Chapter 18

Convergence in Distribution on Polish Spaces

We want to extend the concept of convergence in distribution to probability spaces other than $(\mathbb{R}, \mathcal{B})$. Certain metric spaces, known as 'Polish spaces', play a central role. Particularly important examples of Polish spaces are the real line, the extended real line, d-dimensional Euclidean space, infinite products of intervals, and spaces of continuous functions. Thus, this chapter may be viewed as a mechanism for extending the concepts and results discussed in Chapter 14 to a wide variety of settings. (Basic facts about metric spaces are treated briefly in Appendix B. Some of the topology in Appendix C is also relevant.)

Particular attention will be given to distributions on $\mathbb{R}^d$. A central limit theorem will be proved by using the 'Cramér-Wold Device' which reduces certain problems for $\mathbb{R}^d$ to problems for $\mathbb{R}$.

18.1. Polish spaces

Much of the theory developed in Chapter 14 can be adapted to a certain type of metric space which we now define.

Definition 1. A *Polish space* is a complete metric space that has a countable dense subset.

The real line $\mathbb{R}$ (with the usual Euclidean metric) is a Polish space, with the rational numbers constituting a countable dense subset. As described in Appendix B, the extended real line $\overline{\mathbb{R}}$ is a complete metric space with the distance between x and y defined as $|\arctan y - \arctan x|$. It is a Polish space because the rational numbers constitute a countable dense set.

Remark 1. Because the definition of Polish space requires completeness of the metric, the correct choice of metric is important. For example, if we used the metric $\rho(x, y) = |\arctan x - \arctan y|$ in $\mathbb{R}$, we would not have completeness, despite the fact that the open sets arising from this metric are the same as those arising from the usual metric.

Another important Polish space is $\mathbb{R}^d$ with the usual (Euclidean) metric: the distance between x and y equals

$$\sqrt{\sum_{j=1}^{d}(y_j - x_j)^2}\,.$$

The set of points having rational coordinates is a countable dense set. Three other metrics for $\mathbb{R}^d$ that give the same open sets as this metric and also make $\mathbb{R}^d$ into a Polish space are

$$\sum_{j=1}^{d}|y_j - x_j|\,, \quad \sum_{j=1}^{d}\big(|y_j - x_j| \wedge 1\big)\,,$$

and

$$\sum_{j=1}^{d}\frac{|y_j - x_j| \wedge 1}{2^j}\,. \tag{18.1}$$

Although (18.1) is perhaps the most complicated of the alternate metrics for $\mathbb{R}^d$, it has the advantage that it generalizes easily to $\mathbb{R}^\infty$, as shown by the following example.

Example 1. We use $\mathbb{R}^\infty$ to denote the space of all sequences $(x_1, x_2, \dots)$ of real numbers, with the product topology. We want to *metrize* this topological space; that is, we want to make it into a metric space in such a way that the metric gives the same open sets as does the product topology. We define the distance between two sequences x and y to be

$$\rho(x,y) = \sum_{j=1}^{\infty}\frac{|y_j - x_j| \wedge 1}{2^j}\,. \tag{18.2}$$

It is straightforward to check that this definition gives a metric for $\mathbb{R}^\infty$. We will check that this metric gives the same open sets as does the product topology.

We must show (i) if O is an open set in the product topology, then for each $x \in O$, there is an open set U in the topology given by the metric ρ such that $x \in U$ and $U \subseteq O$, and (ii) if U is an open set in the topology given by ρ, then for each $x \in U$, there is an open set O in the product topology such that $x \in O$ and $O \subseteq U$.

Every open set O in the product topology is the union of sets of the form

$$\{(x_1, x_2, \dots) \colon (x_1, \dots, x_d) \in O_d\} \tag{18.3}$$

for some positive integer d, where O_d is an open set in $\mathbb{R}^d$. It follows that in proving (i) and (ii), we may restrict our attention to sets O of the form (18.3). Similarly, every open set U in the topology given by the metric ρ is a union of sets that are open balls in that metric, so in (i) and (ii) we only need to consider open balls U.

We will prove (ii), and leave the proof of (i) to the reader. Let U be an open ball in $\mathbb{R}^\infty$ with the metric ρ, and pick a point $x = (x_1, x_2, \dots) \in U$. By a standard argument involving the triangle inequality, there exists an $\varepsilon > 0$ such that the open ball $U' = \{y\colon \rho(x,y) < \varepsilon\}$ is contained in U. We will find a set of the form (18.3) that contains x and is contained in U'. Choose d such that $2^{d-1} > \frac{1}{\varepsilon}$, and let O_d be the open ball centered at the point $(x_1, \dots, x_d)$ in $\mathbb{R}^d$ with radius $\frac{\varepsilon}{2}$, using the metric (18.1). Let O be defined in terms of O_d as in (18.3). Clearly $x \in O$. Moreover, it is easy to check from the definition of ρ that if $y \in O$, then $\rho(x,y) < \varepsilon$, so $O \subseteq U'$, as desired.

Those sequences containing only rational terms and only finitely many terms different from 0 constitute a countable dense set in $\mathbb{R}^\infty$. The metric space $\mathbb{R}^\infty$ is also complete, a consequence of the second problem below and the completeness of $\mathbb{R}$. Therefore $\mathbb{R}^\infty$ is a Polish space.

Problem 1. Decide if the following sequence in $\mathbb{R}^\infty$ converges and if so to what:

$$\big((1,0,0,0,0,\dots),\,(0,2,0,0,0,\dots),\,(0,0,4,0,0,\dots),\,(0,0,0,8,0,\dots),\,\dots\big)\,.$$

Problem 2. Let x_n, $n = 1, 2, \dots$, and y be members of $\mathbb{R}^\infty$. Prove that $x_n \to y$ as $n \to \infty$ if and only if $x_{j,n} \to y_j$ where $x_{j,n}$ is the j^{th} term of x_n and y_j is the j^{th} term of y. (Comment: One says that the topology that we have given $\mathbb{R}^\infty$ is the topology of *coordinate-wise convergence*.)

The method used in Example 1 to find a suitable metric for $\mathbb{R}^\infty$ generalizes easily to countable products of arbitrary Polish spaces.

Proposition 2. *For $j = 1, 2, \dots$, let (Ψ_j, ρ_j) be a Polish space, and let $\Psi = \bigotimes_{j=1}^\infty \Psi_j$, with the topology on Ψ being the product topology. For $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ in Ψ, define*

$$\rho(x,y) = \sum_{j=1}^\infty \frac{\rho_j(x_j, y_j) \wedge 1}{2^j}\,.$$

Then (Ψ, ρ) is a Polish space.

In Example 1 and the preceding proposition, the setting is one in which the space of interest already has a natural topology attached to it. Therefore the problem was to construct a metric consistent with the topology so that the resulting metric space is a Polish space. The following problem presents a situation where there is already a natural choice for the metric.

Problem 3. Let $\mathbf{C}[0,1]$ denote the metric space of continuous $\mathbb{R}$-valued functions on $[0,1]$ with the distance between two functions f and g being defined as the quantity

$$\max\{|f(t) - g(t)|\colon t \in [0,1]\}\,. \tag{18.4}$$

Prove that $\mathbf{C}[0,1]$ is a Polish space and that the Borel σ-field equals

$$\sigma(\{f\colon f(t) \in B\}\colon t \in [0,1], \text{Borel } B \subseteq \mathbb{R}).$$

The following proposition provides further examples of Polish spaces.

