Chapter V

Conditional Expectations and Martingales

1 Conditional Expectations

'Access to the martingale concept is afforded by one of the truely basic ideas of probability theory, that of conditional expectation.', see Bauer (1996, p. 109).

Soweit nichts anderes gesagt, betrachten wir die Borelsche σ-Algebra $\mathfrak{B}$ auf $\mathbb{R}$; $\mathfrak{G}$-$\mathfrak{B}$-meßbare Abbildungen werden kurz $\mathfrak{G}$-meßbar genannt.

Erinnerung: *Elementare bedingte Wahrscheinlichkeit*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad A, B \in \mathfrak{A}, \quad P(B) > 0.$$

Das Wahrscheinlichkeitsmaß $P(\cdot \mid B)$ besitzt die P-Dichte $1/P(B) \cdot 1_B$.

Gegeben: Zufallsvariable X auf $(\Omega, \mathfrak{A}, P)$ mit $X \in \mathfrak{L}^1$. *Elementare bedingte Erwartung*

$$\mathrm{E}(X \mid B) = \frac{1}{P(B)} \cdot \mathrm{E}(1_B \cdot X) = \int X \, dP(\cdot \mid B).$$

Klar für $A \in \mathfrak{A}$

$$\mathrm{E}(1_A \mid B) = P(A \mid B).$$

Nun etwas allgemeiner: Sei I höchstens abzählbar. Betrachte Partititon

$$B_i \in \mathfrak{A}, \qquad i \in I,$$

von Ω mit

$$\forall\, i \in I: \quad P(B_i) > 0.$$

Durch

$$\mathfrak{G} = \left\{ \bigcup_{j \in J} B_j : J \subset I \right\}$$

ist eine σ-Algebra $\mathfrak{G} \subset \mathfrak{A}$ gegeben. In der Tat gilt $\mathfrak{G} = \sigma(\{B_i : i \in I\})$. Definiere Zufallsvariable $\mathrm{E}(X \mid \mathfrak{G})$ durch

$$\mathrm{E}(X \mid \mathfrak{G})(\omega) = \sum_{i \in I} \mathrm{E}(X \mid B_i) \cdot 1_{B_i}(\omega), \qquad \omega \in \Omega. \tag{1}$$

Dann ist $\mathrm{E}(X \mid \mathfrak{G})$ $\mathfrak{G}$-meßbar und gehört zu $\mathfrak{L}^1$. Ferner

$$\int_{B_j} \mathrm{E}(X \mid \mathfrak{G})\, dP = \mathrm{E}(X \mid B_j) \cdot P(B_j) = \int_{B_j} X\, dP,$$

und somit für jedes $G \in \mathfrak{G}$

$$\int_G \mathrm{E}(X \mid \mathfrak{G})\, dP = \int_G X\, dP.$$

Der Übergang von X zu $\mathrm{E}(X \mid \mathfrak{G})$ ist eine 'Vergröberung'. Man vergleiche diese Konstruktion mit dem Beweis der ersten Variante des Satzes von Radon und Nikodym.

Example 1. Extremfälle. Einerseits $|I| = 1$, also $\mathfrak{G} = \{\emptyset, \Omega\}$ und

$$\mathrm{E}(X \mid \mathfrak{G}) = \mathrm{E}(X).$$

Andererseits Ω höchstens abzählbar und $I = \Omega$ mit $B_i = \{i\}$, also $\mathfrak{G} = \mathfrak{P}(\Omega)$ und

$$\mathrm{E}(X \mid \mathfrak{G}) = X.$$

Allgemein: Gegeben: Zufallsvariable X auf $(\Omega, \mathfrak{A}, P)$ mit $X \in \mathfrak{L}^1$ und σ-Algebra $\mathfrak{G} \subset \mathfrak{A}$.

Definition 1. Jede Zufallsvariable Z auf $(\Omega, \mathfrak{A}, P)$ mit $Z \in \mathfrak{L}^1$ sowie

(i) Z ist $\mathfrak{G}$-meßbar,

(ii) $\forall\, G \in \mathfrak{G} : \int_G Z\, dP = \int_G X\, dP$

heißt *(Version der) bedingte(n) Erwartung von X gegeben* $\mathfrak{G}$. Bez.: $Z = \mathrm{E}(X \mid \mathfrak{G})$.
Im Falle $X = 1_A$ mit $A \in \mathfrak{A}$ heißt Z *(Version der) bedingte(n) Wahrscheinlichkeit von A gegeben* $\mathfrak{G}$. Bez.: $Z = P(A \mid \mathfrak{G})$.

Theorem 1. Die bedingte Erwartung existiert und ist $P|_{\mathfrak{G}}$-f.s. eindeutig bestimmt.

Proof. Spezialfall: $X \geq 0$. Durch

$$Q(G) = \int_G X\, dP, \qquad G \in \mathfrak{G},$$

wird gemäß Theorem II.7.1 ein endliches Maß auf $(\Omega, \mathfrak{G})$ definiert. Es gilt $Q \ll P|_{\mathfrak{G}}$. Nach dem Satz von Radon-Nikodym existiert eine $\mathfrak{G}$-meßbare Abbildung $Z : \Omega \to [0, \infty[$ mit

$$\forall\, G \in \mathfrak{G} : \quad Q(G) = \int_G Z\, dP.$$

Der allgemeine Fall wird durch Zerlegung in Positiv- und Negativteil erledigt.
Zur Eindeutigkeit: Aus $\int_G Z_1\, dP|_{\mathfrak{G}} = \int_G Z_2\, dP|_{\mathfrak{G}}$ für alle $G \in \mathfrak{G}$ folgt $Z_1 = Z_2$ $P|_{\mathfrak{G}}$-f.s. Siehe Beweis von Theorem II.7.3. $\square$

Im folgenden oft kurz $X = Y$ oder $X \geq Y$, falls diese Eigenschaften f.s. gelten. Ebenso identifizieren wir Abbildungen, die f.s. übereinstimmen.

Die 'explizite' Bestimmung von bedingten Erwartungen ist i.a. nicht-trivial.

Remark 1. Klar: $\mathrm{E}(X \,|\, \mathfrak{G}) = X$, falls X $\mathfrak{G}$-meßbar. Ferner:

$$\mathrm{E}(X) = \int_\Omega \mathrm{E}(X \,|\, \mathfrak{G})\, dP = \mathrm{E}(\mathrm{E}(X \,|\, \mathfrak{G})).$$

Im Spezialfall (1) für $X = 1_A$ mit $A \in \mathfrak{A}$ ist dies die Formel von der totalen Wahrscheinlichkeit, also

$$P(A) = \sum_{i \in I} P(A \,|\, B_i) \cdot P(B_i).$$

Lemma 1. Die bedingte Erwartung ist positiv und linear.

