Stochastics II

Notes of a lecture held in winter 2022/2023
by Prof. Dr. Henryk Zähle

Nils Müller

Saarbrücken, March 1, 2023

Contents

1 Conditional Expectations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 1.1 (Conditional expectation) *Let $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.*

A random variable $Z \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{G}, \mathbb{P}_{|\mathcal{G}})$ is called ***conditional expectation of X given $\mathcal{G}$*** *if $\mathbb{E}[Z\mathbb{1}_G] = \mathbb{E}[X\mathbb{1}_G]$ for all $G \in \mathcal{G}$.*

Theorem 1.1 (Existence and uniqueness) *Let $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ be sub-σ-algebra of $\mathcal{F}$. Then*

(i) there exists a conditional expectation of X given $\mathcal{G}$, and

(ii) if Z_1, Z_2 are conditional expectations of X given $\mathcal{G}$, then $Z_1 = Z_2$ $\mathbb{P}$-a.s..

Remark 1.1 (Notation) *Let $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. If Z is a conditional expectation of X given $\mathcal{G}$, then we also write $\mathbb{E}[X|\mathcal{G}]$ instead of Z.*

Remark 1.2 (Relation to conditional expected value) *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. For each $G \in \mathcal{G}$ the conditional expected value of $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ given the event G can be determined by just knowing the conditional expectation $\mathbb{E}[X|\mathcal{G}]$.*

More precisely, for each $G \in \mathcal{G}$ with $\mathbb{P}[G] > 0$, we have

$$\mathbb{E}[X|G] = \frac{\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_G]}{\mathbb{P}[G]},$$

as $\mathbb{E}[X|G] = \mathbb{E}[X\mathbb{1}_G]/\mathbb{P}[G] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_G]/\mathbb{P}[G]$.

Example 1.1 *Let $\{G_i\}_{i\in I}$ be a partition of Ω consisting of countably many sets from $\mathcal{F}$. Then, for every $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ a conditional expectation of X give $\sigma(\{G_i\}_{i\in I})$ is given by $\sum_{i\in I} \mathbb{E}[X|G_i]\mathbb{1}_{G_i}$. In particular,*

$$\mathbb{E}[X|\sigma(\{G_i\}_{i\in I})] = \sum_{i\in I} \mathbb{E}[X|G_i]\mathbb{1}_{G_i} \qquad \mathbb{P}\text{-a.s..}$$

Proposition 1.1 *Let $X, X_1, X_2 \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$, $\alpha, \alpha_1, \alpha_2 \in \mathbb{R}$, and $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2$ be sub-σ-algebras of $\mathcal{F}$. Then the following assertions hold true, where in (ii) we tacitly assume (w.l.o.g.) that the pointwise additions are well defined.*

(i) $\mathbb{E}[X|\mathcal{G}] = \alpha$ $\mathbb{P}$-a.s. if $X = \alpha$ $\mathbb{P}$-a.s.

(ii) $\mathbb{E}[\alpha_1 X_1 + \alpha_2 X_2|\mathcal{G}] = \alpha_1\mathbb{E}[X_1|\mathcal{G}] + \alpha_2\mathbb{E}[X_2|\mathcal{G}]$ $\mathbb{P}$-a.s.

(iii) $\mathbb{E}[X|\mathcal{G}]^{\pm} \leq \mathbb{E}[X^{\pm}|\mathcal{G}]$ $\mathbb{P}$-a.s.

(iv) $\mathbb{E}[X|\mathcal{G}] = X$ $\mathbb{P}$-a.s. if $X \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{G}, \mathbb{P}_{|\mathcal{G}})$

(v) $\|\mathbb{E}[X|\mathcal{G}]\|_1 \leq \|X\|_1$ *(i.e.,* $\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|] \leq \mathbb{E}[|X|]$*)*

(vi) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$ $\mathbb{P}$*-a.s. if* $\mathcal{G}_1 \subseteq \mathcal{G}_2$

(vii) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$

(viii) $\mathbb{E}[X_1|\mathcal{G}] \leq \mathbb{E}[X_2|\mathcal{G}]$ $\mathbb{P}$*-a.s. if* $X_1 \leq X_2$ $\mathbb{P}$*-a.s.*

(ix) $|E[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ $\mathbb{P}$*-a.s.*

(x) $\mathbb{E}[X'X|\mathcal{G}] = X'\mathbb{E}[X|\mathcal{G}]$ *for* $X' \in \mathcal{L}_{\overline{\mathbb{R}}}(\Omega, \mathcal{G})$ *with* $X'X \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$

(xi) $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ $\mathbb{P}$*-a.s. if* $\sigma(X)$ *and* $\mathcal{G}$ *are independent*

(xii) $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ $\mathbb{P}$*-a.s. if* $\mathbb{P}[G] \in \{0, 1\}$ *for all* $G \in \mathcal{G}$

Definition 1.2 (Conditional probability) *Let* $A \in \mathcal{F}$ *and* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{F}$*. Every conditional expectation of* $\mathbb{1}_A$ *given* $\mathcal{G}$ *is called* ***conditional probability*** *of* A *given* $\mathcal{G}$*.*

Remark 1.3 (Notation) *Let* $A \in \mathcal{F}$ *and* $\mathcal{G}$ *be sub-*σ*-algebra of* $\mathcal{F}$*. If* Z *is a conditional expectation of* A *given* $\mathcal{G}$*, then we also write* $\mathbb{P}[A|\mathcal{G}]$ *instead of* Z *or* $\mathbb{E}[\mathbb{1}_A|\mathcal{G}]$*.*

Remark 1.4 (Conditioning on random variables) *Let* $X \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$ *and* $A \in \mathcal{F}$*. Moreover, let* Y *be a random variable on* $(\Omega, \mathcal{F}, \mathbb{P})$*.*

(i) If Z *is a conditonal expectation of* X *given* Y*, then we also write* $\mathbb{E}[X|Y]$ *instead of* Z *or* $\mathbb{E}[X|\sigma(Y)]$*.*

(ii) If Z *is a conditional probability of* A *given* Y*, then we also write* $\mathbb{P}[A|Y]$ *instead of* Z *or* $\mathbb{E}[\mathbb{1}_A|Y]$*.*

Let Y be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a measurable space $(\Omega', \mathcal{F}')$.

Definition 1.3 (Factorised conditioning) *Let* $X \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$ *and* $A \in \mathcal{F}$*.*

(i) A function $g \in \mathcal{L}_{\overline{\mathbb{R}}}(\Omega', \mathcal{F}')$ *is called* ***factorised conditional expectation of*** X ***given*** Y *if* $g(Y) = \mathbb{E}[X|Y]$ $\mathbb{P}$*-a.s..*

(ii) A function $g \in \mathcal{L}_{\overline{\mathbb{R}}}(\Omega', \mathcal{F}')$ *is called* ***factorised conditional probability of*** A ***given*** Y *if* $g(Y) = \mathbb{P}[A|Y]$ $\mathbb{P}$*-a.s..*

Theorem 1.2 (Existence and uniqueness) *Let* $X \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$ *and* $A \in \mathcal{F}$*.*

(i) There exists a $\mathbb{P}_Y$*-a.s. unique factorised conditional expectation of* X *given* Y*.*

(ii) There exists a $\mathbb{P}_Y$*-a.s. unique factorised conditional probability of* A *given* Y*.*

Remark 1.5 (Notation) *Let* $X \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$ *and* $A \in \mathcal{F}$*.*

(i) If g is a factorised conditional expectation of X given Y, then we also write $\mathbb{E}[X||Y = \cdot]$ instead of $g(\cdot)$.

(ii) If g is a factorised conditional probability of A given Y, then we also write $\mathbb{P}[A||Y = \cdot]$ instead of $g(\cdot)$

Theorem 1.3 (Insertion rule) *Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a measurable space $(E, \mathcal{E})$. Let $f \in \mathcal{L}_{\overline{\mathbb{R}}}(E \times \Omega', \mathcal{E} \times \mathcal{F}')$ and assume that $f(X, Y) \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$.*

If X and Y are independent, then there exists a $\mathbb{P}_Y$-null set $N' \in \mathcal{F}'$ such that $g(\omega') := \mathbb{E}[f(X, \omega')]$, $\omega' \in \Omega' \setminus N'$, and $g(\omega') := 0$, $\omega' \in N'$, defines a function $g \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega', \mathcal{F}', \mathbb{P}_Y)$ that is a factorised conditional expectation of $f(X, Y)$ given Y.

In this case, we have in particular that

$$\mathbb{E}[f(X, Y)||Y = \omega'] = \mathbb{E}[f(X, \omega')] \qquad \mathbb{P}_Y\text{-a.a. } \omega' \in \Omega'.$$

2 Conditional Distributions

Definition 2.1 (Probability kernel) *Let* $(\Omega_1, \mathcal{F}_1)$ *and* $(\Omega_2, \mathcal{F}_2)$ *be mb. spaces. A map* $\mathfrak{p} : \Omega_1 \times \mathcal{F}_2 \to \overline{\mathbb{R}}$ *is called* ***(probability) kernel from*** $(\Omega_1, \mathcal{F}_1)$ ***to*** $(\Omega_2, \mathcal{F}_2)$ *if it satistfies*

(K1) $\mathfrak{p}(\cdot, A_2) \in \mathcal{L}_{\overline{\mathbb{R}}}(\Omega_1, \mathcal{F}_1)$ *for all* $A_2 \in \mathcal{F}_2$*, and*

(K2) $\mathfrak{p}(\omega_1, \cdot) \in \mathcal{M}_1(\Omega_2, \mathcal{F}_2)$ *for all* $\omega_1 \in \Omega_1$.

Proposition 2.1 (Product of kernels) *Let* $(\Omega_j, \mathcal{F}_j)$, $j = 1, 2, 3$*, be mb. spaces. Let* $\mathfrak{p}_{2|1}$ *be a kernel from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2, \mathcal{F}_2)$*, and* $\mathfrak{p}_{3|1,2}$ *be a kernel from* $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ *to* $(\Omega_3, \mathcal{F}_3)$*. Then the right-hand side of*

$$\begin{aligned}
&\mathfrak{p}_{2|1} \otimes \mathfrak{p}_{3|1,2}(\omega_1, A_{2,3}) \\
&\qquad := \int_{\Omega_2} \int_{\Omega_3} \mathbb{1}_{A_{2,3}}\big((\omega_2, \omega_3)\big)\, \mathfrak{p}_{3|1,2}\big((\omega_1, \omega_2), \mathrm{d}\omega_3\big)\, \mathfrak{p}_{2|1}(\omega_1, \mathrm{d}\omega_2) \\
&\qquad = \int_{\Omega_2} \mathfrak{p}_{3|1,2}\big((\omega_1, \omega_2), (A_{2,3})_{\omega_2}\big)\, \mathfrak{p}_{2|1}(\omega_1, \mathrm{d}\omega_2)
\end{aligned}$$

is well defined for all $\omega_1 \in \Omega_1$, $A_{2,3} \in \mathcal{F}_2 \otimes \mathcal{F}_3$*, and the resulting map*

$$\mathfrak{p}_{2|1} \otimes \mathfrak{p}_{3|1,2} : \Omega_1 \times (\mathcal{F}_2 \otimes \mathcal{F}_3) \to \overline{\mathbb{R}}_+$$

is a kernel from $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2 \times \Omega_3, \mathcal{F}_2 \otimes \mathcal{F}_3)$.

