Chapter 8

Central limit theorems

8.1 Characteristic functions

We will start with the definition of a characteristic function. To this end recall that by a complex number x we mean an ordered pair of real numbers. The set of all complex numbers is denoted by $\mathbb{C}$. Thus, if $x = (x_1, x_2)$ is a complex number, we say that x_1 is the real part of x and x_2 is the imaginary part of x. If $x, y \in \mathbb{C}$ we define $x + y = (x_1 + y_1, x_2 + y_2)$ and $xy = (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1)$. We write $x = y$ if, and only if, $x_1 = y_1$ and $x_2 = y_2$. The complex number $(0, 1)$ is denoted by i and is called the imaginary unit. Given the definition of product of complex numbers, $i^2 = -1$ (or $i^2 = (-1, 0)$).

Every complex number x can be written as $x = x_1 + ix_2$. To see this, let $x_1 = (x_1, 0)$ and $x_2 = (x_2, 0)$. Then, $ix_2 = (0, 1)(x_2, 0) = (0, x_2)$ and $x_1 + ix_2 = (x_1, 0) + (0, x_2) = (x_1, x_2) = x$. The complex number $\bar{x} = x_1 - ix_2$ is called the complex conjugate of x and $x\bar{x} = (x_1^2 + x_2^2, 0)$. The "absolute value" of a complex number is defined by $|x| = (x_1^2 + x_2^2)^{1/2}$ and if $x \neq (0, 0)$ then $x^{-1} = (x_1/(x_1^2 + x_2^2), -x_2/(x_1^2 + x_2^2))$ so that $x^{-1}x = (1, 0)$.

If $x = x_1 + ix_2$ we define $e^x = e^{x_1 + ix_2} := e^{x_1}(\cos(x_2) + i\,\sin(x_2))$ (Euler's formula). This definition gives the following desirable properties of complex exponentials,

$$e^x e^y = e^{x+y},\ e^x \neq 0, |e^{ix_2}| = |\cos(x_2) + i\,\sin(x_2)| = (\cos(x_2)^2 + \sin(x_2)^2)^{1/2} = 1.$$

If $X_1,\ X_2 : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ are random variables, we say that $X = X_1 + iX_2$ is a complex valued random variable and its distribution F_X is defined as usual in terms of the joint distribution of X_1 and X_2, i.e.,

$$F_X(x_1, x_2) = P(\{\omega : X_1(\omega) \leq x_1\} \cap \{\omega : X_2(\omega) \leq x_2\}) = P_X((-\infty, x_1] \times (-\infty, x_2]).$$

Since, $|X| = (X_1^2 + X_2^2)^{1/2}$ we have that $E(|X|^2) = E(X_1^2) + E(X_2^2)$. Thus, if $X_1,\ X_2 \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ then $E(|X|^2) < \infty$. Also, we naturally write $E(X) = E(X_1) + iE(X_2)$.

Note that algebraically $|X|$ is the Euclidean norm for vectors in $\mathbb{R}^2$ and, therefore, it is a convex function. By Jensen's Inequality, for any Borel measurable convex function g and integrable random variable Z we have that $g(E(Z)) \leq E(g(Z))$. Consequently, $|E(X)| \leq E(|X|)$.

Definition 8.1. *The characteristic function of a random variable* $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ *with distribution* F_X *is the complex valued function*

$$\phi_X(t) := E(e^{itX}) \text{ for } t \in \mathbb{R}.$$

Remark 8.1. *1. By definition (or Euler's formula)* $e^{itx} = \cos(tx) + i\sin(tx)$. *Hence,*

$$\begin{aligned}\phi_X(t) = E(\cos(tX) + i\sin(tX)) &= \int_\Omega \cos(tX)dP + i\int_\Omega \sin(tX)dP \\ &= \int_{\mathbb{R}} \cos(tX)dP_X + i\int_{\mathbb{R}} \sin(tX)dP_X \\ &= \int_{\mathbb{R}} \cos(tx)dF_X(x) + i\int_{\mathbb{R}} \sin(tx)dF_X(x).\end{aligned}$$

2. $|\phi_X(t)| = |E(e^{itX})| \leq E(|e^{itX}|) = E(|\cos(tX) + i\sin(tX)|) = E\left((\cos^2(tX) + \sin^2(tX))^{1/2}\right) = 1$. *Hence,* $E(e^{itX})$ *always exists and* $\phi_X(0) = 1$.

