SF3953: Markov Chains and Processes **Spring 2017**

Lecture 2: Stopping Times and the Strong Markov Property

Lecturer: Jimmy Olsson *February 17*

Goals of this lecture

- Discuss the existence of Markov chains,
- Introduce some special stopping times of relevance for coming developments,
- Establish the strong Markov property.

Canonical chains

Assume that we are given a measurable space $(\mathsf{X}, \mathcal{X})$, an initial distribution $\nu \in \mathsf{M}_1(\mathcal{X})$, and a Markov kernel P on $\mathsf{X} \times \mathcal{X}$. Let $\mathsf{X}^{\mathbb{N}}$ be the set of X-valued sequences $\omega = (\omega_0, \omega_1, \omega_2, \ldots)$. The σ-field $\mathcal{F} = \mathcal{X}^{\otimes \mathbb{N}}$ is generated by the algebra $\mathcal{A}$ of cylindrical sets of form

$$\prod_{n=0}^{\infty} A_n,$$

where $A_n \in \mathcal{X}$ for all $n \in \mathbb{N}$ and $A_n \neq \mathsf{X}$ for at most finitely many n.

Definition 2.1 (coordinate process)**.** *The* coordinate process $\{X_k : k \in \mathbb{N}\}$ *is the stochastic process defined on* $(\mathsf{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}})$ *by*

$$X_k(\omega) = \omega_k, \quad \omega \in \mathsf{X}^{\mathbb{N}}.$$

A point $\omega \in \mathsf{X}^{\mathbb{N}}$ *is called a* trajectory *or* path*.*

With $\{X_k : k \in \mathbb{N}\}$ being the coordinate process, we set, for all $n \in \mathbb{N}$, $\mathcal{F}_n = \sigma(X_m : m \leq n)$.

Theorem 2.2. *Let* $(\mathsf{X}, \mathcal{X})$ *be a measurable space and* P *a Markov kernel on* $\mathsf{X} \times \mathcal{X}$*. For every* $\mu \in \mathsf{M}_1(\mathcal{X})$ *there exists a unique probability measure* $\mathbb{P}_\mu$ *on the canonical space* $(\Omega, \mathcal{F}) = (\mathsf{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}})$ *such that the coordinate process* $\{X_n : n \in \mathbb{N}\}$ *is a Markov chain with kernel* P *and initial distribution* μ*.*

The proof, which is based on Caratheodory's extension theorem (Theorem 2.17), is rather technical and is hence beyond the scope of the course. Still, for the sake of completeness, the interested reader find a version of the proof Appendix A.

Definition 2.3. *The* canonical Markov chain with kernel P *on* $\mathsf{X} \times \mathcal{X}$ *is the coordinate process* $\{X_n : n \in \mathbb{N}\}$ *on the canonical filtered space* $(X^{\mathbb{N}}, \mathcal{X}^{\otimes\mathbb{N}}, \{\mathcal{F}_k^X : k \in \mathbb{N}\})$ *endowed with the family* $\{\mathbb{P}_\nu : \nu \in \mathsf{M}_1(\mathcal{X})\}$ *of probability measures given by Theorem 2.2.*

Note that with the canonical Markov chain comes a *family* of probability measures.

In the following, "a Markov chain with kernel P on $\mathsf{X} \times \mathcal{X}$" will always refer to the canonical chain.

For $x \in \mathsf{X}$, we introduce the short-hand notation $\mathbb{P}_x = \mathbb{P}_{\delta_x}$ (and similarly for expectations).