Definition 2.2 (Product of kernels) *In the framework of Proposition 2.1, the kernel* $\mathfrak{p}_{2|1} \otimes \mathfrak{p}_{3|1,2}$ *from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2 \times \Omega_3, \mathcal{F}_2 \otimes \mathcal{F}_3)$ *defined is called* ***product of*** $\mathfrak{p}_{2|1}$ ***and*** $\mathfrak{p}_{3|1,2}$.

Corollary 2.0.1 (Concatenation of kernels) *Let* $(\Omega_j, \mathcal{F}_j)$, $j = 1, 2, 3$*, be mb. spaces. Let* $\mathfrak{p}_{2|1}$ *be a kernel from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2, \mathcal{F}_2)$*, and* $\mathfrak{p}_{3|2}$ *be a kernel from* $(\Omega_2, \mathcal{F}_2)$ *to* $(\Omega_3, \mathcal{F}_3)$.

sThen the map $\mathfrak{p}_{2|1}\mathfrak{p}_{3|2} : \Omega_1 \times \mathcal{F}_3 \to \overline{\mathbb{R}}_+$ *defined by*

$$\mathfrak{p}_{2|1}\mathfrak{p}_{3|2}(\omega_1, A_3) := \int_{\Omega_2} \mathfrak{p}_{3|2}(\omega_2, A_3)\, \mathfrak{p}_{2|1}(\omega_1, \mathrm{d}\omega_2)$$

is a kernel from $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_3, \mathcal{F}_3)$.

Definition 2.3 (Concatenation of kernels) *In the framework of Corollary 2.0.1, the kernel* $\mathfrak{p}_{2|1}\mathfrak{p}_{3|2}$ *from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_3, \mathcal{F}_3)$ *defined is called* ***concatenation of*** $\mathfrak{p}_{2|1}$ ***and*** $\mathfrak{p}_{3|2}$.

Corollary 2.0.2 (Concatenation with a measure) *Let* $(\Omega_1, \mathcal{F}_1)$ *and* $(\Omega_2, \mathcal{F}_2)$ *be measurable spaces. Let* μ_1 *be a probability measure on* $(\Omega_1, \mathcal{F}_1)$*, and* $\mathfrak{p}_{2|1}$ *be a kernel* *from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2, \mathcal{F}_2)$.

Then the map $\mu_1 \mathfrak{p}_{2|1} : \mathcal{F}_2 \to \overline{\mathbb{R}}_+$ *defined by*

$$\mu_1 \mathfrak{p}_{2|1}[A_2] := \int_{\Omega_1} \mathfrak{p}_{2|1}(\omega_1, A_2)\, \mu_1(\mathrm{d}\omega_1)$$

is a probability measure on $(\Omega_2, \mathcal{F}_2)$.

Definition 2.4 (Concatenation with a measure) *In the framework of Corollary 2.0.2, the probability measure* $\mu_1 \mathfrak{p}_{2|1}$ *from* $(\Omega_2, \mathcal{F}_2)$ *defined is called* ***concatenation of*** μ_1 ***and*** $\mathfrak{p}_{2|1}$.

Corollary 2.0.3 (Product with a measure) *Let* $(\Omega_1, \mathcal{F}_1)$ *and* $(\Omega_2, \mathcal{F}_2)$ *be measurable spaces. Let* μ_1 *be a probability measure on* $(\Omega_1, \mathcal{F}_1)$*, and* $\mathfrak{p}_{2|1}$ *be a kernel from* $(\Omega_1, \mathcal{F}_1)$ *to* $(\Omega_2, \mathcal{F}_2)$*. Then the right-hand side of*

$$\begin{aligned} \mu_1 \otimes \mathfrak{p}_{2|1}[A_{1,2}] &:= \int_{\Omega_1} \int_{\Omega_2} \mathbb{1}_{A_{1,2}}\big((\omega_1, \omega_2)\big)\, \mathfrak{p}_{2|1}(\omega_1, \mathrm{d}\omega_2)\, \mu_1(\mathrm{d}\omega_1) \\ &= \int_{\Omega_2} \mathfrak{p}_{2|1}\big(\omega_1, (A_{1,2})_{\omega_1}\big)\, \mu_1(\mathrm{d}\omega_1) \end{aligned}$$

is well defined for all $A_{1,2} \in \mathcal{F}_1 \otimes \mathcal{F}_2$*, and the resulting map*

$$\mu_1 \otimes \mathfrak{p}_{2|1} : \mathcal{F}_1 \otimes \mathcal{F}_2 \to \overline{\mathbb{R}}_+$$

is a probability measure on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$.

Definition 2.5 (Product with a measure) *In the framework of Corollary 2.0.3, the probability measure* $\mu_1 \otimes \mathfrak{p}_{2|1}$ *on* $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ *defined is called* ***product of*** μ_1 ***and*** $\mathfrak{p}_{2|1}$.

Remark 2.1 (Tonelli) *Integration and the product/concatenation of a kernel is commutative.*

Remark 2.2 (Associativity) *Products and concatenations of kernels are associative.*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(E, \mathcal{E})$.

Definition 2.6 (Conditional distributions) *A kernel* P *from* $(\Omega, \mathcal{G})$ *to* $(E, \mathcal{E})$ *is called* ***conditional distribution of*** X ***given*** $\mathcal{G}$ *if for every fixed* $B \in \mathcal{E}$,

$$P(\cdot, B) \text{ is a conditional probability of } \{X \in B\} \text{ given } \mathcal{G}.$$

If $\mathcal{G}$ *is generated by a random variable* Y *on* $(\Omega, \mathcal{F}, \mathbb{P})$*, then we speak of a* ***conditional distribution of*** X ***given*** Y.

Remark 2.3 (Notation) *If P is a conditional distribution of X of $\mathcal{G}$, then we also write $\mathbb{P}_{X|\mathcal{G}}$ instead of P.*

Theorem 2.1 (Uniqueness) *For any two conditional distributions P_1, P_2 of X given $\mathcal{G}$, we have*

(i) $P_1(\omega, B) = P_2(\omega, B)$ $\mathbb{P}$-a.e. $\omega \in \Omega$, for all $B \in \mathcal{E}$, and

(ii) $P_1(\omega, B) = P_2(\omega, B)$ for all $B \in \mathcal{E}$, $\mathbb{P}$-a.e. $\omega \in \Omega$, if $\mathcal{E} = \sigma(\mathcal{E}_0)$

for some countable system $\mathcal{E}_0 \subseteq \mathcal{E}$ being closed under intersections.

Theorem 2.2 (Existence) *If E is a complete and separable metric space and $\mathcal{E} = \mathcal{B}(E)$, then a conditional distribution of X given $\mathcal{G}$ exists.*

Theorem 2.3 *Let P be a conditional distribution of X given $\mathcal{G}$, and X' be a $(\mathcal{G}, \mathcal{E}')$-measurable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(E', \mathcal{E}')$. Moreover, let $f \in \mathcal{L}_{\overline{\mathbb{R}}}(E \times E', \mathcal{E} \otimes \mathcal{E}')$ with $f(X, X') \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$.*

Then there exists a $\mathbb{P}_{|\mathcal{G}}$-null set $N \in \mathcal{G}$, such that,

$$Z(\omega) := \int_E f\big(x, X'(\omega)\big)\, P(\omega, \mathrm{d}x)\,,\ \omega \in N^c\,,\ \text{and}\quad Z(\omega) := 0\,,\ \omega \in N\,,$$

defines a conditional expectation of $f(X, X')$ given $\mathcal{G}$. In particular,

$$\mathbb{E}[f(X, X')|\mathcal{G}](\omega) = \int_E f\big(x, X'(\omega)\big)\, \mathbb{P}_{X|\mathcal{G}}(\omega, \mathrm{d}x) \qquad \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Let now $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(E, \mathcal{E})$.
Let Y be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(\Omega', \mathcal{F}')$.

Definition 2.7 (Factorised conditional distributions) *A kernel $\mathfrak{p}$ from $(\Omega', \mathcal{F}')$ to $(E, \mathcal{E})$ is called **factorised conditional distribution of** X **given** Y if*

$$\mathfrak{p}(Y(\cdot), \cdot) \text{ is a conditional distribution of } X \text{ given } Y.$$

Remark 2.4 (Relation to factorised conditional probabilities) *A kernel $\mathfrak{p}$ from $(\Omega', \mathcal{F}')$ to $(E, \mathcal{E})$ is a factorised conditional distribution of X given Y if and only if for every $B \in \mathcal{E}$,*

$$\mathfrak{p}(\cdot, B) \text{ is a factorised conditional probabilty of } X \in B \text{ given } Y.$$

Remark 2.5 (Uniqueness) *Uniqueness trivially follows from the uniqueness of conditional distributions.*

Theorem 2.4 (Existence) *If E is a complete and separable metric space and $\mathcal{E} = \mathcal{B}(E)$, then a factorised conditional distribution of X given Y exists.*

Remark 2.6 (Notation) *If $\mathfrak{p}$ is a factorised conditional distribution of X given Y, then we also write $\mathbb{P}_{X||Y}$ instead of $\mathfrak{p}$.*

Corollary 2.4.1 (Case distinction formula) *For every factorised conditional distribution $\mathbb{P}_{X||Y}$ of X given Y we have*

$$\mathbb{P}[\{X \in B\}] = \int_{\Omega'} \mathbb{P}_{X||Y}(\omega', B)\, \mathbb{P}_Y(\mathrm{d}\omega') \qquad \text{for all } B \in \mathcal{E}.$$

Proposition 2.2 (Construction from densities) *Let X' be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E', \mathcal{E}')$. Let μ and μ' be σ-finite measures on $(E, \mathcal{E})$ and $(E', \mathcal{E}')$, respectively. Assume that $\mathbb{P}_{(X,X')}$ possesses a $\mu \otimes \mu'$-density $f_{(X,X')}$.*

Define the map $f_{X||X'} : E' \times E \to \overline{\mathbb{R}}_+$ by

$$f_{X||X'}(x', x) := \begin{cases} \frac{f_{(X,X')}(x,x')}{f_{X'}(x')} & \text{if } x' \in \{f_{X'} > 0\} \\ 0 & \text{else.} \end{cases}$$

Then the map $\mathfrak{p} : E' \times \mathcal{E} \to \overline{\mathbb{R}}_+$ defined by

$$\mathfrak{p}(x', B) := \begin{cases} \big(f_{X||X'}(x', \cdot)\mu\big)[B] & \text{if } x' \in \{f_{X'} > 0\} \\ \nu & \text{else} \end{cases}$$

(for arbitrary $\nu \in \mathcal{M}_1(E, \mathcal{E})$) is a factorised conditional distribution of X given X'.

