Characteristic Functions and the Central Limit Theorem

Probability Theory, MATH 5451
Fall 2023

The goal of this note is to prove the Central Limit Theorem through the use of characteristic functions. Since we do not assume a general background in measure theory, we will only prove the needed statements for continuous and discrete distributions. The same results apply to more general probability measures on $\mathbb{R}$ as well.

1. Characteristic Functions

Definition 1.1. *The* **characteristic function** *of a random variable X is defined by,*

$$\phi_X(t) = E[e^{itX}] \quad \text{for all } t \in \mathbb{R},$$

where i denotes the imaginary unit.

Lemma 1.2. *For any random variable X, $\phi_X(t)$ exists, is bounded by 1 and continuous for all $t \in \mathbb{R}$.*

Proof. Since $e^{itx} = \cos(tx) + i\sin(tx)$, we have $|e^{itx}| \le 1$ so that $|\phi_X(t)| \le E[1]$ exists and is bounded by 1 for every $t \in \mathbb{R}$. Moreover, for $t, h \in \mathbb{R}$,

$$|\phi_X(t) - \phi_X(t+h)| = \left|E[e^{itX}(1 - e^{ihX})]\right| \le E\left[\left|1 - e^{ihX}\right|\right].$$

Since $|1-e^{ihx}| \to 0$ as $h \to 0$ for each $x \in \mathbb{R}$, and $|1-e^{ihx}| \le 2$, by the Bounded Convergence Theorem, we have $E[|1 - e^{ihX}|] \to 0$ as $h \to 0$, proving continuity. ☐

The first important fact about characteristic functions we will prove is that a characteristic function completely determines the distribution and vice versa. This is due to the following lemma.

Lemma 1.3 (Levy's Inversion Formula)**.** *Suppose that X is a random variable with characteristic function ϕ. Then for all $a < b$,*

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \phi(t)\, dt = P(a < X < b) + \frac{1}{2}(P(X = a) + P(X = b)).$$

Proof. Let's write the proof in the continuous case first and then discuss modifications necessary for the discrete case. Note that for a continuous distribution, $P(X = a) = P(X = b) = 0$. Now for any $T > 0$, letting f denote the p.d.f. of X,

$$\begin{aligned}
\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \phi(t)\, dt &= \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \int_{-\infty}^{\infty} e^{itx} f(x)\, dx\, dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it}\, dt\, dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \int_{-T}^{T} \frac{\cos t(x-a) - \cos t(x-b)}{it} + \frac{\sin t(x-a) - \sin t(x-b)}{t}\, dt\, dx
\end{aligned}$$

Now for t near 0, $\cos t(x-a) - \cos t(x-b) = \mathcal{O}(t^2)$ and $\sin t(x-a) - \sin t(x-b) = \mathcal{O}(t)$ so that both fractions remain bounded (and integrable) despite the t in the denominator.

Since $(\cos t(x-a) - \cos t(x-b))/t$ is an odd function of t and $(\sin t(x-a) - \sin t(x-b))/t$ is an even function of t, we have

$$\int_{-T}^{T} \frac{\cos t(x-a) - \cos t(x-b)}{it}\, dt = 0 \quad \text{and}$$

$$\int_{-T}^{T} \frac{\sin t(x-a) - \sin t(x-b)}{t}\, dt = 2\int_{0}^{T} \frac{\sin t(x-a) - \sin t(x-b)}{t}\, dt.$$

Now we take $T \to \infty$. For any $k \in \mathbb{R}$, (using that sine is odd) we have

$$\int_0^\infty \frac{\sin(kt)}{t}\,dt = \mathrm{sign}(k)\int_0^\infty \frac{\sin|k|t}{t}\,dt = \mathrm{sign}(k)\int_0^\infty \frac{\sin(u)}{u}\,du = \mathrm{sign}(k)\frac{\pi}{2}\,,$$

where we have made the substitution, $u = |k|t$, $du = |k|dt$.

