

Chapter 2

Weak Convergence of Markov Kernels

As indicated in the previous chapter, stable convergence of random variables can be seen as suitable convergence of Markov kernels given by conditional distributions. Let (Ω, \mathcal{F}, P) be a probability space and let \mathcal{X} be a separable metrizable topological space equipped with its Borel σ -field $\mathcal{B}(\mathcal{X})$. In this chapter we briefly describe the weak topology on the set of Markov kernels (transition kernels) from (Ω, \mathcal{F}) to $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.

Let us first recall the weak topology on the set $\mathcal{M}^1(\mathcal{X})$ of all probability measures on $\mathcal{B}(\mathcal{X})$. It is the topology generated by the functions

$$\nu \mapsto \int h d\nu, \quad h \in C_b(\mathcal{X}),$$

where $C_b(\mathcal{X})$ denotes the space of all continuous, bounded functions $h : \mathcal{X} \rightarrow \mathbb{R}$ equipped with the sup-norm $\|h\|_{\sup} := \sup_{x \in \mathcal{X}} |h(x)|$. The weak topology on $\mathcal{M}^1(\mathcal{X})$ is thus the weakest topology for which each function $\nu \mapsto \int h d\nu$ is continuous. Consequently, weak convergence of a net $(\nu_\alpha)_\alpha$ in $\mathcal{M}^1(\mathcal{X})$ to $\nu \in \mathcal{M}^1(\mathcal{X})$ means

$$\lim_\alpha \int h d\nu_\alpha = \int h d\nu$$

for every $h \in C_b(\mathcal{X})$ (here and elsewhere we omit the directed set on which a net is defined from the notation). Because $\int h d\nu_1 = \int h d\nu_2$ for $\nu_1, \nu_2 \in \mathcal{M}^1(\mathcal{X})$ and every $h \in C_b(\mathcal{X})$ implies that $\nu_1 = \nu_2$, this topology is Hausdorff and the limit is unique. Moreover, the weak topology is separable metrizable e.g. by the Prohorov metric, see e.g. [69], Theorem II.6.2, and polish if \mathcal{X} is polish; see e.g. [69], Theorem II.6.5, [26], Corollary 11.5.5. The relatively compact subsets of $\mathcal{M}^1(\mathcal{X})$ are exactly the tight ones, provided \mathcal{X} is polish, where $\Gamma \subset \mathcal{M}^1(\mathcal{X})$ is called *tight* if for every $\varepsilon > 0$ there exists a compact set $A \subset \mathcal{X}$ such that $\sup_{\nu \in \Gamma} \nu(\mathcal{X} \setminus A) \leq \varepsilon$; see e.g. [69], Theorem II.6.7, [26], Theorem 11.5.4.

A map $K : \Omega \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ is called a *Markov kernel from (Ω, \mathcal{F}) to $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$* if $K(\omega, \cdot) \in \mathcal{M}^1(\mathcal{X})$ for every $\omega \in \Omega$ and $K(\cdot, B)$ is \mathcal{F} -measurable for every $B \in \mathcal{B}(\mathcal{X})$. Let $\mathcal{K}^1 = \mathcal{K}^1(\mathcal{F}) = \mathcal{K}^1(\mathcal{F}, \mathcal{X})$ denote the set of all such Markov kernels. If $\mathcal{M}^1(\mathcal{X})$ is equipped with the σ -field $\Sigma(\mathcal{M}^1(\mathcal{X})) := \sigma(\nu \mapsto \nu(B), B \in \mathcal{B}(\mathcal{X}))$, then Markov kernels $K \in \mathcal{K}^1$ can be viewed as $(\mathcal{M}^1(\mathcal{X}), \Sigma(\mathcal{M}^1(\mathcal{X})))$ -valued random variables $\omega \mapsto K(\omega, \cdot)$. Furthermore, $\Sigma(\mathcal{M}^1(\mathcal{X}))$ coincides with the Borel σ -field of $\mathcal{M}^1(\mathcal{X})$ (see Lemma A.2).