Theorem 2.5 *Let X' be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E', \mathcal{E}')$. Let $\mathfrak{p}$ be a factorised conditional distribution of X given X'. Let $f \in \mathcal{L}_{\overline{\mathbb{R}}}(E \times E', \mathcal{E} \otimes \mathcal{E}')$ be such that $f(X, X') \in \mathcal{L}^1_{\overline{\mathbb{R}}}(\Omega, \mathcal{F}, \mathbb{P})$.*

Then there exists a $\mathbb{P}_{X'}$-null set $N' \in \mathcal{E}'$ such that

$$g(x') := \int_E f(x, x')\, \mathfrak{p}(x', \mathrm{d}x)\,,\ x' \in N'^c\,,\ \text{and} \quad g(x') := 0\,,\ x' \in N'\,,$$

defines a factorised conditional expectation of $f(X, X')$ given X'. In particular,

$$\mathbb{E}[f(X, X')||X' = x'] = \int_E f(x, x')\, \mathbb{P}_{X||X'}(x', \mathrm{d}x) \qquad \mathbb{P}_{X'}\text{-a.e. } x' \in E'.$$

Theorem 2.6 (Multilevel models) *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and for every $j = 1, \ldots, k$ let X_j be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(E_j, \mathcal{E}_j)$.*

Assume that for every $j = 2, \ldots, k$ there exists a factorised conditional distribution of X_j given $(X_1, \ldots, X_{j-1})$. Then we have

(i) $\mathbb{P}_{(X_1,\ldots,X_k)} = \mathbb{P}_{(X_1,\ldots,X_{k-1})} \otimes \mathbb{P}_{X_k||(X_1,\ldots,X_{k-1})}$, *and*

(ii) $\mathbb{P}_{(X_1,\ldots,X_k)} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2||X_1} \otimes \mathbb{P}_{X_3||(X_1,X_2)} \otimes \cdots \otimes \mathbb{P}_{X_k||(X_1,\ldots,X_{k-1})}$.

3 Probability Measures on Infinite Product Spaces

Let T be a non-empty set. Let $(E_t, \mathcal{E}_t)$ be a measurable space, $t \in T$.

The **cartesian product** of E_t, $t \in T$, is defined by

$$\bigtimes_{t\in T} E_t := \{(x_t)_{t\in T} : x_t \in E_t \text{ for all } t \in T\}.$$

For two subsets $S \subseteq U \subseteq T$, we defined the map

$$\begin{cases} \pi_{U;S} : \bigtimes_{t\in U} E_t \longrightarrow \bigtimes_{t\in S} E_t \\ (x_t)_{t\in U} \longmapsto (x_t)_{t\in S}. \end{cases}$$

If $S = \{t\}$, then we also write $\pi_{U;t}$ instead of $\pi_{U;\{t\}}$. If $U = T$, then we also write π_S and π_t instead of $\pi_{T;S}$ and $\pi_{T;\{t\}}$, respectively.

For any set Ω, we define

$$\begin{aligned} \mathfrak{P}_{\text{fin}}(\Omega) &:= \text{system of all finite subsets of } \Omega, \\ \mathfrak{P}_{\text{cou}}(\Omega) &:= \text{system of all countable subsets of } \Omega, \\ \mathfrak{P}_{\circ,\text{fin}}(\Omega) &:= \text{system of all non-empty and finite subsets of } \Omega, \text{ and} \\ \mathfrak{P}_{\circ,\text{cou}}(\Omega) &:= \text{system of all non-empty and countable subsets of } \Omega. \end{aligned}$$

Definition 3.1 (Product σ-algebra) *Let $S, U \in \mathfrak{P}_\circ(T)$ with $S \subseteq U$. The **product σ-algebra** $\bigotimes_{U;t\in S} \mathcal{E}_t$ of $\mathcal{E}_t$, $t \in S$, on $\bigtimes_{t\in U} E_t$ is defined by*

$$\bigotimes_{U;t\in S} \mathcal{E}_t := \sigma\left(\bigcup_{t\in S} \pi_{U;t}^{-1}(\mathcal{E}_t)\right).$$

Theorem 3.1 *Let $\mu, \mu' \in \mathcal{M}_1(\bigtimes_{t\in T} E_t, \bigotimes_{T;t\in T} \mathcal{E}_t)$. Then we have $\mu = \mu'$ if and only if*

$$\mu \circ \pi_{T;S}^{-1} = \mu' \circ \pi_{T;S}^{-1} \quad \text{for all } S \in \mathfrak{P}_{\circ,\text{fin}}(T).$$

Definition 3.2 (Projective family) *For every $S \in \mathfrak{P}_{\circ,\text{fin}}(T)$, let*

$$\mu_S \in \mathcal{M}_1\left(\bigtimes_{t\in S} E_t, \bigotimes_{S;t\in S} \mathcal{E}_t\right).$$

*Then $\{\mu_S\}_{S\in\mathfrak{P}_{\circ,\text{fin}}(T)}$ is called **projective family** if it satisfies the consistency condition*

$$\mu_S = \mu_U \circ \pi_{U;S}^{-1} \quad \text{for all } S, U \in \mathfrak{P}_{\circ,\text{fin}}(T) \text{ with } S \subseteq U. \tag{C1}$$

Theorem 3.2 (Kolmogorov's extension theorem) *For every $t \in T$, let E_t be a complete and separable metric space and $\mathcal{E}_t$ be the corresponding Borel σ-algebra.*

Then, for every projective family $\{\mu_S\}_{S\in\mathfrak{P}_{\circ,\mathrm{fin}}(T)}$ *of probability measures*

$$\mu_S \in \mathcal{M}_1\Big(\bigtimes_{t\in S} E_t, \bigotimes_{S;t\in S} \mathcal{F}_t\Big)$$

there exists exactly one probability measure

$$\mu \in \mathcal{M}_1\Big(\bigtimes_{t\in T} E_t, \bigotimes_{T;t\in T} \mathcal{E}_t\Big)$$

such that

$$\mu \circ \pi_{T;S}^{-1} = \mu_S \qquad \text{for all } S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T).$$

Definition 3.3 (Product measure) *For every* $t \in T$*, let* $\mu_t \in \mathcal{M}_1(E_t, \mathcal{E}_t)$*.*

Then, a probability measure $\mu \in \mathcal{M}_1(\bigtimes_{t\in T} E_t, \bigotimes_{t\in T} \mathcal{E}_t)$ *is called* ***product measure of*** μ_t***,*** $t \in T$*, if it satistfies*

$$\mu \circ \pi_{T;S}^{-1} = \bigotimes_{t\in S} \mu_t \quad \text{for all } S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T).$$

In this case, we also write $\bigotimes_{t\in T} \mu_t$ *instead of* μ*.*

Theorem 3.3 (Uniqueness) *For every* $t \in T$*, let* $\mu_t \in \mathcal{M}_1(E_t, \mathcal{E}_t)$*.*
Then there exists at most one product measure of $\mu_t, t \in T$*.*

Theorem 3.4 (Existence) *For every* $t \in T$*, let* E_t *be a complete and separable metric space,* $\mathcal{E}_t$ *be the corresponding Borel* σ*-algebra and* $\mu_t \in \mathcal{M}_1(E_t, \mathcal{E}_t)$*.*
Then there exists (exactly) one product measure of $\mu_t, t \in T$*.*

Corollary 3.4.1 *For every* $t \in T$*, let* E_t *be a complete and separable metric space,* $\mathcal{E}_t$ *be the corresponding Borel* σ*-algebra and* $\mu_t \in \mathcal{M}_1(E_t, \mathcal{E}_t)$*.*

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *and independent random variables* $X_t, t \in T$*, on* $(\Omega, \mathcal{F}, \mathbb{P})$ *whose distributions* $\mathbb{P}_{X_t}, t \in T$*, coincide with* μ_t *on* $(E_t, \mathcal{E}_t), t \in T$*.*

Corollary 3.4.2 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. For every* $t \in T$*, let* X_t *be random variables on* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking vlaues in* $(E_t, \mathcal{E}_t)$*. Then the random variables* $X_t, t \in T$*, are independent if and only if* $\mathbb{P}_{(X_t)_{t\in T}} = \bigotimes_{t\in T} \mathbb{P}_{X_t}$*.*

4 Foundations of Stochastic Processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
Let $(E, \mathcal{E})$ be a measurable space and $T \neq \varnothing$ be a set.

Definition 4.1 (Stochastic process) *A map $X : T \times \Omega \to E$ is called **stochastic process** on $(\Omega, \mathcal{F}, \mathbb{P})$ with **state space** $(E, \mathcal{E})$ and **index set** T if for every $t \in T$ the map $\omega \mapsto X(t, \omega)$ is $(\mathcal{F}, \mathcal{E})$-measurable.*

*In this case, one also speaks of an $(E, \mathcal{E})$-**valued process** on $(\Omega, \mathcal{F}, \mathbb{P})$ with **index set** T.*

Remark 4.1 (Terminology) *Let $X = (X_t)_{t \in T}$ be an $(E, \mathcal{E})$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T. We will use the following terminology.*

*For every fixed $(t, \omega) \in T \times \Omega$, $X_t(\omega)$ is called **state** of X at t given outcome ω.*

*For every fixed $t \in T$, X_t is called t-**coordinate** of X.*

*For every fixed $\omega \in \Omega$, $(X_t(\omega))_{t \in T}$ is called **path** (or **trajectory**) of X given outcome ω.*

*In particular, E^T is called **paths space** of X.*

Theorem 4.1 (Random variables and processes) *For a map $X : T \times \Omega \to E$ the following conditions are equivalent.*

(i) The map $\omega \mapsto (X_t(\omega))_{t \in T}$ is $(\mathcal{F}, \mathcal{E}^{\otimes T})$-measurable.

(ii) The map $\omega \mapsto \pi_S(X(\omega)) = (X_t(\omega))_{t \in S}$ is $(\mathcal{F}, \mathcal{E}^{S;\otimes S})$-measurable for every $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$.

(iii) The map $\omega \mapsto \pi_t(X(\omega)) = X_t(\omega)$ is $(\mathcal{F}, \mathcal{E})$-measurable for every $t \in T$.