Proposition 3. *A closed subset of a Polish space is a Polish space with the inherited metric.*

PROOF. Let (Ψ, ρ) denote the Polish space and C the closed subset. By Problem 1 of Appendix B, (C, ρ) is a metric space.

Consider a Cauchy sequence in C. It converges to a member of ψ of Ψ. This point ψ must be a member of C; otherwise C would not be closed. Thus C is complete.

Let D be a countable dense subset of Ψ. For each positive integer n let D_n consist of those members of D whose distance from C is less than $\frac{1}{n}$. For each member ψ of D_n choose a member of C, conceivably ψ itself, whose distance from ψ is less than $\frac{1}{n}$ and let E_n denote the set of such chosen points. Every point in C is within a distance of $\frac{2}{n}$ of some member of E_n. Hence, $\bigcup_{n=1}^{\infty} E_n$ is a countable dense subset of C. □

Problem 4. Explain why $\{f \in \mathbf{C}[0,1]\colon f(0) = 0\}$ is a Polish space, with the distance between f and g being specified by (18.4).

* **Problem 5.** Give the set $\mathbf{C}[0,\infty)$ of continuous $\mathbb{R}$-valued functions on $[0,\infty)$ a metric so that it becomes a Polish space with the topology of uniform convergence on bounded sets.

Example 2. [Infinite-dimensional cube] Consider the space $[0,1]^\infty$. This is the set of all sequences $(x_1, x_2, \dots)$ of real numbers belonging to the interval $[0,1]$. It is a subset of the space $\mathbb{R}^\infty$ introduced in Example 1, and it is easy to check that it is closed, since it is a product of closed sets. By Proposition 3, it is itself a Polish space, with distance function ρ given by (18.2).

We may also take the point of view that $[0,1]^\infty$ is a countable product of Polish spaces, and hence is itself a Polish space by Proposition 2. And since it is also the product of compact sets, it is compact by the Tychonoff Theorem (Theorem 3 of Appendix C). This fact gives importance to the next result, which says that an arbitrary Polish space is topologically equivalent to a Borel subset of the infinite-dimensional cube.

Lemma 4. *For any Polish space Ψ, there exists a function φ from Ψ onto a Borel subset of $[0,1]^\infty$ such that φ is continuous and one-to-one on Ψ and φ^{-1} is continuous (and one-to-one) on $\varphi(\Psi)$.*

PROOF. Let Ψ be an arbitrary Polish space with metric σ, and let $(\psi_n : n = 1, 2, \dots)$ be a countable dense subset of Ψ. Consider the function $\varphi : \Psi \to [0,1]^\infty$ defined by

$$\psi \rightsquigarrow \big(\sigma(\psi, \psi_1) \wedge 1\,,\, \sigma(\psi, \psi_2) \wedge 1\,,\, \dots\big)\,.$$

We claim that φ is one-to-one and that both φ and φ^{-1} (defined on the image of φ) are continuous.

The continuity of φ is an immediate consequence of the continuity of the functions $\psi \rightsquigarrow \sigma(\psi, \psi_n)$ for all n. To prove that φ is one-to-one, let ψ and η be two distinct members of Ψ, and let $\varepsilon = \sigma(\psi, \eta) \wedge 2$. By the definition of a metric, $\varepsilon > 0$. Using the fact that the set $(\psi_n : n = 1, 2, \dots)$ is dense, choose k so that $\sigma(\psi, \psi_k) < \varepsilon/2$. It follows from the triangle inequality that $\sigma(\eta, \psi_k) > \varepsilon/2$, so $\varphi(\eta)$ and $\varphi(\psi)$ necessarily differ at the k^{th} coordinate. Therefore, $\varphi(\eta) \neq \varphi(\psi)$, and we may conclude that φ is one-to-one.

To show that φ^{-1} is continuous at an arbitrary point $(x_1, x_2, \dots)$ in $\varphi(\Psi)$, set $\varphi^{-1}(x_1, x_2, \dots) = \psi$, fix $\varepsilon \in (0, \frac{1}{3})$, and then consider an arbitrary $(y_1, y_2, \dots) \in \varphi(\Psi)$ for which

$$\rho((x_1, x_2, \dots), (y_1, y_2, \dots)) < \frac{\varepsilon}{3 \cdot 2^k}\,,$$

where ρ is given by (18.2) and k is chosen so that $\sigma(\psi, \psi_k) < \varepsilon/3$. Let $\eta = \varphi^{-1}(y_1, y_2, \dots)$. Note that, by the definition of φ, $\sigma(\eta, \psi_k) = y_k$ and $\sigma(\psi, \psi_k) = x_k$. It follows from the definition of the metric ρ that

$$|\sigma(\eta, \psi_k) - \sigma(\psi, \psi_k)| < 2^k \rho((x_1, x_2, \dots), (y_1, y_2, \dots)) < \frac{\varepsilon}{3}\,,$$

so $\sigma(\eta, \psi_k) < 2\varepsilon/3$. Thus

$$\sigma(\psi, \eta) \le \sigma(\psi, \psi_k) + \sigma(\eta, \psi_k) < \varepsilon\,.$$

The continuity of φ^{-1} follows.

We will prove that $\varphi(\Psi)$ is a Borel set by writing it in terms of countably many operations involving open and closed subsets of $[0,1]^\infty$. Let D be a countable dense subset of Ψ. For each $d \in D$ and each positive integer k, let $B(d, 1/k)$ be the open ball of radius $1/k$ centered at d. By the continuity of φ^{-1}, the set $\varphi(B(d, 1/k))$ is an open subset of $\varphi(\Psi)$ in the relative topology, so there exists a set $V(d, k)$ which is open in the topology of $[0,1]^\infty$ such that

$$\varphi(B(d, 1/k)) = V(d, k) \cap \varphi(\Psi)\,.$$

Let $V(k) = \bigcup_{d \in D} V(d, k)$. We claim that

$$\varphi(\Psi) = \overline{\varphi(\Psi)} \cap \left(\bigcap_{k=1}^{\infty} V(k)\right), \tag{18.5}$$

where $\overline{\varphi(\Psi)}$ is the closure in $[0,1]^\infty$ of $\varphi(\Psi)$. Clearly, (18.5) implies that $\varphi(\Psi)$ is Borel.