Applying this result to $k = x - a$ and $k = x - b$ and substituting it into our inversion formula yields,

$$\begin{aligned} \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi(t)\,dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\,2\,[\mathrm{sign}(x-a)-\mathrm{sign}(x-b)]\,\frac{\pi}{2}\,dx \\ &= \int_{-\infty}^{\infty} f(x)\,\frac{1}{2}\,[\mathrm{sign}(x-a)-\mathrm{sign}(x-b)]\;dx\,. \end{aligned} \tag{1.1}$$

Now if $x < a$, then $\mathrm{sign}(x-a) - \mathrm{sign}(x-b) = -1 - (-1) = 0$. Similarly, if $x > b$, then $\mathrm{sign}(x-a) - \mathrm{sign}(x-b) = 1 - 1 = 0$. On the other hand, if $a < x < b$, then $\mathrm{sign}(x-a) - \mathrm{sign}(x-b) = 1 - (-1) = 2$. So for a continuous distribution, (1.1) implies

$$\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi(t)\,dt = \int_a^b f(x)\,dx = P(a < X < b)\,,$$

which is precisely the statement of the lemma.

On the other hand, suppose we have a distribution for which $P(X = a) > 0$. Since when $x = a$, $\mathrm{sign}(x - a)$ is not defined, we must reason a little differently. But here notice that if $x = a$, then $\sin(t(x - a)) = \sin(0) = 0$ so the first term simply vanishes. In this case, at $x = a$, we have only the contribution from $\frac{1}{2}[-\mathrm{sign}(a - b)] = \frac{1}{2}$ and we get $\frac{1}{2}P(X = a)$. A similar argument produces the $\frac{1}{2}P(X = b)$ term when $P(X = b) > 0$. □

Corollary 1.4. *Two random variables have the same cumulative distribution function if and only if they have the same characteristic function on* $\mathbb{R}$.

Proof. Clearly, from the definition of characteristic function, if $F_X = F_Y$ for two random variables X and Y, then $\phi_X = \phi_Y$. Conversely, if $\phi_X(t) = \phi_Y(t)$ for all $t \in \mathbb{R}$, then applying Lemma 1.3, for arbitrary $a < b$ and letting $a \to -\infty$, we have

$$P(X < b) + \frac{1}{2}P(X = b) = P(Y < b) + \frac{1}{2}P(Y = b) \implies F_X(b) - \frac{1}{2}P(X = b) = F_Y(b) - \frac{1}{2}P(Y = b)\,.$$

Now using the fact that distribution functions are right-continuous, we fix $b_0 \in \mathbb{R}$, and choose a sequence $b \downarrow b_0$ such that $P(X = b) = P(Y = b) = 0$ along the sequence. Then by continuity, the above equation implies

$$F_X(b_0) = \lim_{b\downarrow b_0} F_X(b) = \lim_{b\downarrow b_0} F_Y(b) = F_Y(b_0).$$

Since b_0 was arbitrary, this completes the proof of the corollary. □

2. Convergence in Distribution

Definition 2.1. *Given random variables X, X_n, $n \geq 1$, with distribution functions F_X and F_{X_n}, we say that X_n* **converges in distribution** *to X if*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

at each point of continuity x of F_X. We write $X_n \xrightarrow{d} X$.

Remark 2.2. *The requirement that convergence is required only at continuity points of F_X is needed to avoid situations like the following: Suppose $X_n = 1/n$ with probability 1 and $X = 0$ with probability 1. We would like to say that X_n converges in some sense to X, but clearly $F_X(0) = 1$ while $lim_{n\to\infty} F_{X_n}(0) = 0$. In this case, $x = 0$ is not a continuity point of F_X.*

Since distribution functions are nondecreasing, right continuous and bounded, it is easy to construct convergent subsequences.

Theorem 2.3 (Helly's Selection Theorem)**.** *Given a sequence of distribution functions F_n, there is a subsequence $(F_{n_k})_{k\geq 1}$ and a right continuous nondecreasing function F so that $\lim_{k\to\infty} F_{n_k}(x) = F(x)$ at all continuity points x of F.*

Proof. We start by enumerating the rationals $r_1, r_2, \ldots$. Since for each r_j, $\{F_n(r_j) : n \geq 1\}$ is bounded between 0 and 1, by the Bolzano-Weierstrass theorem, there exists a convergent subsequence $(n_{k(j)})_{k\geq 1}$ such that $\lim_{k\to\infty} F_{n_{k(j)}}(r_j) = b_j$ for some $b_j \in \mathbb{R}$. Using a diagonalization procedure, we may choose a single subsequence $(n_k)_{k\geq 1}$ such that $F_{n_k}(r_j) \to b_j$ for all $j \in \mathbb{N}$.