3. *Now, for* $h \in \mathbb{R}$

$$\begin{aligned}
|\phi_X(t+h) - \phi_X(t)| &= |E\left(e^{i(t+h)X}\right) - E\left(e^{itX}\right)| = |E(e^{itX+ihX} - e^{itX})| \\
&= |E(e^{itX}(e^{ihX} - 1))| \\
&\leq E\left(|e^{itX}||e^{ihX} - 1|\right) \\
&\leq E\left(|e^{ihX} - 1|\right) = \int_{\mathbb{R}} |e^{ihX} - 1| dP_X.
\end{aligned}$$

Now, $e^{ihx} - 1 = \cos(hx) - 1 + i\sin(hx)$ *and*

$$|e^{ihx} - 1| = \left((\cos(hx) - 1)^2 + \sin^2(hx)\right)^{1/2} = (2(1 - \cos(hx)))^{1/2} \leq 2.$$

Hence, as $|h| \to 0$, $|e^{ihx} - 1| \to 0$. *Consequently, by Lebesgue's Dominated Convergence Theorem,* $\int_{\mathbb{R}} |e^{ihX} - 1| dP_X \to 0$ *as* $|h| \to 0$. *Thus,* $\phi_X(t)$ *is uniformly (the bound is independent of* t*) continuous.*

4. *Let* $Y = \frac{X-\mu}{\sigma}$, *for* $\mu \in \mathbb{R}$ *and* $\sigma > 0$. *Then,*

$$\begin{aligned}
\phi_Y(t) &= E(e^{itY}) = E(e^{it(\frac{X-\mu}{\sigma})}) = E(e^{\frac{-it\mu}{\sigma}} e^{\frac{itX}{\sigma}}) \\
&= e^{-\frac{it\mu}{\sigma}} E(e^{\frac{itX}{\sigma}}) = e^{-\frac{it\mu}{\sigma}} \phi_X\left(\frac{t}{\sigma}\right).
\end{aligned}$$

5. *The characteristic function of* $-X$ *is* $\phi_{-X}(t) = E(e^{i(-t)X}) = \phi_X(-t)$.

$$\begin{aligned}
\phi_X(-t) &= \int_{\mathbb{R}} \cos(-tX) dP_X + i \int_{\mathbb{R}} \sin(-tX) dP_X \\
&= \int_{\mathbb{R}} \cos(tX) dP_X - i \int_{\mathbb{R}} \sin(tX) dP_X, \ \text{because } \cos(x) \text{ is even and } \sin(x) \text{ is odd.} \\
&= \bar{\phi}_X(t), \ \text{the complex conjugate of } \phi_X(t).
\end{aligned}$$

Since the imaginary part of a complex number x *is* $(x - \bar{x})/2$ *and* $\phi_X(t) - \bar{\phi}_X(t) = i\,2\int_{\mathbb{R}} \sin(tX) dP_X$, $\phi_X(t)$ *is real valued if, and only if,* $\int_{\mathbb{R}} \sin(tX) dP_X = 0$. *In this case,* $-X$ *and* X *have the same characteristic function.*

6. *If there exists a density f_X associated with P_X, e.g.,*

$$F_X(x) = \int\limits_{(-\infty,x]} f_X d\lambda$$

such that f_X is even, then

$$\begin{aligned}
\phi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x)dx &= \int_{-\infty}^{0} e^{itx} f_X(x)dx + \int_{0}^{\infty} e^{itx} f_X(x)dx \\
&\text{changing variables in the first integral by setting } -y = x, \\
&= \int_{\infty}^{0} e^{-ity} f_X(-y)(-1)dy + \int_{0}^{\infty} e^{itx} f_X(x)dx \\
&= \int_{0}^{\infty} e^{-itx} f_X(x)dx + \int_{0}^{\infty} e^{itx} f_X(x)dx \\
&= \int_{0}^{\infty} (e^{-itx} + e^{itx}) f_X(x)dx \\
&= \int_{0}^{\infty} (\cos(tx) - i\sin(tx) + \cos(tx) + i\sin(tx)) f_X(x)dx \\
&= 2\int_{0}^{\infty} \cos(tx) f_X(x)dx.
\end{aligned}$$

Hence, symmetric densities give real-valued characteristic functions.

7. *If X and Y are independent, then $\phi_{X+Y}(t)$ is $E(e^{it(X+Y)}) = E(e^{itX})E(e^{itY}) = \phi_X(t)\phi_Y(t)$.*

8. *Let $\{X_j\}_{j=1,2,\dots,n}$ be a sequence of IID random variables and $S_n = \sum_{j=1}^{n} X_j$.*

$$E(e^{itS_n}) = \prod_{j=1}^{n} E(e^{itX_j}) = \left(\phi_{X_1}(t)\right)^n.$$

Theorem 8.1. *Let $\phi_X(t)$ be a characteristic function. If $E(|X|^s) < \infty$ for $s = 1, 2, \cdots$*

$$\frac{d^s}{dt^s}\phi_X(t) = \int_{\mathbb{R}} (iX)^s e^{itX} dP_X = E((iX)^s e^{itX}).$$