For a Markov kernel $K \in \mathcal{K}^1$ and a probability distribution Q on \mathcal{F} we define the *product measure* (which is a probability distribution again) on the product σ -field $\mathcal{F} \otimes \mathcal{B}(\mathcal{X})$ by

$$Q \otimes K(C) := \int \int 1_C(\omega, x) K(\omega, dx) dQ(\omega)$$

for $C \in \mathcal{F} \otimes \mathcal{B}(\mathcal{X})$ and its *marginal* on $\mathcal{B}(\mathcal{X})$ by

$$QK(B) := Q \otimes K(\Omega \times B) = \int K(\omega, B) dQ(\omega)$$

for $B \in \mathcal{B}(\mathcal{X})$. For functions $f : \Omega \rightarrow \mathbb{R}$ and $h : \mathcal{X} \rightarrow \mathbb{R}$ let $f \otimes h : \Omega \times \mathcal{X} \rightarrow \mathbb{R}$, $f \otimes h(\omega, x) := f(\omega)h(x)$, be the *tensor product*.

Lemma 2.1 (a) (Fubini's theorem for Markov kernels) *Let $K \in \mathcal{K}^1$ and $g : (\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B}(\mathcal{X})) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ be measurable such that g^- (or g^+) $\in \mathcal{L}^1(P \otimes K)$. Then*

$$\int g dP \otimes K = \int \int g(\omega, x) K(\omega, dx) dP(\omega) .$$

(b) (Uniqueness) *For $K_1, K_2 \in \mathcal{K}^1$, we have $\{\omega \in \Omega : K_1(\omega, \cdot) = K_2(\omega, \cdot)\} \in \mathcal{F}$, and $K_1(\cdot, B) = K_2(\cdot, B)$ P -almost surely for every $B \in \mathcal{B}(\mathcal{X})$ implies $P(\{\omega \in \Omega : K_1(\omega, \cdot) = K_2(\omega, \cdot)\}) = 1$, that is, $K_1 = K_2$ P -almost surely.*

Proof (a) For $g = 1_C$ with $C \in \mathcal{F} \otimes \mathcal{B}(\mathcal{X})$ this is the definition of $P \otimes K$. The formula extends as usual by linearity, monotone convergence and the decomposition $g = g^+ - g^-$.

(b) Note that $\mathcal{B}(\mathcal{X})$ is countably generated. Let \mathcal{C} be a countable generator of $\mathcal{B}(\mathcal{X})$ and let \mathcal{C}_0 be the (countable) system of all finite intersections of sets from \mathcal{C} . Then by measure uniqueness

$$\{\omega \in \Omega : K_1(\omega, \cdot) = K_2(\omega, \cdot)\} = \bigcap_{B \in \mathcal{C}_0} \{\omega \in \Omega : K_1(\omega, B) = K_2(\omega, B)\} .$$

Hence the assertion. □

Exercise 2.1 Let $\mathcal{C} \subset \mathcal{B}(\mathcal{X})$ be closed under finite intersections with $\sigma(\mathcal{C}) = \mathcal{B}(\mathcal{X})$ and let $K : \Omega \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ satisfy $K(\omega, \cdot) \in \mathcal{M}^1(\mathcal{X})$ for every $\omega \in \Omega$ and $K(\cdot, B)$ is \mathcal{F} -measurable for every $B \in \mathcal{C}$. Show that $K \in \mathcal{K}^1$.

Definition 2.2 The topology on \mathcal{K}^1 generated by the functions

$$K \mapsto \int f \otimes h dP \otimes K, \quad f \in \mathcal{L}^1(P), \quad h \in C_b(\mathcal{X})$$

is called the *weak topology* and is denoted by $\tau = \tau(P) = \tau(\mathcal{F}, P)$. Accordingly, *weak convergence of a net $(K_\alpha)_\alpha$ in \mathcal{K}^1 to $K \in \mathcal{K}^1$* means

$$\lim_\alpha \int f \otimes h dP \otimes K_\alpha = \int f \otimes h dP \otimes K$$

for every $f \in \mathcal{L}^1(P)$ and $h \in C_b(\mathcal{X})$.

The dependence of τ on P is usually not explicitly indicated. This topology is well known e.g. in statistical decision theory where \mathcal{K}^1 corresponds to all randomized decision rules and in areas such as dynamic programming, optimal control, game theory or random dynamical systems; see [7, 13, 18, 61, 62, 87].

