7. Convergence in distribution

7.1. **Compactness of distribution functions.** Recall that a distribution is a non-decreasing right-continuous function $F : \mathbb{R} \to [0, 1]$ satisfying

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

Throughout this section, we use $\mathcal{M}$ to denote the set of all non-decreasing right-continuous functions F satisfying $F(-\infty) \geq 0$ and $F(\infty) \leq 1$.

Definition 7.1. Let f, f_n be functions defined on $\mathbb{R}$. f_n is said to **converge in distribution** (or **$\mathcal{D}$-converge**) to f if

$$f_n(x) \to f(x) \quad \forall x \in \mathcal{C}(f)$$

where $\mathcal{C}(f)$ is the set of continuous points of f.

Remark 7.1. If $f_n \to f$ and $f_n \to g$ in distribution and f, g are monotonic and right-continuous, then $f = g$.

Lemma 7.1. *Let f_n, f be functions defined on $\mathbb{R}$. Then, the following are equivalent.*

1. *f_n converges to f in distribution.*
2. *For any subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$, f_{n_k} converges to f in distribution.*
3. *For any subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$, there is a further subsequence $(f_{m_k})_{k=1}^{\infty}$ such that f_{m_k} converges to f in distribution.*

Lemma 7.2. *The set $\mathcal{M}$ is sequentially compact under $\mathcal{D}$-convergence.*

Proof. Let $(f_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{M}$ and $\mathbb{Q}$ be the set of all rational numbers. For $x \in \mathbb{Q}$, since $\{f_n(x) : n = 1, 2, ...\}$ is bounded, one may choose a convergent subsequence. By Cantor's diagonalization method, we may choose a subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ such that f_{n_k} converges on $\mathbb{Q}$. Set, for $x \in \mathbb{Q}$,

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

and define, by the monotonicity of f on $\mathbb{Q}$,

$$f(x) = \lim_{y\downarrow x} f(y) \quad \forall y \in \mathbb{R} \setminus \mathbb{Q}.$$

Clearly, f is non-decreasing in $\mathbb{R}$. Let $\mathcal{C}(f)$ be the set of continuous points of f and $x \in \mathcal{C}(f)$. Since f is continuous at x, we may choose, for each $\epsilon > 0$, $y, z \in \mathbb{Q}$ such that $y < x < z$ and

$$f(x) - \epsilon/2 < f(y) \leq f(z) < f(x) + \epsilon/2.$$

Since $f_{n_k}(y) \to f(y)$ and $f_{n_k}(z) \to f(z)$, we may choose an integer $K = K(\epsilon)$ such that

$$f_{n_k}(y) > f(y) - \epsilon/2, \quad f_{n_k}(z) < f(z) + \epsilon/2 \quad \forall k \geq K.$$

This implies

$$f(x) - \epsilon < f_{n_k}(y) \leq f_{n_k}(x) \leq f_{n_k}(z) < f(x) + \epsilon, \quad \forall k \geq K,$$

which proves that f_n converges to f in distribution. By setting g as a right-continuous version of f, one has $g \in \mathcal{M}$, $f = g$ on $\mathcal{C}(f)$ and $\mathcal{C}(g) = \mathcal{C}(f)$. Thus, f_n converges to g in distribution. □

Remark 7.2. The set of all distribution functions is not sequentially compact. Consider the following sequence of functions

$$F_n(x) = \begin{cases} 0 & \text{for } x < n \\ 1 & \text{for } x \geq n \end{cases}$$

$F_n \to 0$ in distribution.

For any $F \in \mathcal{M}$, let $F(B)$, $B \in \mathcal{B}_1$, be the measure induced by F.

Definition 7.2. Let $\mathcal{N}$ be the set of all distribution functions. A set $\mathcal{A} \subset \mathcal{N}$ is

(1) **mass-preserving** if, for all $\epsilon > 0$, there exists a finite interval I_ϵ such that $F(I_\epsilon^c) < \epsilon$ for all $F \in \mathcal{A}$.

(2) **conditionally compact** in $\mathcal{N}$ under $\mathcal{D}$-convergence if any sequence $(f_n)_{n=1}^\infty$ in $\mathcal{A}$ contains a subsequence that $\mathcal{D}$-converges to a distribution function in $\mathcal{N}$.

