8

Weak Convergence of Measures

8.1 Definition of Weak Convergence

In this chapter we consider the fundamental concept of weak convergence of probability measures. This will lay the groundwork for the precise formulation of the Central Limit Theorem and other Limit Theorems of probability theory (see Chap. 10).

Let (X, d) be a metric space, $\mathcal{B}(X)$ the σ-algebra of its Borel sets and P_n a sequence of probability measures on $(X, \mathcal{B}(X))$. Recall that $C_b(X)$ denotes the space of bounded continuous functions on X.

Definition 8.1. *The sequence* P_n *converges weakly to the probability measure* P *if, for each* $f \in C_b(X)$,

$$\lim_{n\to\infty} \int_X f(x) d\mathrm{P}_n(x) = \int_X f(x) d\mathrm{P}(x).$$

The weak convergence is sometimes denoted as $\mathrm{P}_n \Rightarrow \mathrm{P}$.

Definition 8.2. *A sequence of real-valued random variables* ξ_n *defined on probability spaces* $(\Omega_n, \mathcal{F}_n, \overline{\mathrm{P}}_n)$ *is said to converge in distribution if the induced measures* P_n, $\mathrm{P}_n(A) = \overline{\mathrm{P}}_n(\xi_n \in A)$, *converge weakly to a probability measure* P.

In Definition 8.1 we could omit the requirement that P_n and P are probability measures. We then obtain the definition of the weak convergence for arbitrary finite measures on $\mathcal{B}(X)$. The following lemma provides a useful criterion for the weak convergence of measures.

Lemma 8.3. *If a sequence of measures* P_n *converges weakly to a measure* P, *then*

$$\limsup_{n\to\infty} \mathrm{P}_n(K) \le \mathrm{P}(K) \tag{8.1}$$

for any closed set K. *Conversely, if (8.1) holds for any closed set* K, *and* $\mathrm{P}_n(X) = \mathrm{P}(X)$ *for all* n, *then* P_n *converge weakly to* P.

L. Koralov and Y.G. Sinai, *Theory of Probability and Random Processes*, Universitext, DOI 10.1007/978-3-540-68829-7_8,

Proof. First assume that P_n converges to P weakly. Let $\varepsilon > 0$ and select $\delta > 0$ such that $\mathrm{P}(K_\delta) < \mathrm{P}(K) + \varepsilon$, where K_δ is the δ-neighborhood of the set K. Consider a continuous function f_δ such that $0 \le f_\delta(x) \le 1$ for $x \in X$, $f_\delta(x) = 1$ for $x \in K$, and $f_\delta(x) = 0$ for $x \in X \backslash K_\delta$. For example, one can take $f_\delta(x) = \max(1 - \mathrm{dist}(x, K)/\delta, 0)$.

Note that $\mathrm{P}_n(K) = \int_K f_\delta d\mathrm{P}_n \le \int_X f_\delta d\mathrm{P}_n$ and $\int_X f_\delta d\mathrm{P} = \int_{K_\delta} f_\delta d\mathrm{P} \le \mathrm{P}(K_\delta) < \mathrm{P}(K) + \varepsilon$. Therefore,

$$\lim_{n\to\infty} \sup \mathrm{P}_n(K) \le \lim_{n\to\infty} \int_X f_\delta d\mathrm{P}_n = \int_X f_\delta d\mathrm{P} < \mathrm{P}(K) + \varepsilon,$$

which implies the result since ε was arbitrary.

Let us now assume that $\mathrm{P}_n(X) = P(X)$ for all n and $\limsup_{n\to\infty} \mathrm{P}_n(K) \le \mathrm{P}(K)$ for any closed set K. Let $f \in C_b(X)$. We can find $a > 0$ and b such that $0 < af + b < 1$. Since $\mathrm{P}_n(X) = P(X)$ for all n, if the relation

$$\lim_{n\to\infty} \int_X g(x) d\mathrm{P}_n(x) = \int_X g(x) d\mathrm{P}(x)$$

is valid for $g = af + b$, then it is also valid for f instead of g. Therefore, without loss of generality, we can assume that $0 < f(x) < 1$ for all x. Define the closed sets $K_i = \{x : f(x) \ge i/k\}$, where $0 \le i \le k$. Then

$$\frac{1}{k}\sum_{i=1}^{k} \mathrm{P}_n(K_i) \le \int_X f d\mathrm{P}_n \le \frac{\mathrm{P}_n(X)}{k} + \frac{1}{k}\sum_{i=1}^{k} \mathrm{P}_n(K_i),$$

$$\frac{1}{k}\sum_{i=1}^{k} \mathrm{P}(K_i) \le \int_X f d\mathrm{P} \le \frac{\mathrm{P}(X)}{k} + \frac{1}{k}\sum_{i=1}^{k} \mathrm{P}(K_i).$$