Simple characterizations of weak convergence are as follows. For a sub- σ -field \mathcal{G} of \mathcal{F} , let $\mathcal{K}^1(\mathcal{G}) = \mathcal{K}^1(\mathcal{G}, \mathcal{X})$ denote the subset of \mathcal{K}^1 consisting of all \mathcal{G} -measurable Markov kernels, that is of Markov kernels from (Ω, \mathcal{G}) to $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. For $F \in \mathcal{F}$ with $P(F) > 0$ let $P_F := P(\cdot|F) = P(\cdot \cap F)/P(F)$ denote the *conditional probability measure given F* , and let E_F and Var_F denote expectation and variance, respectively, under P_F . Further recall that for a net $(y_\alpha)_\alpha$ in $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$

$$\limsup_\alpha y_\alpha := \inf_\alpha \sup_{\beta \geq \alpha} y_\beta \quad \text{and} \quad \liminf_\alpha y_\alpha := \sup_\alpha \inf_{\beta \geq \alpha} y_\beta.$$

Theorem 2.3 Let $\mathcal{G} \subset \mathcal{F}$ be a sub- σ -field, $(K_\alpha)_\alpha$ a net in $\mathcal{K}^1(\mathcal{G})$, $K \in \mathcal{K}^1(\mathcal{G})$ and let $\mathcal{E} \subset \mathcal{G}$ be closed under finite intersections with $\Omega \in \mathcal{E}$ such that $\sigma(\mathcal{E}) = \mathcal{G}$. Then the following statements are equivalent:

- (i) $K_\alpha \rightarrow K$ weakly,
- (ii) $\lim_\alpha \int f \otimes h dP \otimes K_\alpha = \int f \otimes h dP \otimes K$ for every $f \in \mathcal{L}^1(\mathcal{G}, P)$ and $h \in C_b(\mathcal{X})$,
- (iii) $QK_\alpha \rightarrow QK$ weakly (in $\mathcal{M}^1(\mathcal{X})$) for every probability distribution Q on \mathcal{F} such that $Q \ll P$,
- (iv) $P_F K_\alpha \rightarrow P_F K$ weakly for every $F \in \mathcal{E}$ with $P(F) > 0$.

Proof (i) \Rightarrow (iii). Let $Q \ll P$. Setting $f := dQ/dP$ and using Fubini's theorem for Markov kernels 2.1 (a), we obtain for $h \in C_b(\mathcal{X})$

$$\begin{aligned} \int h dQ K_\alpha &= \int \int h(x) K_\alpha(\omega, dx) dQ(\omega) = \int f \otimes h dP \otimes K_\alpha \\ &\rightarrow \int f \otimes h dP \otimes K = \int h dQ K. \end{aligned}$$

(iii) \Rightarrow (iv) is obvious because $P_F \ll P$.

(iv) \Rightarrow (ii). Let

$$\mathcal{L} := \left\{ f \in \mathcal{L}^1(\mathcal{G}, P) : \lim_\alpha \int f \otimes h dP \otimes K_\alpha = \int f \otimes h dP \otimes K \right. \\ \left. \text{for every } h \in C_b(\mathcal{X}) \right\}.$$

Then \mathcal{L} is a vector subspace of $\mathcal{L}^1(\mathcal{G}, P)$ with $\{1_G : G \in \mathcal{E}\} \subset \mathcal{L}$, in particular $1_\Omega \in \mathcal{L}$, and if $f_k \in \mathcal{L}$, $f \in \mathcal{L}^1(\mathcal{G}, P)$, $f_k \geq 0$, $f \geq 0$ such that $f_k \uparrow f$, then $f \in \mathcal{L}$. In fact,

$$\begin{aligned} &\left| \int f \otimes h dP \otimes K_\alpha - \int f \otimes h dP \otimes K \right| \\ &\leq \int |f \otimes h - f_k \otimes h| dP \otimes K_\alpha + \left| \int f_k \otimes h dP \otimes K_\alpha - \int f_k \otimes h dP \otimes K \right| \\ &\quad + \int |f_k \otimes h - f \otimes h| dP \otimes K \\ &\leq 2\|h\|_{\sup} \int (f - f_k) dP + \left| \int f_k \otimes h dP \otimes K_\alpha - \int f_k \otimes h dP \otimes K \right| \end{aligned}$$

and hence

$$\limsup_\alpha \left| \int f \otimes h dP \otimes K_\alpha - \int f \otimes h dP \otimes K \right| \leq 2\|h\|_{\sup} \int (f - f_k) dP.$$

Letting $k \rightarrow \infty$ yields by monotone convergence

$$\lim_\alpha \int f \otimes h dP \otimes K_\alpha = \int f \otimes h dP \otimes K.$$

Thus $f \in \mathcal{L}$. One can conclude that $\mathcal{D} := \{G \in \mathcal{G} : 1_G \in \mathcal{L}\}$ is a Dynkin-system so that $\mathcal{D} = \sigma(\mathcal{E}) = \mathcal{G}$. This clearly yields $\mathcal{L} = \mathcal{L}^1(\mathcal{G}, P)$, hence (ii).