Lemma 7.3. *A set $\mathcal{A} \subset \mathcal{N}$ is conditionally compact under $\mathcal{D}$-convergence if and only if $\mathcal{A}$ is mass-preserving.*

Proof. Assume first that $\mathcal{A}$ is mass-preserving and let $(F_n)_{n=1}^\infty$ be a sequence of distribution functions in $\mathcal{A}$. By Lemma 7.2, there exists a subsequence $(F_{n_k})_{k=1}^\infty$ and $F \in \mathcal{M}$ such that F_{n_k} converges to F in distribution. We will show in the following that $F \in \mathcal{N}$ and it remains to prove $F(-\infty) = 0$ and $F(\infty) = 1$.

Let $\epsilon > 0$ and choose an interval I such that $f(I^c) < \epsilon$ for all $f \in \mathcal{A}$. Let $a, b \in \mathcal{C}(F)$ be such that $I \subset (a, b]$. This implies

$$F_{n_k}(b) - F_{n_k}(a) = F_{n_k}((a,b]) > 1 - \epsilon \quad \forall k \geq 1,$$

and, then,

$$F((a,b]) = F(b) - F(a) = \lim_{k\to\infty} [F_{n_k}(b) - F_{n_k}(a)] \geq 1 - \epsilon.$$

Letting $a \to -\infty$ and $b \to \infty$ yields $1 \geq F(\infty) - F(-\infty) \geq 1 - \epsilon$ for all $\epsilon > 0$. As it is known that $F(-\infty) \geq 0$ and $F(\infty) \leq 1$, we have $F \in \mathcal{N}$.

For the converse, assume that $\mathcal{A}$ is conditionally compact and assume the inverse that $\mathcal{A}$ is not mass-preserving. Equivalently, the latter means that there is $\epsilon > 0$ such that, for any finite interval I, one may select $F \in \mathcal{A}$ such that $F(I) \leq 1 - \epsilon$. For $n \geq 1$, let $F_n \in \mathcal{A}$ be such that $F_n([-n, n]) \leq 1 - \epsilon$. Since $\mathcal{A}$ is conditionally compact, we may choose a subsequence $(F_{n_k})_{k=1}^\infty$ and $F \in \mathcal{N}$ such that

$$F_{n_k} \to F \quad \text{in distribution.}$$

This implies that, for $a, b \in \mathcal{C}(F)$,

$$F((a,b]) = \lim_{k\to\infty} F_{n_k}((a,b]) \leq \lim_{k\to\infty} F_{n_k}([-n_k, n_k]) \leq 1 - \epsilon.$$

Letting $a \to -\infty$ and $b \to \infty$ yields $F(\infty) - F(-\infty) \leq 1 - \epsilon < 1$. This contradicts the fact $F \in \mathcal{N}$. □

Proposition 7.4. *Suppose $F_n \in \mathcal{N}$ converges to F in distribution. Then, $F \in \mathcal{N}$ if and only if $\{F_n | n = 1, 2, ...\}$ is mass-preserving.*

Proof. First, assume that $F \in \mathcal{N}$ and set $\mathcal{A} = \{F_n | n \geq 1\} \subset \mathcal{N}$. Let $(f_n)_{n=1}^\infty$ be a sequence in $\mathcal{A}$ and write $f_n = F_{\ell_n}$. Set $\mathcal{I} = \{\ell_n | n \geq 1\}$. If $|\mathcal{I}| < \infty$, then there is $M \in \mathcal{I}$ and a subsequence $(m_k)_{k=1}^\infty$ of $\mathbb{N}$ such that $f_{m_k} = F_M$ for all $k \geq 1$. Clearly, f_{m_k} converges to $F_M \in \mathcal{N}$ in distribution. When $|\mathcal{I}| = \infty$, we may choose a subsequence $(m_k')_{k=1}^\infty$ of $\mathbb{N}$ such that $(\ell_{m_k'})_{k=1}^\infty$ is increasing. This implies that $f_{m_k'}$ converges to F in distribution. Consequently, $\mathcal{A}$ is conditionally compact and, by Lemma 7.3, $\mathcal{A}$ is mass-preserving.

Next, assume that $(F_n)_{n=1}^\infty$ is mass-preserving. By Lemma 7.3, there is a subsequence $(F_{n_k})_{k=1}^\infty$ and $G \in \mathcal{N}$ such that $F_{n_k} \to G$ in distribution. Since $F_{n_k} \to F$, by Remark 7.1, $F = G$. □

Exercise 7.1. For $a < b$, show that the set of all distribution functions F satisfying $F(b) = 1$ and $F(a - \epsilon) = 0$ for all $\epsilon > 0$ is sequentially compact.