Proof. Folgt aus den entsprechenden Eigenschaften des Integrals und den Abschlußeigenschaften von Mengen meßbarer Abbildungen. □

Lemma 2. Sei Y $\mathfrak{G}$-meßbar mit $X \cdot Y \in \mathfrak{L}^1$. Dann

$$\mathrm{E}(X \cdot Y \,|\, \mathfrak{G}) = Y \cdot \mathrm{E}(X \,|\, \mathfrak{G}).$$

Proof. Klar: $Y \cdot \mathrm{E}(X \,|\, \mathfrak{G})$ ist $\mathfrak{G}$-meßbar.

Spezialfall: $Y = 1_C$ mit $C \in \mathfrak{G}$. Dann gilt für $G \in \mathfrak{G}$:

$$\int_G Y \cdot \mathrm{E}(X \,|\, \mathfrak{G})\, dP = \int_{G \cap C} \mathrm{E}(X \,|\, \mathfrak{G})\, dP = \int_{G \cap C} X\, dP = \int_G X \cdot Y\, dP.$$

Jetzt algebraische Induktion. □

Lemma 3. Für σ-Algebren $\mathfrak{G}_1 \subset \mathfrak{G}_2 \subset \mathfrak{A}$ gilt

$$\mathrm{E}(\mathrm{E}(X \,|\, \mathfrak{G}_1) \,|\, \mathfrak{G}_2) = \mathrm{E}(X \,|\, \mathfrak{G}_1) = \mathrm{E}(\mathrm{E}(X \,|\, \mathfrak{G}_2) \,|\, \mathfrak{G}_1).$$

Proof. Die erste Identität folgt mit Remark 1. Zur zweiten Identität beachte man für $G \in \mathfrak{G}_1 \subset \mathfrak{G}_2$

$$\int_G \mathrm{E}(\mathrm{E}(X \,|\, \mathfrak{G}_2) \,|\, \mathfrak{G}_1)\, dP = \int_G \mathrm{E}(X \,|\, \mathfrak{G}_2)\, dP = \int_G X\, dP.$$

□

Terminologie: X und $\mathfrak{G}$ heißen unabhängig, falls $(\sigma(X), \mathfrak{G})$ unabhängig ist.

Lemma 4. Seien X und $\mathfrak{G}$ unabhängig. Dann

$$\mathrm{E}(X \,|\, \mathfrak{G}) = \mathrm{E}(X).$$

Proof. Klar: $\mathrm{E}(X)$ $\mathfrak{G}$-meßbar. Sei $G \in \mathfrak{G}$. Nach Voraussetzung sind X und 1_G unabhängig. Also

$$\int_G \mathrm{E}(X)\, dP = \mathrm{E}(X) \cdot \mathrm{E}(1_G) = \mathrm{E}(X \cdot 1_G) = \int_G X\, dP.$$

□

Theorem 2 (Jensensche Ungleichung)**.** Sei $J \subset \mathbb{R}$ ein Intervall, und gelte $X(\omega) \in J$ für alle $\omega \in \Omega$. Sei $\varphi : J \to \mathbb{R}$ konvex mit $\varphi \circ X \in \mathfrak{L}^1$. Dann:

$$\varphi \circ \mathrm{E}(X \,|\, \mathfrak{G}) \leq \mathrm{E}(\varphi \circ X \,|\, \mathfrak{G}).$$

Proof. Gänssler, Stute (1977, Kap. V.4). □

Remark 2. Spezialfall: $J = \mathbb{R}$ und $\varphi(u) = |u|^{p/q}$ mit $1 \leq q \leq p$. Dann:

$$(\mathrm{E}(|X|^q \,|\, \mathfrak{G}))^{1/q} \leq (\mathrm{E}(|X|^p \,|\, \mathfrak{G}))^{1/p}$$

für $X \in \mathfrak{L}^p$ sowie

$$\left(\int_\Omega |\, \mathrm{E}(X \,|\, \mathfrak{G})|^p\, dP \right)^{1/p} \leq \left(\int_\Omega |X|^p\, dP \right)^{1/p}. \tag{2}$$

Also ist $\mathrm{E}(\cdot \,|\, \mathfrak{G})$ ein idempotenter beschränkter linearer Operator auf $L^p(\Omega, \mathfrak{A}, P)$ mit Norm 1. Speziell für $p = 2$: orthogonale Projektion auf den Unterraum $L^2(\Omega, \mathfrak{G}, P)$.

Oft liegt folgende Situation vor. Gegeben: Meßraum $(\Omega', \mathfrak{A}')$ und Zufallselement $Y : \Omega \to \Omega'$. Betrachte die von Y erzeugte σ-Algebra $\mathfrak{G} = \sigma(Y)$.

Definition 2. *Bedingte Erwartung von X gegeben Y*:

$$\mathrm{E}(X \,|\, Y) = \mathrm{E}(X \,|\, \sigma(Y)).$$

Anwendung von Theorem II.2.8 auf obige Situation: Faktorisierung der bedingten Erwartung: Es existiert eine $\mathfrak{A}'$-meßbare Abbildung $g : \Omega' \to \mathbb{R}$ mit

$$\mathrm{E}(X \,|\, Y) = g \circ Y.$$

Je zwei solche Abbildungen stimmen P_Y-f.s. überein.

Definition 3. *Bedingte Erwartung von X gegeben $Y = y$*:

$$\mathrm{E}(X \,|\, Y = y) = g(y),$$

wobei g wie oben gewählt.

Analoge Begriffsbildung für *bedingte Wahrscheinlichkeiten*, wobei $X = 1_A$ mit $A \in \mathfrak{A}$.

Example 2. Gelte $(\Omega, \mathfrak{A}, P) = ([0,1], \mathfrak{B}([0,1]), \lambda)$ und $(\Omega', \mathfrak{A}') = (\mathbb{R}, \mathfrak{B})$. Ferner

$$X(\omega) = \omega^2, \qquad Y(\omega) = \begin{cases} 1, & \text{falls } \omega \in [0, 1/2], \\ \omega - 1/2, & \text{falls } \omega \in]1/2, 1]. \end{cases}$$

Dann

$$\sigma(Y) = \{A \cup B : A \in \{\emptyset, [0, 1/2]\},\ B \subset]1/2, 1]\,,\ B \in \mathfrak{A}\}$$

sowie

$$\mathrm{E}(X \,|\, Y)(\omega) = \begin{cases} 1/12, & \text{falls } \omega \in [0, 1/2], \\ \omega^2, & \text{falls } \omega \in]1/2, 1] \end{cases}$$

und

$$\mathrm{E}(X \,|\, Y = y) = \begin{cases} 1/12, & \text{falls } y = 1, \\ (y + 1/2)^2, & \text{falls } y \in]0, 1/2]. \end{cases}$$

Beachte, daß $P(\{Y = y\}) = 0$ für alle $y \in]0, 1/2]$.