(ii) \Rightarrow (i). For $f \in \mathcal{L}^1(P)$ we have $E(f|\mathcal{G}) \in \mathcal{L}^1(\mathcal{G}, P)$ and thus in view of the \mathcal{G} -measurability of K_α and K

$$\begin{aligned}
\lim_{\alpha} \int f \otimes h dP \otimes K_{\alpha} &= \lim_{\alpha} \int E(f|\mathcal{G}) \otimes h dP \otimes K_{\alpha} \\
&= \int E(f|\mathcal{G}) \otimes h dP \otimes K = \int f \otimes h dP \otimes K
\end{aligned}$$

for every $h \in C_b(\mathcal{X})$. □

Exercise 2.2 Prove that weak convergence $K_{\alpha} \rightarrow K$ is also equivalent to $QK_{\alpha} \rightarrow QK$ weakly for every probability distribution Q on \mathcal{F} such that $Q \equiv P$, where \equiv means mutual absolute continuity.

Exercise 2.3 Show that weak convergence is preserved under an absolutely continuous change of measure, that is, $\tau(Q) \subset \tau(P)$, if $Q \ll P$, and hence $\tau(Q) = \tau(P)$, if $Q \equiv P$.

Exercise 2.4 One may consider $\mathcal{M}^1(\mathcal{X})$ as a subset of \mathcal{K}^1 . Show that $\tau \cap \mathcal{M}^1(\mathcal{X})$ is the weak topology on $\mathcal{M}^1(\mathcal{X})$.

The weak topology on \mathcal{K}^1 is not necessarily Hausdorff and the weak limit kernel is not unique, but it is P -almost surely unique. In fact, if $\int f \otimes h dP \otimes K_1 = \int f \otimes h dP \otimes K_2$ for $K_1, K_2 \in \mathcal{K}^1$ and every $f \in \mathcal{L}^1(P)$ and $h \in C_b(\mathcal{X})$, then $\int h dP_F K_1 = \int h dP_F K_2$ for every $h \in C_b(\mathcal{X})$ so that $P_F K_1 = P_F K_2$ for every $F \in \mathcal{F}$ with $P(F) > 0$. This implies $K_1(\cdot, B) = K_2(\cdot, B)$ P -almost surely for every $B \in \mathcal{B}(\mathcal{X})$ and thus $K_1 = K_2$ P -almost surely by Lemma 2.1 (b).

The following notion is sometimes useful.

Definition 2.4 Assume that \mathcal{X} is polish. Let $K \in \mathcal{K}^1$ and $\mathcal{G} \subset \mathcal{F}$ be a sub- σ -field. Then by disintegration of measures there exists a (P -almost surely unique) kernel $H \in \mathcal{K}^1(\mathcal{G})$ such that

$$P \otimes H|_{\mathcal{G}} \otimes \mathcal{B}(\mathcal{X}) = (P|_{\mathcal{G}}) \otimes H = P \otimes K|_{\mathcal{G}} \otimes \mathcal{B}(\mathcal{X})$$

(see Theorem A.6). The Markov kernel H is called the *conditional expectation of K w.r.t. \mathcal{G}* and is denoted by $E(K|\mathcal{G})$.

For a sub- σ -field $\mathcal{G} \subset \mathcal{F}$, the weak topology on $\mathcal{K}^1(\mathcal{G})$ is denoted by $\tau(\mathcal{G}) = \tau(\mathcal{G}, P)$. We will see that the map $\mathcal{K}^1 \mapsto \mathcal{K}^1(\mathcal{G})$ or $\mathcal{K}^1, K \mapsto E(K|\mathcal{G})$, is weakly continuous.

Corollary 2.5 Let $(K_{\alpha})_{\alpha}$ be a net in \mathcal{K}^1 , $K \in \mathcal{K}^1$ and $\mathcal{G} \subset \mathcal{F}$ a sub- σ -field.

- (a) $\tau(\mathcal{G})$ coincides with the topology induced by τ on $\mathcal{K}^1(\mathcal{G})$, that is $\tau(\mathcal{G}) = \tau \cap \mathcal{K}^1(\mathcal{G})$.
- (b) Assume that \mathcal{X} is polish. If $K_{\alpha} \rightarrow K$ weakly, then $E(K_{\alpha}|\mathcal{G}) \rightarrow E(K|\mathcal{G})$ weakly (in \mathcal{K}^1 and $\mathcal{K}^1(\mathcal{G})$).
- (c) Assume that \mathcal{X} is polish. If $\{N \in \mathcal{F} : P(N) = 0\} \subset \mathcal{G}$, then $\mathcal{K}^1(\mathcal{G})$ is τ -closed in \mathcal{K}^1 .