Exercise 7.2. Let $F_n, F \in \mathcal{N}$ and assume that $F_n \to F$ in distribution. Prove that, if $B \subset \mathbb{R}$ is closed and $F_n(B) = 1$ for all $n \geq 1$, then $F(B) = 1$. Find an example that $F_n(B) = 1$ for $n \geq 1$ but $F(B) < 1$.

Exercise 7.3. Let $g : \mathbb{R} \to \mathbb{R}$ be a random variable satisfying $|g(x)| \to \infty$ as $|x| \to \infty$ and $A \subset \mathcal{N}$ be a set satisfying $\sup_{F \in A} \int |g(x)| dF(x) < \infty$. Show that A is mass-preserving.

Remark 7.3. For $n \geq 1$, let X_n be a random variable with distribution function F_n. By the above exercise, if $\limsup_n \mathbb{E}|X_n|^r < \infty$ for some $r > 0$, then $\{F_n | n \in \mathbb{N}\}$ is mass-preserving.

7.2. **Preserving the integrations.** In this subsection, we consider the following question. Suppose $F_n \in \mathcal{N}$ converges to $F \in \mathcal{N}$ in distribution. Under what assumption does the limit $\int f dF_n \to \int f dF$ hold? For $n \geq 1$, let F_n be defined by

$$F_n(x) = \begin{cases} 0 & \text{for } x < -\frac{1}{n} \\ 1 & \text{for } x \geq -\frac{1}{n} \end{cases}$$

Then, $F_n \to F$ in distribution, where

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

Let f be a function defined by $f(x) = 1$ for $x < 0$ and $f(x) = 0$ for $x \geq 0$. Clearly, one may compute that $\int f dF_n = 1$ and $\int f dF = 0$.

Proposition 7.5. *Let X and X_n with $n, \geq 1$ be random variables with distribution functions F and F_n. Then, the following are equivalent.*

1. *$F_n \to F$ in distribution.*
2. *$\mathbb{E}\phi(X_n) \to \mathbb{E}\phi(X)$ for any bounded continuous function ϕ.*
3. *$\mathbb{E}\phi(X_n) \to \mathbb{E}\phi(X)$ for any bounded $\mathcal{B}(\mathbb{R})$-measurable function ϕ satisfying $F(\mathcal{C}(\phi)) = 1$.*
4. *There are Y_n and Y such that $X_n \stackrel{d}{=} Y_n$, $X \stackrel{d}{=} Y$ and $Y_n \to Y$.*

Remark 7.4. For any $\mathcal{B}(\mathbb{R})$-measurable function ϕ, $\mathcal{C}(\phi) \in \mathcal{B}(\mathbb{R})$.

Proof of Proposition 7.5. For (4)⇒(3), let $Y_n \stackrel{d}{=} X_n$ and $Y \stackrel{d}{=} X$ be random variables on $(\Omega, \mathcal{F}, P)$ such that $Y_n \to Y$. By Theorem 4.15, we have

$$\mathbb{E}\phi(X_n) = \mathbb{E}\phi(Y_n), \quad \mathbb{E}\phi(X) = \mathbb{E}\phi(Y).$$

Set $A = \{Y \in \mathcal{C}(\phi)\}$. Note that $\mathbb{P}(A) = F(\mathcal{C}(\phi)) = 1$ and $\phi(Y_n)$ converges to $\phi(Y)$ on A. By the Lebesgue dominant convergence theorem, $\mathbb{E}\phi(Y_n) \to \mathbb{E}\phi(Y)$, as desired.

(3)⇒(2) is clear. For (2)⇒(1), let $x \in \mathcal{C}(F)$ and, for $\epsilon > 0$, set

$$\phi(t) = \begin{cases} 1 & \text{for } t < x \\ \dfrac{x + \epsilon - t}{\epsilon} & \text{for } x \leq t \leq x + \epsilon \\ 0 & \text{for } t > x + \epsilon \end{cases}$$

It is easy to see that $\mathbb{E}\phi(X_n) \geq F_n(x)$ and this implies

$$\limsup_{n \to \infty} F_n(x) \leq \mathbb{E}\phi(X) \leq F(x + \epsilon).$$

Similarly, one may use $\varphi(t) = \phi(t + \epsilon)$ to derive

$$\liminf_{n \to \infty} F_n(x) \geq F(x - \epsilon).$$

By the continuity of F at x, letting $\epsilon \to 0$ implies $F_n(x) \to F(x)$.

