3 Convergence in Distribution

Given: a metric space (M, ρ). Put

$$C^b(M) = \{f : M \to \mathbb{R} : f \text{ bounded, continuous}\},$$

and consider the Borel-σ-algebra $\mathfrak{B}(M)$ in M. Moreover, let $\mathfrak{M}(M)$ denote the set of all probability measures on $\mathfrak{B}(M)$.

Definition 1.

(i) A sequence $(Q_n)_{n\in\mathbb{N}}$ in $\mathfrak{M}(M)$ *converges weakly* to $Q \in \mathfrak{M}(M)$ if

$$\forall\, f \in C^b(M) : \lim_{n\to\infty} \int f\, dQ_n = \int f\, dQ.$$

Notation: $Q_n \xrightarrow{\text{w}} Q$.

(ii) A sequence $(X_n)_{n\in\mathbb{N}}$ of random elements with values in M *converges in distribution* to a random element X with values in M iff $\mathfrak{L}(X_n) \xrightarrow{\text{w}} \mathfrak{L}(X)$ (recall $\mathfrak{L}(X)$ is the distribution of X).
Notation: $X_n \xrightarrow{\text{d}} X$.

Remark 1. For convergence in distribution the random elements need not be defined on a common probability space.

In the sequel: Q_n, $Q \in \mathfrak{M}(M)$ for $n \in \mathbb{N}$.

Example 1.

(i) For $x_n,\ x \in M$
$$\delta_{x_n} \xrightarrow{\text{w}} \delta_x \quad \Leftrightarrow \quad \lim_{n\to\infty} \rho(x_n, x) = 0.$$
For the proof of '⇐', note that
$$\int f\, d\varepsilon_{x_n} = f(x_n), \qquad \int f\, d\varepsilon_x = f(x).$$
For the proof of '⇒', suppose that $\limsup_{n\to\infty} \rho(x_n, x) > 0$. Take
$$f(y) = \min(\rho(y, x), 1), \qquad y \in M,$$
and observe that $f \in C^b(M)$ and
$$\limsup_{n\to\infty} \int f\, d\delta_{x_n} = \limsup_{n\to\infty} (\min(\rho(x_n, x), 1)) > 0$$
while $\int f\, d\delta_x = 0$.

(ii) Let $(M, \rho) = (\mathfrak{R}, |\cdot|)$, $Q_n = N(\mu_n, \sigma_n^2)$ with $\mu_n \in \mathfrak{R}$, $\sigma_n > 0$.
Claim:
$$Q_n \xrightarrow{\text{w}} Q \Leftrightarrow Q = N(\mu, \sigma), \mu_n \to \mu, \sigma_n \to \sigma\ .$$
(Here, $N(\mu, 0) = \delta_\mu$).
Proof: '⇐': For $f \in C_b(\mathfrak{R})$,
$$\int f\, dQ_n = 1/\sqrt{2\pi} \cdot \int_{\mathbb{R}} \underbrace{f(\sigma_n \cdot x + \mu_n)}_{\to f(\sigma x+\mu), \leq b} \cdot \exp(-1/2 \cdot x^2)\, \lambda_1(dx) \to \int_{\mathbb{R}} f(\sigma \cdot x + \mu) \cdot \exp(-1/2 \cdot x^2)\, \lambda_1(dx)$$
'⇒': Übung 8.4.

Remark 2. Note that $Q_n \xrightarrow{\text{w}} Q$ does not imply
$$\forall\, A \in \mathfrak{B}(M) : \lim_{n\to\infty} Q_n(A) = Q(A).$$
For instance, assume $\lim_{n\to\infty} \rho(x_n, x) = 0$ with $x_n \neq x$ for every $n \in \mathbb{N}$. Then
$$\delta_{x_n}(\{x\}) = 0, \qquad \delta_x(\{x\}) = 1.$$

Theorem 1 (Portmanteau Theorem). The following properties are equivalent:

(i) $Q_n \xrightarrow{\text{w}} Q$,

(ii) $\forall f \in C^b(M)$ uniformly continuous : $\lim_{n\to\infty} \int f\, dQ_n = \int f\, dQ$,

(iii) $\forall A \subset M$ closed : $\limsup_{n\to\infty} Q_n(A) \leq Q(A)$,

(iv) $\forall A \subset M$ open : $\liminf_{n\to\infty} Q_n(A) \geq Q(A)$,

(v) $\forall A \in \mathfrak{B}(M)$: $Q(\partial A) = 0 \;\Rightarrow\; \lim_{n\to\infty} Q_n(A) = Q(A)$.