Each member of $\varphi(\Psi)$ belongs to each $V(k)$ and thus to the set on the right side of (18.5). It remains to show that if v belongs to the set on the right side of

(18.5), then v belongs to $\varphi(\Psi)$. For each k, the fact that $v \in V(k)$ implies the existence of $d_k \in D$ such that $v \in V(d_k, k)$. Since v is in the closure of $\varphi(\Psi)$, every open neighborhood of v contains a member of $\varphi(\Psi)$. In particular, for each k, we may choose $v_k \in \varphi(\Psi)$ such that

$$v_k \in V(d_1, 1) \cap \cdots \cap V(d_k, k) \cap \{u \in [0,1]^\infty : \rho(u, v) < \tfrac{1}{k}\}.$$

Note that $v_k \to v$ as $k \to \infty$. Also note that $(\varphi^{-1}(v_k) : k = 1, 2, \dots)$ is a Cauchy sequence in Ψ, since for $j, k \geq m$, v_j and v_k are both members of $V(d_m, m)$ and hence $\sigma(\varphi^{-1}(v_j), \varphi^{-1}(v_k)) < \frac{2}{m}$. Since Ψ is a Polish space, $\psi = \lim_k \varphi^{-1}(v_k)$ exists. By the continuity of φ, $\varphi(\psi) = \lim_k v_k = v$. Thus $v \in \varphi(\Psi)$, as desired. □

Problem 6. [Infinite-dimensional space-filling curve] It is a well-known fact, often discussed in topology texts, that there exists a continuous function h from $[0,1]$ onto $[0,1]^2$. Such a function is a *Peano curve*, named after its discoverer. Use h to construct a continuous function from $[0,1]$ onto $[0,1]^\infty$. *Hint:* Since h is a continuous $[0,1]^2$-valued function, it can be expressed in terms of two continuous $[0,1]$-valued functions as $h = (h_1, h_2)$. Consider the function

$$(h_1, h_1 \circ h_2, h_1 \circ h_2 \circ h_2, h_1 \circ h_2 \circ h_2 \circ h_2, \dots).$$

18.2. Definition of and criteria for convergence

When we view Polish spaces as measurable spaces we follow our customary convention that the σ-field is the Borel σ-field unless something to the contrary is explicitly stated. In particular, throughout this section this convention is in force. The following definition is motivated by Proposition 7 of Chapter 14.

Definition 5. Let Q and Q_n, $n = 1, 2, \dots$, be probability measures on a Polish space Ψ. Then $(Q_n : n = 1, 2, \dots)$ *converges* to Q, denoted by $Q_n \to Q$ as $n \to \infty$, if, for every $\mathbb{R}$-valued bounded continuous function g on Ψ, $\int g\, dQ_n \to \int g\, dQ$ as $n \to \infty$.

For random variables having values in a Polish space, we say that a sequence (X_n) *converges* to X *in distribution* and write

$$X_n \xrightarrow{\mathcal{D}} X \quad \text{as } n \to \infty,$$

if $Q_n \to Q$, where Q_n and Q denote the distributions of X_n and X, respectively.

The following theorem generalizes Proposition 7 of Chapter 14 to the setting of Polish spaces. The name of the theorem refers to the fact that it contains so many conditions. Note the new condition (vi) and the change in condition (ii). We leave it to the reader to check that the set which appears in condition (vi) is a Borel set. As in Proposition 7 of Chapter 14 we use ∂A for the boundary of a set A.

Theorem 6. [Portmanteau] *Let Q and $Q_n, n = 1, 2, \ldots$, be probability measures on a Polish space Ψ. Then the following conditions are equivalent:*

(i) $Q_n \to Q$ *as* $n \to \infty$;

(ii) $\lim_{n\to\infty} \int g\, dQ_n = \int g\, dQ$ *for each bounded uniformly continuous function* g *on* Ψ;

(iii) $\limsup_{n\to\infty} Q_n(C) \leq Q(C)$ *for each closed subset* C *of* Ψ;

(iv) $\liminf_{n\to\infty} Q_n(O) \geq Q(O)$ *for each open subset* O *of* Ψ;

(v) $\lim_{n\to\infty} Q_n(A) = Q(A)$ *for each Borel subset* A *of* Ψ *for which* $Q(\partial A) = 0$;

(vi) $\lim_{n\to\infty} \int g\, dQ_n = \int g\, dQ$ *for each bounded measurable function* g *for which* $Q(\{\psi \in \Psi : g \text{ is discontinuous at } \psi\}) = 0$.

PROOF. That (i) $\Longrightarrow$ (ii) and (vi) $\Longrightarrow$ (i) are both obvious.

The proof that (ii) $\Longrightarrow$ (iii) is essentially the same as the proof of the corresponding part of Proposition 7 of Chapter 14, since the functions introduced in that proof are uniformly continuous and are defined in a manner that works equally well in a metric space.

The proofs that (iii) $\Longleftrightarrow$ (iv) and {(iii), (iv)} $\Longrightarrow$ (v) are also the same as the corresponding parts of the proof of Proposition 7 of Chapter 14.

Finally we prove that (v) $\Longrightarrow$ (vi). Let g be a bounded measurable function from Ψ to $\mathbb{R}$, and let

$$D = \{\psi \in \Psi : g \text{ is discontinuous at } \psi\}.$$

Assume that $Q(D) = 0$. For each n, g may be regarded as a bounded random variable from $(\Psi, \mathcal{A}, Q_n)$ to $(\mathbb{R}, \mathcal{B})$, where $\mathcal{A}$ and $\mathcal{B}$ are the respective Borel σ-fields. Let R_n be the distribution of this random variable. Similarly, g is a bounded random variable from $(\Psi, \mathcal{A}, Q)$ to $(\mathbb{R}, \mathcal{B})$ whose distribution will be denoted by R. We will first show that $R_n \to R$ as $n \to \infty$.

Let B be a Borel subset of $\mathbb{R}$ for which $R(\partial B) = 0$. Then $Q(g^{-1}(\partial B)) = 0$. It is easily checked that if $\psi \in \partial g^{-1}(B)$, then either g is discontinuous at ψ, or $g(\psi) \in \partial B$. Hence $\partial g^{-1}(B) \subseteq D \cup g^{-1}(\partial B)$, so that $Q(\partial g^{-1}(B)) = 0$. Thus, $Q_n(g^{-1}(B)) \to Q(g^{-1}(B))$ as $n \to \infty$, or equivalently, $R_n(B) \to R(B)$ as $n \to \infty$. By (v) of Proposition 7 of Chapter 14, $R_n \to R$ as $n \to \infty$.

By (ii) of Proposition 7 of Chapter 14, $\int h\, dR_n \to \int h\, dR$ as $n \to \infty$ for each bounded continuous $\mathbb{R}$-valued function h on $\mathbb{R}$. We apply this fact with

$$h(x) = \begin{cases} x & \text{if } |x| \leq c \\ -c & \text{if } x < -c \\ c & \text{if } x > c \end{cases}$$

where $c = \sup\{|g(\psi)| : \psi \in \Psi\}$. Doing so completes the proof since $\int g\, dQ_n = \int h\, dR_n$ and $\int g\, dQ = \int h\, dR$. □

Corollary 7. *Let Q and R be two distributions on a Polish space. If*

$$\int g\,dR = \int g\,dQ \tag{18.6}$$

for each bounded uniformly continuous function g on Ψ then $R = Q$.

PROOF. Suppose that (18.6) holds. By letting $R = Q_n$ for all n in the Portmanteau Theorem one gets $R(O) \geq Q(O)$ for every open set O. Interchanging the roles of Q and R gives $Q(O) \geq R(O)$. Since $R(O) = Q(O)$ for every open set O, it follows by the Uniqueness Theorem that $R = Q$. □

Corollary 8. *Let Q and R be two distributions on a Polish space. If both R and Q are limits of the same sequence of distributions, then $R = Q$.*

PROOF. If R and Q are both limits of the same sequence, then by the Portmanteau Theorem, (18.6) holds. □

It is not immediately apparent from the definition of convergence for a sequence of distributions on a Polish space that a given sequence can converge to only one distribution, but the preceding corollary shows that to be the case. (Of course, a sequence that does not converge might have various subsequences that converge to different limits.)