Remark 4.2 *Thus, every $(E, \mathcal{E})$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T is an $(E^T, \mathcal{E}^{\otimes T})$-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, and vice versa.*

Definition 4.2 (Distribution of a process) *Let $X = (X_t)_{t \in T}$ be an $(E, \mathcal{E})$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T.*

*Then, $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ is called **distribution of** X. Moreover, for every $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$, $\mathbb{P}_{\pi_S(X)} = \mathbb{P}_{(X_t)_{t \in S}}$ is called **finite-dimensional distribution of** X **with base** S.*

Theorem 4.2 (Characterisation by finite projections) *Let $X = (X_t)_{t \in T}$ and $X' = (X'_t)_{t \in T}$ be two $(E, \mathcal{E})$-valued processes with index set T on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$, respectively.*
Then, $\mathbb{P}_X = \mathbb{P}'_{X'}$, if and only if $\mathbb{P}_{\pi_S(X)} = \mathbb{P}'_{\pi_S(X')}$ for all $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$.

Theorem 4.3 (Existence) *Let E be a complete and separable metric space and $\mathcal{E}$ be the corresponding Borel σ-algebra.*

For every projective family $\{\mu_S\}_{S\in\mathfrak{P}_{\circ,\mathrm{fin}}(T)\}}$ there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an $(E,\mathcal{E})$-valued process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T such that $\mathbb{P}_{\pi_S(X)} = \mu_S$ for all $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$.

Definition 4.3 *A process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to possess* ***independent coordinates*** *if the family $\{X_t\}_{t\in T}$ of random variables is independent.*

Corollary 4.3.1 *For a process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ the following statements are equivalent.*

(i) X possesses independent coordinates.

(ii) $\mathbb{P}_{(X_t)_{t\in T}} = \bigotimes_{t\in T} \mathbb{P}_{X_t}$.

(iii) $\mathbb{P}_{(X_t)_{t\in S}} = \bigotimes_{t\in S} \mathbb{P}_{X_t}$ for all $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$.

Corollary 4.3.2 *Let E be a complete and separable metric space and $\mathcal{E}$ be the corresponding Borel σ-algebra. Moreover, let $\mu_t \in \mathcal{M}_1(E, \mathcal{E}), t \in T$.*

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an $(E,\mathcal{E})$-valued process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T that possesses independent coordinates and that satisfies $\mathbb{P}_{X_t} = \mu_t$ for all $t \in T$.

Example 4.1 *For every $\mu_1 \in \mathcal{M}_1(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we can find by Corollary 4.3.2 a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a real-valued process $X = (X_j)_{j\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with independent coordinates and $\mathbb{P}_{X_j} = \mu_1$ for all $j \in \mathbb{N}$.*

Obviously, this process is nothing but a sequence of i.i.d. random variables with $\mathbb{P}_{X_1} = \mu_1$.

If in addition $\int |x|\,\mu_1(\mathrm{d}x) < \infty$, then each of the random variables X_j, $j \in \mathbb{N}$, is contained in $\mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathbb{E}[X_1] = 0$, then the real-valued process $S = (S_n)_{n\in\mathbb{N}_0}$ defined by

$$S_0 := 0 \qquad \text{and} \qquad S_n := \sum_{j=1}^{n} X_j\ ,\ n \in \mathbb{N}$$

is called ***symmetric random walk*** *with innovation distribution μ_1.*

For a random variable with a "simple" co-domain as, e.g., $\mathbb{R}$, the distribution $\mathbb{P}_X$ provides "full" information on the behavior of X.

The situation is different when the co-domain of X is more complicated. Some qualitative properties of paths may differ significatly for two processes, even with the same distribution.

Example 4.2 *Let* $(\Omega, \mathcal{F}, \mathbb{P}) \overset{!}{=} ([0,1], \mathcal{B}_{[0,1]}(\mathbb{R}), \ell_{|[0,1]})$ *and* $T := [0,1]$. *Let two real-valued processes* $X = (X_t)_{t \in T}$ *and* $X' = (X'_t)_{t \in T}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *be defined by* $X_t(\omega) := 0$ *and* $X'_t(\omega) := \mathbb{1}_{\{\omega\}}(t)$.
Then, $\mathbb{P}_X = \mathbb{P}_{X'}$ *but each path of* X *is continuous, whereas each path of* X' *has a discontinuity.*

Theorem 4.4 (Inadequacy of the product σ-algebra) *Let* $(E, \mathcal{E})$ *be a measurable space and* T *be a non-empty set. Then, for every* $A \in \mathcal{E}^{\otimes T}$ *there exists an* $S \in \mathfrak{P}_{\circ,\mathrm{cou}}(T)$ *such that for all* $f, g \in E^T$,

$$f = g \text{ on } S\,,\ f \in A \quad \Longrightarrow \quad g \in A\,.$$

Remark 4.3 *For each countable set* $S \subseteq [0,1]$ *and each* $f \in C([0,1])$ *there are of course plenty of functions* $g \in \mathbb{R}^{[0,1]} \setminus C([0,1])$ *with* $f = g$ *on* S. *Thus, by contraposition of Theorem 4.4, the set* $C([0,1])$ *can indeed not be contained in* $\mathcal{B}(\mathbb{R})^{\otimes[0,1]}$.
The same is true, e.g., for the set of all monotone functions from $\mathbb{R}^{[0,1]}$.

Definition 4.4 (Modifications, indistiguishability) *Let* $X = (X_t)_{t \in T}$ *and* $Y = (Y_t)_{t \in T}$ *be two* $(E, \mathcal{E})$*-valued processes on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* T.

The processes X *and* Y *are called* ***modifications*** *of each other if*

$$X_t(\omega) = Y_t(\omega) \quad \text{for } \mathbb{P}\text{-a.a. } \omega \in \Omega, \text{ for all } t \in T.$$

The processes X *and* Y *are called* ***indistiguishable*** *if*

$$X_t(\omega) = Y_t(\omega) \quad \text{for all } t \in T, \text{ for } \mathbb{P}\text{-a.a. } \omega \in \Omega.$$

Theorem 4.5 *Let* $X = (X_t)_{t \in T}$ *and* $Y = (Y_t)_{t \in T}$ *be two* $(E, \mathcal{E})$*-valued processes on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* T. *Then*

(i) X, Y *indistiguishable* $\Longrightarrow$ X, Y *modifications of each other.*

(ii) X, Y *modifications of each other* $\Longrightarrow$ $\mathbb{P}_X = \mathbb{P}_Y$.

(iii) X, Y *modifications of each other* $\Longrightarrow$ X, Y *indistinguishable, if one of the following two conditions is met*

  (a) T *is countable.*

  (b) T *is an interval in* $\mathbb{R}$, E *is a separable metric space,* $\mathcal{E}$ *is the corresponding Borel* σ*-algebra, and* $\mathbb{P}$*-a.a. paths of* X *and* Y *are left-continuous of* $\mathbb{P}$*-a.a. paths of* X *and* Y *are right-continuous.*

Theorem 4.6 (Kolmogorov-Chentsov) *Let* (E, d) *be a complete and separable metric space and* $\mathcal{E}$ *be the corresponding Borel* σ*-algebra. Moreover, let* T *be a finite union of bounded intervals in* $\mathbb{R}^m$ *and* $X = (X_t)_{t \in T}$ *be an* $(E, \mathcal{E})$*-valued process on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* T.

If we can find constants $q, \delta, C \in \mathbb{R}_{>0}$ *such that*

$$\mathbb{E}[d(X_s, X_t)^q] \leq C \|s-t\|^{m+\delta} \qquad \text{for all } s, t \in T,$$

then there exists a modification of X *whose paths are all continuous.*

The modification can even be chosen such that for every $\gamma \in (0, \delta/q)$ *the paths of the modification are all Hölder-*γ*-continuous.*

Definition 4.5 (Product measurable) *Assume that* T *is equipped with a* σ*-algebra* $\mathcal{T}$*, and let* X *be an* $(E, \mathcal{E})$*-valued process* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* T*.*

The process X *is called* ***product measurable*** *if the map* $X : T \times \Omega \to E$ *is* $(\mathcal{T} \otimes \mathcal{F}, \mathcal{E})$*-measurable.*

Proposition 4.1 *Let* T *be an interval in* $\mathbb{R}$ *and* $X = (X_t)_{t \in T}$ *be a real-valued process* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* T*.*

If the paths of X *are all left-continuous or all right-continuous, and* T *is equipped with the Borel* σ*-algebra* $\mathcal{B}(T) = \mathcal{B}(T) \cap T$*, then* X *is product measurable.*

Definition 4.6 (Strictly stationary) *Let* $(E, \mathcal{E})$ *be a measurable space and* $T \in \mathfrak{P}_\circ(\mathbb{R})$ *be closed under addtion. Then, an* $(E, \mathcal{E})$*-valued process* $X = (X_t)_{t \in T}$ *is called* ***strictly stationary*** *if*

$$\mathbb{P}_{(X_{t_1+s}, \ldots, X_{t_k+s})} = \mathbb{P}_{(X_{t_1}, \ldots, X_{t_k})}$$

for all $k \in \mathbb{N}$ *and* $s, t_1, \ldots, t_k \in T$*.*

Definition 4.7 (Independent increments) *Let* $T \in \mathfrak{P}_\circ(\mathbb{R})$*. Then, an* $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$*-valued process* $X = (X_t)_{t \in T}$ *is said to possess* ***independent increments*** *if for every* $k \geq 2$ *and* $t_0, \ldots, t_k \in T$ *with* $t_0 < t_1 < \cdots < t_k$ *the random variables*

$$X_{t_1} - X_{t_0}, \ldots, X_{t_k} - X_{t_k-1} \quad \text{are independent.}$$

Definition 4.8 (Increments independent of the intitial state) *Let* $T \in \mathfrak{P}_\circ(\mathbb{R})$ *be such that* $\sigma := \inf T$ *is contained in* T*. Then, an* $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$*-valued process* $X = (X_t)_{t \in T}$ *is said to possess* ***independent increments that are independent independent of the initial state*** *if for every* $k \in \mathbb{N}$ *and* $t_0, \ldots, t_k \in T$ *with* $\sigma < t_1 < t_1 < \cdots < t_k$ *the random variables*

$$X_\sigma, X_{t_1} - X_\sigma, \ldots, X_{t_k} - X_{t_k-1} \quad \text{are independent.}$$

Definition 4.9 (Stationary increments) *Let* $T \in \mathfrak{P}_\circ(\mathbb{R})$ *be closed under addtion. Then, an* $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$*-valued process* $X = (X_t)_{t \in T}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is said to possess* ***stationary increments*** *if for every* $k \in \mathbb{N}$ *and* $s, t_0, \ldots, t_k \in T$ *with* $t_0 < \cdots < t_k$*, we have*

$$\mathbb{P}_{(X_{t_1+s} - X_{t_0+s}, \ldots, X_{t_k+s} - X_{t_{k-1}+s}} = \mathbb{P}_{(X_{t_1} - X_{t_0}, \ldots, X_{t_k} - X_{t_{k-1}})}.$$

Lemma 4.1 *Let $T \in \mathfrak{P}_\circ(\mathbb{R})$ be closed under addition and $\Delta T := \{t - s : s, t \in T \text{ with } s < t\}$. Moreover, let $X = (X_t)_{t\in T}$ be an $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$-valued process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

If X possesses independent increments and for every $h \in \Delta T$ there exists a $\mu_h \in \mathcal{M}_1(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, such that, $\mathbb{P}_{X_t - X_s} = \mu_h$ for all $s, t \in T$ with $t - s = h$, then X possesses stationary increments.