Proposition 2.4. *For all* $A \in \mathcal{X}^{\otimes\mathbb{N}}$*,*

(i) the function $\mathsf{X} \ni x \mapsto \mathbb{P}_x(A)$ *belongs to* $\mathsf{F}_\mathrm{b}(\mathcal{X})$*,*

(ii) for all $\mu \in \mathsf{M}_1(\mathcal{X})$*,* $\mathbb{P}_\mu(A) = \int \mathbb{P}_x(A)\, \mu(\mathrm{d}x)$*.*

Exercise 2.5. *In order to prove Proposition 2.4, show that the family* $\mathcal{M}$ *of sets* $A \in \mathcal{X}^{\otimes\mathbb{N}}$ *satisfying (i) and (ii) is a monotone class (see Definition 2.20). Now, conclude the proof of the proposition using the monotone class theorem, Theorem 2.21, in combination with Theorem 1.14 (last time).*

Stopping times

In the following, consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_k : k \in \mathbb{N}\}, \mathbb{P})$ and an adapted process $\{(X_n, \mathcal{F}_n) : n \in \mathbb{N}\}$. Define $\mathcal{F}_\infty$ as the σ-field generated by the union of all the $\{\mathcal{F}_k : k \in \mathbb{N}\}$.

We recall the following definition.

Definition 2.6 (stopping time)**.**

(i) A random variable τ *from* Ω *to* $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ *is called a* stopping time *if for all* $k \in \mathbb{N}$*,* $\{\tau \leq k\} \in \mathcal{F}_k$*.*

(ii) The stopping time σ-field $\mathcal{F}_\tau$ *is generated by the sets* $A \subset \Omega$ *such that for all* $k \in \mathbb{N}$*,* $A \cap \{\tau \leq k\} \in \mathcal{F}_k$*.*

It is easily checked that

- $\mathcal{F}_\tau$ is indeed a σ-field,
- a constant $\tau(\omega) = n \in \mathbb{N}$ is a stopping time (in which case $\mathcal{F}_\tau = \mathcal{F}_n$),
- the event $\{\tau = \infty\}$ belongs to $\mathcal{F}_\infty$.

Given the stochastic process $\{X_n : n \in \mathbb{N}\}$ and some arbitrary $\mathcal{F}_\infty$-measurable random variable X_∞ we define

$$X_\tau = X_k \quad \text{on} \quad \{\tau = k\}, \quad k \in \bar{\mathbb{N}}.$$

Note that X_τ is $\mathcal{F}_\tau$-measurable, since for $A \in \mathcal{F}_\tau$,

$$\begin{aligned}\{X_\tau \in A\} \cap \{\tau \le k\} = \bigcup_{\ell=0}^{k} \{X_\tau \in A\} \cap \{\tau = \ell\} &= \bigcup_{\ell=0}^{k} \{X_\ell \in A\} \cap \{\tau = \ell\} \\ &= \bigcup_{\ell=0}^{k} \{X_\ell \in A\} \cap (\{\tau \le \ell\} \setminus \{\tau \le \ell - 1\}) \in \mathcal{F}_k.\end{aligned}$$

2.1 The shift operator and the strong Markov property

Definition 2.7. *The mapping* $\mathsf{X}^{\mathbb{N}} \to \mathsf{X}^{\mathbb{N}}$ *defined by*

$$\theta : \omega = (\omega_0, \omega_2, \ldots) \mapsto \theta(\omega) = (\omega_1, \omega_2, \ldots)$$

is called the shift operator*.*

Multiple shift operators are defined recursively by letting θ_0 be the identity function (i.e., $\theta_0(\omega) = \omega$ for all $\omega \in \Omega$) and for $k \in \mathbb{N}^*$,

$$\theta_k = \theta \circ \theta_{k-1}.$$

Thus, each θ_k shifts a sequence k times.

Now, let $\{X_k : k \in \mathbb{N}\}$ be the coordinate process on $\mathsf{X}^{\mathbb{N}}$. Then for all $(j, k) \in \mathbb{N}^2$,

$$X_k \circ \theta_j = X_{j+k}.$$

Exercise 2.8. *Let* $\{\mathcal{F}_k : k \in \mathbb{N}\}$ *be the natural filtration of the coordinate process. Show that for all* $(m, k) \in \mathbb{N}^2$ *such that* $m \ge k$*,* θ_k *is measurable from* $(\mathsf{X}^{\mathbb{N}}, \mathcal{F}_m)$ *to* $(\mathsf{X}^{\mathbb{N}}, \mathcal{F}_{m-k})$*.*