In Chapter 14 we found a connection between convergence in distribution on $\mathbb{R}$ and convergence in probability of sequences of $\mathbb{R}$-valued random variables. This connection will carry over to Polish spaces, once we have appropriately generalized the definition of convergence in probability.

Let $(X_n : n = 1, 2, \dots)$ be a sequence of random variables with values in a Polish space (Ψ, ρ). The sequence (X_n) *converge in probability* to a Ψ-valued random variable X if, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} P\big[\rho(X, X_n) > \varepsilon\big] = 0\,.$$

It is *Cauchy in probability* if, for every $\varepsilon > 0$, there is an integer l such that

$$P\big[\rho(X_n, X_m) > \varepsilon\big] < \varepsilon$$

whenever $m, n \geq l$.

Since a Polish space is a complete metric space, a sequence of random variables with values in a Polish space converges almost surely if and only if it is almost surely Cauchy. The statements and proofs of Theorem 2 and Lemma 24, both of Chapter 12, carry over to the present setting along with Lemma 3 and Problem 39 of that same chapter. Thus, a sequence is Cauchy in probability if and only if it converges in probability. Moreover, almost sure convergence implies convergence in probability. Also, convergence in probability implies almost sure convergence for an appropriate subsequence.

Proposition 9. *Let X and X_n, $n = 1, 2, \ldots$, be random variables on a common probability space having values in a common Polish space. Suppose that the sequence $(X_n : n = 1, 2, \ldots)$ converges to X in probability as $n \to \infty$. Then $X_n \xrightarrow{\mathcal{D}} X$ as $n \to \infty$.*

Problem 7. Prove the preceding proposition. *Hint:* Let g be a uniformly continuous bounded $\mathbb{R}$-valued function defined on the Polish space, and show that the sequence $(g \circ X_n : n = 1, 2, \ldots)$ of $\mathbb{R}$-valued random variables converges in probability to the $\mathbb{R}$-valued random variable $g \circ X$.

The next result says that if a sequence of probability measures on a Polish space converges in distribution, then so does any sequence induced from it by a continuous function into another Polish space.

Proposition 10. *Let Q and $Q_n, n = 1, 2, \ldots$, be probability measures on a common Polish space Ψ. Let h be a continuous function from Ψ to a Polish space Υ, and let R and R_n be the measures induced by h from Q and Q_n, respectively. If $Q_n \to Q$ as $n \to \infty$, then $R_n \to R$ as $n \to \infty$.*

* **Problem 8.** Prove the preceding proposition.

Problem 9. Let X and X_n, $n = 1, 2, \ldots$, be $\mathbf{C}[0,1]$-valued random variables and suppose that $X_n \xrightarrow{\mathcal{D}} X$ as $n \to \infty$. Prove that

$$\max\{X_n(t) : 0 \le t \le 1\} \xrightarrow{\mathcal{D}} \max\{X(t) : 0 \le t \le 1\}$$

as $n \to \infty$.

18.3. Relative sequential compactness

In this section we prove a basis fact about compactness in Polish spaces, introduce the concept of relative sequential compactness for families of probability distributions on a Polish space, and finally prove that any such family is relatively sequentially compact if the Polish space itself is compact.

As described in Appendix B, a set is *totally bounded* if for every $\varepsilon > 0$, it is contained in the union of a finite collection of balls of radius less than ε.

Proposition 11. *A Polish space is compact if and only if it is totally bounded.*

Problem 10. Prove the preceding proposition. *Hint:* Use Proposition 2 and Proposition 3, both of Appendix B.

Problem 11. Either by using the preceding proposition or by a direct argument, show that $\{x : |x(t)| \le 1 \text{ for all } t \in [0,1]\}$ is not a compact subset of the Polish space $\mathbf{C}[0,1]$ described in Problem 3.

Problem 12. Without using Proposition 11, give a direct proof of the total boundedness of the infinite-dimensional cube defined in Example 2.

Definition 12. A family $\mathcal{Q}$ of probability distributions on a Polish space Ψ is *relatively sequentially compact* if every sequence $(Q_n : n = 1, 2, \dots)$ of members of $\mathcal{Q}$ has a convergent subsequence.

The following lemma constitutes a major step in the identification of the relatively sequentially compact families of distributions on a Polish space.

Lemma 13. *Every family of probability distributions on the infinite-dimensional cube* $[0,1]^\infty$ *is relatively sequentially compact.*

PROOF. Let $(Q_n : n = 1, 2, \dots)$ be a sequence of probability measures on $[0,1]^\infty$. By Problem 6 there exists a continuous function g from $[0,1]$ onto $[0,1]^\infty$. Define $f : [0,1]^\infty \to [0,1]$ by

$$f(x) = \inf\{t \in [0,1] : g(t) = x\}\,, \quad x \in [0,1]^\infty\,.$$

It follows from the continuity of g that $g \circ f$ is the identity function. The measurability of f follows from the fact that $\{x : f(x) \leq a\}$ is compact, being the image under g of the compact set $[0,a]$ (see Problem 7 of Appendix B). For $n = 1, 2, \dots$, let R_n be the sequence of probability measures induced on $[0,1]$ by f from Q_n. By Theorem 13 of Chapter 14, there exists a convergent subsequence $(R_{n_k} : k = 1, 2, \dots)$ with limit equal to some probability measure R on $[0,1]$. Let Q be the measure induced on $[0,1]^\infty$ by g. Note that for each n, Q_n is the measure induced by g from R_n. By Proposition 10, the subsequence $(Q_{n_k} : k = 1, 2, \dots)$ converges to Q. $\square$

The following result is worth remembering, even though it is a special case of the forthcoming Theorem 17.

Proposition 14. *Every family of probability distributions on a compact Polish space is relatively sequentially compact.*

Problem 13. Use Problem 7 of Appendix B and Lemma 4 and Lemma 13 of this chapter to prove the preceding proposition.

We conclude this section with a simple result that is quite useful for proving convergence in distribution in Polish spaces.

Proposition 15. *Let* $(Q_n : n = 1, 2, \dots)$ *be a relatively sequentially compact sequence of probability measures on a Polish space such that every convergent subsequence has the same limiting probability measure* Q. *Then* $Q_n \to Q$ *as* $n \to \infty$.

Problem 14. Prove the preceding proposition. *Hint:* See the solution of Proposition 4 of Appendix B.

18.4. Uniform tightness and the Prohorov Theorem

In this section we identify necessary and sufficient conditions for a family of probability measures on a Polish space to be relatively sequentially compact.

Definition 16. A probability distribution on a Polish space Ψ is *tight* if, for every $\varepsilon > 0$, there exists a compact subset K of Ψ such that $Q(K^c) < \varepsilon$. A family $\mathcal{Q}$ of probability distributions on Ψ is *uniformly tight* if, for every $\varepsilon > 0$, there exists a compact subset K of Ψ such that $Q(K^c) < \varepsilon$ for every $Q \in \mathcal{Q}$.

Some use the term 'tight' for a family to mean 'uniformly tight', but we will not use the abbreviated term.