Since $\limsup_{n\to\infty} \mathrm{P}_n(K_i) \le \mathrm{P}(K)$ for each i, and $P_n(X) = P(X)$, we obtain

$$\limsup_{n\to\infty} \int_X f d\mathrm{P}_n \le \frac{\mathrm{P}(X)}{k} + \int_X f d\mathrm{P}.$$

Taking the limit as $k \to \infty$, we obtain

$$\limsup_{n\to\infty} \int_X f d\mathrm{P}_n \le \int_X f d\mathrm{P}.$$

By considering the function $-f$ instead of f we can obtain

$$\liminf_{n\to\infty} \int_X f d\mathrm{P}_n \ge \int_X f d\mathrm{P}.$$

This proves the weak convergence of measures. □

The following lemma will prove useful when proving the Prokhorov Theorem below.

Lemma 8.4. *Let X be a metric space and $\mathcal{B}(X)$ the σ-algebra of its Borel sets. Any finite measure P on $(X, \mathcal{B}(X))$ is regular, that is for any $A \in \mathcal{B}(X)$ and any $\varepsilon > 0$ there are an open set U and a closed set K such that $K \subseteq A \subseteq U$ and $\mathrm{P}(U) - \mathrm{P}(K) < \varepsilon$.*

Proof. If A is a closed set, we can take $K = A$ and consider a sequence of open sets $U_n = \{x : \mathrm{dist}(x, A) < 1/n\}$. Since $\bigcap_n U_n = A$, there is a sufficiently large n such that $\mathrm{P}(U_n) - \mathrm{P}(A) < \varepsilon$. This shows that the statement is true for all closed sets.

Let $\mathcal{K}$ be the collection of sets A such that for any ε there exist K and U with the desired properties. Note that the collection of all closed sets is a π-system. Clearly, $A \in \mathcal{K}$ implies that $X \backslash A \in \mathcal{K}$. Therefore, due to Lemma 4.13, it remains to prove that if $A_1, A_2, \ldots \in \mathcal{K}$ and $A_i \bigcap A_j = \emptyset$ for $i \neq j$, then $A = \bigcup_n A_n \in \mathcal{K}$.

Let $\varepsilon > 0$. Find n_0 such that $\mathrm{P}(\bigcup_{n=n_0}^{\infty} A_n) < \varepsilon/2$. Find open sets U_n and closed sets K_n such that $K_n \subseteq A_n \subseteq U_n$ and $\mathrm{P}(U_n) - \mathrm{P}(K_n) < \varepsilon/2^{n+1}$ for each n. Then $U = \bigcup_n U_n$ and $K = \bigcup_{n=1}^{n_0} K_n$ have the desired properties, that is $K \subseteq A \subseteq U$ and $\mathrm{P}(U) - \mathrm{P}(K) < \varepsilon$. □

8.2 Weak Convergence and Distribution Functions

Recall the one-to-one correspondence between the probability measures on $\mathbb{R}$ and the distribution functions. Let F_n and F be the distribution functions corresponding to the measures P_n and P respectively. Note that x is a continuity point of F if and only if $\mathrm{P}(x) = 0$. We now express the condition of weak convergence in terms of the distribution functions.