Definition 4.10 (Process, centered) *Let T be a non-empty set and $p \in [1, \infty)$.*

An $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process $(X_t)_{t\in T}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called ***$\mathcal{L}^p_\mathbb{R}$-process*** *if $X_t \in \mathcal{L}^p_\mathbb{R}(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in T$.*

An $\mathcal{L}^p_\mathbb{R}$-process $(X_t)_{t\in T}$ is called ***centered*** *if $\mathbb{E}[X_t] = 0$ for all $t \in T$.*

Definition 4.11 *Let $X = (X_t)_{t\in T}$ be a real-valued process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

If X is an $\mathcal{L}^1_\mathbb{R}$-process, the its ***mean function*** *$m_X : T \to \mathbb{R}$ is defined by*

$$m_X(t) := \mathbb{E}[X_t]\,.$$

If X is an $\mathcal{L}^2_\mathbb{R}$-process, then its ***covariation function*** *$\gamma_X : T \times T \to \mathbb{R}$ and its variance function $v_X : T \to \mathbb{R}$ are defined by*

$$\gamma_X(s,t) := \operatorname{Cov}(X_s, X_t) \quad \textit{and} \quad v_X(t) := \operatorname{Var}(X_t)\,, \ \textit{respectively.}$$

Lemma 4.2 *Let $T \in \mathfrak{P}_\circ(\mathbb{R})$ such that $\sigma := \inf T \in T$. Then, if an $\mathcal{L}^2_\mathbb{R}$-process $X = (X_t)_{t\in T}$ possesses independent increments that are independent of the initial state, we have*

$$\gamma_X(s,t) = v_X(\min\{s,t\}) \quad \textit{for all } s,t \in T.$$

Definition 4.12 (Symmetric, positive semi-definite) *A map $\gamma : T \times T \to \mathbb{R}$ is* ***symmetric*** *if $\gamma(s,t) = \gamma(t,s)$ for all $s, t \in T$, and* ***positive semi-definite*** *if $\sum_{i=1}^k \sum_{k=1}^k \lambda_i \lambda_j \gamma(t_i, t_j) \geq 0$ for $k \in \mathbb{N}$, $t_1, \ldots, t_k \in T$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$.*

Lemma 4.3 *The covariance function γ_X of any $\mathcal{L}^2_\mathbb{R}$-process $X = (X_t)_{t\in T}$ is symmetric and positive semi-definite.*

Lemma 4.4 *The mean function m_X of any strictly stationary $\mathcal{L}^1_\mathbb{R}$-process $X = (X_t)_{t\in T}$ is constant.*

For the covariance function γ_X of any strictly stationary $\mathcal{L}^2_\mathbb{R}$-process $X = (X_t)_{t\in\mathbb{R}_+}$, we have

$$\gamma_X(s,t) = \gamma_X(0, |s-t|) \qquad \textit{for all } s, t\mathbb{R}_{\geq 0}.$$

5 Poisson Processes

Definition 5.1 (Poisson process) *For every $\lambda \in \mathbb{R}_{>0}$, a $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process $N = (N_t)_{t \in \mathbb{R}_{\geq 0}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with index set $\mathbb{R}_{\geq 0}$ is called* ***Poisson process with intensity*** *λ if it possesses the following properties.*

(P1) $N_0 = 0$ $\mathbb{P}$-*a.s.*

(P2) N has independent increments.

(P3) $\mathbb{P}_{N_t - N_s} = \mathrm{Poiss}_{\lambda(t-s)}$ *for all $s, t \in \mathbb{R}_{\geq 0}$ with $s < t$.*

(P4) $\mathbb{P}$-a.a. paths of N are monotonically increasing and right-continuous.

Proposition 5.1 *If $N = (N_t)_{t \in \mathbb{R}_{\geq 0}}$ is a Poisson process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then $\mathbb{P}$-a.a. paths of N take values only in $\mathbb{N}_0$.*

Remark 5.1 *Two Poisson processes with the same intensity possess the same finite dimensional distributions.*

Theorem 5.1 (Existence) *Let $\lambda \in \mathbb{R}_{>0}$ and $(W_j)_{j \in \mathbb{N}}$ be a sequence of i.i.d. real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}_{W_1} = \mathrm{Exp}_\lambda$.*

Then, using the convention $\max \varnothing := \max \mathbb{N} := 0$, we can define by

$$N_t := \max\{k \in \mathbb{N} : \sum_{j=1}^{k} W_j \leq t\}\,,\ t \in \mathbb{R}_{\geq 0}$$

a Poisson process $N = (N_t)_{t \in \mathbb{R}_{\geq 0}}$ with intensity λ.

Proposition 5.2 *Let $N = (N_t)_{t \in \mathbb{R}_{\geq 0}}$ be an $(\mathbb{R}_{\geq 0}, \mathcal{B}(\mathbb{R}_{\geq 0}))$-valued process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that satisfies conditions (P1), (P2), (P4), and assume that $\mathbb{P}$-a.a. of its paths take values only in $\mathbb{N}_0$.*

Then, N satisfies condition (P3) for a given $\lambda \in \mathbb{R}_{>0}$ if and only if it satisfies the following conditions (P5)-(P7).

(P5) N has stationary increments.

(P6) N is an $\mathcal{L}^1_{\mathbb{R}}$-process and $\mathbb{E}[N_1] = \lambda$.

(P7) $\lim_{h \searrow 0} \mathbb{P}[\{N_h \geq 2\}]/h = 0$.

Corollary 5.1.1 *Let $N = (N_t)_{t \in \mathbb{R}_{\geq 0}}$ be a Poisson process with intensity $\lambda \in \mathbb{R}_{>0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for $\mathbb{P}$-a.a. $\omega \in \Omega$, there does **not** exist any $t \in \mathbb{R}_{>0}$ such that $N_t(\omega) - \lim_{s \nearrow t} N_s(\omega) \geq 2$.*

6 Brownian Motion

Definition 6.1 *Let T be a nonempty set. A $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process $X = (X_t)_{t\in T}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T is called a **Gaussian process** if all of its finite-dimensional distributions are normal distributions, i.e., if for every $S \in \mathfrak{P}_{\circ,\mathrm{fin}}(T)$,*

$$\mathbb{P}_{\pi_S(X)} \quad \text{is a } \#S\text{-variate normal distribution.}$$

Theorem 6.1 *The distribution of a Gaussian process $X = (X_t)_{t\in T}$ is uniquely determined by the mean function and the covariance function of X.*

Theorem 6.2 (Existence) *For a non-empty set T, let $m : T \to \mathbb{R}$ be an arbitrary function and $\gamma : T \times T \to \mathbb{R}$ be a symmetric and positive semi-definite function.*

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Gaussian process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with $m_X \equiv m$ and $\gamma_X = \gamma$.

Definition 6.2 (Brownian motion) *An $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with index set $\mathbb{R}_{\geq 0}$ is called 1-dimensional **(standard) Brownian motion** if it possesses the following properties.*

(B1) $B_0 = 0$ $\mathbb{P}$-*a.s.*.

(B2) B *has independent increments.*

(B3) $\mathbb{P}_{B_t - B_s} = \mathrm{N}_{0,t-s}$ *for all* $s, t \in \mathbb{R}_{\geq 0}$ *with* $s < t$.

(B4) $\mathbb{P}$-*a.a. paths of* B *are continuous.*

Remark 6.1 *Two Brownian motions possess the same finite-dimensional distributions.*

Proposition 6.1 *Let $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

(i) $\mathbb{P}_{B_t} = \mathrm{N}_{0,t}$ *for all* $t \in \mathbb{R}_{\geq 0}$.

(ii) B *is a* $\mathcal{L}^p_{\mathbb{R}}$-*process for every* $p \in [1, \infty)$.

(iii) $m_B(t) := \mathbb{E}[B_t] = 0$ *and* $v_B(t) := \mathrm{Var}[B_t] = t$ *for all* $t \in \mathbb{R}_{\geq 0}$.

(iv) $\gamma_B(s, t) := \mathrm{Cov}(B_s, B_t) = \min\{s, t\}$ *for all* $s, t \in \mathbb{R}_{\geq 0}$.

All of these four statements are valid even without assumption (B4).

Proposition 6.2 *Let $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ be an $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that satisfies condition (B1). Then the process B satisfies (B2) & (B3) if and only if it satisfies the following condition (B5).*

$$\mathbb{P}_{(B_{t_1},\ldots,B_{t_k})} = \mathrm{N}_{\mathbf{0},V_{t_1,\ldots,t_k}} \quad \text{for all } k \in \mathbb{N} \text{ and } t_1, \ldots, t_k \in \mathbb{R}_{\geq 0} \text{ with } t_1 < \cdots < t_k, \tag{B5}$$

where $\mathbf{0} := (0, \ldots, 0)$ and $V_{t_1,\ldots,t_k} := (\min\{t_i, t_j\})_{1\leq i,j\leq k}$.

Remark 6.2 *In the framework of Proposition 6.2, condition (B5) can be written as*

B *is a centered Gaussian process with covariance funciton* $\gamma_B(s,t) = \min\{s,t\}$.

In particular, every Brownian motion possesses property (B6).

Theorem 6.3 (Existence) *There exists a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a Brownian motion* $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$.

B *can be chosen such that* $\mathbb{P}$*-a.a. of its paths are locally Hölder-*γ*-continuous for every* $\gamma \in (0, 1/2)$.

Proposition 6.3 (Transformations) *Let* $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ *be a Brownian motion on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Moreover, let* $c \in \mathbb{R} \setminus \{0\}$ *and* $s \in \mathbb{R}_{>0}$*. Then also the process* $\tilde{B} = (\tilde{B}_t)_{t\in\mathbb{R}_{\geq 0}}$ *is a Brownian motion if it is defined by*

(i) $\tilde{B}_t := \frac{1}{c} B_{c^2 t}, t \in \mathbb{R}_{\geq 0}$.

(ii) $\tilde{B}_t := -B_t, t \in \mathbb{R}_{\geq 0}$.

(iii) $\tilde{B}_t := B_{t+s} - B_s, t \in \mathbb{R}_{\geq 0}$.

(iv) $\tilde{B}_t := tB_{1/t}, t \in \mathbb{R}_{>0}$*, and* $\tilde{B}_0 := B_0$.

Theorem 6.4 (Paley-Wiener-Zygmund) *Let* $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ *be a Brownian motion on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $\gamma \in (1/2, \infty)$.