Remark 3. Klar: für $A' \in \mathfrak{A}'$ gilt

$$\int_{\{Y \in A'\}} X \, dP = \int_{A'} \mathrm{E}(X \,|\, Y = y) \, P_Y(dy) \tag{3}$$

und insbesondere

$$P(A \cap \{Y \in A'\}) = \int_{A'} P(A \,|\, Y = y) \, P_Y(dy)$$

für $A \in \mathfrak{A}$. Durch (3) für alle $A' \in \mathfrak{A}'$ und die Forderung der $\mathfrak{A}'$-Meßbarkeit ist $\mathrm{E}(X \,|\, Y = \cdot)$ P_Y-f.s. eindeutig bestimmt.

Wie das folgende Theorem zeigt, ist die bedingte Erwartung der beste Schätzer im Quadratmittel. Vgl. Übung 10.4 und Lemma 4.

Theorem 3. Gelte $X \in \mathfrak{L}^2$. Dann gilt für jede $\mathfrak{A}'$-meßbare Abbildung $\varphi : \Omega' \to \mathbb{R}$

$$\int_\Omega (X - \mathrm{E}(X \,|\, Y))^2 \, dP \leq \int_\Omega (X - \varphi \circ Y)^2 \, dP$$

mit Gleichheit gdw. $\varphi = \mathrm{E}(X \,|\, Y = \cdot)$ P_Y-f.s.

Proof. Setze $Z^* = \mathrm{E}(X \,|\, Y)$ und $Z = \varphi \circ Y$. Die Jensensche Ungleichung liefert $Z^* \in \mathfrak{L}^2$, siehe (2) mit $p = 2$. OBdA: $Z \in \mathfrak{L}^2$. Dann

$$\mathrm{E}(X - Z)^2 = \mathrm{E}(X - Z^*)^2 + \underbrace{\mathrm{E}(Z^* - Z)^2}_{\geq 0} + 2 \cdot \mathrm{E}((X - Z^*)(Z^* - Z)).$$

Mit Lemma 1 und 2 folgt

$$\begin{aligned} \mathrm{E}((X - Z^*)(Z^* - Z)) &= \int_\Omega \mathrm{E}((X - Z^*)(Z^* - Z) \,|\, Y) \, dP \\ &= \int_\Omega (Z^* - Z) \cdot \mathrm{E}((X - Z^*) \,|\, Y) \, dP \\ &= \int_\Omega (Z^* - Z) \cdot \underbrace{(\mathrm{E}(X \,|\, Y) - Z^*)}_{=0} \, dP. \end{aligned}$$

□

Nun: der Zusammenhang zwischen Markov-Kernen und bedingten Wahrscheinlichkeiten.

Gegeben: Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ sowie Meßräume $(\Omega', \mathfrak{A}')$ und $(\Omega'', \mathfrak{A}'')$. Ferner eine $\mathfrak{A}$-$\mathfrak{A}'$ meßbare Abbildung $Y : \Omega \to \Omega'$ und eine $\mathfrak{A}$-$\mathfrak{A}''$ meßbare Abbildung $X : \Omega \to \Omega''$. Spezialfall

$$(\Omega'', \mathfrak{A}'') = (\Omega, \mathfrak{A}), \qquad X = \mathrm{id}\,. \tag{4}$$

Lemma 5. Für jede Abbildung $P_{X\,|\,Y} : \Omega' \times \mathfrak{A}'' \to \mathbb{R}$ sind äquivalent:

(i) $P_{X\,|\,Y}$ Markov-Kern von $(\Omega', \mathfrak{A}')$ nach $(\Omega'', \mathfrak{A}'')$ und

$$P_{(Y,X)} = P_Y \times P_{X\,|\,Y}, \tag{5}$$

(ii) für jedes $y \in \Omega'$ ist $P_{X\,|\,Y}(y, \cdot)$ ein Wahrscheinlichkeitsmaß auf $(\Omega'', \mathfrak{A}'')$ und für alle $A'' \in \mathfrak{A}''$ gilt

$$P_{X\,|\,Y}(\cdot, A'') = P(\{X \in A''\} \,|\, Y = \cdot).$$

Gilt (i), so sind X und Y genau dann unabhängig, wenn

$$P_{X\,|\,Y}(y, \cdot) = P_X$$

für P_Y-f.a. $y \in \mathbb{R}$ gilt.

Proof. Definitionsgemäß gilt für $A' \in \mathfrak{A}'$ und $A'' \in \mathfrak{A}''$

$$P_Y \times P_{X\,|\,Y}(A' \times A'') = \int_{A'} P_{X\,|\,Y}(y, A'')\, P_Y(dy)$$

und

$$P_{(Y,X)}(A' \times A'') = \int_{A'} P(\{X \in A''\} \,|\, Y = y)\, P_Y(dy).$$

Dies zeigt die Äquivalenz von (i) und (ii). Charakterisierung der Unabhängigkeit: Übung 14.2. $\square$

Definition 4. Jeder Markov-Kern $P_{X\,|\,Y}$ von $(\Omega', \mathfrak{A}')$ nach $(\Omega'', \mathfrak{A}'')$ mit der Eigenschaft (5) heißt eine *reguläre bedingte Verteilung von X gegeben Y* und im Spezialfall (4) auch *reguläre bedingte Wahrscheinlichkeit gegeben Y*. Ferner heißt (5) *Desintegration der gemeinsamen Verteilung von Y und X*.

Remark 4. Betrachte im Spezialfall (4) paarweise disjunkte Mengen $A_1, A_2, \dots \in \mathfrak{A}$. Für jede Menge $A' \in \mathfrak{A}'$ gilt

$$\begin{aligned}
\int_{A'} P\left(\bigcup_{i=1}^{\infty} A_i \,|\, Y = y\right) P_Y(dy) &= P\left(\bigcup_{i=1}^{\infty} A_i \cap \{Y \in A'\}\right) \\
&= \sum_{i=1}^{\infty} P(A_i \cap \{Y \in A'\}) \\
&= \sum_{i=1}^{\infty} \int_{A'} P(A_i \,|\, Y = y)\, P_Y(dy) \\
&= \int_{A'} \sum_{i=1}^{\infty} P(A_i \,|\, Y = y)\, P_Y(dy).
\end{aligned}$$

Es folgt P_Y-f.s.

$$P\left(\bigcup_{i=1}^{\infty} A_i \,|\, Y = \cdot\right) = \sum_{i=1}^{\infty} P(A_i \,|\, Y = \cdot).$$

Beachte: die entsprechende Nullmenge in $\mathfrak{A}'$ kann von der Wahl der Mengen A_i abhängen.