Theorem 8.5. *The sequence of probability measures P_n converges weakly to the probability measure P if and only if $\lim_{n\to\infty} F_n(x) = F(x)$ for every continuity point x of the function F.*

Proof. Let $\mathrm{P}_n \Rightarrow \mathrm{P}$ and let x be a continuity point of F. We consider the functions f, f_δ^+ and f_δ^-, which are defined as follows:

$$f(y) = \begin{cases} 1, & y \leq x, \\ 0, & y > x, \end{cases}$$

$$f_\delta^+(y) = \begin{cases} 1, & y \leq x, \\ 1 - (y - x)/\delta, & x < y \leq x + \delta, \\ 0, & y > x + \delta, \end{cases}$$

$$f_\delta^-(y) = \begin{cases} 1, & y \le x - \delta, \\ 1 - (y - x + \delta)/\delta, & x - \delta < y \le x, \\ 0, & y > x. \end{cases}$$

The functions f_δ^+ and f_δ^- are continuous and $f_\delta^- \le f \le f_\delta^+$. Using the fact that x is a continuity point of F we have, for any $\varepsilon > 0$ and $n \ge n_0(\varepsilon)$,

$$F_n(x) = \int_{\mathbb{R}} f(y)dF_n(y) \le \int_{\mathbb{R}} f_\delta^+(y)dF_n(y)$$

$$\le \int_{\mathbb{R}} f_\delta^+(y)dF(y) + \frac{\varepsilon}{2} \le F(x+\delta) + \frac{\varepsilon}{2} \le F(x) + \varepsilon,$$

if δ is such that $|F(x \pm \delta) - F(x)| \le \frac{\varepsilon}{2}$. On the other hand, for such n we also have

$$F_n(x) = \int_{\mathbb{R}} f(y)dF_n(y) \ge \int_{\mathbb{R}} f_\delta^-(y)dF_n(y)$$

$$\ge \int_{\mathbb{R}} f_\delta^-(y)dF(y) - \frac{\varepsilon}{2} \ge F(x-\delta) - \frac{\varepsilon}{2} \ge F(x) - \varepsilon.$$

In other words, $|F_n(x) - F(x)| \le \varepsilon$ for all sufficiently large n.

Now we prove the converse. Let $F_n(x) \to F(x)$ at every continuity point of F. Let f be a bounded continuous function. Let ε be an arbitrary positive constant. We need to prove that

$$|\int_{\mathbb{R}} f(x)dF_n(x) - \int_{\mathbb{R}} f(x)dF(x)| \le \varepsilon \tag{8.2}$$

for sufficiently large n.

Let $M = \sup|f(x)|$. Since the function F is non-decreasing, it has at most a countable number of points of discontinuity. Select two points of continuity A and B for which $F(A) \le \frac{\varepsilon}{10M}$ and $F(B) \ge 1 - \frac{\varepsilon}{10M}$. Therefore $F_n(A) \le \frac{\varepsilon}{5M}$ and $F_n(B) \ge 1 - \frac{\varepsilon}{5M}$ for all sufficiently large n.

Since f is continuous, it is uniformly continuous on $[A,B]$. Therefore we can partition the half-open interval $(A,B]$ into finitely many half-open subintervals $I_1 = (x_0, x_1], I_2 = (x_1, x_2], \ldots, I_n = (x_{n-1}, x_n]$ such that $|f(y) - f(x_i)| \le \frac{\varepsilon}{10}$ for $y \in I_i$. Moreover, the endpoints x_i can be selected to be continuity points of $F(x)$. Let us define a new function f_ε on $(A,B]$ which is equal to $f(x_i)$ on each of the intervals I_i.

In order to prove (8.2), we write

$$|\int_{\mathbb{R}} f(x)dF_n(x) - \int_{\mathbb{R}} f(x)dF(x)|$$

$$\le \int_{(-\infty,A]} |f(x)|dF_n(x) + \int_{(-\infty,A]} |f(x)|dF(x)$$

$$+ \int_{(B,\infty)} |f(x)|dF_n(x) + \int_{(B,\infty)} |f(x)|dF(x)$$

$$+|\int_{(A,B]} f(x)dF_n(x) - \int_{(A,B]} f(x)dF(x)|.$$

The first term on the right-hand side is estimated from above for large enough n as follows:

$$\int_{(-\infty,A]} |f(x)|dF_n(x) \le MF_n(A) \le \frac{\varepsilon}{5}$$

Similarly, the second, third and fourth terms are estimated from above by $\frac{\varepsilon}{10}$, $\frac{\varepsilon}{5}$ and $\frac{\varepsilon}{10}$ respectively.

