (Zu regulären Versionen bedingter Verteilungen) Wenn $Y = 1_B$ für ein Ereignis $B \in \mathscr{A}$, so schreibt man gelegentlich auch

$$\mathbb{P}(B \mid \mathscr{G}) = \mathbb{E}[Y \mid \mathscr{G}]$$

Man muss allerdings etwas vorsichtig sein bei der Interpretation von $\mathbb{P}(\cdot \mid \mathscr{G})$ als ein (zufälliges) Maß, da i.A. überabzählbar viele B in Frage kommen und damit die Kompatibilität der in der Definition der bedingten Erwartung implizit vorkommenden Nullmengen (vgl. Def. 6.3) wenigstens a priori unklar bleibt.

In „gutartigen" Fällen ist eine konsistente Wahl möglich, das Stichwort dazu lautet „reguläre bedingte Verteilung einer Zufallsvariable".

Skizze für den reellwertigen Fall:
$\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra.
Dann gibt es einen stochastischen Kern $\kappa_{X|\mathscr{G}}$ von $(\Omega, \mathscr{G})$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit

$$\kappa_{X|\mathscr{G}}(\omega, B) = \mathbb{E}\Big[\mathbf{1}_{\{X \in B\}} \Big| \mathscr{G}\Big](\omega) \ \text{ f.s. für alle } B \in \mathcal{B}(\mathbb{R}),$$

d.h. (vgl. Def. 5.4)

> für jedes $B \in \mathcal{B}(\mathbb{R})$ ist $\omega \mapsto \kappa_{X|\mathscr{G}}(\omega, B)$ eine Version von $\mathbb{E}\Big[\mathbf{1}_{\{X \in B\}} \Big| \mathscr{G}\Big]$ und
>
> für jedes ω ist $\kappa_{X|\mathscr{G}}(\omega, \cdot)$ ein W'maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Skizze für den reellwertigen Fall:

$\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra. Es gibt einen stochastischen Kern $\kappa_{X|\mathscr{G}}$ von $(\Omega, \mathscr{G})$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit

$$\kappa_{X|\mathscr{G}}(\omega, B) = \mathbb{E}\Big[\mathbf{1}_{\{X\in B\}}\Big|\mathscr{G}\Big](\omega) \text{ f.s. für alle } B \in \mathcal{B}(\mathbb{R}),$$

Hauptidee: charakterisiere das gewünschte Maß $\kappa_{X|\mathscr{G}}(\omega, \cdot)$ auf $\mathbb{R}$ anhand seiner Verteilungsfunktion (vgl. Satz 1.27); die zielführende Beobachtung ist dann, dass eine Verteilungsfunktion (wegen der Monotonie) bereits durch ihre Werte an abzählbar vielen Stellen festgelegt ist.

Man betrachtet also $B = (-\infty, r], r \in \mathbb{Q}$ und setzt

$$F_r := \mathbb{E}[\mathbf{1}_{(-\infty,r]}(X) \mid \mathscr{G}].$$

Dann gilt (mit den Eigenschaften der bedingten Erwartung aus Satz 6.6) wie gewünscht $\mathbb{P}$-f.s.:
$F_r \le F_{r'}$, für $r < r'$, $(r, r' \in \mathbb{Q})$, $\lim_{n\to\infty} F_{r+\frac{1}{n}} = F_r$ für $r \in \mathbb{Q}$, $\lim_{n\to\infty} F_n = 1$, $\lim_{n\to-\infty} F_n = 0$.

Wegen der Abzählbarkeit von $\mathbb{Q}$ gibt es $N \in \mathscr{F}$ mit $\mathbb{P}(N) = 0$, so dass obiges für $\omega \in \Omega \setminus N$ und alle $r, r' \in \mathbb{Q}$ gilt. Dann definiert

$$\widetilde{F}_s := \begin{cases} \inf\{F_r : r \ge s, r \in \mathbb{Q}\}, & \omega \in \Omega \setminus N, \\ \overline{F}_s, & \omega \in N, \end{cases}$$

wobei $\overline{F}$ irgendeine Verteilungsfunktion ist, die (zufällige) Verteilungsfunktion von $k_{X,\mathscr{G}}$. Details finden sich beispielsweise in [Kl, Kap. 8.3, speziell Satz 8.28].

Bericht 6.9 (Fortsetzung)**.** Man kann dieses Argument relativ leicht erweitern auf die Situation, dass der Wertebereich $(E, \mathscr{B})$ von X ein sogenannter Standard-Borel-Raum ist (auch der Name Borel'scher Raum ist üblich), d.h. wenn es ein $A \in \mathcal{B}(\mathbb{R})$ und eine Bijektion $\phi : E \to A$ gibt, so dass ϕ und ϕ^{-1} jeweils messbar sind (dann sind $(E, \mathscr{B})$ und $(A, \mathcal{B}(A))$ isomorph als messbare Räume). Dann ist nämlich $X' := \phi \circ X$ eine reellwertige ZV, und die Argumentation oben greift (vgl. auch [Kl, Satz 8.36]).

Schließlich kann man zeigen, dass jeder separable und vollständige metrische Raum E, versehen mit seiner Borel-σ-Algebra, ein Standard-Borel-Raum ist (siehe z.B. L.C.G. Rogers, D. Williams, *Diffusions, Markov processes and martingales*, Bd. 1, Ch. II.82; L. Breiman, *Probability*, Appendix 7). Solche Wertebereiche heißen *polnische Räume*, sie spielen in der allgemeinen Theorie der Stochastik eine wichtige Rolle (beispielsweise sind $\mathbb{R}^d$ oder $C([0, 1])$ mit Supremumsnorm polnisch).