Theorem 17. [Prohorov] *A family of probability measures on a Polish space Ψ is relatively sequentially compact if and only if it is uniformly tight.*

PROOF. Let $\mathcal{Q}$ be a uniformly tight family of probability measures on Ψ, and let $(Q_n : n = 1, 2, \dots)$ be a sequence of members of $\mathcal{Q}$. By definition, there exists, for each $\varepsilon > 0$, a compact set $K_\varepsilon \subseteq \Psi$ with the property that $Q_n(K_\varepsilon) > 1 - \varepsilon$ for all n. We use the function φ defined in Lemma 4 to transfer everything to the Polish space $[0,1]^\infty$. Let $C_\varepsilon = \varphi(K_\varepsilon)$. By Problem 7 of Appendix B, each C_ε is a compact subset of $[0,1]^\infty$. Let $(R_n : n = 1, 2, \dots)$ be the sequence of probability measures on $[0,1]^\infty$ induced by φ from the sequence $(Q_n : n = 1, 2, \dots)$. Note that $R_n(C_\varepsilon) > 1 - \varepsilon$ for all n and ε.

By Lemma 13 there is a subsequence $(R_{n_k} : k = 1, 2, \dots)$ that converges to a probability measure R on $[0,1]^\infty$. By the Portmanteau Theorem, $R(C_\varepsilon) \geq 1 - \varepsilon$ for all ε. Since $C_\varepsilon \subseteq \varphi(\Psi)$ for all ε, it follows that $R(\varphi(\Psi)) = 1$. Let Q be the measure induced by φ^{-1} on Ψ from R. It follows from the continuity of φ^{-1} and Proposition 10 that $Q_{n_k} \to Q$ as $k \to \infty$.

To prove the converse, suppose that $\mathcal{Q}$ is a relatively sequentially compact family of probability measures on Ψ. Let $(\psi_n : n = 1, 2, \dots)$ be a countable dense subset of Ψ, and for each $\delta > 0$, let $B(\psi_n, \delta)$ be the open ball of radius δ about the point ψ_n. Let $\overline{B(\psi_n, \delta)}$ be the closure of $B(\psi_n, \delta)$.

We now show that for each δ, there exists an integer $p(\delta)$ such that for all $Q \in \mathcal{Q}$,

$$Q\left(\bigcup_{n=1}^{p(\delta)} B(\psi_n, \delta)\right) > 1 - \delta\,.$$

Suppose that such an integer $p(\delta)$ does not exist for some particular choice of $\delta > 0$. Then for each positive integer m, there exists a probability measure

$Q_m \in \mathcal{Q}$ such that

$$Q_m\Big(\bigcup_{n=1}^{m} B(\psi_n, \delta)\Big) \leq 1 - \delta\,.$$

By relative sequential compactness, the sequence $(Q_m \colon m = 1, 2, \dots)$ has a convergent subsequence with limit equal to some probability measure Q. By the Portmanteau Theorem,

$$Q\Big(\bigcup_{n=1}^{m} B(\psi_n, \delta)\Big) \leq 1 - \delta$$

for all m. Since the collection of balls $(B(\psi_n, \delta) \colon n = 1, 2, \dots)$ covers the space Ψ, it follows, by Continuity of Measure, that $Q(\Psi) \leq 1 - \delta$. Since Q is a probability measure, we have derived a contradiction, so the integer $p(\delta)$ must exist for all $\delta > 0$ as asserted.

For each $\delta > 0$, let

$$C_\delta = \bigcup_{n=1}^{p(\delta)} \overline{B(\psi_n, \delta)}\,.$$

Each set C_δ is the union of a finite number of closed sets, and hence is closed (but not necessarily compact!). Fix $\varepsilon > 0$ and let

$$K = \bigcap_{n=1}^{\infty} C_{\varepsilon/2^n}\,.$$

Since K is an intersection of closed sets, it is closed. By construction, K is totally bounded, so K is compact by Proposition 11. Since $Q(C_\delta) > 1 - \delta$ for all $\delta > 0$ and all $Q \in \mathcal{Q}$, an elementary calculation shows that $Q(K) > 1 - \varepsilon$ for all $Q \in \mathcal{Q}$. $\square$

Corollary 18. *Every probability measure on a Polish space is tight.*

* **Problem 15.** Prove that if Q is a probability measure on a Polish space Ψ, then for any open set $A \subseteq \Psi$,

$$Q(A) = \sup\{Q(K) \colon K \text{ is compact and } K \subseteq A\}\,.$$

18.5. Convergence in product spaces

We begin with some terminology. For $j = 1, 2, \dots$, let (Ψ_j, ρ_j) be Polish spaces, and let (Ψ, ρ) be the corresponding product Polish space (see Proposition 2). Let $A = \{j_1, j_2, \dots\}$ be any (finite or countably infinite) set of positive integers. For simplicity, assume that $j_1 < j_2 < \dots$. Let $h_A \colon \Psi \to \Psi_{j_1} \times \Psi_{j_2} \times \dots$ be the function

$$x = (x_1, x_2, \dots) \rightsquigarrow (x_{j_1}, x_{j_2}, \dots)\,,$$

known as the *projection of* Ψ *onto the coordinates indexed by* A. It is easily checked that any such projection is continuous.

If Q is a distribution on the Polish space Ψ, then the measure induced from Q by h_A is called the *marginal of* Q *corresponding to* A. If A is finite, then the measure induced from Q by h_A is known as a *finite-dimensional marginal.* If A has cardinality n, then the corresponding marginal is sometimes called an *n-dimensional marginal.* If $A = \{j\}$, then the corresponding 1-dimensional marginal is called the j^{th} *coordinate marginal.*

It is easy to use the Uniqueness Theorem to show that a probability measure on a countable product of Polish spaces is determined by its finite-dimensional marginals. The next theorem says that convergence in distribution on a countable product of Polish spaces is equivalent to convergence in distribution of each of the finite-dimensional marginal distributions.

Theorem 19. *Let* Q_n, $n = 1, 2, \ldots$, *and* Q *be distributions on the Polish space* $\Psi = \bigotimes_{j=1}^{\infty} \Psi_j$ *that was defined in* Proposition 2. *If* $Q_n \to Q$ *as* $n \to \infty$, *then for each set* $A \subseteq \{1, 2, \ldots\}$, $Q_n^A \to Q^A$ *as* $n \to \infty$, *where* Q_n^A *and* Q^A *are the measures induced from* Q_n *and* Q *by the projection* h_A. *On the other hand, if for all finite sets* A, *there exists a measure* $\tilde{Q}^A$ *such that* $Q_n^A \to \tilde{Q}^A$ *as* $n \to \infty$, *then there exists a measure* Q *such that* $Q_n \to Q$ *as* $n \to \infty$, *and* $Q^A = \tilde{Q}^A$.

PROOF. The first part of the theorem follows immediately from Proposition 10 and the fact that each of the functions h_A is continuous.