Since $|f_\varepsilon - f| \le \frac{\varepsilon}{10}$ on $(A, B]$, the last term can be estimated as follows:

$$|\int_{(A,B]} f(x)dF_n(x) - \int_{(A,B]} f(x)dF(x)|$$

$$\le |\int_{(A,B]} f_\varepsilon(x)dF_n(x) - \int_{(A,B]} f_\varepsilon(x)dF(x)| + \frac{\varepsilon}{5}.$$

Note that

$$\lim_{n\to\infty} |\int_{I_i} f_\varepsilon(x)dF_n(x) - \int_{I_i} f_\varepsilon(x)dF(x)|$$

$$= \lim_{n\to\infty} (|f(x_i)||F_n(x_i) - F_n(x_{i-1}) - F(x_i) + F(x_{i-1})|) = 0,$$

since $F_n(x) \to F(x)$ at the endpoints of the interval I_i. Therefore,

$$\lim_{n\to\infty} |\int_{(A,B]} f_\varepsilon(x)dF_n(x) - \int_{(A,B]} f_\varepsilon(x)dF(x)| = 0,$$

and thus

$$|\int_{(A,B]} f_\varepsilon(x)dF_n(x) - \int_{(A,B]} f_\varepsilon(x)dF(x)| \le \frac{\varepsilon}{5}$$

for large enough n. □

8.3 Weak Compactness, Tightness, and the Prokhorov Theorem

Let X be a metric space and P_α a family of probability measures on the Borel σ-algebra $\mathcal{B}(X)$. The following two concepts, weak compactness (sometimes also referred to as relative compactness) and tightness, are fundamental in probability theory.

Definition 8.6. *A family of probability measures* $\{\mathrm{P}_\alpha\}$ *on* $(X, \mathcal{B}(X))$ *is said to be weakly compact if from any sequence* P_n, $n = 1, 2, \ldots$, *of measures from the family, one can extract a weakly convergent subsequence* P_{n_k}, $k = 1, 2, \ldots$, *that is* $\mathrm{P}_{n_k} \Rightarrow \mathrm{P}$ *for some probability measure* P.

Remark 8.7. Note that it is not required that $\mathrm{P} \in \{\mathrm{P}_\alpha\}$.

Definition 8.8. *A family of probability measures* $\{\mathrm{P}_\alpha\}$ *on* $(X, \mathcal{B}(X))$ *is said to be tight if for any* $\varepsilon > 0$ *one can find a compact set* $K_\varepsilon \subseteq X$ *such that* $\mathrm{P}(K_\varepsilon) \geq 1 - \varepsilon$ *for each* $\mathrm{P} \in \{\mathrm{P}_\alpha\}$.

In the case when $(X, \mathcal{B}(X)) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we have the following theorem.

Theorem 8.9 (Helly Theorem). *A family of probability measures* $\{\mathrm{P}_\alpha\}$ *on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *is tight if and only if it is weakly compact.*

The Helly Theorem is a particular case of the following theorem, due to Prokhorov.

Theorem 8.10 (Prokhorov Theorem). *If a family of probability measures* $\{\mathrm{P}_\alpha\}$ *on a metric space* X *is tight, then it is weakly compact. On a separable complete metric space the two notions are equivalent.*

The proof of the Prokhorov Theorem will be preceded by two lemmas. The first lemma is a general fact from functional analysis, which is a consequence of the Alaoglu Theorem and will not be proved here.

Lemma 8.11. *Let* X *be a compact metric space. Then from any sequence of measures* μ_n *on* $(X, \mathcal{B}(X))$, *such that* $\mu_n(X) \leq C < \infty$ *for all* n, *one can extract a weakly convergent subsequence.*

We shall denote an open ball of radius r centered at a point $a \in X$ by $B(a, r)$. The next lemma provides a criterion of tightness for families of probability measures.

Lemma 8.12. *A family* $\{\mathrm{P}_\alpha\}$ *of probability measures on a separable complete metric space* X *is tight if and only if for any* $\varepsilon > 0$ *and* $r > 0$ *there is a finite family of balls* $B(a_i, r)$, $i = 1, \ldots, n$, *such that*

$$\mathrm{P}_\alpha\Big(\bigcup_{i=1}^{n} B(a_i, r)\Big) \geq 1 - \varepsilon$$

for all α.