Define $F(r_j) = b_j$. F is nondecreasing and bounded between 0 and 1. Extend F to all real numbers by defining

$$F(x) = \inf\{F(r) : r \in \mathbb{Q}, r > x\} = \lim_{\substack{r\to x^+ \\ r\in\mathbb{Q}}} F(r)\,. \tag{2.1}$$

By construction, F is right continuous. It remains to check that F_{n_k} converges to F at all (irrational) points of continuity.

Let x be a point of continuity of F and let $r > x$ be rational. Then $F_{n_k}(x) \leq F_{n_k}(r)$ so

$$\limsup_{k\to\infty} F_{n_k}(x) \leq \lim_{k\to\infty} F_{n_k}(r) = F(r)\,.$$

Since this is true for every $r > x$, by the definition in (2.1) we have

$$\limsup_{k\to\infty} F_{n_k}(x) \leq F(x)\,. \tag{2.2}$$

Now for any $y < x$ we can find $r \in \mathbb{Q}$ such that $y < r < x$. Then using monotonicity of F_n yields,

$$\liminf_{k\to\infty} F_{n_k}(x) \geq \lim_{k\to\infty} F_{n_k}(r) = F(r) \geq F(y)\,.$$

Since this is true for each $y < x$, we have

$$\liminf_{k\to\infty} F_{n_k}(x) \geq \sup_{y<x} F(y) = \lim_{y\to x^-} F(y) = F(x)\,, \tag{2.3}$$

where we have used monotonicity and continuity of F at x. Combining (2.2) and (2.3) implies that the limit of $F_{n_k}(x)$ exists and equals $F(x)$. □

Helly's Selection Theorem does not guarantee that the limiting F is a distribution function. For example, let $F_n(x) = 0$ if $x < n$ and $F_n(x) = 1$ if $x \geq n$. Then $(F_n)_{n\geq 1}$ is a sequence of distribution functions and $\lim_{n\to\infty} F_n(x) = 0$ for all $x \in \mathbb{R}$, but $F \equiv 0$ is not a distribution function.

Yet combining this theorem with what we know about characteristic functions allows us to prove the following connection to convergence in distribution.

Theorem 2.4. *Given random variables X, X_n, $n \geq 1$ with distribution functions F_X, F_{X_n}, and characteristic functions ϕ_X, ϕ_{X_n}, the following are equivalent:*

1) $X_n \xrightarrow{d} X$, *i.e.,* X_n *converges to* X *in distribution;*
2) *for any bounded continuous function* g, $\lim_{n\to\infty} \int_{-\infty}^{\infty} g(x)\,dF_n(x) = \int_{-\infty}^{\infty} g(x)\,dF(x)$*;*
3) $\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t)$ *for all* $t \in \mathbb{R}$.

Proof. (1) $\implies$ (2): Let $\varepsilon > 0$ be arbitrary. Since F_X is a nondecreasing function with $F(-\infty) = 0$ and $F(\infty) = 1$, we can choose $a < b$ such that a and b are continuity points of F_X, $F_X(a) < \varepsilon$ and $F_X(b) > 1 - \varepsilon$. Then since F_{X_n} converges to F_X at every point of continuity by assumption, we can find $N \in \mathbb{N}$ such that for all $n \geq N$, $F_{X_n}(a) < 2\varepsilon$ and $F_{X_n}(b) > 1 - 2\varepsilon$.

Now let g be a continuous function on $\mathbb{R}$. Then g is uniformly continuous on the finite interval $[a, b]$. Fix $\delta > 0$ and choose a partition $a < a_1 < \cdots < a_{M+1} = b$ such that each a_j is a continuity point of F_X and on each interval $I_j = [a_j, a_{j+1})$, we have $|g(x) - g(a_j)| < \delta$. Define $\psi = \sum_{j=1}^{M} g(a_j)\chi_{I_j}$, where χ_{I_j} denotes the indicator function of I_j, i.e. ψ is constant and equal to $g(a_j)$ on each I_j.