For the proof of the second part, assume that for each finite A, there exists a measure $\tilde{Q}^A$ such that $Q_n^A \to \tilde{Q}^A$ as $n \to \infty$. We use the convergence of the 1-dimensional marginals to prove that the sequence (Q_n) is uniformly tight. For each $j = 1, 2, \ldots$, let Q_n^j be the j^{th} coordinate marginal of Q_n. Since $Q_n^j \to \tilde{Q}^{\{j\}}$ as $n \to \infty$, the sequence (Q_n^j) is uniformly tight for each fixed j. Fix $\varepsilon > 0$, and for each j, choose a compact set $K_j \subseteq \Psi_j$ such that $Q_n^j(K_j) > 1 - \varepsilon/2^j$ for all n. Let $K = K_1 \times K_2 \times \ldots$. By the Tychonoff Theorem (see Appendix C), K is a compact subset of Ψ. For each n,

$$Q_n(K) = Q_n\Big(\bigcap_{j=1}^{\infty}\{(x_1, x_2, \ldots): x_j \in K_j\}\Big) \geq 1 - \sum_{j=1}^{\infty} Q_n^j(K_j^c) > 1 - \varepsilon\,.$$

Thus the sequence (Q_n) is uniformly tight.

Since (Q_n) is uniformly tight, there exists a measure R and a subsequence (Q_{n_k}) such that $Q_{n_k} \to R$ as $k \to \infty$. By the first part of the theorem, the finite-dimensional marginals of the terms in the sequence (Q_{n_k}) converge to the finite-dimensional marginals of R. Thus, the finite-dimensional marginals of R are the measures $\tilde{Q}^A$. Any other subsequential limit must have these same marginals, and thus be equal to R. An application of Proposition 15 completes the proof. $\square$

Problem 16. Let $\Psi = \Psi_1 \times \Psi_2 \times \dots$ be a countable product of Polish spaces, as in Theorem 19, and let (Q_n) be a sequence of probability measures on Ψ. Show that (Q_n) is uniformly tight if and only if (Q_n^j) is uniformly tight for each j, where Q_n^j is the j^{th} coordinate marginal of Q_n.

Problem 17. Let Ψ be as in the preceding problem, and for each n, let X_n be a Ψ_n-valued random variable. Show that as $n \to \infty$,

$$(X_1, X_2, \dots, X_n, X_n, X_n, \dots) \xrightarrow{\mathcal{D}} (X_1, X_2, X_3, \dots).$$

Problem 18. Describe how to use the previous result to quickly obtain Theorem 16 of Chapter 9 as a corollary of Theorem 7 of that same chapter in the Polish space setting.

In general, the hypothesis of the second part of Theorem 19 cannot be weakened in any significant way. For example, it would not be enough for all the n-dimensional marginals to converge for some fixed n. But there is one important special case in which the convergence of the 1-dimensional marginals is sufficient.

Theorem 20. *Let J be a countable set, and for $j \in J$, let Ψ_j be a Polish space. Set*

$$\Psi = \bigotimes_{j \in J} \Psi_j ,$$

viewed as a Polish space via Proposition 2. *For $n = 1, 2, \dots$, let*

$$Q_n = \bigotimes_{j \in J} Q_n^j$$

be a product measure on Ψ, where for each j and n, Q_n^j is a probability distribution on the Borel subsets of Ψ_j. Then the sequence (Q_n) converges to a distribution Q as $n \to \infty$ if and only if for each j, the sequence (Q_n^j) converges to a distribution Q^j on Ψ_j, in which case

$$Q = \bigotimes_{j \in J} Q^j .$$

Partial Proof. For the 'only if' aspect, note that for each j and n, the probability measure Q_n^j is the j^{th} coordinate marginal of Q_n. Thus, if $Q_n \to Q$ as $n \to \infty$, it follows from Theorem 19 that for each j, the sequence (Q_n^j) converges as $n \to \infty$ to the j^{th} coordinate marginal of Q, as desired.

For the proof of the 'if' portion we focus on the case $J = \{1, 2\}$ and leave the rest to the reader. Let g be an $\mathbb{R}$-valued bounded continuous function on $\Psi_1 \times \Psi_2$. By the Fubini Theorem, it is enough to show that

$$\lim_{n \to \infty} \int_{\Psi_2} h_n \, dQ_n^2 = \int_{\Psi_2} h \, dQ^2 , \tag{18.7}$$

where

$$h_n(y) = \int_{\Psi_1} g(x,y)\, Q_n^1(dx) \quad \text{and} \quad h(y) = \int_{\Psi_1} g(x,y)\, Q^1(dy)\,.$$

Let $\varepsilon > 0$ and note that

$$\begin{aligned} \Big| \int_{\Psi_2} h_n\, dQ_n^2 - \int_{\Psi_2} h\, dQ^2 \Big| \le \Big| \int_{\Psi_2} h\, dQ_n^2 - \int_{\Psi_2} h\, dQ^2 \Big| & \\ + \int_{\Psi_2 \setminus C} |h_n - h|\, dQ_n^2 & \\ + \int_C |h_n - h|\, dQ_n^2 & \end{aligned} \tag{18.8}$$

for any compact $C \subseteq \Psi_2$. There exists l such that the first term on the right is less than $\varepsilon/3$ for $n > l$, since h inherits continuity and boundedness from g and since $Q_n^2 \to Q^2$ as $n \to \infty$. The second term on the right is less than $\varepsilon/3$ for all n and some C, since all $|h_n|$ and $|h|$ inherit a common bound from g and since the sequence (Q_n^2) is uniformly tight.

For the last term on the right of (18.8), cover C by a finite number of sets $B_1, \dots, B_k$ having the property that $|g(x_1, u) - g(x_1, v)| < \varepsilon/9$ whenever u and v are in a common B_i. Fix $u_i \in B_i$. There exist l_i such that for $n > l_i$, the following calculation is valid for all $v_i \in B_i$:

$$|h_n(v_i) - h(v_i)| < |h_n(v_i) - h_n(u_i)| + |h_n(u_i) - h(u_i)| + |h(u_i) - h(v_i)| < 3(\tfrac{\varepsilon}{9}) = \tfrac{\varepsilon}{3}\,.$$

Therefore $|h_n(v) - h(v)| < \varepsilon/3$ for $v \in C$ and $n > \max\{l_i \colon 1 \le i \le k\}$, and thus the left side of (18.8) is less than ε for $n > l \vee \max\{l_i \colon 1 \le i \le k\}$. □

Problem 19. Complete the proof of the preceding theorem by first treating the case of $\#J < \infty$ by mathematical induction and then treating the case of J being countably infinite.

18.6. The Continuity Theorem for $\mathbb{R}^d$

The following result generalizes Theorem 15 of Chapter 14.

Theorem 21. [Continuity (for Characteristic Functions in $\mathbb{R}^d$)] *A sequence of probability distributions on* $\mathbb{R}$ *converges to a probability distribution* Q *if and only if the sequence of corresponding characteristic functions converges pointwise to a function* γ *which is continuous at* 0*, in which case the convergence to* γ *is uniform on each compact subset of* $\mathbb{R}^d$*, and* γ *is the characteristic function of* Q.

Partial Proof. We leave it as a problem to prove that if $(Q_n \colon n = 1, 2, \dots)$ is a sequence of probability measures on $\mathbb{R}^d$ that converges to a probability measure Q, then $(\beta_n \colon n = 1, 2, \dots)$ converges to β uniformly on each compact set, where β_n and β are the characteristic functions of Q_n and Q, respectively.