Then, $\mathbb{P}$*-a.a. paths of* B *are not Hölder-*γ*-continuous at any point.*
In particular, $\mathbb{P}$*-a.a. paths of* B *are not differentiable at any point.*

Definition 6.3 (d-dimensional Brownian motion) *An* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*-valued process* $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with index set* $\mathbb{R}_{\geq 0}$ *is called* ***d-dimensional (standard) Brownian motion*** *if it possesses the following properties.*

(B1) $B_0 = 0$ $\mathbb{P}$*-a.s..*

(B2) B *has independent increments.*

(B3) $\mathbb{P}_{B_t - B_s} = \mathrm{N}_{0,(t-s)I_d}$ *for all* $s, t \in \mathbb{R}_{\geq 0}$ *with* $s < t$.

(B4) $\mathbb{P}$*-a.a. paths of* B *are continuous.*

Theorem 6.5 *If* $B = (B^{(1)}, \ldots, B^{(d)})$ *is a d-dimensional Brownian motion on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, then its coordinates* $B^{(1)}, \ldots, B^{(d)}$ *are independent* 1*-dimensional Brownian motions.*

If, conversely, $B^{(1)}, \ldots, B^{(d)}$ *are independent* 1*-dimensional Brownian motions on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, then* $B := (B^{(1)}, \ldots, B^{(d)})$ *is a d-dimensional Brownian motion.*

Corollary 6.5.1 (Existence) *There exists a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a d-dimensional Brownian motion* $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$.

7 Markov Processes

Let $(E, \mathcal{E})$ be a measurable space.
Let $T \in \mathfrak{P}_\circ(\mathbb{R})$.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 7.1 (Filtration) *For every $t \in T$, let $\mathcal{F}_t$ be a sub-σ-algebra of $\mathcal{F}$. Then the family $(\mathcal{F}_t)_{t\in T}$ is called a **filtration in** $\mathcal{F}$ if*

$$\mathcal{F}_s \subseteq \mathcal{F}_t \qquad \text{for all } s, t \in T \text{ with } s \leq t.$$

For two filtrations $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ and $\mathbb{F}' = (\mathcal{F}'_t)_{t\in T}$ in $\mathcal{F}$, we henceforth write $\mathbb{F} \subseteq \mathbb{F}'$ if $\mathcal{F}_t \subseteq \mathcal{F}'_t$ for all $t \in T$.

Definition 7.2 (Adaptedness) *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and $X = (X_t)_{t\in T}$ be an $(E, \mathcal{E})$-valued process $(\Omega, \mathcal{F}, \mathbb{P})$.*

*The process X is said to be **adapted to the filtration** $\mathbb{F}$ (short: $\mathbb{F}$**-adapted**) if X_t is $(\mathcal{F}_t, \mathcal{E})$-measurable for every $t \in T$.*

Definition 7.3 (Natural filtration) *Let $X = (X_t)_{t\in T}$ be a $(E, \mathcal{E})$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$, and set*

$$\begin{aligned}
\mathcal{F}^X_S &:= \sigma(X_t, t \in S) \quad \text{for all } S \in \mathfrak{P}_\circ(T),\\
\mathcal{F}^X_t &:= \mathcal{F}^X_{\leq t} := \mathcal{F}^X_{(\infty,t]\cap T} \quad \text{for all } t \in T,\\
\mathcal{F}^X_{\geq t} &:= \mathcal{F}^X_{[t,\infty)\cap T} \quad \text{for all } t \in T.
\end{aligned}$$

*The filtration $\mathbb{F}^X := (\mathcal{F}^X_T)_{t\in T}$ is called **natural filtration of** X.*

Definition 7.4 (Markov process) *Let $X = (X_t)_{t\in T}$ be an $(E, \mathcal{E})$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$.*

*The process X is called **Markov process w.r.t.** $\mathbb{F}$ (short: $\mathbb{F}$**-Markov process**) if it is $\mathbb{F}$-adapted and satistfies*

$$\mathbb{P}[\{X_t \in A\}|\mathcal{F}_s] = \mathbb{P}[\{X_t \in A\}|X_s] \quad \mathbb{P}\text{-a.s. for all } s, t \in T \text{ with } s \leq t \text{ and } A \in \mathcal{E}.$$

*If $\mathbb{F}$ coincides with the natural filtration $\mathbb{F}^X$ of X, then X is simply referred to as **Markov process**.*

Proposition 7.1 (Equivalent properties) *For a $(E, \mathcal{E})$-valued process $X = (X_t)_{\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$, the following five conditions are equivalent.*

(i) For all $k \in \mathbb{N}$, $t_1, \ldots, t_k, t \in T$ with $t_1 < \cdots < t_k < t$ and $A \in \mathcal{E}$

$$\mathbb{P}[\{X_t \in A\}|(X_{t_1}, \ldots, X_{t_k})] = \mathbb{P}[\{X_t \in A\}|X_{t_k}] \quad \mathbb{P}\text{-a.s.}.$$

(ii) X *possesses the Markov property for* $\mathbb{F} := \mathbb{F}^X$.

(iii) *For all* $s \in T$ *and* $B \in \mathcal{F}^X_{\geq s}$

$$\mathbb{P}[B|\mathcal{F}^X_s] = \mathbb{P}[B|X_s] \quad \mathbb{P}\text{-}as..$$

(iv) *For all* $s \in T$, $T_s := [s, \infty) \cap T$, *and bounded* $f \in \mathcal{L}_{\mathbb{R}}(E^{T_s}, \mathcal{E}^{\otimes T_s})$

$$\mathbb{E}[f(\pi_{t_s}(X))|\mathcal{F}^X_s] = \mathbb{E}[f(\pi_{T_s}(X))|X_s] \quad \mathbb{P}\text{-}a.s..$$

(v) *For all* $s \in T$, $A \in \mathcal{F}^X_s$ *and* $B \in \mathcal{F}^X_{\geq s}$

$$\mathbb{P}[A \cap B|X_s] = \mathbb{P}[A|X_s]\mathbb{P}[B|X_s] \quad \mathbb{P}\text{-}a.s..$$

Definition 7.5 (Transition probability) *Let* $X = (X_t)_{t\in T}$ *be an* $(E, \mathcal{E})$*-valued Markov process on* $(\Omega, \mathcal{F}, \mathbb{P})$.

If, for $s, t \in T$ *with* $s \leq t$, *there exists a factorised conditional distribution of* X_t *given* X_s, *then it is referred to as* ***transition probability*** *of the Markov process* X ***from*** s ***to*** t.

Proposition 7.2 *Let* $X = (X_t)_{t\in T}$ *be an* $(E, \mathcal{E})$*-valued Markov process on* $(\Omega, \mathcal{F}, \mathbb{P})$. *Assume that for all* $s, t \in T$ *with* $s \leq t$ *there exists a factorised conditional distribution* $\mathbb{P}_{X_t||X_s}$ *of* X_t *given* X_s. *Then*

(i) $\mathbb{P}_{X_t||X_t}(x, \cdot) = \delta_x[\cdot]$ *for all* $t \in T$ *and* $x \in E$.

(ii) $\mathbb{P}_{X_t||X_s} = \mathbb{P}_{X_u||X_s}\mathbb{P}_{X_t||X_u}$ *for all* $s, u, t \in T$ *with* $s \leq u \leq t$.

Definition 7.6 *A family* $\mathfrak{P} = (\mathfrak{p}_{s,t})_{s,t\in T: s\leq t}$ *of kernels from* $(E, \mathcal{E})$ *to* $(E, \mathcal{E})$ *is called a* ***Markov transition function on*** $(E, \mathcal{E})$ *if*

(i) $\mathfrak{p}_{t,t}(x, \cdot) = \delta_x[\cdot]$ *for all* $t \in T$ *and* $x \in E$,

(ii) $\mathfrak{p}_{s,t} = \mathfrak{p}_{s,u}\mathfrak{p}_{u,t}$ *for* $s, u, t \in T$ *with* $s \leq u \leq t$.

In the framework of Definition 7.5, $\mathfrak{P}^X := (\mathbb{P}_{X_t||X_s})_{s,t\in T: s\leq t}$ *is called* ***Markov transition function of the Markov process*** X.

Lemma 7.1 (Markov property as transition probability) *Let* $X = (X_t)_{t\in T}$ *be an* $(E, \mathcal{E})$*-valued Markov process on* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $t_1, \ldots, t_k, t \in T$ *with* $t_1 \leq \cdots \leq t_k \leq t$.

Then, if there exists a factorised conditional distribution $\mathbb{P}_{X_t||X_{t_k}}$ *of* X_t *given* X_{t_k},

$$\mathbb{P}_{X_t||(X_{t_1},\ldots,X_{t_k})} := \mathbb{P}_{X_t||X_{t_k}}$$

is a factorised conditional distribution of X_t *given* $(X_{t_1}, \ldots, X_{t_k})$.

Proposition 7.3 (Finite-dimensional distributions) *Let $X = (X_t)_{t\in T}$ be an $(E, \mathcal{E})$-valued Markov process on $(\Omega, \mathcal{F}, \mathbb{P})$ and $t_0, \ldots, t_k \in T$ with $t_0 \leq \cdots \leq t_k$.*

Assume that there exists a factorised conditional distribution $\mathbb{P}_{X_{t_j}||X_{t_{j-1}}}$ of X_{t_j} given $X_{t_{j-1}}$ for every $j = 2, \ldots, k$

(i) $\mathbb{P}_{(X_{t_1},\ldots,X_{t_k})} = \mathbb{P}_{X_{t_1}} \otimes \mathbb{P}_{X_{t_2}||X_{t_1}} \otimes \cdots \otimes \mathbb{P}_{X_{t_k}||X_{t_{k-1}}}$,

and for every $j = 1, \ldots, k$

(ii) $\mathbb{P}_{(X_{t_1},\ldots,X_{t_k})} = (\mathbb{P}_{X_{t_0}}\mathbb{P}_{X_{t_1}||X_{t_0}}) \otimes \mathbb{P}_{X_{t_2}||X_{t_1}} \otimes \cdots \otimes \mathbb{P}_{X_{t_k}||X_{t_{k-1}}}$, *and*

(iii) $\mathbb{P}_{(X_{t_1},\ldots,X_{t_k})||X_{t_0}} = \mathbb{P}_{X_{t_1}||X_{t_0}} \otimes \cdots \otimes \mathbb{P}_{X_{t_k}||X_{t_{k-1}}}$.

Theorem 7.1 (Existence) *Asumme that E is a complete and separable metric space and $\mathcal{E}$ is the corresponding Borel-σ algebra. Moreover, assume that $\sigma := \inf T \in T$.*

Then, for every $\mu_\sigma \in \mathcal{M}_1(E, \mathcal{E})$ and every Markov transition function $\mathfrak{P} = (\mathfrak{p}_{s,t})_{s,t\in T:s\leq t}$ on $(E, \mathcal{E})$, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an $(E, \mathcal{E})$-valued Markov process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with initial distribution $\mathbb{P}_{X_\sigma} = \mu_\sigma$ and Markov transition function $\mathfrak{P}$.