For (1)⇒(4), we define, for $n \geq 1$,

$$f_n(t) = \inf\{x|F_n(x) \geq t\} = \min\{x|F_n(x) \geq t\}, \quad \forall t \in (0,1).$$

In the above setting, f_n is non-decreasing and satisfies

$$f_n(t) \leq x \quad \Leftrightarrow \quad F_n(x) \geq t. \tag{7.1}$$

To see the left-continuity of f_n, observe that

$$F_n(f_n(t)) \geq t, \quad \forall t \in (0,1).$$

Fix $t \in (0,1)$. By the monotonicity of f_n and F_n, the above inequality implies

$$F_n\left(\lim_{s\to t^-} f_n(s)\right) \geq \lim_{s\to t^-} F_n\left(f_n(s)\right) \geq t,$$

which yields

$$f_n(t) \leq \lim_{s\to t^-} f_n(s) \leq f_n(t).$$

Let f be the function defined by $f(t) = \inf\{x|F(x) \geq t\}$.
Step 1: $f_n \to f$ on $\mathcal{C}(f)$. Let $t \in \mathcal{C}(f)$ and $x = f(t)$. Since f is right-continuous at t, $F(x+\epsilon) > t$ for all $\epsilon > 0$, otherwise $f(s) \geq x + \epsilon$ for all $s > t$. This implies that, for $\epsilon > 0$ and $x + \epsilon \in \mathcal{C}(F)$, one may choose $\delta > 0$ such that $F(x+\epsilon) \geq t + 2\delta$. Since $F_n(x+\epsilon) \to F(x+\epsilon)$, we may choose $N \in \mathbb{N}$ such that $F_n(x+\epsilon) \geq t + \delta$ for $n \geq N$. This implies $f_n(t) \leq f_n(t+\delta) \leq x + \epsilon = f(t) + \epsilon$ for $n \geq N$. Letting $n \to \infty$ and $\epsilon \to 0$ yields $\limsup_n f_n(t) \leq x$.

Next, observe that $F(x-\epsilon) < t$ for $\epsilon > 0$. This implies, for $\epsilon > 0$ and $x - \epsilon \in \mathcal{C}(F)$, one may select $\delta > 0$ such that $F(x-\epsilon) < t - 2\delta$. Since $F_n(x-\epsilon) \to F(x-\epsilon)$, we may choose $N' \in \mathbb{N}$ such that $F_n(x-\epsilon) < t - \delta$ for $n \geq N'$. This implies $f_n(t) \geq f_n(t-\delta) > x - \epsilon = f(t) - \epsilon$ for $n \geq N'$. Letting $n \to \infty$ and $\epsilon \to 0$ yields $\liminf_n f_n(t) \geq x$.
Step 2: Let μ be the Lebesgue measure on $(0,1)$. By (7.1), one has

$$\mu(\{t \in (0,1) : f_n(t) \leq x\}) = \mu(\{t \in (0,1) : F_n(x) \geq t\}) = F_n(x).$$

This implies $X_n \overset{d}{=} f_n$ and similarly $X \overset{d}{=} f$. To achieve the pointwise convergence, one only needs to set $Y_n = f_n \mathbf{1}_A$ and $Y = f\mathbf{1}_A$, where $A = (0,1) \cap \mathcal{C}(f)$. □

Definition 7.3. A sequence of random variables $(X_n)_{n=1}^\infty$ is said to converge to X in distribution if the distribution of X_n $\mathcal{D}$-converges to that of X.

Corollary 7.6. *Let $X_1, X_2, \ldots$ be random variables converging in distribution to X. If ϕ is a nonnegative function satisfying $\mathbb{P}(X \in \mathcal{C}(\phi)) = 1$, then $\mathbb{E}[\phi(X)] \leq \liminf_n \mathbb{E}[\phi(X_n)]$.*

Exercise 7.4. Let $F_n, F \in \mathcal{N}$ and suppose $F_n \to F$ in distribution. Show that if $B \in \mathcal{B}(\mathbb{R})$ satisfies $F(\partial B) = 0$, where $\partial B = \overline{B} \cap \overline{B^c}$ is the boundary of B, then $F_n(B) \to F(B)$.

Exercise 7.5. Let $X_1, X_2, \ldots$ be a sequence of random variables converging to X in distribution. Let g, h be continuous functions satisfying

$$\lim_{|x|\to\infty} |g(x)| = \infty, \quad \lim_{|x|\to\infty} \frac{|h(x)|}{|g(x)|} = 0.$$

Prove that if $\limsup_n \mathbb{E}|g(X_n)| < \infty$, then $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$.