Proof. Let $\{\mathrm{P}_\alpha\}$ be tight, $\varepsilon > 0$, and $r > 0$. Select a compact set K such that $\mathrm{P}(K) \geq 1-\varepsilon$ for all $\mathrm{P} \in \{\mathrm{P}_\alpha\}$. Since any compact set is totally bounded, there is a finite family of balls $B(a_i, r)$, $i = 1, \ldots, n$, which cover K. Consequently, $\mathrm{P}(\bigcup_{i=1}^n B(a_i, r)) \geq 1 - \varepsilon$ for all $\mathrm{P} \in \{\mathrm{P}_\alpha\}$.

Let us prove the converse statement. Fix $\varepsilon > 0$. Then for any integer $k > 0$ there is a family of balls $B^{(k)}(a_i, \frac{1}{k})$, $i = 1, \ldots, n_k$, such that $\mathrm{P}(A_k) \geq 1 - 2^{-k}\varepsilon$ for all $\mathrm{P} \in \{\mathrm{P}_\alpha\}$, where $A_k = \bigcup_{i=1}^{n_k} B^{(k)}(a_i, \frac{1}{k})$. The set $A = \bigcap_{k=1}^\infty A_k$ satisfies $\mathrm{P}(A) \geq 1 - \varepsilon$ for all $\mathrm{P} \in \{\mathrm{P}_\alpha\}$ and is totally bounded. Therefore, its closure is compact since X is a complete metric space. □

Proof of the Prokhorov Theorem. Assume that a family $\{\mathrm{P}_\alpha\}$ is weakly compact but not tight. By Lemma 8.12, there exist $\varepsilon > 0$ and $r > 0$ such that for any family $B_1, \ldots, B_n$ of balls of radius r, we have $\mathrm{P}(\bigcup_{1 \le i \le n} B_i) \le 1 - \varepsilon$ for some $\mathrm{P} \in \{\mathrm{P}_\alpha\}$. Since X is separable, it can be represented as a countable union of balls of radius r, that is $X = \bigcup_{i=1}^\infty B_i$. Let $A_n = \bigcup_{1 \le i \le n} B_i$. Then we can select $\mathrm{P}_n \in \{\mathrm{P}_\alpha\}$ such that $\mathrm{P}_n(A_n) \le 1 - \varepsilon$. Assume that a subsequence P_{n_k} converges to a limit P. Since A_m is open, $\mathrm{P}(A_m) \le \liminf_{k\to\infty} \mathrm{P}_{n_k}(A_m)$ for every fixed m due to Lemma 8.3. Since $A_m \subseteq A_{n_k}$ for large k, we have $\mathrm{P}(A_m) \le \liminf_{k\to\infty} \mathrm{P}_{n_k}(A_{n_k}) \le 1 - \varepsilon$, which contradicts $\bigcup_{m=1}^\infty A_m = X$. Thus, weak compactness implies tightness.

Now assume that $\{\mathrm{P}_\alpha\}$ is tight. Consider a sequence of compact sets K_m such that

$$\mathrm{P}(K_m) \ge 1 - \frac{1}{m} \ \text{ for all } \ \mathrm{P} \in \{\mathrm{P}_\alpha\}, \quad m = 1, 2, \ldots$$

Consider a sequence of measures $\mathrm{P}_n \in \{\mathrm{P}_\alpha\}$. By Lemma 8.11, using the diagonalization procedure, we can construct a subsequence P_{n_k} such that, for each m, the restrictions of P_{n_k} to $\widetilde{K}_m = \bigcup_{i=1}^m K_i$ converge weakly to a measure μ_m. Note that $\mu_m(\widetilde{K}_m) \ge 1 - \frac{1}{m}$ since $\mathrm{P}_{n_k}(\widetilde{K}_m) \ge 1 - \frac{1}{m}$ for all k.