Theorem 7.2 *Assume that $\sigma := \inf T \in T$. Let $X = (X_t)_{t\in T}$ be a $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$.*

If X possesses independent increments that are independent of the initial state, then X is a Markov process.

Example 7.1 *Let $B = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathfrak{P} = (\mathfrak{p}_{s,t})_{s,t\in\mathbb{R}_{\geq 0}:s\leq t}$ be a family of kernels from $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that are given by*

$$\mathfrak{p}_{s,t}(x, A) := \mathrm{N}_{x,t-s}[A]\,,\ (x, A) \in \mathbb{R} \times \mathcal{B}(\mathbb{R})\,.$$

Then, B is a Markov process with Markov transition function $\mathfrak{P}$.

Example 7.2 *Let $N = (N_t)_{t\in\mathbb{R}_{\geq 0}}$ be a Poisson process with intensity $\lambda \in \mathbb{R}_{>0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathfrak{P} = (\mathfrak{p}_{s,t})_{s,t\in\mathbb{R}_{\geq 0}:s\leq t}$ be a family of kernels from $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that are given by*

$$\mathfrak{p}_{s,t}(x, A) := \mathrm{Poiss}_{\lambda(t-s)}[A - x]\,,\ (x, A) \in \mathbb{R} \times \mathcal{B}(\mathbb{R})\,.$$

Then, N is a Markov process with Markov transition function $\mathfrak{P}$.

8 Martingales, Sub- and Supermartingales

Let $T \in \mathfrak{P}_\circ(\mathbb{R})$.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 8.1 (Martingale) *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and $X = (X_t)_{t\in T}$ be an $\mathcal{L}^1_{\mathbb{R}}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$. If the process X is $\mathbb{F}$-adapted then it is called*

(i) $\mathbb{F}$***-martingale*** *if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ $\mathbb{P}$-a.s. $\forall s, t \in T$ with $s \leq t$.*

(ii) $\mathbb{F}$***-submartingale*** *if $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$ $\mathbb{P}$-a.s. $\forall s, t \in T$ with $s \leq t$.*

(iii) $\mathbb{F}$***-supermartingale*** *if $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$ $\mathbb{P}$-a.s. $\forall s, t \in T$ with $s \leq t$.*

*If X is in addition an $\mathcal{L}^p_{\mathbb{R}}$-process for some $p \in (1, \infty)$. then one speaks of an $\mathcal{L}^p_{\mathbb{R}}$-$\mathbb{F}$-**(sub-, super-) martingale**.*

*If $\mathbb{F}$ is the natural filtration $\mathbb{F}^X$ of X, then one simply speaks of a **(sub-, super-) martingale**.*

Proposition 8.1 (Transformations) *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$, and $X = (X_t)_{t\in T}$ and $(Y_t)_{t\in T}$ be two $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued processes on $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

(i) X is a $\mathbb{F}$-supermartingale if and only if $-X := (-X_t)_{t\in T}$ is a $\mathbb{F}$-submartingale.

(ii) If X and Y are $\mathbb{F}$-martingales and $a, b \in \mathbb{R}$, then $aX + bY := (aX_t + bY_t)_{t\in T}$ is a $\mathbb{F}$-martingale.

(iii) If X and Y are $\mathbb{F}$-supermartingales and $a, b \in \mathbb{R}_{\geq 0}$, then $aX + bY := (aX_t + bY_t)_{t\in T}$ is a $\mathbb{F}$-supermartingale.

(iv) If X and Y are $\mathbb{F}$-submartingales and $a, b \in \mathbb{R}_{\geq 0}$, then $aX + bY := (aX_t + bY_t)_{t\in T}$ is a $\mathbb{F}$-submartingale.

(v) If X and Y are $\mathbb{F}$-supermartingales, then $X \wedge Y := (\min\{X_t, Y_t\})_{t\in T}$ is a $\mathbb{F}$-supermartingale.

(vi) If X and Y are $\mathbb{F}$-submartingales, then $X \wedge Y := (\max\{X_t, Y_t\})_{t\in T}$ is a $\mathbb{F}$-submartingale.

(vii) If X is an $\mathbb{F}$-martingale and $f : \mathbb{R} \to \mathbb{R}$ is a convex function such that the process $Y := (Y_t)_{t\in T}$ defined by $Y_t := f(X_t), t \in T$, is an $\mathcal{L}^1_{\mathbb{R}}$-process, then Y is an $\mathbb{F}$-submartingale.

Lemma 8.1 *Let $\xi \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \Omega)$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$.*
Then, the process $X = (X_t)_{t\in T}$ defined by $X_t := \mathbb{E}[\xi|\mathcal{F}_t], t \in T$ is a $\mathbb{F}$-martingale.

Proposition 8.2 *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and $X = (X_t)_{t\in T}$ be a $\mathbb{F}$-adapted $\mathcal{L}^1_{\mathbb{R}}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$. Then* *(i) If $\mathbb{P}$-a.a. paths of X are monotonically increasing, then X is a $\mathbb{F}$-submartingale.*

(ii) If $\mathbb{P}$-a.a. paths of X are monotonically decreasing, then X is a $\mathbb{F}$-supermartingale.

Example 8.1 *Every Poisson process is a submartingale.*

Proposition 8.3 *Assume that $\sigma := \inf T \in T$, and let $X = (X_t)_{t\in T}$ be a $\mathcal{L}^1_{\mathbb{R}}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Then if X possesses a constant mean function as well as independent increments that are independent of the initial state, X is a martingale.

Example 8.2 *Every 1-dimensional Brownian motion is a martingale.*

Example 8.3 (Symmetric random walk) *Let $S = (S_n)_{n\in\mathbb{N}_0}$ be the symmetric random walk defined by*

- $S_0 := 0$,
- $S_n := \sum_{j=1}^{n} X_j, n \in \mathbb{N}$,

where $(X_j)_{j\in\mathbb{N}}$ is a sequence of independent random variables in $\mathcal{L}^1_{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}[X_j] = 0$ for all $j \in \mathbb{N}$. Moreover, let the filtration $\mathbb{F}^X = (\mathcal{F}^X_n)_{n\in\mathbb{N}_0}$ be given by

- $\mathcal{F}^X_0 := \{\varnothing, \Omega\}$,
- $\mathcal{F}^X_n := \sigma(X_1, \dots, X_n), n \in \mathbb{N}$.

Then, S is a $\mathbb{F}^X$-martingale.

Definition 8.2 (Predictable process) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$ and $H = (H_n)_{n\in\mathbb{N}}$ be a process on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(E, \mathcal{E})$. Then, H is called **$\mathbb{F}$-predictable** if H_n is $(\mathcal{F}_{n-1}, \mathcal{E})$-measurable for every $n \in \mathbb{N}$.*

Definition 8.3 (Martingale transformation) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$, and $X = (X_n)_{n\in\mathbb{N}_0}$ and $H = (H_n)_{n\in\mathbb{N}}$ are two $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued processes on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Then, the process $H \bullet X = (H \bullet X_n)_{n\in\mathbb{N}_0}$ defined by

$$H \bullet X_0 := 0 \quad \text{and} \quad H \bullet X_n := \sum_{j=1}^{n} H_j (X_j - X_{j-1}), \quad n \in \mathbb{N}$$

*is called **martingale transformation of X w.r.t. H**.*

Remark 8.1 *The process $H = (H_n)_{n\in\mathbb{N}}$ will be called **locally bounded** if for every $n \in \mathbb{N}$ there exists some $c_n \in \mathbb{R}_{>0}$ such that $|H_n| \le c_n$ $\mathbb{P}$-a.s..*

Proposition 8.4 (Properties) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$, and $X = (X_n)_{n\in\mathbb{N}_0}$ and $H = (H_n)_{n\in\mathbb{N}}$ be two $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued processes on $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, let $H \bullet X$ be the martingale transformation of X w.r.t. H.*

(i) If H is $\mathbb{F}$-predictable, $\mathbb{R}_{\geq 0}$-valued and locally bounded, then $H \bullet X$ is a $\mathbb{F}$-submartingale if X is a $\mathbb{F}$-submartingale.

(ii) If H is $\mathbb{F}$-predictable, $\mathbb{R}_{\geq 0}$-valued and locally bounded, then $H \bullet X$ is a $\mathbb{F}$-supermartingale if X is a $\mathbb{F}$-supermartingale.

(iii) If H is $\mathbb{F}$-predictable, $\mathbb{R}_{\geq 0}$-valued $\mathcal{L}^2_{\mathbb{R}}$-process, then $H \bullet X$ is a $\mathbb{F}$-submartingale if X is a $\mathcal{L}^2_{\mathbb{R}}$-$\mathbb{F}$-submartingale.

(iv) If H is $\mathbb{F}$-predictable, $\mathbb{R}_{\geq 0}$-valued $\mathcal{L}^2_{\mathbb{R}}$-process, then $H \bullet X$ is a $\mathbb{F}$-supermartingale if X is a $\mathcal{L}^2_{\mathbb{R}}$-$\mathbb{F}$-supermartingale.

(v) If H is $\mathbb{F}$-predictable and locally bounded, then $H \bullet X$ is a $\mathbb{F}$-martingale if X is a $\mathbb{F}$-martingale.

(vi) If H is a $\mathbb{F}$-predictable $\mathcal{L}^2_{\mathbb{R}}$-process, then $H \bullet X$ is a $\mathbb{F}$-martingale if X is a $\mathcal{L}^2_{\mathbb{R}}$-$\mathbb{F}$-martingale.

Theorem 8.1 (Doob decomposition) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$, and $X = (X_n)_{n\in\mathbb{N}_0}$ be a $\mathbb{F}$-adapted $\mathcal{L}^1_{\mathbb{R}}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

(i) On $(\Omega, \mathcal{F}, \mathbb{P})$, there exist a $\mathbb{F}$-martingale $M = (M_n)_{n\in\mathbb{N}_0}$ with $M_0 = 0$ $\mathbb{P}$-a.s. and a $\mathbb{F}$-predictable $\mathcal{L}^1_{\mathbb{R}}$-process $A = (A_n)_{n\in\mathbb{N}}$ such that $X_n = X_0 + M_n + A_n$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}_0$, where $A_0 := 0$.