7.3. **Classes of functions that separate.**

Definition 7.4. Let $\mathcal{N}$ be the set of all distribution functions. A class $\mathcal{C}$ of bounded continuous functions defined on $\mathbb{R}$ (not necessarily real-valued) is called *$\mathcal{N}$-separating* if, for any $F, G \in \mathcal{N}$,

$$\int_{\mathbb{R}} f(x)dF(x) = \int_{\mathbb{R}} f(x)dG(x), \quad \forall f \in \mathcal{C},$$

implies $F = G$.

Proposition 7.7. *Let $\mathcal{C}$ be an $\mathcal{N}$-separating class. Then, $F_n \in \mathcal{N}$ converges in distribution to some distribution $F \in \mathcal{N}$ if and only if $(F_n)_{n=1}^{\infty}$ is mass-preserving and*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x)dF_n(x) \quad \text{exists}, \quad \forall f \in \mathcal{C}.$$

Proof. By Proposition 7.5, if F_n $\mathcal{D}$-converges to F, then $\int f dF_n \to \int f dF$ for all $f \in \mathcal{C}$. Conversely, assume that $(F_n)_{n=1}^{\infty}$ is mass-preserving and $\int f dF_n$ converges for all $f \in \mathcal{C}$. By Lemma 7.2 and Proposition 7.4, one may choose a $\mathcal{D}$-convergent subsequence $(F_{n_k})_{k=1}^{\infty}$ with limit $F \in \mathcal{N}$ and, then, $\int f dF_n \to \int f dF$. Let $(F_{m_k})_{k=1}^{\infty}$ be any subsequence of $(F_n)_{n=1}^{\infty}$ and $(F_{m'_k})_{k=1}^{\infty}$ be a further subsequence that is $\mathcal{D}$-convergent with limit $G \in \mathcal{N}$. As a result, this implies

$$\int f dF = \lim_{n\to\infty} \int f dF_n = \lim_{k\to\infty} \int f dF_{m'_k} = \int f dG, \quad \forall f \in \mathcal{C}.$$

Since $\mathcal{C}$ separates $\mathcal{N}$, $F = G$. By Lemma 7.1, $F_n \to F$ in distribution. □

Proposition 7.8. *Let $\mathcal{C}_0$ be the set of functions of the following form.*

$$f(x) = \begin{cases} 0 & x < a - \epsilon \\ 1 + \frac{x-a}{\epsilon} & a - \epsilon \le x < a \\ 1 & a \le x < b \\ 1 - \frac{x-b}{\epsilon} & b \le x < b + \epsilon \\ 0 & x \ge b + \epsilon \end{cases} \tag{7.2}$$

where $a < b$ and $\epsilon > 0$. Then $\mathcal{C}_0$ is $\mathcal{N}$-separating.

Proof. Let $F, G \in \mathcal{N}$ and $a < b$ with $a \in \mathcal{C}(F) \cap \mathcal{C}(G)$. Let f be the function defined by (7.2). Then,

$$\int_{\mathbb{R}} f(x)dF(x) = \int_{\mathbb{R}} f(x)dG(x).$$

This implies

$$F((a,b]) \le \int_{\mathbb{R}} f(x)dF(x) = \int_{\mathbb{R}} f(x)dG(x) \le G((a-\epsilon, b+\epsilon]).$$

Similarly, one has $G((a,b]) \le F((a-\epsilon, b+\epsilon])$. Letting $\epsilon \to 0$ yields $F((a,b]) = G((a,b])$ for all $a \in \mathcal{C}(F) \cap \mathcal{C}(G)$ and, then, for all $a \in \mathbb{R}$. This proves $F = G$. □

Corollary 7.9. *Let $\mathcal{C}_0$ be the class in Proposition 7.8 and $\mathcal{C}$ be a class of bounded continuous functions on $\mathbb{R}$ satisfying the property that, for any $f \in \mathcal{C}_0$, there exists a sequence $f_n \in \mathcal{C}$ such that f_n converges to f boundedly. Then $\mathcal{C}$ is $\mathcal{N}$-separating.*

Proof. Let $F, G \in \mathcal{N}$ and assume that $\int_{\mathbb{R}} f(x)dF(x) = \int_{\mathbb{R}} f(x)dG(x)$ for all $f \in \mathcal{C}$. For $g \in \mathcal{C}_0$, let $g_n \in \mathcal{C}$ converge to g boundedly. By the Lebesgue dominant convergence theorem, we have

$$\int_{\mathbb{R}} g(x)dF(x) = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)dF(x) = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)dG(x) = \int_{\mathbb{R}} g(x)dG(x).$$