Now if g is bounded by L, then

$$\left|\int_{\mathbb{R}} g(x)dF_{X_n}(x) - \sum_{j=1}^{M} g(a_j)[F_{X_n}(a_j) - F_{X_n}(a_{j+1})]\right|$$
$$\leq \left|\int_a^b g(x)dF_{X_n}(x) - \int_{-\infty}^{\infty} \psi(x)dF_{X_n}(x)\right| + LF_{X_n}(a) + L(1 - F_{X_n}(b)) \leq \delta + 4L\varepsilon\,.$$

Similarly,

$$\left|\int_{\mathbb{R}} g(x)dF_X(x) - \sum_{j=1}^{M} g(a_j)[F_X(a_j) - F_X(a_{j+1})]\right| \leq \delta + 2L\varepsilon\,.$$

Since a_j is a continuity point for F_X, $\lim_{n\to\infty} F_{X_n}(a_j) = F_X(a_j)$ for each j. Thus,

$$\lim_{n\to\infty}\left|\int_{\mathbb{R}} g(x)dF_{X_n}(x) - \int_{\mathbb{R}} g(x)dF_X(x)\right| \leq 2\delta + 6L\varepsilon\,,$$

and since δ and ε were arbitrary, the proof is complete.

(2) $\implies$ *(3):* This is immediate since e^{itx} is a bounded, continuous function of x for each $t \in \mathbb{R}$.

(3) $\implies$ *(1):* Using Theorem 2.3, we choose a subsequence F_{n_k} which converges to a nondecreasing, right continuous function G at every point of continuity of G. We will first show that G is a distribution function, using the continuity of ϕ_X at $t = 0$.

This will use the following claim: For any characteristic function ϕ of a distribution F, and all $T > 0$,

$$1 - F(\tfrac{2}{T}) + F(-\tfrac{2}{T}) \leq 2\left(1 - \frac{1}{2T}\int_{-T}^{T}\phi(t)\,dt\right)\,. \tag{2.4}$$

To prove this, we argue as in the proof of Lemma 1.3, using that cosine is even and sine is odd,

$$\frac{1}{2T}\int_{-T}^{T}\phi(t)\,dt = \int_{\mathbb{R}}\frac{1}{2T}\int_{-T}^{T} e^{itx}\,dt\,f(x)dx = \int_{\mathbb{R}}\frac{1}{T}\int_0^T \cos(tx)\,dt\,f(x)dx$$
$$= \int_{\mathbb{R}}\frac{\sin(Tx)}{Tx}f(x)\,dx \leq \int_{|x|<\frac{2}{T}}\left|\frac{\sin(Tx)}{Tx}\right| f(x)\,dx + \int_{|x|\geq\frac{2}{T}}\left|\frac{\sin(Tx)}{Tx}\right| f(x)\,dx$$
$$\leq P(|X| < \tfrac{2}{T}) + \frac{1}{2}P(|X| \geq \tfrac{2}{T})\,,$$

where in the last line we have used $|\frac{\sin(Tx)}{Tx}| \leq 1$ for the first term and $|\sin(Tx)| \leq 1$ for the second. Using this, we write,

$$1 - \frac{1}{2T}\int_{-T}^{T}\phi(t)\,dt \geq 1 - P(|X| < \tfrac{2}{T}) - \frac{1}{2}P(|X| \geq \tfrac{2}{T})$$
$$= \frac{1}{2}P(|X| \geq \tfrac{2}{T}) \geq \frac{1}{2}[1 - F(\tfrac{2}{T}) + F(-\tfrac{2}{T})]\,,$$

which implies (2.4).

Now we apply (2.4) to each distribution F_{n_k} and its characterictic function $\phi_{X_{n_k}}$. Choosing T such that $\pm 2/T$ are points of continuity of G, we take the limit as $k \to \infty$. Then $F_{X_{n_k}}(\pm 2/T)$ converges to $G(\pm 2/T)$ while $\phi_{X_{n_k}}$ converges to ϕ_X on the bounded interval $[-T, T]$. This yields,

$$1 - G(\tfrac{2}{T}) + G(-\tfrac{2}{T}) \leq 2\left(1 - \frac{1}{2T}\int_{-T}^{T}\phi_X(t)\,dt\right)\,.$$

Since ϕ_X is continuous at $t = 0$, if we take the limit $T \to 0$, $\frac{1}{2T}\int_{-T}^{T}\phi_X(t)\,dt \to \phi_X(0) = 1$. This implies,

$$1 - G(\infty) + G(-\infty) \leq 0\,,$$

which forces $G(\infty) = 1$ and $G(-\infty) = 0$, i.e. G is a distribution.