Exercise 2.9. *Let* $\{\mathcal{F}_k : k \in \mathbb{N}\}$ *be the natural filtration of the coordinate process* $\{X_k : k \in \mathbb{N}\}$ *and* τ *and* σ *stopping times with respect to* $\{\mathcal{F}_k : k \in \mathbb{N}\}$*. Show that*

(a) for each positive integer $k \in \mathbb{N}$*,* $k + \tau \circ \theta_k$ *is a stopping time.*

(b) the random variable $\rho = \sigma + \tau \circ \theta_\sigma$ *is a stopping time. Moreover, if* σ *and* τ *are finite, then* $X_\tau \circ \theta_\sigma = X_\rho$*.*

From now on, a Markov kernel P on $\mathsf{X} \times \mathcal{X}$ is given, and we let $\{X_k : k \in \mathbb{N}\}$ be the canonical chain with Markov kernel P on $(\Omega, \mathcal{F}) = (\mathsf{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}})$. This means that $\{\mathcal{F}_k : k \in \mathbb{N}\}$ is the natural filtration of $\{X_k : k \in \mathbb{N}\}$. In addition, $\mathbb{P}_\mu$ denotes the law induced by $\mu \in \mathsf{M}_1(\mathcal{X})$ and $\mathbb{E}_\mu$ stands for the associated expectation operator.

Theorem 2.10 (Markov property). *For all $\mathcal{F}$-measurable positive or bounded random variables Y, initial distributions $\mu \in \mathsf{M}_1(\mathcal{X})$ and $k \in \mathbb{N}$, it holds, $\mathbb{P}_\mu$-a.s., that*

$$\mathbb{E}_\mu\left[Y \circ \theta_k \mid \mathcal{F}_k\right] = \mathbb{E}_{X_k}\left[Y\right]. \tag{2.11}$$

Proof. We use the functional monotone class theorem (see previous lecture). Let $\mathcal{H}$ be set of bounded $\mathcal{F}$-measurable random variables Y such that (2.11) holds. One proves straightforwardly that $\mathcal{H}$ is a vector space (as the property (2.11) involves expectations). Now, let $\mathcal{E}$ be the set of finite rectangles in $\Omega = \mathsf{X}^{\mathbb{N}}$, which is closed under finite intersection. To prove that $\mathbb{1}_A \in \mathcal{H}$ for all $A \in \mathcal{E}$, we prove more generally that every Y of form $Y = g(X_0, \dots, X_p)$, where $g \in \mathsf{F}_{\mathrm{b}}(\mathcal{X}^{p+1})$, belongs to $\mathcal{H}$. We want to prove an identity for conditional expectations; thus, pick arbitrarily a bounded $\mathcal{F}_k$-measurable random variable Z and prove that

$$\mathbb{E}_\mu\left[Z(Y \circ \theta_k - \mathbb{E}_{X_k}[Y])\right] = 0.$$

By the factorisation lemma there exists a measurable function z such that $Z = z(X_0, \dots, X_k)$. Moreover, by Theorem 1.14 and the Fubini-Tonelli theorems,

$$\begin{aligned}
\mathbb{E}_\mu[Z(Y \circ \theta_k)] &= \mathbb{E}_\mu\left[z(X_0, \dots, X_k) g(X_k, \dots, X_{k+p})\right] \\
&= \int \cdots \int z(x_0, \dots, x_k) g(x_k, \dots, x_{k+p})\, \mu(\mathrm{d}x_0) \prod_{\ell=0}^{k+p-1} P(x_\ell, \mathrm{d}x_{\ell+1}) \\
&= \int \cdots \int (\delta_{x_k} \otimes P^{\otimes p} g) z(x_0, \dots, x_k)\, \mu(\mathrm{d}x_0) \prod_{\ell=0}^{k-1} P(x_\ell, \mathrm{d}x_{\ell+1}) \\
&= \int \cdots \int \mathbb{E}_{x_k}[Y] z(x_0, \dots, x_k)\, \mu(\mathrm{d}x_0) \prod_{\ell=0}^{k-1} P(x_\ell, \mathrm{d}x_{\ell+1}) \\
&= \mathbb{E}_\mu\left[Z\, \mathbb{E}_{X_k}[Y]\right],
\end{aligned}$$