(ii) The decomposition in (i) is $\mathbb{P}$-a.s. unique and has the form

$$A_n = \sum_{j=1}^{n} \mathbb{E}[X_j - X_{j-1}|\mathcal{F}_{j-1}] = \sum_{j=1}^{n} \mathbb{E}[X_j|\mathcal{F}_{j-1}] - X_{j-1} \quad \mathbb{P}\text{-a.s.}$$

$$M_n = X_n - A_n - X_0 = \sum_{j=1}^{n} X_j - \mathbb{E}[X_j|\mathcal{F}_{j_1}] \quad \mathbb{P}\text{-a.s.}$$

(iii) X is a $\mathbb{F}$-submartingale if and only if $\mathbb{P}$-a.a. paths of A are monotonically increasing.

(iv) X is a $\mathbb{F}$-supermartingale if and only if $\mathbb{P}$-a.a. paths of A are monotonically decreasing.

Definition 8.4 (Compensator) *In the framework of Theorem 8.1, the $\mathbb{P}$-a.s. unique $\mathbb{F}$-predictable process A is called* ***compensator of X (w.r.t. $\mathbb{F}$)****.*

Corollary 8.1.1 *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$ and $X = (X_n)_{n\in\mathbb{N}_0}$ be a $\mathcal{L}^2_\mathbb{R}$-$\mathbb{F}$-martingale on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Then, there exist a centered $\mathbb{F}$-martingale $M = (M_n)_{n\in\mathbb{N}_0}$ with $M_0 = 0$ $\mathbb{P}$-a.s. and a $\mathbb{F}$-predictable $\mathcal{L}^1_\mathbb{R}$-process $A = (A_n)_{n\in\mathbb{N}}$ such that

$$X_n^2 = X_0^2 + M_n + A_n \quad \mathbb{P}\text{-a.s. for all } n \in \mathbb{N}_0,$$

where $A_0 := 0$.
This decomposition of the process $X^2 := (X_n^2)_{n\in\mathbb{N}}$ is $\mathbb{P}$-a.s. unique and $\mathbb{P}$-a.a. paths of A are monotonically increasing.

Definition 8.5 (Quadratic variation process) *In the framework of Corollary 8.1.1, the $\mathbb{P}$-a.s. unique monotonically increasing and $\mathbb{F}$-predictable process A is called* ***quadratic variation process of*** *X. We also write $\langle X\rangle$ and $\langle X\rangle_n$ instead of A and A_n, respectively.*

Proposition 8.5 *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$ and $X = (X_n)_{n\in\mathbb{N}_0}$ be a $\mathcal{L}^2_\mathbb{R}$-$\mathbb{F}$-martingale on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the following assertions hold true for the quadratic variation process $\langle X\rangle = (\langle X\rangle_n)_{n\in\mathbb{N}_0}$ of X.*

(i) $\mathbb{E}[\langle X\rangle_n] = \mathbb{E}[X_n^2] - \mathbb{E}[X_0^2] = \mathbb{E}[(X_n - X_0)^2] = \mathrm{Var}[X_n - X_0]$ *for all* $n \in \mathbb{N}_0$.

(ii) $\langle X\rangle_n - \langle X\rangle_{n-1} = \mathbb{E}[(X_n - X_{n-1})^2|\mathcal{F}_{n-1}]$ $\mathbb{P}$-*a.s. for all* $n \in \mathbb{N}$.

(iii) $\langle X\rangle_n = \sum_{j=1}^n \mathbb{E}[(X_j - X_{j-1})^2|\mathcal{F}_{j-1}]$ $\mathbb{P}$-*a.s. for all* $n \in \mathbb{N}$.

Example 8.4 (Symmetric random walk) *Let $S = (S_n)_{n\in\mathbb{N}_0}$ be the symmetric random walk defined by $S_0 := 0$ and $S_n := \sum_{j=1}^n X_j, n \in \mathbb{N}$, where $(X_j)_{j\in\mathbb{N}}$ is a sequence of independent random variables from $\mathcal{L}^1_\mathbb{R}(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[X_j] = 0, j \in \mathbb{N}$.*

According to Example 8.3, S is a ($\mathcal{L}^2_\mathbb{R}$-) $\mathbb{F}^X$-martingale.

For the quadratic variation process $\langle S\rangle$ of S, we have

$$\langle S\rangle_n = \sum_{j=1}^n \mathbb{E}[X_j^2] \quad \mathbb{P}\text{-a.s. for all } n \in \mathbb{N}.$$

If the innovations $X_j, j \in \mathbb{N}$, are identically distributed (or if, at least, all of them possess the same 2-nd moment), then specifically

$$\langle S\rangle_n = n\mathbb{E}[X_1^2] = n\mathrm{Var}[X_1] \quad \mathbb{P}\text{-a.s. for all } n \in \mathbb{N}.$$

Definition 8.6 (Stopping time) *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$.*
A map $\tau : \Omega \to T \cup \{+\infty\}$ is called ***stopping time w.r.t.*** *$\mathbb{F}$ (short: $\mathbb{F}$-**stopping time**) if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in T$.*

Proposition 8.6 *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and τ and σ be two $\mathbb{F}$-stopping times. Then also the maps $\tau \wedge \sigma, \tau \vee \sigma : \Omega \to T \cup \{\infty\}$ defined by*

$$\tau \wedge \sigma(\omega) := \min\{\tau(\omega), \sigma(\omega)\}\,, \quad \omega \in \Omega$$
$$\tau \vee \sigma(\omega) := \max\{\tau(\omega), \sigma(\omega)\}\,, \quad \omega \in \Omega$$

are $\mathbb{F}$-stopping times. If T is contained in $\mathbb{R}_{\geq 0}$ and closed under addition, then also the map $\tau + \sigma : \Omega \to T \cup \{\infty\}$ defined by

$$(\tau + \sigma)(\omega) := \tau(\omega) + \sigma(\omega)\,, \quad \omega \in \Omega$$

is a $\mathbb{F}$-stopping time.

Lemma 8.2 *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and τ be an $\mathbb{F}$-stopping time. Then, the following system of sets is a sub-σ-algebra of $\mathcal{F}$*

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in T\}\,.$$

Definition 8.7 (σ-algebra of the τ-history) *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration in $\mathcal{F}$ and τ be an $\mathbb{F}$-stopping time.*
*Then, $\mathcal{F}_\tau$ defined by Lemma 8.2 is called **σ-algebra of the τ-history**.*

Example 8.5 (Hitting time) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$, and $X = (X_n)_{n\in\mathbb{N}_0}$ be an $\mathbb{F}$-adapted process taking values in a measurable space $(E, \mathcal{E})$.*

Then, for every $A \in \mathcal{E}$,

$$\tau_A(\omega) := \min\{n \in \mathbb{N}_0 : X_n(\omega) \in A\}\,, \quad \omega \in \Omega$$

defines an $\mathbb{F}$-stopping time τ_A, where we set $\min \varnothing := \infty$.

Remark 8.2 (Notation) *For any process $X = (X_n)_{n\in\mathbb{N}_0}$ and any stopping time τ w.r.t. a filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$, we will henceforth use the notation*

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega)\,, \quad \omega \in \{\tau \in \mathbb{N}_0\}\,,$$

and as before $\mathcal{F}_\tau$ denotes the σ-algebra of the τ-history.

Theorem 8.2 (Optional sampling) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$, let $X = (X_n)_{n\in\mathbb{N}_0}$ be an $\mathbb{F}$-adapted $\mathcal{L}^1_\mathbb{R}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$, and let σ and τ be two $\mathbb{F}$-stopping time with $\sigma \leq \tau$. If X is an $\mathbb{F}$-martingale (-submartingale, -supermartingale) and τ is $\mathbb{P}$-a.s. bounded, then*

$$\mathbb{E}[X_\tau | \mathcal{F}_\sigma] = (\geq, \leq) X_\sigma \quad \mathbb{P}\text{-a.s.}$$

and in particular $\mathbb{E}[X_\tau] = (\geq, \leq)\mathbb{E}[X_\sigma]$.

Corollary 8.2.1 (Martingale test) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration $\mathcal{F}$ and $X = (X_n)_{n\in\mathbb{N}_0}$ be an $\mathbb{F}$-adapted $\mathcal{L}^1_\mathbb{R}$-process on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Then, X is an $\mathbb{F}$-martingale if and only if $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ for every $\mathbb{P}$-a.s. bounded $\mathbb{F}$-stopping time τ.

Definition 8.8 *For any filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ in $\mathcal{F}$ and any $n \in \mathbb{N}_0$, we set*

$$\mathcal{F}_n^\tau := \mathcal{F}_{\tau\wedge n}\,.$$

Definition 8.9 (Stopped process) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$. For any $\mathbb{F}$-stopping time τ and any $\mathbb{F}$-adapted process $X = (X_n)_{n\in\mathbb{N}_0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(E, \mathcal{E})$, the process $X^\tau = (X_n^\tau)_{n\in\mathbb{N}_0}$ defined by*

$$X_n^\tau(\omega) := X_{\tau\wedge n}(\omega) = X_{\min\{n,\tau(\omega)\}}(\omega)$$

*is said to be the corresponding **stopped process**.*

Theorem 8.3 (Optional stopping) *Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration in $\mathcal{F}$ and τ be a $\mathbb{F}$-stopping time. Moreover, let $X = (X_n)_{n\in\mathbb{N}_0}$ be a $\mathbb{F}$-adapted $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued process on $(\Omega, \mathcal{F}, \mathbb{P})$.*

If X is a martingale (submartingale, supermartingale) w.r.t. $\mathbb{F}$, then the stopped process $X^\tau = (X_n^\tau)_{n\in\mathbb{N}_0}$ is a martingale (submartingale, supermartingale) w.r.t. both $\mathbb{F}$ and $\mathbb{F}^\tau$.

Example 8.6 *Let $(X_j)_{j\in\mathbb{N}}$ be a sequence of i.i.d. real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}_{X_1} = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, and $\mathbb{F}^X = (\mathcal{F}_n^X)_{n\in\mathbb{N}_0}$ be given by $\mathcal{F}_0^X := \{\varnothing, \Omega\}$, $\mathcal{F}_n^X := \sigma(X_1, \dots, X_n), n \in \mathbb{N}$. Set $S_0 := 0, S_n := \sum_{j=1}^n X_j, n \in \mathbb{N}$. $S = (S_n)_{n\in\mathbb{N}_0}$ is an $\mathcal{F}^X$-martingale. For $a, b \in \mathbb{Z}$ with $a < 0$ and $b > 0$, let*

- $\tau_a := \min\{n \in \mathbb{N}_0 : S_n = a\}$,
- $\tau_b := \min\{n \in \mathbb{N}_0 : S_n = b\}$,
- $\tau_{a,b} := \tau_a \wedge \tau_b$.

We know that $\tau_a, \tau_b, \tau_{a,b}$ are $\mathbb{F}^X$-stopping times. For these stopping times, we have

(i) $\mathbb{P}[\{\tau_{a,b} = \tau_a\}] = b/(|a| + b)$,

(ii) $\mathbb{E}[\tau_{a,b}] = |a|b$,

(iii) $\mathbb{E}[\tau_a] = \infty$.