For the converse, assume that the sequence $(\beta_n : n = 1, 2, \dots)$ converges pointwise to a function γ that is continuous at $0 \in \mathbb{R}^d$. It follows from Problem 58 of Chapter 13 that, for each j, the characteristic functions of the j^{th} coordinate marginals converge to a function that is continuous at $0 \in \mathbb{R}$. By the Continuity Theorem for $\mathbb{R}$, the 1-dimensional marginals converge. By Problem 16, the sequence (Q_n) is uniformly tight. By the first part of this theorem, every convergent subsequence of (Q_n) has a limit with characteristic function γ. By Theorem 16 of Chapter 13, these convergent subsequences all have the same limit. An application of Proposition 15 completes the proof. $\square$

Problem 20. Complete the proof of the preceding theorem by doing the first portion of the proof.

Problem 21. Find a sequence $((X_n, Y_n) : n = 1, 2, \dots)$ which does not converge in distribution (as a sequence of $\mathbb{R}^2$-valued random variables), but for which both of the sequences $(X_n : n = 1, 2, \dots)$ and $(Y_n : n = 1, 2, \dots)$ converge in distribution. For your example, find a $w \in \mathbb{R}^2$ for which $E(\exp(i\langle w, (X_n, Y_n)\rangle))$ does not converge.

Problem 22. Let $(X_n : n = 1, 2, \dots)$ be an iid sequence of $\mathbb{R}$-valued random variables. Show that as $n \to \infty$,

$$(X_n\,,\; X_{n+1} + \tfrac{1}{n} X_n) \xrightarrow{\mathcal{D}} (X_1, X_2)\,.$$

The preceding problem shows that the lack of independence is not necessarily preserved when passing to a limit. Theorem 20 says that independence is preserved.

In the proof of Theorem 21, we found it useful to analyze a sequence of distributions on $\mathbb{R}^d$ in terms of related distributions on $\mathbb{R}$. The following theorem extends this idea.

Theorem 22. [Cramér-Wold Device] *Let $d < \infty$. A sequence $(X_n : n = 1, 2, \dots)$ of $\mathbb{R}^d$-valued random variables converges in distribution to a random variable X if and only if each sequence $(\langle w, X_n\rangle : n = 1, 2, \dots)$, $w \in \mathbb{R}^d$, converges in distribution in which case*

$$\langle w, X_n\rangle \xrightarrow{\mathcal{D}} \langle w, X\rangle \tag{18.9}$$

for each $w \in \mathbb{R}^d$.

PROOF. The convergence in (18.9) follows immediately from Proposition 10.

For the converse, suppose that $(\langle w, X_n\rangle)$ converges in distribution for each w. It is clear that if (X_n) converges in distribution or has a convergent subsequence, the limit must have that unique distribution whose characteristic function is

$$w \rightsquigarrow \lim_{n\to\infty} E(e^{i\langle w, X_n\rangle})\,.$$

In view of the Prohorov Theorem, we only need show uniform tightness of $\{Q_n : n = 1, 2, \dots\}$, where Q_n is the distribution of X_n, because then an appeal to Proposition 15 completes the proof. By setting $w = (0, \dots, 0, 1, 0, \dots, 0)$, with 1 in the j^{th} position, we see that the sequence $(X_{j,n} : n = 1, 2, \dots)$ of j^{th} coordinates converges in distribution. Thus, for each j, the sequence of j^{th} coordinate marginals of (Q_n) is uniformly tight, so (Q_n) is uniformly tight by Problem 16. □

Problem 23. Show that the hypothesis of convergence of $(\langle w, X_n \rangle)$ for all w in the preceding theorem cannot be replaced by the hypothesis of convergence for all w in some basis of $\mathbb{R}^d$.

Theorem 23. [Multi-dimensional Central Limit] *Let d be a positive integer. Let $(X_1, X_2, \dots)$ be an iid sequence of $\mathbb{R}^d$-valued random variables having finite mean vector $\mu = E(X_1)$ and finite covariance matrix*

$$\Sigma = E\big([X_1 - \mu]^T [X_1 - \mu]\big)\,,$$

where $[X_1 - \mu]$ denotes the row matrix corresponding to the vector $X_1 - \mu$ and $[\cdot]^T$ denotes transpose. Then

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}} \xrightarrow{\mathcal{D}} Z\,,$$

where Z is a normally distributed $\mathbb{R}^d$-valued random variable with mean vector 0 and covariance matrix Σ.

* **Problem 24.** Use the Classical Central Limit Theorem and the Cramér-Wold Device to prove the preceding theorem.

Problem 25. Apply the Multi-dimensional Central Limit Theorem to the iid sequence $(Z_1, Z_2, \dots)$ where each Z_j has the same distribution as the random vector Z of Problem 60 of Chapter 13.

18.7. † The Prohorov metric

Let $\mathcal{Q}(\Psi)$ denote the family of all probability distributions on a Polish space (Ψ, ρ). Our final goal of this chapter is to turn $\mathcal{Q}(\Psi)$ into a Polish space with a metric $\hat{\rho}$ so that convergence of sequences in the Polish space $(\mathcal{Q}(\Psi), \hat{\rho})$ is equivalent to convergence of sequences of distributions on Ψ. The metric $\hat{\rho}$ we will use is called the *Prohorov metric* and is defined by

$$\hat{\rho}(Q, R) = \inf\{\varepsilon : R(A) \le Q(A_\varepsilon) + \varepsilon \text{ for every Borel } A \subseteq \Psi\}\,, \tag{18.10}$$

where

$$A_\varepsilon = \{x : \rho(x, y) < \varepsilon \text{ for some } y \in A\}\,.$$

(Note: It is important that A_ε be Borel; indeed, it is open, since it is the union of open balls.)

* **Problem 26.** Choose $\delta, \varepsilon > 0$. Prove that if $R(A) \le Q(A_\varepsilon) + \delta$ for all Borel sets $A \subseteq \Psi$, then $Q(A) \le R(A_\varepsilon) + \delta$ for all Borel sets $A \subseteq \Psi$.

Problem 27. Prove that $\hat{\rho}$ defined by (18.10) is a metric.

From the preceding problem we see that $(\mathcal{Q}(\Psi), \hat{\rho})$ is a metric space. Let C be a countable dense subset of Ψ. Then it is easy to show that the set $\mathcal{C}$ of probability distributions whose values are rational and whose supports are finite subsets of C is a countable dense subset of $\mathcal{Q}(\Psi)$. We have almost proved the following theorem.

Theorem 24. *Let (Ψ, ρ) be a Polish space and let $\mathcal{Q}(\Psi)$ denote the family of all probability measures on (Ψ, ρ). Define $\hat{\rho}$ by (18.10). Then $(\mathcal{Q}(\Psi), \hat{\rho})$ is a Polish space.*

PARTIAL PROOF. In view of the discussion preceding the theorem, we need only show that $(\mathcal{Q}(\Psi), \hat{\rho})$ is complete. Let $(Q_n : n = 1, 2, \dots)$ be an arbitrary Cauchy sequence. To prove the convergence of a Cauchy sequence, it is enough to find a subsequence that converges. By the Prohorov Theorem, it is enough to find a subsequence that is uniformly tight. Using the definition of Cauchy sequence, a routine argument shows that there exists a subsequence $(R_n : n = 1, 2, \dots)$ with the property that the $\hat{\rho}(R_n, R_{n+1}) < 1/2^{n+1}$ for $n = 1, 2, \dots$. We will prove that this subsequence is uniformly tight.