Example 3. Betrachte ein Wahrscheinlichkeitsmaß μ auf $(\Omega', \mathfrak{A}')$ und einen Markov-Kern $P_{X\,|\,Y}$ von $(\Omega', \mathfrak{A}')$ nach $(\Omega'', \mathfrak{A}'')$. Auf dem Produktraum

$$(\Omega, \mathfrak{A}) = (\Omega' \times \Omega'', \mathfrak{A}' \otimes \mathfrak{A}'')$$

betrachten wir das Wahrscheinlichkeitsmaß

$$P = \mu \times P_{X\,|\,Y}$$

und die Projektionen

$$Y(\omega', \omega'') = \omega', \qquad X(\omega', \omega'') = \omega''.$$

Unmittelbar aus den Definitionen folgt $\mu = P_Y$ und $P = P_{(Y,X)}$. Lemma 5 sichert für jedes $A'' \in \mathfrak{A}''$, daß

$$P_{X\,|\,Y}(y, A'') = P(\{X \in A''\} \,|\, Y = y)$$

für P_Y-f.a. $y \in \Omega'$ gilt. Analog für Folgen von Markov-Kernen.

Example 4. Sei $(\Omega', \mathfrak{A}') = (\Omega'', \mathfrak{A}'') = (\mathbb{R}, \mathfrak{B})$ und gelte

$$P_{(Y,X)} = f \cdot \lambda_2$$

mit einer Wahrscheinlichkeitsdichte f auf $(\mathbb{R}^2, \mathfrak{B}_2)$. Für $A', A'' \in \mathfrak{B}$ sichert der Satz von Fubini (beachte: Meßbarkeit als Teilaussage)

$$P_{(Y,X)}(A' \times A'') = \int_{A' \times A''} f \, d(\lambda_1 \times \lambda_1) = \int_{A'} \int_{A''} f(y, x) \, \lambda_1(dx) \, \lambda_1(dy).$$

Die Wahl von $A'' = \mathbb{R}$ zeigt

$$P_Y = h \cdot \lambda_1$$

mit der Wahrscheinlichkeitsdichte

$$h(y) = \int_{\mathbb{R}} f(y, \cdot) \, d\lambda_1, \qquad y \in \mathbb{R}.$$

Definiere für $y, x \in \mathbb{R}$

$$f(x \,|\, y) = \begin{cases} f(y,x)/h(y) & \text{falls } h(y) > 0 \\ 1_{[0,1]}(x) & \text{sonst} \end{cases}$$

und eine Abbildung $P_{X\,|\,Y} : \Omega' \times \mathfrak{A}'' \to [0,1]$ durch

$$P_{X\,|\,Y}(y, A'') = \int_{A''} f(x \,|\, y) \, \lambda_1(dx),$$

also $P_{X\,|\,Y}(y,\cdot) = f(\cdot\,|\,y)\cdot\lambda_1$. Man erhält

$$\begin{aligned} P_{(Y,X)}(A'\times A'') &= \int_{A'}\int_{A''} f(x\,|\,y)\,\lambda_1(dx)\cdot h(y)\,\lambda_1(dy) \\ &= \int_{A'} P_{X\,|\,Y}(y,A'')\,P_Y(dy). \end{aligned}$$

Lemma 5 zeigt, daß $P_{X\,|\,Y}$ eine reguläre bedingte Verteilung von X gegeben Y ist. Wesentlich für die Richtigkeit dieser Aussage ist nur, daß die gemeinsame Verteilung von X und Y eine Dichte bzgl. des Produktes zweier σ-endlicher Maße besitzt. Schließlich gilt für P_Y-f.a. $y\in\mathbb{R}$

$$\mathrm{E}(X\,|\,Y=y) = \int_{\mathbb{R}} x\,P_{X\,|\,Y}(y,dx) = \int_{\mathbb{R}} x\cdot f(x\,|\,y)\,\lambda_1(dx). \tag{6}$$

Beweis: Übung 14.2.

Example 5. In der Situation von Example 2 gilt

$$P_{X\,|\,Y}(y,\cdot) = \begin{cases} g\cdot\lambda_1 & \text{falls } y=1,\\ \varepsilon_{(y+1/2)^2} & \text{falls } y\in\,]0,1/2], \end{cases}$$

wobei die Dichte g durch

$$g(x) = 1/\sqrt{x}\cdot 1_{]0,1/4]}(x)$$

gegeben ist. Ferner gilt

$$P_{\mathrm{id}\,|\,Y}(y,\cdot) = \begin{cases} \nu & \text{falls } y=1,\\ \varepsilon_{y+1/2} & \text{falls } y\in\,]0,1/2], \end{cases}$$

wobei ν die Gleichverteilung auf $[0,1/2]$ bezeichnet. Beweis: Übung 14.2.

Theorem 4. Gelte $(\Omega',\mathfrak{A}') = (M,\mathfrak{B}(M))$ mit einem vollständigen und separablen metrischen Raum (M,ρ). Dann existiert eine reguläre bedingte Verteilung von X gegeben Y. Für je zwei solche Verteilungen $P^{(i)}_{X\,|\,Y}$ existiert eine Menge $A'\in\mathfrak{A}'$ mit $P_Y(A')=1$ und

$$\forall\, y\in A'\ \forall\, A''\in\mathfrak{A}'':\quad P^{(1)}_{X\,|\,Y}(y,A'') = P^{(2)}_{X\,|\,Y}(y,A'').$$

Proof. Siehe Gänssler, Stute (1977, Kap. V.3) oder Yeh (1995, App. C). □

Im folgenden sei $V\in\mathfrak{L}^1(\Omega,\mathfrak{A},P)$ und $P_{\mathrm{id}\,|\,Y}$ eine reguläre bedingte Wahrscheinlichkeit gegeben Y.

Theorem 5.

(i) $\displaystyle\int_\Omega V(\omega)\,P_{\mathrm{id}\,|\,Y}(\cdot,d\omega)$ $\mathfrak{A}'$-meßbar,

(ii) $\displaystyle\int_{A'}\int_\Omega V(\omega)\,P_{\mathrm{id}\,|\,Y}(y,d\omega)\,P_Y(dy) = \int_{\{Y\in A'\}} V\,dP$ für $A'\in\mathfrak{A}'$.