By Proposition 7.8, $F = G$. □

Exercise 7.6. Let $C_0^\infty(\mathbb{R})$ be the set of all infinitely differentiable functions on $\mathbb{R}$ with compact support. Prove that $C_0^\infty(\mathbb{R})$ is $\mathcal{N}$-separating. (*Hint*: Let ϕ be a function defined by

$$\phi(x) = \begin{cases} c\exp\left\{-\frac{1}{1-x^2}\right\} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

where c is a constant such that $\int_{-\infty}^{\infty} \phi(x)dx = 1$. For $\epsilon > 0$, set $\phi_\epsilon(x) = \epsilon^{-1}\phi(x\epsilon^{-1})$. Consider the convolution $f * \phi_\epsilon$ and show that, for any compactly supported function f, $f * \phi_\epsilon \in C_0^\infty(\mathbb{R})$ and converges to f uniformly as $\epsilon \to 0$.)

7.4. **Characteristic functions.** Let $\mathcal{C} = \{e^{inx} : n \in \mathbb{Z}\}$ and $C(S^1)$ be the set of all continuous functions f defined on $[0, 2\pi]$ satisfying $f(0) = f(2\pi)$. A trigonometric polynomial is a finite linear combination of functions in $\mathcal{C}$. The following theorem is a well-known result about the denseness of trigonometric polynomials in $C(S^1)$ under the sup-norm.

Theorem 7.10. *For $f \in C(S^1)$ and $\epsilon > 0$, there exists a trigonometric polynomial g such that*

$$\sup_{0\leq x\leq 2\pi} |f(x) - g(x)| < \epsilon.$$

Proof. See Rudin's "Real and complex analysis" (Chapter 4). □

Theorem 7.11. *The class $\{e^{iux} : u \in \mathbb{R}\}$ is $\mathcal{N}$-separating.*

Proof. Let $\mathcal{C}_0$ be the function in Proposition 7.8. We prove this theorem by the following two steps.

Step 1: Let $f \in \mathcal{C}_0$ and choose $N \in \mathbb{N}$ such that $f = 0$ on $[-N\pi, N\pi]^c$. By Theorem 7.10, we may select, for each $n \geq N$, a trigonometric polynomial $f_n(x) = \sum_k c_{n,k}e^{ikx}$ such that $\sup_{|x|\leq\pi} |f_n(x) - f(nx)| < 1/n$. By setting $g_n(x) = f_n(x/n) = \sum_k c_{n,k}e^{ikx/n}$, this is equivalent to $\sup_{|x|\leq n\pi} |g_n(x) - f(x)| < 1/n$, which implies $g_n \to f$ on $\mathbb{R}$. Furthermore, since f is bounded and g_n is periodic, g_n is uniformly bounded. This implies that any function in $\mathcal{C}_0$ is the limit of some boundedly convergent sequence of trigonometric polynomials.

Step 2: Let F, G be two distribution functions satisfying

$$\int_{\mathbb{R}} e^{iux}dF(x) = \int_{\mathbb{R}} e^{iux}dG(x), \quad \forall u \in \mathbb{R}. \tag{7.3}$$

Let $f \in \mathcal{C}_0$ and, by Step 1, let $(g_n)_{n=1}^\infty$ be a sequence of trigonometric polynomials converging to f boundedly. By (7.3), it is clear that $\int_{\mathbb{R}} g_n(x)dF(x) = \int_{\mathbb{R}} g_n(x)dG(x)$ for all $n \geq 1$. The LDCT then implies $\int_{\mathbb{R}} f(x)dF(x) = \int_{\mathbb{R}} f(x)dG(x)$. □

Definition 7.5. Let F be a distribution function and X be a random variable.

(1) The *characteristic function* of F is defined to be the following complex-valued function

$$f(u) = \int_{\mathbb{R}} e^{iux}dF(x).$$

(2) The characteristic functions of X is defined by

$$f(u) = \mathbb{E}[e^{iuX}].$$

Proposition 7.12. *Let F be a distribution function with characteristic function f. Then,*

(1) $f(0) = 1$,
(2) $|f(u)| \leq 1$,
(3) $f(-u) = \overline{f}(u)$,
(4) *f is uniformly continuous on $\mathbb{R}$,*

(5) *There exists a constant $A \in (0,\infty)$ independent of F such that*

$$F\left([-1/u,1/u]^c\right) \le \frac{A}{u}\int_0^u (1-\Re f(v))dv \quad \forall u>0$$

where $\Re a$ is the real part of a.