which was to be established. To check the second condition of the functional monotone class theorem, let $\{Y_n : n \in \mathbb{N}\}$ be an increasing sequence of random variables in $\mathcal{H}$ and denote by $Y = \lim_{n\to\infty} Y_n$ the pointwise limit. Then, using monotone convergence (also for conditional expectations) and (2.11), $\mathbb{P}_\mu$-a.s.,

$$\mathbb{E}_\mu\left[Y \circ \theta_k \mid \mathcal{F}_k\right] = \lim_{n\to\infty} \mathbb{E}_\mu[Y_n \circ \theta_k \mid \mathcal{F}_k] = \lim_{n\to\infty} \mathbb{E}_{X_k}\left[Y_n\right] = \mathbb{E}_{X_k}\left[Y\right],$$

which shows that $Y \in \mathcal{H}$. We may now conclude the proof by applying the monotone class theorem. □

Importantly, the Markov property of Theorem 2.10 can be extended to random time shifts.

Theorem 2.12 (strong Markov property)**.** *For all $\mathcal{F}$-measurable positive or bounded random variables Y, initial distribution $\nu \in \mathsf{M}_1(\mathcal{X})$ and stopping time τ, it holds, $\mathbb{P}$-a.s., that*

$$\mathbb{E}_\mu\left[(Y \circ \theta_\tau)\mathbb{1}_{\{\tau<\infty\}} \mid \mathcal{F}_\tau\right] = \mathbb{E}_{X_\tau}\left[Y\right]\mathbb{1}_{\{\tau<\infty\}}. \tag{2.13}$$

Proof. We will show that for all $A \in \mathcal{F}_\tau$,

$$\mathbb{E}_\mu\left[(Y \circ \theta_\tau)\mathbb{1}_{\{\tau<\infty\}}\mathbb{1}_A\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X_\tau}\left[Y\right]\mathbb{1}_{\{\tau<\infty\}}\mathbb{1}_A\right].$$

For this purpose, consider, for $k \in \mathbb{N}$, the related expectation

$$\mathbb{E}_\mu\left[(Y \circ \theta_\tau)\mathbb{1}_{A\cap\{\tau=k\}}\right] = \mathbb{E}_\mu\left[(Y \circ \theta_k)\mathbb{1}_{A\cap\{\tau=k\}}\right]$$

Since τ is a stopping time, $\{\tau = k\} \in \mathcal{F}_k$. Thus, by the tower property and Theorem 2.10,

$$\mathbb{E}_\mu\left[(Y \circ \theta_k)\mathbb{1}_{A\cap\{\tau=k\}}\right] = \mathbb{E}_\mu\left[\mathbb{E}_\mu\left[Y \circ \theta_k \mid \mathcal{F}_k\right]\mathbb{1}_{A\cap\{\tau=k\}}\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X_k}\left[Y\right]\mathbb{1}_{A\cap\{\tau=k\}}\right].$$

Now, decompose, using monotone convergence, the original expectation of interest according to

$$\begin{aligned}\mathbb{E}_\mu\left[(Y \circ \theta_\tau)\mathbb{1}_{\{\tau<\infty\}}\mathbb{1}_A\right] &= \sum_{k=0}^{\infty}\mathbb{E}_\mu\left[(Y \circ \theta_k)\mathbb{1}_{A\cap\{\tau=k\}}\right] = \sum_{k=0}^{\infty}\mathbb{E}_\mu\left[\mathbb{E}_{X_k}\left[Y\right]\mathbb{1}_{A\cap\{\tau=k\}}\right]\\ &= \sum_{k=0}^{\infty}\mathbb{E}_\mu\left[\mathbb{E}_{X_\tau}\left[Y\right]\mathbb{1}_{A\cap\{\tau=k\}}\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X_\tau}\left[Y\right]\mathbb{1}_A\sum_{k=0}^{\infty}\mathbb{1}_{\{\tau=k\}}\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X_\tau}\left[Y\right]\mathbb{1}_{\{\tau<\infty\}}\mathbb{1}_A\right],\end{aligned}$$

which establishes the claim. □

Hitting times and return times

The following stopping times will play critical roles in the following:

Definition 2.14. *The* first hitting time τ_A *and* return time σ_A *to some set $A \in \mathcal{X}$ of the process $\{X_n : n \in \mathbb{N}\}$ are defined by*

$$\tau_A = \inf\{n \in \mathbb{N} : X_n \in A\},$$
$$\sigma_A = \inf\{n \in \mathbb{N}^* : X_n \in A\},$$

respectively, where, by convention, $\inf \emptyset = \infty$*. The* successive return times $\{\sigma_A^{(n)} : n \in \mathbb{N}\}$, *are defined inductively by $\sigma_A^{(0)} = 0$ and for all $n \in \mathbb{N}$,*

$$\sigma_A^{(n+1)} = \inf\left\{k \in \mathbb{N} : k > \sigma_A^{(n)}, X_k \in A\right\}.$$

One shows easily that τ_A is a stopping time; indeed, write for $m \in \mathbb{N}$,

$$\{\tau_A \leq m\} = \bigcup_{\ell=0}^{m} \{X_\ell \in A\},$$

where each event $\{X_\ell \in A\}$ belongs to $\mathcal{F}_\ell$. Thus, as $\mathcal{F}_\ell \subset \mathcal{F}_m$, this shows that τ_A is a stopping time. The return time σ_A is treated similarly. That $\{\sigma_A^{(n)} : n \in \mathbb{N}\}$ are stopping times follows by inductive use of Exercise 2.9, as, if $\sigma_A^{(n-1)}$ is finite, for all $n \in \mathbb{N}^*$,

$$\sigma_A^{(n)} = \sigma_A^{(n-1)} + \sigma_A \circ \theta_{\sigma_A^{(n-1)}}.$$

A Proof of Theorem 2.2

Definition 2.15. *A set function μ defined on an algebra $\mathcal{A}$ is said to be* σ-additive *if for all collections $\{A_n : n \in \mathbb{N}\}$ of pairwise disjoint sets in $\mathcal{A}$ such that $\cup_{n=0}^{\infty} A_n \in \mathcal{A}$, it holds that*

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n). \tag{2.16}$$

The proof of Theorem 2.2 relies on the following result, which is fundamental in measure theory.

Theorem 2.17 (Caratheodory's extension theorem)**.** *Let μ be some σ-additive, non-negative set function on an algebra $\mathcal{A}$ on some set X. Then there exists a measure $\bar{\mu}$ on the σ-field generated by $\mathcal{A}$. If μ is σ-finite, this extension is unique.*

Proof of Theorem 2.2. Let $\mathcal{A}$ be the algebra of cylindrical sets generating $\mathcal{X}^{\otimes \mathbb{N}}$ (i.e., $\mathcal{A}$ is the family of sets of form $A = A_n \times \mathsf{X}^{\mathbb{N}}$, where A_n is a product of $n+1$ sets in $\mathcal{X}$). We define the set function

$$\mathcal{A} \ni A \mapsto \mu \otimes P^{\otimes n}(A_n).$$

Let $\{A_n : n \in \mathbb{N}\}$ be a collection of pairwise disjoint sets of $\mathcal{A}$ such that $\cup_{n=0}^{\infty} A_n \in \mathcal{A}$. Using Caratheodory's extension theorem, we must establish the σ-additivity (2.16). Since $\cup_{n=0}^{\infty} A_n \in \mathcal{A}$, there must exist $n_0 \in \mathbb{N}$ such that $\cup_{n=0}^{\infty} A_n \in \mathcal{F}_{n_0}$ (since the filtration is generated by the coordinate process). For $n \in \mathbb{N}$, set $B_n = \cup_{k=n+1}^{\infty} A_k$. For all $n \in \mathbb{N}$, there exists $k_n \in \mathbb{N}$ such that $\{A_j\}_{j=0}^{n} \subset \mathcal{F}_{k_n}$. As the sets are pairwise disjoint, this implies that