Also für P_Y-f.a. $y \in \Omega'$

$$\mathrm{E}(V \mid Y = y) = \int_\Omega V(\omega)\, P_{\mathrm{id}\,|\,Y}(y, d\omega).$$

Proof. Algebraische Induktion. □

Theorem 6. Für P_Y-f.a. $y \in \Omega'$

$$P_{\mathrm{id}\,|\,Y}(y, Y^{-1}(\{y\})) = 1.$$

Proof. Siehe Yeh (1995, p. 486). □

Die Ergebnisse dieses Abschnittes beantworten die im einführenden Example I.4 gestellten Fragen.

2 Discrete-Time Martingales

Gegeben: Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$.

Definition 1. Folge $\widetilde{\mathfrak{A}} = (\mathfrak{A}_n)_{n\in\mathbb{N}_0}$ von σ-Algebren $\mathfrak{A}_n \subset \mathfrak{A}$ mit

$$\forall\, n \in \mathbb{N}_0 :\ \mathfrak{A}_n \subset \mathfrak{A}_{n+1}$$

heißt *Filtration*.

Gegeben: Folge $\widetilde{X} = (X_n)_{n\in\mathbb{N}_0}$ von Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$.

Example 1. *Kanonische Filtration*:

$$\mathfrak{A}_n = \sigma(\{X_0, \dots, X_n\}), \qquad n \in \mathbb{N}_0, \tag{1}$$

definiert die 'kleinste' Filtration $\widetilde{\mathfrak{A}}$, so daß X_n $\mathfrak{A}_n$-meßbar ist. Bei dieser Wahl ist $Y : \Omega \to \mathbb{R}$ genau dann $\mathfrak{A}_n$-meßbar, wenn eine $\mathfrak{B}_{n+1}$-meßbare Abbildung $g : \mathbb{R}^{n+1} \to \mathbb{R}$ existiert, die $Y = g \circ (X_0, \dots, X_n)$ erfüllt. Siehe Theorem II.2.8 und Corollary II.3.1.(i).

Definition 2. $\widetilde{X}$ heißt *Martingal (bzgl.* $\widetilde{\mathfrak{A}}$*)*, falls $X_n \in \mathfrak{L}^1$ für alle $n \in \mathbb{N}_0$ und

$$\forall\, n, m \in \mathbb{N}_0 : \quad n < m \quad \Rightarrow \quad \mathrm{E}(X_m \mid \mathfrak{A}_n) = X_n.$$

Zur Interpretation: Theorem 1.3.

Remark 1. Für jedes Martingal $\widetilde{X}$ und $n < m$

$$\mathrm{E}(X_m) = \mathrm{E}(\mathrm{E}(X_m \mid \mathfrak{A}_n)) = \mathrm{E}(X_n).$$

Klar: Aus der Konstanz der Erwartungswerte folgt i.a. nicht die Martingaleigenschaft.

Remark 2. $\widetilde{X}$ Martingal gdw.

$$\forall\, n \in \mathbb{N}_0 : \quad \mathrm{E}(X_{n+1} \,|\, \mathfrak{A}_n) = X_n.$$

Zum Beweis verwende man Lemma 1.3.

Example 2. Sei $(Y_i)_{i\in\mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit $\mathrm{E}(Y_i) = a$ für alle $i \in \mathbb{N}$. Ferner sei $\mathfrak{A}_0 = \{\emptyset, \Omega\}$ und $\mathfrak{A}_n = \sigma(\{Y_1, \ldots, Y_n\})$ für $n \geq 1$. Setze $X_0 = 0$ und

$$X_n = \sum_{i=1}^{n} Y_i, \qquad n \in \mathbb{N}.$$

Es gilt (1) sowie

$$\mathrm{E}(X_{n+1} \,|\, \mathfrak{A}_n) = \mathrm{E}(X_n \,|\, \mathfrak{A}_n) + \mathrm{E}(Y_{n+1} \,|\, \mathfrak{A}_n) = X_n + \mathrm{E}(Y_{n+1}) = X_n + a.$$

Somit ist $\widetilde{X}$ genau im Falle $a = 0$ ein Martingal. Bsp: random walk, coin tossing.

Mögliche Interpretation: Y_i Gewinn bei einfachem Spiel in Runde i und X_n akkumulierter Gewinn nach n Runden. Martingal heißt: 'faires Spiel'.

Frage: Kann man im Martingalfall durch

(i) eine geeignete Wahl der Einsätze und

(ii) einen geeigneten Abbruch des Spiels

im Mittel einen positiven Gesamtgewinn erreichen?

Example 3. Das *Cox-Ross-Rubinstein-Modell* für Aktienkurse X_n zu Zeiten $n \in \mathbb{N}_0$. Wähle reelle Zahlen

$$X_0 > 0, \quad 0 < p < 1, \quad 0 < d < u,$$

und betrachte $(Y_i)_{i\in\mathbb{N}}$ i.i.d. mit

$$P(\{Y_i = u\}) = p = 1 - P(\{Y_i = d\}).$$

Setze $\mathfrak{A}_0 = \{\emptyset, \Omega\}$ und definiere

$$X_n = X_0 \cdot \prod_{i=1}^{n} Y_i$$

sowie $\mathfrak{A}_n = \sigma(\{Y_1, \ldots, Y_n\})$ für $n \in \mathbb{N}$. Für $n < m$ zeigen Lemma 1.2 und Lemma 1.4

$$\mathrm{E}(X_m \,|\, \mathfrak{A}_n) = X_n \cdot \mathrm{E}\left(\prod_{\ell=n+1}^{m} Y_\ell\right) = X_n \cdot \mathrm{E}(Y_1)^{m-n}.$$

Also

$$\widetilde{X} \text{ Martingal} \qquad \Leftrightarrow \qquad \mathrm{E}(Y_1) = 1,$$

bzw., da $\mathrm{E}(Y_1) = pu + (1-p)d$,

$$\widetilde{X} \text{ Martingal} \qquad \Leftrightarrow \qquad d < 1 < u \ \wedge\ p = \frac{1-d}{u-d}.$$

Frage: Wie in Example 2 ('Handelsstrategie', 'Verkaufsstrategie').

Im folgenden:

(i) $\widetilde{X} = (X_n)_{n\in\mathbb{N}_0}$ Martingal bzgl. $\widetilde{\mathfrak{A}}$,

(ii) $\widetilde{H} = (H_n)_{n\in\mathbb{N}_0}$ Folge von Zufallsvariablen, so daß

$$\forall\, n \in \mathbb{N}_0 : \; H_n \; \mathfrak{A}_n\text{-meßbar} \wedge \; H_n \cdot (X_{n+1} - X_n) \in \mathfrak{L}^1.$$

Definition 3. Die Folge $\widetilde{Z} = (Z_n)_{n\in\mathbb{N}_0}$ von Zufallsvariablen $Z_0 = 0$ und

$$Z_n = \sum_{i=0}^{n-1} H_i \cdot (X_{i+1} - X_i), \qquad n \geq 1,$$

heißt *Martingaltransformation von* $\widetilde{X}$ *mittels* $\widetilde{H}$. Bez.: $\widetilde{Z} = \widetilde{H} \bullet \widetilde{X}$.