Proof. The proofs of (1)-(3) are clear. For (4), observe that, when $h \in \mathbb{R}$,

$$|f(u+h)-f(u)| = \left|\int_{\mathbb{R}} e^{iux}(e^{ihx}-1)dF(x)\right| \le \int_{\mathbb{R}} |e^{ihx}-1|dF(x).$$

This implies that, for $\delta > 0$,

$$\Delta(\delta) := \sup_{u,|h|<\delta} |f(u+h)-f(u)| \le \sup_{|h|\le\delta}\int_{\mathbb{R}} |e^{ihx}-1|dF(x)$$

Since $\int_{\mathbb{R}} |e^{ihx}-1|dF(x) \to 0$ when $|h| \to 0$, we have $\Delta(\delta)\to 0$ when $\delta \to 0$.

For (5), note that

$$\frac{1}{u}\int_0^u [1-\Re f(v)]dv = \frac{1}{u}\int_0^u \left(\int_{\mathbb{R}}[1-\cos(vx)]dF(x)\right)dv$$
$$= \int_{\mathbb{R}}\left(\frac{1}{u}\int_0^u[1-\cos(vx)]dv\right)dF(x) = \int_{\mathbb{R}}\left(1-\frac{\sin(ux)}{ux}\right)dF(x).$$

This yields

$$\frac{1}{u}\int_0^u [1-\Re f(v)]dv \ge \int_{|ux|>1}\left(1-\frac{\sin(ux)}{ux}\right)dF(x) \ge A^{-1}F\left([-1/u,1/u]^c\right),$$

where $A^{-1} := \inf\left\{1-\frac{\sin t}{t} : |t|\ge 1\right\} > 0$. □

The next theorem simplifies the assumption of mass-preservation in Proposition 7.7.

Theorem 7.13 (The continuity theorem). *For $n \ge 1$, let F_n be a distribution function and f_n be the corresponding characteristic function f_n. Then, there exists $F \in \mathcal{N}$ such that F_n converges in distribution to F if and only if f_n converges pointwise to some function f, which is continuous at 0. Furthermore, f is the characteristic function of F.*

Proof. By Proposition 7.7, since $\{e^{iux} : u \in \mathbb{R}\}$ is $\mathcal{N}$-separating, it remains to prove that if $f_n \to f$ and f is continuous at 0, then $(F_n)_{n=1}^\infty$ is mass-preserving. By Proposition 7.12(5), one may choose a universal constant $A > 0$ such that

$$F_n([-1/u,1/u]^c) \le \frac{A}{u}\int_0^u [1-\Re f_n(v)]dv$$

By the Lebesgue dominant convergence theorem, we have

$$\limsup_{n\to\infty} F_n([-1/u,1/u]^c) \le \frac{A}{u}\int_0^u [1-\Re f(v)]dv.$$

Since f is continuous at 0, the above inequality implies

$$\lim_{u\downarrow 0}\limsup_{n\to\infty} F_n([-1/u,1/u]^c) = 0.$$

This is equivalent to say that, for any $\epsilon > 0$, there is $u > 0$ and $N \in \mathbb{N}$ such that

$$F_n([-1/u,1/u]^c) \le \epsilon, \quad \forall n \ge N.$$

This implies $(F_n)_{n=1}^\infty$ is mass-preserving. □

Lemma 7.14. *Let $(F_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{N}$ with characteristic functions $(f_n)_{n=1}^{\infty}$. If F_n $\mathcal{D}$-converges to $F \in \mathcal{N}$, then, for any $\epsilon > 0$, there is $\delta > 0$ such that*

$$|f_n(u+h) - f_n(u)| < \epsilon, \quad \forall |h| < \delta, u \in \mathbb{R}, n \geq 1.$$

Proof. Since F_n converges to F in distribution, $(F_n)_{n=1}^{\infty}$ is mass-preserving. For $\epsilon > 0$, we choose a finite interval I such that $F_n(I^c) < \epsilon/3$ for all $n \geq 1$. This implies that, for $u, h \in \mathbb{R}$,

$$|f_n(u+h) - f_n(u)| \leq \int_{\mathbb{R}} |e^{ihx} - 1| dF_n(x) \leq \int_I |e^{ihx} - 1| dF_n(x) + 2F_n(I^c) \leq \Delta(h) + \frac{2\epsilon}{3},$$

where $\Delta(h) = \max_{x \in I} |e^{ihx} - 1|$. It is clear that $\Delta(h) \to 0$ as $|h| \to 0$. This implies that one may choose $\delta > 0$ such that $\Delta(h) < \epsilon/3$ when $|h| < \delta$. □