Proof (a) is an immediate consequence of Theorem 2.3.

(b) is immediate from Theorem 2.3 and

$$\int f \otimes h dP \otimes E(K|\mathcal{G}) = \int f \otimes h dP \otimes K$$

for $K \in \mathcal{K}^1$, $f \in \mathcal{L}^1(\mathcal{G}, P)$ and $h \in C_b(\mathcal{X})$.

(c) Let $(K_\alpha)_\alpha$ be a net in $\mathcal{K}^1(\mathcal{G})$, $K \in \mathcal{K}^1$ and assume $K_\alpha \rightarrow K$ weakly in \mathcal{K}^1 . Then by (b), $K_\alpha = E(K_\alpha|\mathcal{G}) \rightarrow E(K|\mathcal{G})$ weakly in \mathcal{K}^1 and hence, by almost sure uniqueness of limit kernels, we obtain $E(K|\mathcal{G}) = K$ P -almost surely. The assumption on \mathcal{G} now implies $K \in \mathcal{K}^1(\mathcal{G})$. Thus $\mathcal{K}^1(\mathcal{G})$ is τ -closed. \square

We provide further characterizations of weak convergence. Recall that a function $h : \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ on a topological space \mathcal{Y} is said to be lower semicontinuous if $\{h \leq r\}$ is closed for every $r \in \overline{\mathbb{R}}$ or, what is the same, if $h(y) \leq \liminf_\alpha h(y_\alpha)$ for every net $(y_\alpha)_\alpha$ and y in \mathcal{Y} with $y_\alpha \rightarrow y$. The function h is upper semicontinuous if $-h$ is lower semicontinuous. A function which is both upper and lower semicontinuous is continuous.

Theorem 2.6 *For a net $(K_\alpha)_\alpha$ and K in \mathcal{K}^1 the following statements are equivalent:*

- (i) $K_\alpha \rightarrow K$ weakly,
- (ii) $\lim_\alpha \int g dP \otimes K_\alpha = \int g dP \otimes K$ for every measurable, bounded function $g : (\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B}(\mathcal{X})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $g(\omega, \cdot) \in C_b(\mathcal{X})$ for every $\omega \in \Omega$,
- (iii) (For \mathcal{X} polish) $\limsup_\alpha \int g dP \otimes K_\alpha \leq \int g dP \otimes K$ for every measurable function $g : (\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B}(\mathcal{X})) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ which is bounded from above such that $g(\omega, \cdot)$ is upper semicontinuous for every $\omega \in \Omega$,
- (iv) (For \mathcal{X} polish) $\liminf_\alpha \int g dP \otimes K_\alpha \geq \int g dP \otimes K$ for every measurable function $g : (\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B}(\mathcal{X})) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ which is bounded from below such that $g(\omega, \cdot)$ is lower semicontinuous for every $\omega \in \Omega$.

Note that statements (ii)–(iv) say that the function $\mathcal{K}^1 \rightarrow \overline{\mathbb{R}}, K \mapsto \int g dP \otimes K$, is weakly continuous, upper semicontinuous and lower semicontinuous, respectively. Moreover, it is interesting to note that the $\mathcal{F} \otimes \mathcal{B}(\mathcal{X})$ -measurability of the function g in (ii) already follows from the \mathcal{F} -measurability of $g(\cdot, x)$ for every $x \in \mathcal{X}$; see [18], Lemma 1.1.

Proof (i) \Rightarrow (ii) and (i) \Rightarrow (iv). Let $g : \Omega \times \mathcal{X} \rightarrow \overline{\mathbb{R}}$ be as in (iv). Replacing g by $g - \inf g$, we may assume $g \geq 0$. There exists a totally bounded metric d inducing the topology of \mathcal{X} so that the subspace $U_b(\mathcal{X}, d)$ of $C_b(\mathcal{X})$ consisting of all d -uniformly continuous, bounded functions is separable; see [26], Theorem 2.8.2, [69], Lemma II.6.3. Let $\{h_n : n \in \mathbb{N}\}$ be a countable dense subset of $U_b(\mathcal{X}, d)$. We obtain the representation

$$g(\omega, x) = \sup \{h_n^+(x) : h_n \leq g(\omega, \cdot), n \in \mathbb{N}\}$$