$$B_n = \left(\bigcup_{j=0}^{n} A_n\right)^{\complement} \cap \bigcup_{k=0}^{\infty} A_k \in \mathcal{F}_{k_n \vee n_0}.$$

Thus, we have a sequence $\{B_n : n \in \mathbb{N}\}$ of decreasing sets in $\mathcal{A}$ such that $\cap_{n=0}^{\infty} B_n = \emptyset$, and it is enough to show that

$$\lim_{n \to \infty} \mu(B_n) = 0. \tag{2.18}$$

(Indeed, since μ is additive (as $\mu \otimes P^{\otimes n}$ is a probability for all n), we may write

$$\mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \mu\left(B_n \cup \bigcup_{k=0}^{n} A_n\right) = \mu(B_n) + \sum_{k=0}^{n} \mu(A_k);$$

thus, (2.18) implies (2.16).) Since for each $n \in \mathbb{N}$ there exists $k(n) \in \mathbb{N}$ such that $B_n \in \mathcal{F}_{k(n)}$ and since the σ-fields $\mathcal{F}_n$ are increasing, we can assume that the sequence $\{\mathcal{F}_{k(n)} : n \in \mathbb{N}\}$ is also non-decreasing. Moreover, by repeating if necessary certain terms B_n in the sequence, we can assume that $k(n) = n$, i.e., we can assume that B_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$.

Assume that each B_n is of form $C_n \times \mathsf{X}^{\mathbb{N}}$ and define for $(k, n) \in \mathbb{N}^2$ such that $k \leq n-1$,

$$f_k^n : \mathsf{X}^{k+1} \ni x_0^k \mapsto \int \mathbb{1}_{C_n}(x_0^n)\, P^{\otimes(n-k)}(x_k, \mathrm{d}x_{k+1}^n),$$

which is $\mathcal{X}^{k+1}$-measurable, and let $f_n^n = \mathbb{1}_{C_n}$. Recall that the sets $\{B_n : n \in \mathbb{N}\}$ are decreasing; thus, for fixed $k \in \mathbb{N}$, $\{f_k^n : n \in \mathbb{N}, n \geq k\}$ are non-increasing functions. (Indeed, since for all $n \in \mathbb{N}$, $B_{n+1} \subset B_n$, it follows that $C_{n+1} \subset C_n \times \mathsf{X}$. Hence, for all $n \geq k+1$ and $x_0^k \in \mathsf{X}^{k+1}$,

$$\begin{aligned} f_k^{n+1}(x_0^k) = \int \mathbb{1}_{C_{n+1}}(x_0^{n+1})\, P^{\otimes(n+1-k)}(x_k, \mathrm{d}x_{k+1}^{n+1}) \\ \leq \int \mathbb{1}_{C_n}(x_0^n)\, P^{\otimes(n-k)}(x_k, \mathrm{d}x_{k+1}^n) = f_k^n(x_0^k), \end{aligned}$$

and, similarly, $f_k^{k+1}(x_0^k) \leq \mathbb{1}_{C_k}(x_0^k) = f_k^k(x_0^k)$.) Thus, for all $k \in \mathbb{N}$, there are limits $g_k = \lim_{n\to\infty} f_k^n$, which are uniformly bounded by one. Note that by construction, $\mu(B_n) = \nu f_0^n$. Moreover, by dominated convergence, $\nu g_0 = \nu(\lim_{n\to\infty} f_0^n) = \lim_{n\to\infty} \nu f_0^n$, and consequently we must prove that $\nu g_0 = 0$.