Proof. See Gänssler, Stute (1977, Satz 8.4.9). □

In the sequel, we study the particular case $(M, \mathfrak{B}(M)) = (\mathbb{R}, \mathfrak{B})$, i.e., convergence in distribution for random variables. The Central Limit Theorem deals with this notion of convergence, see the introductory Example I.1 and Section IV.5.
Notation: for any $Q \in \mathfrak{M}(\mathbb{R})$

$$F_Q(x) = Q(]-\infty, x]), \qquad x \in \mathbb{R},$$

and for any function $F : \mathbb{R} \to \mathbb{R}$

$$\mathrm{Cont}(F) = \{x \in \mathbb{R} : F \text{ continuous at } x\}.$$

Theorem 2.

$$Q_n \xrightarrow{\text{w}} Q \quad \Leftrightarrow \quad \forall x \in \mathrm{Cont}(F_Q) : \lim_{n\to\infty} F_{Q_n}(x) = F_Q(x).$$

Moreover, if $Q_n \xrightarrow{\text{w}} Q$ and $\mathrm{Cont}(F_Q) = \mathbb{R}$ then

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_{Q_n}(x) - F_Q(x)| = 0.$$

Proof. '⇒': If $x \in \mathrm{Cont}(F_Q)$ and $A =]-\infty, x]$ then $Q(\partial A) = Q(\{x\}) = 0$, see Theorem 1.2. Hence Theorem 1 implies

$$\lim_{n\to\infty} F_{Q_n}(x) = \lim_{n\to\infty} Q_n(A) = Q(A) = F_Q(x).$$

'⇐': Consider a non-empty open set $A \subset \mathbb{R}$. Take pairwise disjoint open intervals $A_1, A_2, \ldots$ such that $A = \bigcup_{i=1}^{\infty} A_i$. Fatous's Lemma implies

$$\liminf_{n\to\infty} Q_n(A) = \liminf_{n\to\infty} \sum_{i=1}^{\infty} Q_n(A_i) \geq \sum_{i=1}^{\infty} \liminf_{n\to\infty} Q_n(A_i).$$

Note that $\mathbb{R} \setminus \mathrm{Cont}(F_Q)$ is countable. Fix $\varepsilon > 0$, and take

$$A_i' =]a_i', b_i'] \subset A_i$$

for $i \in \mathbb{N}$ such that

$$a_i', b_i' \in \mathrm{Cont}(F_Q) \;\wedge\; Q(A_i) \leq Q(A_i') + \varepsilon \cdot 2^{-i}.$$

Then

$$\liminf_{n\to\infty} Q_n(A_i) \geq \liminf_{n\to\infty} Q_n(A_i') = Q(A_i') \geq Q(A_i) - \varepsilon \cdot 2^{-i}.$$

We conclude that

$$\liminf_{n\to\infty} Q_n(A) \geq Q(A) - \varepsilon,$$

and therefore $Q_n \xrightarrow{\mathrm{w}} Q$ by Theorem 1.

Uniform convergence, Übung 9.1. □

Corollary 1.

$$Q_n \xrightarrow{\mathrm{w}} Q \ \wedge\ Q_n \xrightarrow{\mathrm{w}} \widetilde{Q} \quad \Rightarrow \quad Q = \widetilde{Q}.$$

Proof. By Theorem 2 $F_Q(x) = F_{\widetilde{Q}}(x)$ if $x \in D = \mathrm{Cont}(F_Q) \cap \mathrm{Cont}(F_{\widetilde{Q}})$. Since D is dense in $\mathbb{R}$ and F_Q as well as $F_{\widetilde{Q}}$ are right-continuous, we get $F_Q = F_{\widetilde{Q}}$. Apply Theorem 1.3. □

Given: random variables X_n, X on $(\Omega, \mathfrak{A}, P)$ for $n \in \mathbb{N}$.

Theorem 3.

$$X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{\mathrm{d}} X$$

and

$$X_n \xrightarrow{\mathrm{d}} X \ \wedge\ X \text{ constant a.s.} \quad \Rightarrow \quad X_n \xrightarrow{P} X.$$

Proof. Assume $X_n \xrightarrow{P} X$. For $\varepsilon > 0$ and $x \in \mathbb{R}$

$$\begin{aligned}
&P(\{X \leq x - \varepsilon\}) - P(\{|X - X_n| > \varepsilon\}) \\
&\quad \leq P(\{X \leq x - \varepsilon\} \cap \{|X - X_n| \leq \varepsilon\}) \\
&\quad \leq P(\{X_n \leq x\}) \\
&\quad \leq P(\{X_n \leq x\} \cap \{X \leq x + \varepsilon\}) + P(\{X_n \leq x\} \cap \{X > x + \varepsilon\}) \\
&\quad \leq P(\{X \leq x + \varepsilon\}) + P(\{|X - X_n| > \varepsilon\}).
\end{aligned}$$

Thus

$$F_X(x - \varepsilon) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x + \varepsilon).$$

For $x \in \mathrm{Cont}(F_X)$ we get $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$. Apply Theorem 2.