Let us show that for any Borel set A the sequence $\mu_m(A \bigcap \widetilde{K}_m)$ is non-decreasing. Thus, we need to show that $\mu_{m_1}(A \bigcap \widetilde{K}_{m_1}) \le \mu_{m_2}(A \bigcap \widetilde{K}_{m_2})$ if $m_1 < m_2$. By considering $A \bigcap \widetilde{K}_{m_1}$ instead of A we can assume that $A \subseteq \widetilde{K}_{m_1}$. Fix an arbitrary $\varepsilon > 0$. Due to the regularity of the measures μ_{m_1} and μ_{m_2} (see Lemma 8.4), there exist sets $\overline{U}^i, \overline{K}^i \subseteq \widetilde{K}_{m_i}$, $i = 1, 2$, such that $\overline{U}^i$ ($\overline{K}^i$) are open (closed) in the topology of $\widetilde{K}_{m_i}$, $\overline{K}^i \subseteq A \subseteq \overline{U}^i$ and

$$\mu_{m_i}(\overline{U}^i) - \varepsilon < \mu_{m_i}(A) < \mu_{m_i}(\overline{K}^i) + \varepsilon, \quad i = 1, 2.$$

Note that $\overline{U}^1 = \overline{\overline{U}} \cap \widetilde{K}_{m_1}$ for some set $\overline{\overline{U}}$ that is open in the topology of $\widetilde{K}_{m_2}$. Let $U = \overline{\overline{U}} \cap \overline{U}^2$ and $K = \overline{K}^1 \cup \overline{K}^2$. Thus $U \subseteq \widetilde{K}_{m_2}$ is open in the topology of $\widetilde{K}_{m_2}$, $K \subseteq \widetilde{K}_{m_1}$ is closed in the topology of $\widetilde{K}_{m_1}$, $K \subseteq A \subseteq U$ and

$$\mu_{m_1}(U \bigcap \widetilde{K}_{m_1}) - \varepsilon < \mu_{m_1}(A) < \mu_{m_1}(K) + \varepsilon, \tag{8.3}$$

$$\mu_{m_2}(U) - \varepsilon < \mu_{m_2}(A) < \mu_{m_2}(K) + \varepsilon. \tag{8.4}$$

Let f be a continuous function on $\widetilde{K}_{m_2}$ such that $0 \le f \le 1$, $f(x) = 1$ if $x \in K$, and $f(x) = 0$ if $x \notin U$. By (8.3) and (8.4),

$$|\mu_{m_1}(A) - \int_{\widetilde{K}_{m_1}} f d\mu_{m_1}| < \varepsilon,$$

$$|\mu_{m_2}(A) - \int_{\widetilde{K}_{m_2}} f d\mu_{m_2}| < \varepsilon.$$

Noting that $\int_{\widetilde{K}_{m_i}} f d\mu_{m_i} = \lim_{k\to\infty} \int_{\widetilde{K}_{m_i}} f d\mathrm{P}_{n_k}$, $i = 1, 2$, and $\int_{\widetilde{K}_{m_1}} f d\mathrm{P}_{n_k} \leq \int_{\widetilde{K}_{m_2}} f d\mathrm{P}_{n_k}$, we conclude that

$$\mu_{m_1}(A) \leq \mu_{m_2}(A) + 2\varepsilon.$$

Since ε was arbitrary, we obtain the desired monotonicity.

Define

$$\mathrm{P}(A) = \lim_{m\to\infty} \mu_m(A \bigcap \widetilde{K}_m).$$

Note that $\mathrm{P}(X) = \lim_{m\to\infty} \mu_m(\widetilde{K}_m) = 1$. We must show that P is σ-additive in order to conclude that it is a probability measure. If $A = \bigcup_{i=1}^{\infty} A_i$ is a union of non-intersecting sets, then

$$\mathrm{P}(A) \geq \lim_{m\to\infty} \mu_m(\bigcup_{i=1}^{n} A_i \bigcap \widetilde{K}_m) = \sum_{i=1}^{n} \mathrm{P}(A_i)$$

for each n, and therefore $\mathrm{P}(A) \geq \sum_{i=1}^{\infty} \mathrm{P}(A_i)$. If $\varepsilon > 0$ is fixed, then for sufficiently large m

$$\mathrm{P}(A) \leq \mu_m(A \bigcap \widetilde{K}_m) + \varepsilon = \sum_{i=1}^{\infty} \mu_m(A_i \bigcap \widetilde{K}_m) + \varepsilon \leq \sum_{i=1}^{\infty} \mathrm{P}(A_i) + \varepsilon.$$

Since ε was arbitrary, $\mathrm{P}(A) \leq \sum_{i=1}^{\infty} \mathrm{P}(A_i)$, and thus P is a probability measure.