Example 4. In Example 2: H_n Einsatz im $(n+1)$-ten Spiel und Z_n akkumulierter Gewinn nach n Runden mit Einsätzen $H_0, \dots, H_{n-1}$. Spezialfall: $H_n \in \{\pm 1\}$ bei coin tossing.

Theorem 1. $\widetilde{Z} = \widetilde{H} \bullet \widetilde{X}$ ist Martingal bzgl. $\widetilde{\mathfrak{A}}$.

Proof. Klar für $n \in \mathbb{N}_0$: Z_n ist $\mathfrak{A}_n$-meßbar und $Z_n \in \mathfrak{L}^1$. Somit

$$\mathrm{E}(Z_{n+1} \,|\, \mathfrak{A}_n) = Z_n + \mathrm{E}(H_n \cdot (X_{n+1} - X_n) \,|\, \mathfrak{A}_n),$$

und weiter

$$\mathrm{E}(H_n \cdot (X_{n+1} - X_n) \,|\, \mathfrak{A}_n) = H_n \cdot \mathrm{E}((X_{n+1} - X_n) \,|\, \mathfrak{A}_n) = 0.$$

□

3 Optional Sampling

Gegeben: Filtration $\widetilde{\mathfrak{A}}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$.

Definition 1. $\tau : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ heißt *Stoppzeit (bzgl.* $\widetilde{\mathfrak{A}}$*)*, falls

$$\forall\, n \in \mathbb{N}_0 : \; \{\tau \leq n\} \in \mathfrak{A}_n.$$

Lemma 1.

$$\tau \text{ Stoppzeit} \qquad \Leftrightarrow \qquad \forall\, n \in \mathbb{N}_0 : \; \{\tau = n\} \in \mathfrak{A}_n.$$

Proof. Verwende

$$\{\tau \leq n\} = \bigcup_{i=0}^{n} \{\tau = i\}, \qquad \{\tau = n\} = \{\tau \leq n\} \setminus \{\tau \leq n-1\}.$$

□

Example 1. Verkaufsstrategien für eine Aktie mit Preis X_n zur Zeit $n \in \mathbb{N}_0$:

(i) Verkaufe, sobald der Preis a erreicht oder überschritten ist, spätestens jedoch zur Zeit N.

(ii) Verkaufe beim ersten Eintreten des Maximum von $X_0, \dots, X_N$.

Formal heißt (i)

$$\tau = \inf\left(\{i \in \{0, \dots, N\} : X_i \geq a\} \cup \{N\}\right).$$

Dann: τ ist Stoppzeit bzgl. der kanonischen Filtration $\widetilde{\mathfrak{A}}$ zu $\widetilde{X}$, d.h. 'realisierbare Strategie'. Es gilt nämlich für $k = 0, \dots, N-1$

$$\{\tau = k\} = \bigcap_{i=0}^{k-1} \underbrace{\{X_i < a\}}_{\in \mathfrak{A}_i \subset \mathfrak{A}_{k-1}} \cap \underbrace{\{X_k \geq a\}}_{\in \mathfrak{A}_k} \in \mathfrak{A}_k$$

sowie

$$\{\tau = N\} = \bigcap_{i=0}^{N-1} \{X_i < a\} \in \mathfrak{A}_{N-1}.$$

Formal heißt (ii)

$$\tau = \inf\{i \in \{0, \dots, N\} : X_i = M\} \qquad \text{mit} \qquad M = \max_{i=0,\dots,N} X_i.$$

Dies ist i.a. keine Stoppzeit, d.h. eine 'nicht realisierbare Strategie'. Betrachte etwa das Cox-Ross-Rubinstein-Modell mit $d < 1 < u$. Für $N = 1$ gilt

$$\{\tau = 0\} = \{Y_1 = d\} \notin \{\emptyset, \Omega\} = \mathfrak{A}_0.$$

Lemma 2.

$$\sigma, \tau \text{ Stoppzeiten bzgl. } \widetilde{\mathfrak{A}} \quad \Rightarrow \quad \sigma + \tau,\ \min\{\sigma, \tau\},\ \max\{\sigma, \tau\} \text{ Stoppzeiten bzgl. } \widetilde{\mathfrak{A}}.$$

Proof. $\boxed{\text{Übung}}$. □

Gegeben: Folge $\widetilde{X} = (X_n)_{n \in \mathbb{N}_0}$ von Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$, so daß für alle $n \in \mathbb{N}_0$ gilt:

(i) X_n $\mathfrak{A}_n$-meßbar,

(ii) $X_n \in \mathfrak{L}^1$.

Für eine Abbildung $\tau : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ definieren wir

$$X_\tau : \Omega \to \mathbb{R}$$

durch

$$X_\tau(\omega) = \begin{cases} X_{\tau(\omega)}(\omega) & \text{falls } \tau(\omega) < \infty \\ 0 & \text{sonst.} \end{cases}$$

Die folgenden beiden Sätze sind Varianten des *optional sampling theorem*.

Theorem 1.

$$\widetilde{X} \text{ Martingal bzgl. } \widetilde{\mathfrak{A}} \quad \Leftrightarrow \quad \forall\, \tau \text{ beschränkte Stoppzeit bzgl. } \widetilde{\mathfrak{A}} : \mathrm{E}(X_\tau) = \mathrm{E}(X_0).$$

Proof. '$\Rightarrow$' Sei τ eine Stoppzeit mit $\tau(\omega) \le N$ für alle $\omega \in \Omega$. Also

$$X_\tau = \sum_{n=0}^{N} 1_{\{\tau=n\}} \cdot X_n.$$

Also ist X_τ $\mathfrak{A}$-meßbar und $\mathrm{E}(|X_\tau|) \le \sum_{n=0}^{N} \mathrm{E}(|X_n|) < \infty$. Weiter

$$\begin{aligned}\mathrm{E}(X_\tau) &= \sum_{n=0}^{N} \mathrm{E}(1_{\{\tau=n\}} \cdot X_n) = \sum_{n=0}^{N} \mathrm{E}(1_{\{\tau=n\}} \cdot \mathrm{E}(X_N \,|\, \mathfrak{A}_n)) \\ &= \sum_{n=0}^{N} \mathrm{E}(\mathrm{E}(1_{\{\tau=n\}} \cdot X_N \,|\, \mathfrak{A}_n)) = \sum_{n=0}^{N} \mathrm{E}(1_{\{\tau=n\}} \cdot X_N) = \mathrm{E}(X_N) = \mathrm{E}(X_0).\end{aligned}$$