We proceed by contraction. First note that for all large n and $x_0^k \in \mathsf{X}^{k+1}$,

$$\begin{aligned} f_k^n(x_0^k) &= \iint \mathbb{1}_{C_n}(x_0^n)\, P(x_k, \mathrm{d}x_{k+1})\, P^{\otimes(n-k-1)}(x_{k+1}, \mathrm{d}x_{k+2}^n) \\ &= \int f_{k+1}^n(x_0^{k+1})\, P(x_k, \mathrm{d}x_{k+1}). \end{aligned}$$

Thus, by dominated convergence, for all $k \in \mathbb{N}$ and $x_0^k \in \mathsf{X}^{k+1}$,

$$\begin{aligned} g_k(x_0^k) = \lim_{n\to\infty} \int f_{k+1}^n(x_0^{k+1})\, P(x_k, \mathrm{d}x_{k+1}) = \int \lim_{n\to\infty} f_{k+1}^n(x_0^{k+1})\, P(x_k, \mathrm{d}x_{k+1}) \\ = \int g_{k+1}(x_0^{k+1})\, P(x_k, \mathrm{d}x_{k+1}). \quad (2.19) \end{aligned}$$

Now assume that $\nu g_0 > 0$; then, by (2.19),

$$0 < \nu g_0 = \iint g_1(x_0^1)\,\nu(\mathrm{d}x_0)\,P(x_0, \mathrm{d}x_1),$$

which implies that there exists $\tilde{x}_0^1 \in \mathsf{X}^2$ such that $g_1(\tilde{x}_0^1) > 0$. This implies in turn, by (2.19), that

$$0 < g_1(\tilde{x}_0^1) = \iint g_2(\tilde{x}_0^1, x_2)\,P(\tilde{x}_1, \mathrm{d}x_2),$$

which implies that there exists $\tilde{x}_2 \in \mathsf{X}$ such that $g_2(\tilde{x}_0^2) > 0$, etc. In this way we construct a sequence $\tilde{\mathbf{x}} = \{\tilde{x}_n : n \in \mathbb{N}\}$ such that $g_n(\tilde{x}_0^n) > 0$ for all $n \in \mathbb{N}$. Now, since for all $n \in \mathbb{N}$, $\{f_n^\ell : \ell \in \mathbb{N}, n \geq \ell\}$ are non-increasing functions and g_n the corresponding limit,

$$\mathbb{1}_{B_n}(\tilde{\mathbf{x}}) = \mathbb{1}_{C_n}(\tilde{x}_0^n) = f_n^n(\tilde{x}_0^n) \geq g_n(\tilde{x}_0^n) > 0.$$

Thus, for all $n \in \mathbb{N}$, $\mathbb{1}_{B_n}(\tilde{\mathbf{x}}) = 1$, implying that $\tilde{\mathbf{x}} \in \cap_{n\in\mathbb{N}} B_n$, which contradicts $\cap_{n\in\mathbb{N}} B_n = \emptyset$. Hence, $\lim_{n\to\infty} \mu(B_n) = 0$, and since μ is finite we may conclude the proof by applying Caratheodory's theorem. □

B The monotone class theorem

Definition 2.20. *Let Ω be a set. A collection $\mathcal{M}$ of subsets of Ω is called a* monotone class *if*

(i) $\Omega \in \mathcal{M}$,

(ii) for $A \in \mathcal{M}$ *and* $B \in \mathcal{M}$ *such that* $A \subset B$, $B \setminus A \in \mathcal{M}$.

(iii) for all increasing sequences $\{A_n : n \in \mathbb{N}^*\}$ *of sets in* $\mathcal{M}$, $\cup_{n=1}^\infty A_n \in \mathcal{M}$.

Theorem 2.21 (monotone class theorem)**.** *Let $\mathcal{M}$ be a monotone class and assume that $\mathcal{E} \subset \mathcal{M}$ is stable by finite intersection. Then $\sigma(\mathcal{E}) \subset \mathcal{M}$.*

References

[1] J. Jacod and P. Protter. *Probability Essentials.* Springer, 2000.