for every $\omega \in \Omega$ and $x \in \mathcal{X}$. To see this, let $\varepsilon > 0$, fix $\omega \in \Omega$ and $x \in \mathcal{X}$ and consider the functions

$$g_k : \mathcal{X} \rightarrow \mathbb{R}, \quad g_k(y) := \inf_{z \in \mathcal{X}} \{k \wedge g(\omega, z) + kd(y, z)\} - \varepsilon$$

for $k \in \mathbb{N}$. One easily checks that g_k is d -Lipschitz and thus $g_k \in U_b(\mathcal{X}, d)$, $g_k \leq g(\omega, \cdot) - \varepsilon$ and $g_k(y) \uparrow g(\omega, y) - \varepsilon$ for every $y \in \mathcal{X}$. If $g(\omega, x) < \infty$, choose $k \in \mathbb{N}$ such that $g_k(x) \geq g(\omega, x) - 2\varepsilon$ and then $m \in \mathbb{N}$ such that $\|g_k - h_m\|_{\sup} \leq \varepsilon$. This implies $h_m \leq g(\omega, \cdot)$ and $h_m(x) \geq g(\omega, x) - 3\varepsilon$, hence

$$\sup \{h_n^+(x) : h_n \leq g(\omega, \cdot)\} \geq \sup \{h_n(x) : h_n \leq g(\omega, \cdot)\} \geq g(\omega, x) - 3\varepsilon.$$

Since ε was arbitrary, we get the above representation. If $g(\omega, x) = \infty$, for $t > 0$, choose $k \in \mathbb{N}$ such that $g_k(x) > t + \varepsilon$ and $m \in \mathbb{N}$ such that $\|g_k - h_m\|_{\sup} \leq \varepsilon$. Then $h_m \leq g(\omega, \cdot)$ and $h_m(x) > t$ which yields $\sup \{h_n^+(x) : h_n \leq g(\omega, \cdot)\} = \infty$.

Setting $F_n := \{\omega \in \Omega : h_n \leq g(\omega, \cdot)\}$ for $n \in \mathbb{N}$ we obtain $g(\omega, x) = \sup \{1_{F_n} \otimes h_n^+(\omega, x) : n \in \mathbb{N}\}$ for every $\omega \in \Omega$ and $x \in \mathcal{X}$.

Now assume that g is bounded and $g(\omega, \cdot) \in C_b(\mathcal{X})$ for every $\omega \in \Omega$. Then

$$F_n = \bigcap_{x \in \mathcal{X}_0} \{h_n(x) \leq g(\cdot, x)\}$$

for some countable dense subset \mathcal{X}_0 of \mathcal{X} and hence $F_n \in \mathcal{F}$. In view of the rather obvious fact that

$$V := \left\{ \sum_{i=1}^n 1_{H_i} \otimes k_i : H_i \in \mathcal{F} \text{ pairwise disjoint, } k_i \in C_b(\mathcal{X})_+, n \in \mathbb{N} \right\}$$

is a lattice in the pointwise ordering there exists a nondecreasing sequence $(v_n)_{n \geq 1}$ in V such that $g(\omega, x) = \sup_{n \in \mathbb{N}} v_n(\omega, x)$ for every $\omega \in \Omega$ and $x \in \mathcal{X}$. Using monotone convergence we obtain that the map $K \mapsto \int g dP \otimes K = \sup_{n \in \mathbb{N}} \int v_n dP \otimes K$ is lower τ -semicontinuous on \mathcal{K}^1 . This can be applied to the function $-g + \sup g$ and yields that the map $K \mapsto \int g dP \otimes K$ is τ -continuous, hence (ii).

In the setting of (iv) the proof is a bit more involved because F_n is not necessarily in \mathcal{F} . However,

$$F_n^c = \bigcup_{x \in \mathcal{X}} \{\omega \in \Omega : h_n(x) > g(\omega, x)\}$$

is the image of $A_n := \{(\omega, x) \in \Omega \times \mathcal{X} : h_n(x) > g(\omega, x)\} \in \mathcal{F} \otimes \mathcal{B}(\mathcal{X})$ under the projection $\pi_\Omega : \Omega \times \mathcal{X} \rightarrow \Omega$ onto Ω , that is

$$\pi_\Omega(A_n) = \bigcup_{x \in \mathcal{X}} A_{n,x} = F_n^c,$$

and hence, using that \mathcal{X} is polish, it follows from a projection theorem that F_n belongs to the P -completion of \mathcal{F} ; see [83], Theorem 4. Therefore, for every $n \in \mathbb{N}$ there is a set $G_n \in \mathcal{F}$ and a P -null set $N_n \in \mathcal{F}$ such that $G_n \subset F_n$ and $F_n \setminus G_n \subset N_n$. Defining $N := \bigcup_{n \in \mathbb{N}} N_n$ we obtain $g(\omega, x) = \sup \{1_{G_n} \otimes h_n^+(\omega, x) : n \in \mathbb{N}\}$ for every $\omega \in N^c$ and $x \in \mathcal{X}$. As above, this yields the lower τ -semicontinuity of $K \mapsto \int g dP \otimes K$, hence (iv).