Since $F_{X_{n_k}}$ converges to G, by (1) $\implies$ (3) of the current theorem, $\phi_{X_{n_k}}$ converges to ψ, the characteristic function of G. But by assumption, $\phi_{X_{n_k}}$ converges to ϕ_X. Thus in fact $\psi = \phi_X$ and by Corollary 1.4, it must be that $G = F_X$, i.e. $(F_{X_{n_k}})_{k\geq 1}$ converges in distribution to F_X. Since this holds for any limit point of the sequence $(F_{X_n})_{n\geq 1}$ and by the same argument, every subsequence has a convergent subsequence which converges to F_X, we conclude that $(F_{X_n})_n$ converges to F_X at every point of continuity of F_X, as required. □

Remark 2.5. *During the course of the proof of (3) $\implies$ (1) in Theorem 2.4, we actually proved a more general result: Suppose ϕ_n are characteristic functions of distributions F_n. If for each $t \in \mathbb{R}$, $\phi_n(t)$ converges to a function $\phi(t)$ and ϕ is continuous at $t = 0$, then there exists a distribution G with ϕ as its characteristic function and F_n converges in distribution to G.*

3. Central Limit Theorem

We shall prove the Central Limit Theorem by showing that the characteristic function of $\frac{1}{\sqrt{n}}S_n$, where $S_n = \sum_{j=1}^n X_j$ and X_j are independent and identically distributed, converges to the characteristic function of a normal random variable. Before proving the main theorem, we obtain bounds on the error terms that will appear in the proof.

Lemma 3.1. *For all $t, x \in \mathbb{R}$,*

$$e^{itx} = 1 + ixt - \tfrac{1}{2}x^2t^2 + t^2 \min(\tfrac{tx^3}{6}, x^2)\,.$$

Proof. The lemma is a simple consequence of the Taylor expansion for e^θ about $\theta = 0$, with error term. We derive the error term as follows. Suppose a function f is 3 times differentiable. By the Fundamental Theorem of Calculus,

$$f(t) - f(0) = \int_0^t f'(s)\,ds \qquad \implies \qquad f(t) = f(0) + \int_0^t f'(s)\,ds\,. \tag{3.1}$$

Now we integrate by parts using:

$$\begin{aligned} u &= f'(s) & dv &= ds \\ du &= f''(s)ds & v &= -(t-s) \end{aligned}$$

so that

$$\int_0^t f'(s)ds = -f'(s)(t-s)\Big|_0^t + \int_0^t (t-s)f''(s)ds = f'(0)t + \int_0^t (t-s)f''(s)ds \tag{3.2}$$

We continue the expansion by integrating $\int_0^t (t-s)f''(s)ds$ by parts again. Let

$$\begin{aligned} u &= f''(s) & dv &= (t-s)ds \\ du &= f^{(3)}(s)ds & v &= \tfrac{-(t-s)^2}{2} \end{aligned}$$

Then

$$\int_0^t f''(s)(t-s)ds = -f''(s)\tfrac{(t-s)^2}{2}\Big|_0^t + \int_0^t \tfrac{(t-s)^2}{2} f^{(3)}(s)ds = \tfrac{f''(0)}{2}t^2 + \tfrac{1}{2}\int_0^t (t-s)^2 f^{(3)}(s)ds$$

Substituting this and equation (3.2) into (3.1), we obtain

$$f(t) = f(0) + f'(0)t + \tfrac{f''(0)}{2}t^2 + \tfrac{1}{2}\int_0^t (t-s)^2 f^{(3)}(s)\,ds\,. \tag{3.3}$$

If we set $f(t) = e^{itx}$, equation (3.3) implies the lemma once we show the error term has the proper form. On the one hand, notice that by the triangle inequality,

$$\left|\tfrac{1}{2}\int_0^t (t-s)^2 f^{(3)}(s)ds\right| = \left|-\tfrac{i}{2}\int_0^t (t-s)^2 x^3 e^{ixt}\,ds\right| \leq \tfrac{|x|^3}{2}\int_0^t (t-s)^2\,ds = \tfrac{t^3|x|^3}{6}\,. \tag{3.4}$$