Now, assume that $X_n \xrightarrow{\mathrm{d}} X$ and $P_X = \varepsilon_x$. Let $\varepsilon > 0$ and take $f \in C^b(\mathbb{R})$ such that $f \geq 0$, $f(x) = 0$, and $f(y) = 1$ if $|x - y| \geq \varepsilon$. Then

$$P(\{|X - X_n| > \varepsilon\}) = P(\{|x - X_n| > \varepsilon\}) \leq \int 1_{\mathbb{R}\setminus[x-\varepsilon, x+\varepsilon]}\, dP_{X_n} \leq \int f\, dP_{X_n}$$

and

$$\lim_{n\to\infty} \int f\, dP_{X_n} = \int f\, dP_X = 0.$$

□

Example 2. Consider the uniform distribution P on $\Omega = \{0, 1\}$. Put

$$X_n(\omega) = \omega, \qquad X(\omega) = 1 - \omega.$$

Then $P_{X_n} = P_X$ and therefore

$$X_n \xrightarrow{\mathrm{d}} X.$$

However, $\{|X_n - X| < 1/2\} = \emptyset$ and therefore

$$X_n \xrightarrow{P} X \text{ does not hold.}$$

We quote without proof an interesting converse to Theorem 3:

Theorem 4 (Skorohod). Let $\Omega =]0, 1[$ with the uniform distribution, and $Q_n, Q \in \mathfrak{M}(\mathfrak{R})$. If $Q_n \xrightarrow{\mathrm{w}} Q$, then for the random variables

$$X_{Q_n}(\omega) := F^{-}_{Q_n}(z) := \inf\{z \in \mathfrak{R} \ : \ \omega \leq F_{Q_n}(z)\}, \qquad \omega \in]0, 1[$$

we have $P_{X_{Q_n}} = Q_n$ and $X_{Q_n} \to X_Q$ P–a.s..

Remark 3. Skorohod's Theorem is based on a general method to transform uniformly distributed 'random numbers' from $]0, 1[$ into 'random numbers' with distribution Q. Namely, if U is uniformly distributed on $]0, 1[$ and F is some distribution function, $X := F^{-}(U)$ has F as distribution function.

Remark 4. Let $\mu, \mu_n \in \mathfrak{M}(M)$. Then

$$\mu_n \xrightarrow{\mathrm{w}} \mu \quad \Leftrightarrow \quad \forall_{n_k} \, \exists_{n_{k_l}} \, \mu_{n_{k_l}} \xrightarrow{\mathrm{w}} \mu \ ,$$

see `Übung`8.3.

Finally, we present a compactness criterion, which is very useful for construction of probability measures on $\mathfrak{B}(M)$. We need a generalized Bolzano–Weierstrass Theorem:

Lemma 1. Let $x_{n,\ell} \in \mathbb{R}$ for $n, \ell \in \mathbb{N}$ with

$$\forall \, \ell \in \mathbb{N} : \ \sup_{n \in \mathbb{N}} |x_{n,\ell}| < \infty.$$

Then there exists an increasing sequence $(n_i)_{i \in \mathbb{N}}$ in $\mathbb{N}$ such that

$$\forall \, \ell \in \mathbb{N} : \ (x_{n_i,\ell})_{i \in \mathbb{N}} \text{ converges.}$$

Proof. (Sketch, see Billingsley (1979, Thm. 25.13 for details): For fixed l, $x_{n,l}$ is a bounded sequence. Hence, by the original Bolzano–Weierstrass Theorem, there is a subsequence $\pi_1(n)$ (where $\pi_1 : \mathfrak{N} \to \mathfrak{N}$ is a monotone mapping) such that $(x_{\pi(n),1})$ converges. Next, $(x_{\pi_1(n),2})$ is bounded, hence there is $\pi_2 : \mathfrak{N} \to \mathfrak{N}$ monotone such that $x_{\pi_2(\pi_1(n),2)}$ converges. Set $\pi^2 := \pi_2 \circ \pi_1$. Iterating, we can find for each m some $\pi^m = \pi_m \circ \pi^{m-1}$ such that for $l \leq m$, $x_{\pi^m(n),k}$ converges. Define $n_i = \pi^i(i)$. □

Definition 2.