(ii) \Rightarrow (i) is obvious, as is (iv) \Leftrightarrow (iii) \Rightarrow (ii). \square

Finally we mention a characterization of compactness in \mathcal{K}^1 . For this, it is convenient to identify Markov kernels in \mathcal{K}^1 that agree P -almost surely. One arrives at the space $K^1(P) = K^1(\mathcal{F}, P) = K^1(\mathcal{F}, P, \mathcal{X})$ of P -equivalence classes. The weak topology on $K^1(P)$, still denoted by $\tau(P)$, is now Hausdorff. For a sub- σ -field $\mathcal{G} \subset \mathcal{F}$, let $K^1(\mathcal{G}, P)$ denote the subspace of $K^1(P)$ consisting of equivalence classes which contain at least one representative from $\mathcal{K}^1(\mathcal{G})$. By Corollary 2.5 (c), the set $K^1(\mathcal{G}, P)$ is weakly closed in $K^1(P)$ provided \mathcal{X} is polish.

A net in $\mathcal{M}^1(\mathcal{X})$ is called *tight* if the corresponding subset is tight. A weakly convergent sequence in $\mathcal{M}^1(\mathcal{X})$ is tight provided \mathcal{X} is polish. In fact, weak convergence $\nu_n \rightarrow \nu$ in $\mathcal{M}^1(\mathcal{X})$ obviously implies weak compactness of $\{\nu_n : n \in \mathbb{N}\} \cup \{\nu\}$, hence $\{\nu_n : n \in \mathbb{N}\}$ is relatively weakly compact and thus tight.

Theorem 2.7 *Assume that \mathcal{X} is polish. For a subset $\Gamma \subset K^1(P)$,*

(i) Γ *is relatively $\tau(P)$ -compact*

if and only if

(ii) $P\Gamma := \{PK : K \in \Gamma\}$ *is relatively compact in $\mathcal{M}^1(\mathcal{X})$,*

and then

(iii) Γ *is relatively sequentially $\tau(P)$ -compact.*

In particular, if $(K_\alpha)_\alpha$ is a net (sequence) in \mathcal{K}^1 such that $(PK_\alpha)_\alpha$ is tight, then $(K_\alpha)_\alpha$ has a weakly convergent subnet (subsequence).

Proof (i) \Rightarrow (ii) is an immediate consequence of the continuity of the map $K \mapsto PK$.

(ii) \Rightarrow (i). Choose as in the proof of Theorem 2.6 a totally bounded metrization of \mathcal{X} . Then the completion \mathcal{Y} of \mathcal{X} is compact and $\mathcal{X} \in \mathcal{B}(\mathcal{Y})$ because \mathcal{X} is, as a polish subspace of the polish space \mathcal{Y} , a G_δ -set, i.e. a countable intersection of open subsets of \mathcal{Y} . Hence $\mathcal{B}(\mathcal{X}) \subset \mathcal{B}(\mathcal{Y})$. Because $U_b(\mathcal{X})$ and $C_b(\mathcal{Y})$ are isometrically isomorphic, it follows from the Portmanteau theorem that $(K^1(P, \mathcal{X}), \tau(P, \mathcal{X}))$ is homeomorphic to the subspace $\{K \in K^1(P, \mathcal{Y}) : PK(\mathcal{X}) = 1\}$ of $(K^1(P, \mathcal{Y}), \tau(P, \mathcal{Y}))$. Identifying these spaces and because $K^1(P, \mathcal{Y})$ is $\tau(P, \mathcal{Y})$ -compact, see [29], [65], [33], Theorem 3.58, the $\tau(P, \mathcal{Y})$ -closure $\overline{\Gamma}$ of Γ in $K^1(P, \mathcal{Y})$ is compact. Let $K \in \overline{\Gamma}$ and let $(K_\alpha)_\alpha$ be a net in Γ such that $K_\alpha \rightarrow K$ weakly in $K^1(P, \mathcal{Y})$. Because $P\Gamma$ is tight in $\mathcal{M}^1(\mathcal{X})$, for every $\varepsilon > 0$ we find a compact set $A \subset \mathcal{X}$ such that $PK_\alpha(A) \geq 1 - \varepsilon$ for every α . By Theorem 2.3 and the Portmanteau theorem we obtain

$$1 - \varepsilon \leq \limsup_{\alpha} PK_\alpha(A) \leq PK(A) \leq PK(\mathcal{X}).$$