Fix $\varepsilon > 0$. We will define a sequence of compact subsets of Ψ. Choose a positive integer l such that $1/2^l < \varepsilon/2$. By Corollary 18 we can find a compact set K such that $R_n(K) > 1 - \varepsilon/2$ for $n = 1, \dots, l$. Let $K_j = K$ for $j = 1, \dots, l$. We now proceed recursively to define compact sets $K_{l+1}, K_{l+2}, \dots$, such that for all $n > l$,

$$R_n(K_n) > 1 - \Big(1 - \frac{1}{2^{n+1-l}}\Big)\varepsilon$$

and

$$K_{n+1} \subseteq (K_n)_{2^{-(n+1)}}\,, \tag{18.11}$$

where

$$(K_n)_{1/2^{n+1}} = \Big\{x \in \Psi : \rho(x, y) < \frac{1}{2^{n+1}} \text{ for some } y \in K_n\Big\}\,.$$

We have already defined K_l with the desired properties. Assume that sets $K_l, \dots, K_n$ have been defined with the desired properties for some $n \ge l$. By the definition of $\hat{\rho}$ and our choice of the integer l,

$$R_{n+1}((K_n)_{1/2^{n+1}}) > R_n(K_n) - \frac{1}{2^{n+1}} > 1 - \Big(1 - \frac{1}{2^{n+2-l}}\Big)\varepsilon\,.$$

It follows from Problem 15 that there exists a compact set K_{n+1} satisfying (18.11) such that $R(K_{n+1}) > 1 - (1 - 2^{-(n+2-l)})\varepsilon$. The recursive construction of the sequence $(K_n : n = 1, 2, \dots)$ of compact sets is complete.

Let

$$K = \text{the closure of } \bigcup_{n=1}^{\infty} K_n .$$

Clearly $R_n(K) > 1 - \varepsilon$ for all n. The proof of uniform tightness is completed by showing that K is compact. By Proposition 11, it is enough to prove that K is totally bounded. This final step is left as a exercise. □

Problem 28. Prove that the set K defined in the preceding proof is totally bounded. *Hint:* Show that for each $\varepsilon > 0$, a sufficiently large value of n can be found so that any covering of $K_1 \cup \dots \cup K_n$ by ε-balls can be 'inflated' to a covering of K by doubling the diameters of each of the balls.

* **Problem 29.** For the metric space of distributions on $\mathbb{R}$, calculate the distance between the uniform distribution on $[0, 1]$ and the uniform distribution on $[a, a+1]$. Also calculate the distance between the uniform distributions on $[0, 8]$ and $[0, 9]$.

Problem 30. For the metric space of distributions on $\mathbb{R}^2$, calculate the distance between the delta distribution at $(0, 0)$ and the uniform distribution on the square region with vertices at $(\pm a, \pm a)$.

Problem 31. For the metric space of distributions on $\mathbb{R}^\infty$, calculate the distance between the delta distribution at $(0, 0, 0, \dots)$ and the distribution Q_d induced by the uniform distribution on $[-1, 1]$ in $\mathbb{R}$ and the function $\varphi_d : \mathbb{R} \to \mathbb{R}^\infty$ defined by

$$\varphi_d(x) = (0, \dots, 0, x, 0, 0, \dots),$$

where x on the right side is the d^{th} term.

Theorem 25. *Let $\mathcal{Q}(\Psi)$ be the Polish space described in* Theorem 24. *Then $\hat{\rho}(Q, Q_n) \to 0$ as $n \to \infty$ if and only if $Q_n \to Q$ as $n \to \infty$.*

Proof. Suppose that $\hat{\rho}(Q, Q_n) \to 0$ and let C be a closed set. For each $\varepsilon > 0$ there exists an integer l such that

$$Q_n(C) \le Q(C_\varepsilon) + \varepsilon$$

for $n \ge l$. Hence,

$$\limsup_{n\to\infty} Q_n(C) \le Q(C_\varepsilon) + \varepsilon .$$

Now let $\varepsilon \searrow 0$ through a sequence and use the equation

$$C = \bigcap_{\varepsilon > 0} C_\varepsilon ,$$

which is a consequence of the fact that C is closed, to conclude

$$\limsup_{n\to\infty} Q_n(C) \leq Q(C)\,.$$

An appeal to the Portmanteau Theorem completes this half of the proof.

For the converse assume that $\int g\,dQ_n \to \int g\,dQ$ for all continuous bounded functions g. By the Portmanteau Theorem,

$$\lim_{n\to\infty} Q_n(A) = Q(A) \tag{18.12}$$

for every Borel set A for which $Q(\partial A) = 0$.

Let $\varepsilon > 0$ and let $(x_j : j = 1, 2, \dots)$ be a dense sequence in Ψ. For each j, let B_j be a ball centered at x_j with radius strictly between $\frac{\varepsilon}{4}$ and $\frac{\varepsilon}{2}$ and having the additional property that $Q(\partial B_n) = 0$. For each j, set

$$C_j = B_j \setminus \bigcup_{i=1}^{j-1} B_i\,.$$

It is clear that:

- $C_j \cap C_i = \emptyset$ if $i \neq j$;
- $\Psi = \bigcup_{j=1}^{\infty} C_j$;
- $1 = \sum_{j=1}^{\infty} Q(C_j)$;
- $Q(\partial C_j) = 0$ for each j;
- the distance between any two members of any one C_j is less than ε.

Choose k so that $\sum_{j=1}^{k} Q(C_j) > 1 - \frac{\varepsilon}{2}$ and then use (18.12) to deduce the existence of l such that

$$Q(C_j) < Q_n(C_j) + \tfrac{\varepsilon}{2k}\,, \quad \text{for } j \leq k\,,\ n \geq l\,.$$

Let A be any Borel set and denote by A_ε the set of points that each lie no more than distance ε from some point in A. Let $(C_{j_1}, C_{j_2}, \dots, C_{j_r})$ be the subsequence of (C_j) consisting of those C_j for which $j \leq k$ and $C_j \cap A \neq \emptyset$. Then, for $n \geq l$,

$$Q(A) < \frac{\varepsilon}{2} + \sum_{i=1}^{r} Q(C_{j_i}) < \varepsilon + \sum_{i=1}^{r} Q_n(C_{j_i}) \leq \varepsilon + Q_n(A_\varepsilon)\,.$$

Therefore the distance between Q_n and Q is less than ε for $n \geq l$. □

Th
a

In this chapter, we bring
construct one of the mos
'Brownian motion'. In or
first take the point of vie
values in the Polish spac
We will find that when
to a C[0, 1]-valued rand
and scaled so that its dis
consequence, much can
walks by looking at the
on the material concerni
Limit Theorem (Chapte
Appendix B) also play in

We include only a sma
motion on $[0, \infty)$ that ha
include indicate the vari
involve stopping times,
to extend ideas from Cl
continuous time setting.

Random variables X t
Thus, $X(\omega)$ is a continu
be denoted by $X_t(\omega)$. He
notation for X is

When speaking of events
or an $\mathbb{R}$-valued random
tion $(X^{(n)} : n = 1, 2, \dots)$
random variables.