'$\Leftarrow$' Für $n < m$ und $A \in \mathfrak{A}_n$ ist zu zeigen

$$\int_A X_m \, dP = \int_A X_n \, dP.$$

Definiere

$$\tau = n \cdot 1_A + m \cdot 1_{\Omega \setminus A}.$$

Klar: τ ist beschränkte Stoppzeit. Also

$$\mathrm{E}(X_0) = \mathrm{E}(X_\tau) = \mathrm{E}(1_A \cdot X_n + 1_{\Omega\setminus A} \cdot X_m) = \mathrm{E}(X_m) - \mathrm{E}(1_A \cdot X_m) + \mathrm{E}(1_A \cdot X_n).$$

Beachte schließlich, daß n.V. insbesondere $\mathrm{E}(X_0) = \mathrm{E}(X_m)$ gilt. □

Theorem 2.1 und Theorem 1 beantworten die in Example 2.2 gestellten Fragen negativ, solange man eine obere Schranke für die Spieldauer akzeptiert.

Theorem 2. Sei $\widetilde{X}$ Martingal und τ Stoppzeit mit

$$P(\{\tau < \infty\}) = 1 \quad \wedge \quad \mathrm{E}(|X_\tau|) < \infty \quad \wedge \quad \lim_{n\to\infty} \int_{\{\tau > n\}} |X_n| \, dP = 0. \tag{1}$$

Dann

$$\mathrm{E}(X_\tau) = \mathrm{E}(X_0).$$

Proof. Für $\tau_N = \min\{\tau, N\}$ gilt

$$|\,\mathrm{E}(X_\tau) - \mathrm{E}(X_{\tau_N})| \le \int_{\{\tau > N\}} |X_\tau| \, dP + \int_{\{\tau > N\}} |X_N| \, dP$$

und somit

$$\lim_{N\to\infty} \mathrm{E}(X_{\tau_N}) = \mathrm{E}(X_\tau).$$

Theorem 1 und Lemma 2 liefern $\mathrm{E}(X_0) = \mathrm{E}(X_{\tau_N})$. □

Example 2. In Example 2.2 gelte: $(Y_i)_{i\in\mathbb{N}}$ i.i.d. mit $P_{Y_1} = 1/2 \cdot (\varepsilon_1 + \varepsilon_{-1})$. Einsatz $H_n = 2^n$ in $(n+1)$-ten Spiel (Verdopplungsstrategie). Nach Theorem 1 (einfacher: Example 2.2) definiert $Z_0 = 0$ und

$$Z_n = \sum_{i=0}^{n-1} 2^i \cdot Y_{i+1}, \qquad n \in \mathbb{N},$$

ein Martingal. Für die Stoppzeit

$$\tau = \inf\{i \in \mathbb{N} : Y_i = 1\}$$

ergibt sich

(i) $\tau = \inf\{n \in \mathbb{N}_0 : Z_n > 0\}$,

(ii) $Z_\tau = 1$,

(iii) $P(\{\tau = n\}) = 2^{-n}$, also τ f.s. endlich und $\mathrm{E}(\tau) = 2$.

Jedoch ist $\tau > n$ äquivalent zu $Z_n = -1 - \cdots - 2^{n-1} = -(2^n - 1)$, so daß

$$\int_{\{\tau>n\}} |Z_n|\, dP = (2^n - 1) \cdot \sum_{m=n+1}^{\infty} 2^{-m} = 1 - 2^{-n}.$$

Example 3 (Das Ruin-Problem)**.** Betrachte das Glücksspiel aus Example 2.2 mit

$$P_{Y_i} = p \cdot \varepsilon_1 + (1-p) \cdot \varepsilon_{-1}$$

für festes $p \in]0,1[$. Startkapital C. Ziel: Gewinn G, wobei $0 < C < G$. Spiele bis G erreicht oder C verspielt. Also

$$\tau = \inf\{n \in \mathbb{N}_0 : X_n = G \vee X_n = -C\}.$$

Bestimme die Ruin-Wahrscheinlichkeit $P(\{X_\tau = -C\})$ sowie den Erwartungswert der Spieldauer τ.

Dazu zeigt man vorab

$$\exists\, a > 0\ \exists\, \gamma \in]0,1[\ \forall\, j \in \mathbb{N}_0: \quad P(\{\tau > j\}) \le a \cdot \gamma^j, \tag{2}$$

siehe Irle (1998, p. 48).

Mit (2) folgt

$$P(\{\tau = \infty\}) \le \liminf_{j\to\infty} P(\{\tau > j\}) = 0$$

und weiter

$$\mathrm{E}(\tau) = \sum_{j=1}^{\infty} P(\{\tau \ge j\}) < \infty.$$

Also

$$1 = P(\{\tau < \infty\}) = P(\underbrace{\{X_\tau = G\}}_{\text{'Gewinn'}}) + P(\underbrace{\{X_\tau = -C\}}_{\text{'Ruin'}}).$$

Nun Anwendung des optional sampling theorem. Klar: τ ist unbeschränkt, deshalb verwenden wir Theorem 2.

Definiere $M_0 = 0$ und

$$M_n = \sum_{i=1}^{n} (Y_i - \mathrm{E}(Y_i)) = X_n - na,$$

wobei $a = 2p - 1$. Dann ist $\widetilde{M}$ ein Martingal, siehe Example 2.2. Wir verifizieren die weiteren Voraussetzungen von Theorem 2.

Es gilt

$$|M_\tau| \le |X_\tau| + \tau \cdot |a| \le \max\{G, C\} + |a| \cdot \tau,$$

und somit

$$\mathrm{E}(|M_\tau|) \le \max\{G, C\} + |a| \cdot \mathrm{E}(\tau) < \infty.$$

Ferner

$$\begin{aligned} \int_{\{\tau > n\}} |M_n| \, dP &\le \int_{\{\tau > n\}} (|X_n| + |a| \cdot n) \, dP \\ &\le \max\{G, C\} \cdot P(\{\tau > n\}) + |a| \cdot n \cdot P(\{\tau > n\}), \end{aligned}$$

und somit sichert (2)

$$\lim_{n \to \infty} \int_{\{\tau > n\}} |M_n| \, dP = 0.$$

Theorem 2 liefert

$$\begin{aligned} 0 = \mathrm{E}(M_0) = \mathrm{E}(M_\tau) &= \mathrm{E}(X_\tau) - \mathrm{E}(\tau) \cdot a \\ &= G \cdot P(\{X_\tau = G\}) - C \cdot P(\{X_\tau = -C\}) - \mathrm{E}(\tau) \cdot a. \end{aligned} \tag{3}$$