(i) $\mathfrak{P} \subset \mathfrak{M}(M)$ *tight* if

$$\forall \varepsilon > 0 \; \exists K \subset M \text{ compact } \forall P \in \mathfrak{P}: \quad P(K) \geq 1 - \varepsilon.$$

(ii) $\mathfrak{P} \subset \mathfrak{M}(M)$ *relatively compact* if every sequence in $\mathfrak{P}$ contains a subsequence that converges weakly.

Theorem 5 (Prohorov). Assume that M is a complete separable metric space and $\mathfrak{P} \subset \mathfrak{M}(M)$. Then

$$\mathfrak{P} \text{ relatively compact} \quad \Leftrightarrow \quad \mathfrak{P} \text{ tight}.$$

Proof. We only treat the case $M = \mathfrak{R}$; see Parthasarathy (1967, Thm. II.6.7) for the general case.
'$\Rightarrow$': Suppose that $\mathfrak{P}$ is not tight. Then, for some $\varepsilon > 0$, there exists a sequence $(P_n)_{n \in \mathbb{N}}$ in $\mathfrak{P}$ such that

$$P_n([-n, n]) < 1 - \varepsilon.$$

For a suitable subsequence, $P_{n_k} \xrightarrow{\text{w}} P \in \mathfrak{M}(\mathbb{R})$. Take $m > 0$ such that

$$P(]-m, m[) > 1 - \varepsilon.$$

Theorem 1 implies

$$P(]-m, m[) \leq \liminf_{k \to \infty} P_{n_k}(]-m, m[) \leq \liminf_{k \to \infty} P_{n_k}([-n_k, n_k]) < 1 - \varepsilon,$$

which is a contradiction.
'$\Leftarrow$': Consider any sequence $(P_n)_{n \in \mathbb{N}}$ in $\mathfrak{P}$ and the corresponding sequence $(F_n)_{n \in \mathbb{N}}$ of distribution functions. Use Lemma 1 to obtain a subsequence $(F_{n_i})_{i \in \mathbb{N}}$ and a non-decreasing function $G : \mathbb{Q} \to [0, 1]$ with

$$\forall q \in \mathbb{Q}: \lim_{i \to \infty} F_{n_i}(q) = G(q).$$

Put

$$F(x) = \inf\{G(q) : q \in \mathbb{Q} \wedge x < q\}, \qquad x \in \mathbb{R}.$$

Claim (*Helly's Theorem*):

(i) F is non-decreasing and right-continuous,

(ii) $\forall x \in \text{Cont}(F): \lim_{i \to \infty} F_{n_i}(x) = F(x)$.

Proof: Ad (i): Obviously F is non-decreasing. For $x \in \mathbb{R}$ and $\varepsilon > 0$ take $\delta_2 > 0$ such that

$$\forall q \in \mathbb{Q} \cap]x, x + \delta_2[: \quad G(q) \leq F(x) + \varepsilon.$$

Thus, for $z \in]x, x + \delta_2[$,

$$F(x) \leq F(z) \leq F(x) + \varepsilon.$$

Ad (ii): If $x \in \text{Cont}(F)$ take $\delta_1 > 0$ such that

$$F(x) - \varepsilon \leq F(x - \delta_1).$$

Thus, for $q_1, q_2 \in \mathbb{Q}$ with

$$x - \delta_1 < q_1 < x < q_2 < x + \delta_2,$$

we get

$$\begin{aligned} F(x) - \varepsilon \leq F(x - \delta_1) \leq G(q_1) \leq \liminf_{i\to\infty} F_{n_i}(x) \leq \limsup_{i\to\infty} F_{n_i}(x) \\ \leq G(q_2) \leq F(x) + \varepsilon. \end{aligned}$$

Claim:

$$\lim_{x\to-\infty} F(x) = 0 \ \wedge \ \lim_{x\to\infty} F(x) = 1.$$

Proof: For $\varepsilon > 0$ take $m \in \mathbb{Q}$ such that

$$\forall\, n \in \mathbb{N} : \ P_n(]-m, m]) \geq 1 - \varepsilon.$$

Thus

$$G(m) - G(-m) = \lim_{i\to\infty} \big(F_{n_i}(m) - F_{n_i}(-m)\big) = \lim_{i\to\infty} P_{n_i}(]-m, m]) \geq 1 - \varepsilon.$$

Since $F(m) \geq G(m)$ and $F(-m-1) \leq G(-m)$, we obtain

$$F(m) - F(-m-1) \geq 1 - \varepsilon.$$

It remains to apply Theorems 1.3 and 2. □