Theorem 7.15. *For $n \geq 1$, let $F_n \in \mathcal{N}$ and $f_n(u) = \int_{\mathbb{R}} e^{iux} dF_n(x)$. Then, there exists $F \in \mathcal{N}$ such that F_n $\mathcal{D}$-converges to F if and only if f_n converges uniformly to some function f in every finite interval.*

Proof. By Theorem 7.13, it remains to show that if F_n $\mathcal{D}$-converges to F, then f_n converges uniformly on any finite interval. Let I be a finite interval. By Lemma 7.14, we may choose, for each $\epsilon > 0$, a positive constant $\delta > 0$ such that

$$|f_n(u+h) - f_n(u)| < \epsilon/3, \quad \forall |h| < \delta, u \in \mathbb{R}, n \geq 1$$

and

$$|f(u+h) - f(u)| < \epsilon/3, \quad \forall |h| < \delta, u \in \mathbb{R}.$$

Let $u_1 < u_2 < \cdots < u_k$ be points in I such that $I \subset \bigcup_{i=1}^k (u_i - \delta, u_i + \delta)$. Since $f_n(u_i) \to f(u_i)$ for $1 \leq i \leq k$, we may choose $N \in \mathbb{N}$ such that

$$\max_{1 \leq i \leq k} |f_n(u_i) - f(u_i)| < \epsilon/3, \quad \forall n \geq N.$$

This implies that, for $u \in I$ with $|u - u_i| < \delta$,

$$|f_n(u) - f(u)| \leq |f_n(u) - f_n(u_i)| + |f_n(u_i) - f(u_i)| + |f(u_i) - f(u)| < \epsilon.$$

when $n \geq N$. □

Exercise 7.7. A random variable is said to have a symmetric distribution if $\mathbb{P}(X \in B) = \mathbb{P}(-X \in B)$ for all $B \in \mathcal{B}(\mathbb{R})$. Prove that the characteristic function of X is real if and only if X has a symmetric distribution.

Exercise 7.8. Show that any characteristic function f is non-negative definite, that is, for any $\lambda_1, ..., \lambda_n \in \mathbb{C}$ and $u_1, ..., u_n \in \mathbb{R}$,

$$\sum_{k,\ell=1}^{n} f(u_k - u_\ell) \lambda_k \overline{\lambda}_\ell \geq 0.$$

Exercise 7.9. For $n \geq 1$, let $b_n \in \mathbb{R}$ and $X_n = b_n$ a.s. with characteristic function f_n. Prove that the following are equivalent.

(1) $X_n \xrightarrow{\mathcal{D}} X$ for some $X \in \mathcal{N}$.
(2) b_n is Cauchy.
(3) There is $\epsilon > 0$ such that the sequence $(f_n(u))_{n=1}^{\infty}$ is Cauchy for all $|u| < \epsilon$.

Corollary 7.16. *Let X_n be random variables with characteristic function f_n. Then, $X_n \xrightarrow{\mathcal{D}} 0$ if and only if $f_n \to 1$ on $(-a, a)$ for some $a < \infty$.*

Proof. One direction is obvious from Theorem 7.13. For the other direction, assume that f_n converges to 1 on $(-a, a)$. Let F_n be the distribution of X_n. By Proposition 7.12 (5), there is $A < \infty$ such that

$$F_n([-1/a, 1/a]^c) \leq \frac{A}{a} \int_0^a (1 - \Re f_n(v))dv, \quad \forall n \geq 1.$$

Since $f_n \to 1$ on $(-a, a)$, the right hand side converges to 0 and, thus, F_n is mass-preserving.

Let $(F_{k_n})_{n=1}^\infty$ be a subsequence of $(F_n)_{n=1}^\infty$ and $(F_{n'_k})_{k=1}^\infty$ be a $\mathcal{D}$-convergent further subsequence with limit $F \in \mathcal{N}$. Let f be the characteristic function of F. Then, f is identically 1 on $(-a, a)$. This implies

$$F(\{2\ell\pi/u : \ell \in \mathbb{Z}\}) = 1, \quad \forall 0 < |u| < a.$$

Hence, $F(\{0\}) = 1$, as desired. □