On the other hand, integrating this integral by parts, setting,

$$\begin{array}{rclcrcl} u &=& \frac{1}{2}(t-s)^2 & \qquad & dv &=& f^{(3)}(s)ds \\ du &=& -(t-s)ds & \qquad & v &=& f''(s) \end{array}$$

yields,

$$\tfrac{1}{2}\int_0^t (t-s)^2 f^{(3)}(s)ds = \tfrac{1}{2}(t-s)^2 f''(s)|_0^t + \int_0^t (t-s)f''(s)ds = \tfrac{1}{2}t^2x^2 - \int_0^t (t-s)x^2 e^{ixs}ds\,.$$

We again use the triangle inequality to arrive at the upper bound,

$$\left|\tfrac{1}{2}\int_0^t (t-s)^2 f^{(3)}(s)ds\right| \le \tfrac{1}{2}t^2x^2 + x^2\int_0^t (t-s)ds \le t^2x^2\,. \tag{3.5}$$

Since the error term must satisfy the bound in both (3.4) and (3.5), it must be bounded by the minimum of these two, proving the lemma. □

Corollary 3.2. *If $E[X^2] < \infty$, then*

$$\phi_X(t) = 1 + itE[X] - \tfrac{1}{2}t^2E[X^2] + o(t^2)\,,$$

where $\lim_{t\to 0}\frac{o(t^2)}{t^2} = 0$.

Proof. The first 3 terms follow immediately by integrating the first 3 terms of the expansion for e^{itx} in Lemma 3.1. The error term is then $t^2E[\min(\frac{tX^3}{6}, X^2)]$. We must show that $E[\min(\frac{tX^3}{6}, X^2)] \to 0$ as $t \to 0$. This follows from the Dominated Convergence Theorem, since the function $g(x) = \min(\frac{tx^3}{6}, x^2)$ converges to 0 pointwise (that is, for each fixed x as $t \to 0$) and is dominated by x^2, which has finite integral by assumption. □

We are finally ready to state and prove the Central Limit Theorem.

Theorem 3.3 (Central Limit Theorem)**.** *Suppose $(X_j)_{j\ge 1}$ is a sequence of independent, identically distributed random variables with $E[X_j] = \mu$ and $Var(X_j) = \sigma^2 \in (0,\infty)$. Define $S_n = \sum_{j=1}^n X_j$. Then as $n \to \infty$,*

$$\frac{1}{\sqrt{n}}(S_n - n\mu) \xrightarrow{d} \mathcal{N}(0,\sigma^2)\,,$$

where $\mathcal{N}(0,\sigma^2)$ is the normal distribution with mean 0 and variance σ^2.

Equivalently, $\dfrac{1}{\sigma\sqrt{n}}(S_n - n\mu) \xrightarrow{d} \mathcal{N}(0,1)$.

Proof. Let X_j be as in the statement of the theorem. Without loss of generality, by replacing X_j with $X_j - \mu$, we may assume that $\mu = 0$, and so $\mathrm{Var}(X_j) = E[X_j^2] = \sigma^2$. We compute the characteristic function of $\frac{1}{\sqrt{n}}S_n$, using the fact that the X_j are independent.

$$E[e^{i\frac{t}{\sqrt{n}}S_n}] = E[e^{i\frac{t}{\sqrt{n}}\sum_{j=1}^n X_j}] = E\left[\prod_{j=1}^n e^{i\frac{t}{\sqrt{n}}X_j}\right] = \prod_{j=1}^n E[e^{i\frac{t}{\sqrt{n}}X_j}] = \left(E[e^{i\frac{t}{\sqrt{n}}X_j}]\right)^n\,.$$

Now we apply Corollary 3.2 to $E[e^{i\frac{t}{\sqrt{n}}X_j}] = \phi_{X_j}\left(\frac{t}{\sqrt{n}}\right)$, recalling that $E[X_j] = 0$ and $E[X_j^2] = \sigma^2$.

$$E[e^{i\frac{t}{\sqrt{n}}S_n}] = \left(1 - \frac{t^2\sigma^2}{2n} + o\Big(\frac{t^2}{n}\Big)\right)^n \xrightarrow[n\to\infty]{} e^{-\frac{t^2\sigma^2}{2}}\,.$$

We recognize $e^{-\frac{t^2\sigma^2}{2}}$ as the characteristic function of $\mathcal{N}(0,\sigma^2)$. So Theorem 2.4 implies that $\frac{1}{\sqrt{n}}S_n$ converges in distribution to $\mathcal{N}(0,\sigma^2)$. □