This implies $PK(\mathcal{X}) = 1$ and hence $K \in K^1(P, \mathcal{X})$.

(i) \Rightarrow (iii). Let $(K_n)_{n \geq 1}$ be a sequence in Γ and $\mathcal{G} := \sigma(K_n, n \in \mathbb{N})$. If \mathcal{A} denotes a countable generator of $\mathcal{B}(\mathcal{X})$ which is stable under finite intersections, then $\mathcal{G} = \sigma(K_n(\cdot, B), B \in \mathcal{A}, n \in \mathbb{N})$ so that \mathcal{G} is a countably generated sub- σ -field of \mathcal{F} . In view of Corollary 2.5 (a) the set $\{K_n : n \in \mathbb{N}\}$ is relatively $\tau(\mathcal{G}, P)$ -compact and because $(K^1(\mathcal{G}, P), \tau(\mathcal{G}, P))$ is metrizable, see [33], Proposition 3.25, [18], Theorem 4.16, $(K_n)_{n \geq 1}$ has a $\tau(\mathcal{G}, P)$ -convergent subsequence which is again by Corollary 2.5 (a) also $\tau(P)$ -convergent. \square

Exercise 2.5 Show that one can replace in the last part of Theorem 2.7 the tightness of the net $(PK_\alpha)_\alpha$ by its weak convergence in $\mathcal{M}^1(\mathcal{X})$.

Exercise 2.6 Assume that \mathcal{X} is polish and let $\Gamma \subset \mathcal{K}^1$. Regarding each $K \in \mathcal{K}^1$ as an $(\mathcal{M}^1(\mathcal{X}), \mathcal{B}(\mathcal{M}^1(\mathcal{X})))$ -valued random variable, prove that $P\Gamma$ is tight in $\mathcal{M}^1(\mathcal{X})$ if and only if $\{P^K : K \in \Gamma\}$ is tight in $\mathcal{M}^1(\mathcal{M}^1(\mathcal{X}))$. Here P^K denotes the image measure.

Exercise 2.7 Let \mathcal{Y} be a further separable metrizable topological space. Show that the weak topology on $\mathcal{M}^1(\mathcal{X} \times \mathcal{Y})$ is generated by the functions

$$\mu \mapsto \int h \otimes k d\mu, \quad h \in C_b(\mathcal{X}), \quad k \in C_b(\mathcal{Y})$$

and the weak topology on $\mathcal{K}^1(\mathcal{F}, \mathcal{X} \times \mathcal{Y})$ is generated by the functions

$$H \mapsto \int 1_F \otimes h \otimes k dP \otimes H, \quad F \in \mathcal{F}, \quad h \in C_b(\mathcal{X}), \quad k \in C_b(\mathcal{Y}).$$

Exercise 2.8 Let \mathcal{Y} be a further separable metrizable space. Let $(H_\alpha)_\alpha$ be a net in $\mathcal{K}^1(\mathcal{F}, \mathcal{X})$, $H \in \mathcal{K}^1(\mathcal{F}, \mathcal{X})$ and let $(K_\alpha)_\alpha$ be a net in $\mathcal{K}^1(\mathcal{F}, \mathcal{Y})$, $K \in \mathcal{K}^1(\mathcal{F}, \mathcal{Y})$. Assume that $H_\alpha \rightarrow H$ weakly and

$$\int k(y) K_\alpha(\cdot, dy) \rightarrow \int k(y) K(\cdot, dy) \quad \text{in } \mathcal{L}^1(P) \quad \text{for every } k \in C_b(\mathcal{Y}).$$

Show that $H_\alpha \otimes K_\alpha \rightarrow H \otimes K$ weakly in $\mathcal{K}^1(\mathcal{F}, \mathcal{X} \times \mathcal{Y})$.

Stable Convergence and Stable Limit Theorems

Häusler, E.; Luschgy, H.

2015, X, 228 p., Hardcover

ISBN: 978-3-319-18328-2