It remains to show that the measures P_{n_k} converge to the measure P weakly. Let A be a closed set and $\varepsilon > 0$. Then, by the construction of the sets $\widetilde{K}_m$, there is a sufficiently large m such that

$$\limsup_{k\to\infty} \mathrm{P}_{n_k}(A) \leq \limsup_{k\to\infty} \mathrm{P}_{n_k}(A \bigcap \widetilde{K}_m) + \varepsilon \leq \mu_m(A) + \varepsilon \leq \mathrm{P}(A) + \varepsilon.$$

By Lemma 8.3, this implies the weak convergence of measures. Therefore the family of measures $\{\mathrm{P}_\alpha\}$ is weakly compact. □

8.4 Problems

1. Let (X, d) be a separable complete metric space. For $x \in X$, let δ_x be the measure on $(X, \mathcal{B}(X))$ which is concentrated at x, that is $\delta(A) = 1$ if $x \in A$, $\delta(A) = 0$ if $x \notin A$, $A \in \mathcal{B}(X)$. Prove that δ_{x_n} converge weakly if and only if there is $x \in X$ such that $x_n \to x$ as $n \to \infty$.
2. Prove that if P_n and P are probability measures, then P_n converges weakly to P if and only if

$$\liminf_{n\to\infty} \mathrm{P}_n(U) \geq \mathrm{P}(U)$$

for any open set U.

3. Prove that if P_n and P are probability measures, then P_n converges to P weakly if and only if

$$\lim_{n\to\infty} \mathrm{P}_n(A) = \mathrm{P}(A)$$

for all sets A such that $\mathrm{P}(\partial A) = 0$, where ∂A is the boundary of the set A.

4. Let X be a metric space and $\mathcal{B}(X)$ the σ-algebra of its Borel sets. Let μ_1 and μ_2 be two probability measures such that $\int_X f d\mu_1 = \int_X f d\mu_2$ for all $f \in C_b(X)$, $f \geq 0$. Prove that $\mu_1 = \mu_2$.

5. Give an example of a family of probability measures P_n on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathrm{P}_n \Rightarrow \mathrm{P}$ (weakly), P_n, P are absolutely continuous with respect to the Lebesgue measure, yet there exists a Borel set A such that $\mathrm{P}_n(A)$ does not converge to $\mathrm{P}(A)$.

6. Assume that a sequence of random variables ξ_n converges to a random variable ξ in distribution, and a numeric sequence a_n converges to 1. Prove that $a_n\xi_n$ converges to ξ in distribution.

7. Suppose that ξ_n, η_n, $n \geq 1$, and ξ are random variables defined on the same probability space. Prove that if $\xi_n \Rightarrow \xi$ and $\eta_n \Rightarrow c$, where c is a constant, then $\xi_n\eta_n \Rightarrow c\xi$.

8. Prove that if $\xi_n \to \xi$ in probability, then $\mathrm{P}_{\xi_n} \Rightarrow \mathrm{P}_\xi$, that is the convergence of the random variables in probability implies weak convergence of the corresponding probability measures.

9. Let P_n, P be probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Suppose that $\int_{\mathbb{R}} f d\mathrm{P}_n \to \int_{\mathbb{R}} f d\mathrm{P}$ as $n \to \infty$ for every infinitely differentiable function f with compact support. Prove that $\mathrm{P}_n \Rightarrow \mathrm{P}$.

10. Prove that if ξ_n and ξ are defined on the same probability space, ξ is identically equal to a constant, and ξ_n converge to ξ in distribution, then ξ_n converge to ξ in probability.

11. Consider a Markov transition function P on a compact state space X. Prove that the corresponding Markov chain has at least one stationary measure. (Hint: Take an arbitrary initial measure μ and define $\mu_n = (P^*)^n\mu$, $n \geq 0$. Prove that the sequence of measures $\nu_n = (\mu_0 + \ldots + \mu_{n-1})/n$ is weakly compact, and the limit of a subsequence is a stationary measure.)