1. Fall: Faires Spiel, d.h.

$$p = \frac{1}{2}.$$

Dann $a = 0$ und

$$P(\{X_\tau = G\}) = \frac{C}{C + G}, \qquad P(\{X_\tau = -C\}) = \frac{G}{C + G}.$$

Weiterhin ist $(X_n^2 - n)_{n \in \mathbb{N}_0}$ ein Martingal, und die Voraussetzungen von Theorem 2 sind erfüllt. Also

$$0 = \mathrm{E}(X_0^2 - 0) = \mathrm{E}(X_\tau^2 - \tau) = \mathrm{E}(X_\tau^2) - \mathrm{E}(\tau),$$

so daß

$$\mathrm{E}(\tau) = \mathrm{E}(X_\tau^2) = G^2 \cdot \frac{C}{C + G} + C^2 \cdot \frac{G}{C + G} = C \cdot G.$$

2. Fall Unfaires Spiel, d.h.

$$p \ne \frac{1}{2}.$$

Setze

$$q = \frac{p}{1-p}.$$

Man erhält

$$P(\{X_\tau = G\}) = \frac{1-q^C}{(1/q)^G - q^C}, \qquad P(\{X_\tau = -C\}) = \frac{(1/q)^G - 1}{(1/q)^G - q^C},$$

und mit (3) folgt

$$\mathrm{E}(\tau) = \frac{G \cdot P(\{X_\tau = G\}) - C \cdot P(\{X_\tau = -C)\}}{2p-1}.$$

Siehe Irle (1998, p. 50).

Numerische Berechnungen zeigen: kleine Abweichungen von $p = 1/2$ führen zu drastischen Änderungen der Ruin-Wahrscheinlichkeit.

4 Ausblick

. . .

Literature

H. Bauer, *Probability Theory*, de Gruyter, Berlin, 1996.

P. Billingsley, *Probability and Measure*, Wiley, New York, first edition 1979, third edition 1995.

Y. S. Chow, H. Teicher, *Probability Theory*, Springer, New York, first editon 1978, third edition 1997.

R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.

J. Elstrodt, *Maß- und Integrationstheorie*, Springer, Berlin, first edition 1996, fifth edition, 2007.

K. Floret, *Maß- und Integrationstheorie*, Teubner, Stuttgart, 1981.

P. Gänssler, W. Stute, *Wahrscheinlichkeitstheorie*, Springer, Berlin, 1977.

E. Hewitt, K. Stromberg, *Real and Abstract Analysis*, Springer, Berlin, 1965.

A. Irle, *Finanzmathematik*, Teubner, Stuttgart, 1998.

A. Klenke, *Wahrscheinlichkeitstheorie*, Springer, Berlin, first edition 2006, second edition 2008.

K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

A. N. Širjaev, *Wahrscheinlichkeit*, Deutscher Verlag der Wissenschaften, Berlin, 1988.
A. N. Shiryayev, *Probability*, Springer, New York, 1984.

J. Yeh, *Martingales and Stochastic Analysis*, World Scientific, Singapore, 1995.

Index

σ-additive mapping, 17
σ-algebra, 3
  generated by a class of sets, 5
  generated by a family of mappings, 9
σ-continuity at $\emptyset$, 19
σ-continuity from above, 19
σ-continuity from below, 19
σ-finite mapping, 23
σ-subadditivity, 19

absolutely continuous distribution, 51
absolutely continuous measure, 34
abstract integral, 27
additive mapping, 17
algebra, 3
  generated by a class of sets, 5
almost everywhere, 27
almost surely, 27
asymptotically negligible, 91

Bernoulli distribution, 50
binomial distribution, 50
Borel-σ-algebra, 7

Cauchy distribution, 52
characteristic function, 86
closed set, 7
closed w.r.t.
  intersections, 3
  unions, 3
compact set, 7
complete measure space, 24
completion of a measure space, 25
conditional expectation, 97, 98, 100
conditional probability, 97, 98, 100
content, 17
convergence
  almost everywhere, 29
  in $\mathfrak{L}^p$, 29
  in distribution, 56
  in mean, 29
  in mean-square, 29
  in probability, 54
  weak, 56
convolution, 70
counting measure, 18
covariance, 70
Cox-Ross-Rubinstein model, 106
cylinder set, 15

desintegration, 102
Dirac measure, 18
discrete distribution, 50
discrete probability measure, 18
distribution, 49
distribution function, 53
Dynkin class, 4
  generated by a class of sets, 5

empirical distribution, 83
empirical distribution function, 83
essential supremum, 31
essentially bounded function, 31
event, 49
expectation, 52
exponential distribution, 51

Feller condition, 91
filtration, 105
  natural (canonical), 105
finite mapping, 23
Fourier transform
  of a probability measure, 85
  of an integrable function, 85

geometric distribution, 50

i.i.d, 76
identically distributed, 49

image measure, 46
independence
  of a family of classes, 66
  of a family of events, 65
  of a family of random elements, 67
integrable function, 27
  complex-valued, 85
integral, 27
  of a complex-valued function, 85
  of a non-negative function, 26
  of a simple function, 25
  over a subset, 32

joint distribution, 68

kernel, 36
  σ-finite, 36
  Markov, 36

Lévy distance, 60
Lebesgue measurable set, 25
Lebesgue pre-measure, 18
limes inferior, 74
limes superior, 74
Lindeberg condition, 91
Lyapunov condition, 91

marginal distribution, 69
martingale, 105
martingale transform, 107
measurable
  mapping, 8
  rectangle, 13
  set, 8
  space, 8
measure, 17
  with density, 32
measure space, 18
monotonicity, 19
monotonicity of the integral, 26
Monte Carlo algorithm, 82

normal distribution
  multidimensional
    standard, 32
  one-dimensional, 51

open set, 7
outer measure, 21

Poisson distribution, 50
positive semi-definite function, 86
pre-measure, 17
probability density, 32
probability measure, 17
probability space, 18
product σ-algebra, 14
product (measurable) space, 14
product measure, 45
  n factors, 43
  two factors, 40
product measure space, 45
  n factors, 43
  two factors, 40

quasi-integrable mapping, 27

random element, 49
random variable, 49
random vector, 49
random walk, 96
regular conditional probability, 102
relatively compact set of measures, 61

section
  of a mapping, 37
  of a set, 38
semi-algebra, 3
simple function, 11
square-integrable function, 28
standard deviation, 52
stopping time, 107
subadditivity, 19

tail σ-algebra, 73
tail (terminal) event, 73
tightness, 61
topological space, 6
trace-σ-algebra, 7

unbiased estimator, 82
uncorrelated random variables, 70
uniform distribution
  on a finite set, 18, 50
  on a subset of $\mathbb{R}^k$, 32, 51

uniform integrability, 62

variance, 52

with probability one, 27