Proof. For $h \neq 0$ consider

$$\begin{aligned}
\frac{\phi_X(t+h) - \phi_X(t)}{h} &= \frac{1}{h}\left(E(e^{i(t+h)X}) - E(e^{itX})\right) \\
&= \frac{1}{h}\left(\int_{\mathbb{R}} e^{i(t+h)X} dP_X - \int_{\mathbb{R}} e^{itX} dP_X\right) \\
&= \int_{\mathbb{R}} \frac{e^{i(t+h)X} - e^{itX}}{h} dP_X.
\end{aligned}$$

Then, for $x \neq 0$

$$\frac{e^{i(t+h)x} - e^{itx}}{h} = x\frac{\cos(x(t+h)) - \cos(tx)}{hx} + ix\frac{\sin(x(t+h)) - \sin(tx)}{hx}.$$

Taking limits on both sides as $h \to 0$ we have that

$$\frac{d}{dt}e^{itx} = \lim_{h\to 0}\frac{e^{i(t+h)x} - e^{itx}}{h} = -x\sin(tx) + ix\cos(tx) = ix(\cos(tx) + i\sin(tx)) = ixe^{itx}.$$

In addition, $|ixe^{itx}| = (x^2\sin^2(tx) + x^2\cos^2(tx))^{1/2} = |x|$. Hence, if $\int_{\mathbb{R}} |X| dP_X < \infty$ we have by Theorem 3.15

$$\frac{d}{dt}\phi_X(t) = \int_{\mathbb{R}} (iX)e^{itX}dP_X = E((iX)e^{itX}).$$

For $s = 2, 3, \cdots$ use the same argument with integrands $(ix)^{s-1}e^{itx}$. ■

An immediate consequence of this theorem is that $\frac{d^s}{dt^s}\phi_X(0) = i^s E(X^s)$.

Theorem 8.2. *For $x \in \mathbb{R}$ we have*

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \min\left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}.$$

Proof. Note that for $n \geq 0$, $x > 0$ and integration by parts (Riemann-Stieltjes integrals)

$$\begin{aligned}
\int_0^x e^{is}(x-s)^n ds &= \int_0^x e^{is} d\left(-\frac{(x-s)^{n+1}}{n+1}\right) \\
&= -e^{is}\frac{(x-s)^{n+1}}{n+1}\Big|_0^x - \int_0^x \left(-\frac{(x-s)^{n+1}}{(n+1)}\right) de^{is} \\
&= \frac{x^{n+1}}{n+1} + i\int_0^x \frac{(x-s)^{n+1}}{n+1}e^{is}ds. \qquad (8.1)
\end{aligned}$$

For $n = 0$, $\int_0^x e^{is}ds = x + i\int_0^x (x-s)e^{is}ds$. By Taylor's Theorem with remainder in Cauchy form at $x = 0$

$$\begin{aligned}
e^{ix} &= 1 + ix + i^2\int_0^x (x-s)e^{is}ds \\
&= 1 + ix + i^2\left(\frac{x^2}{2!} + \frac{i}{2!}\int_0^x (x-s)^2 e^{is}ds\right) \text{ using equation (8.1) with } n = 1.
\end{aligned}$$

Repeated substitution of the integral inside the parenthesis gives

$$e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \cdots + \frac{(ix)^n}{n!} + \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds$$
$$= \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds. \tag{8.2}$$

Hence,

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| = \left| \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds \right|.$$

But,

$$\left| \int_0^x (x-s)^n e^{is} ds \right| \le \int_0^x (x-s)^n |e^{is}| ds = \int_0^x (x-s)^n ds = -\frac{(x-s)^{n+1}}{n+1}\Big|_0^x = \frac{x^{n+1}}{n+1}.$$

Thus,

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \le \frac{|i^{n+1}|}{n!}\frac{x^{n+1}}{(n+1)} = \frac{x^{n+1}}{(n+1)!}.$$

Now, from equation (8.1)

$$\int_0^x e^{is}(x-s)^{n-1} ds - \frac{x^n}{n} = \frac{i}{n}\int_0^x (x-s)^n e^{is} ds.$$

Multiplying by $\frac{i^n}{(n-1)!}$, we get

$$\frac{i^n}{(n-1)!}\int_0^x e^{is}(x-s)^{n-1} ds - \frac{(ix)^n}{n!} = \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds.$$

Hence, using equation (8.2)

$$\frac{i^n}{(n-1)!}\int_0^x e^{is}(x-s)^{n-1} ds - \frac{(ix)^n}{n!} = e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!}$$

and consequently

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \le \frac{|i^n|}{(n-1)!}\frac{x^n}{n} + \frac{x^n}{n!} = 2\frac{x^n}{n!}.$$

Hence, combining the two bounds we have

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \le \min\left\{ \frac{x^{n+1}}{(n+1)!}, 2\frac{x^n}{n!} \right\}.$$

A similar argument applies for $x < 0$ to give,

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \min \left\{ \frac{|x|^{n+1}}{(n+1)!}, 2\frac{|x|^n}{n!} \right\}.$$

If $x = 0$ the two sides of the weak inequality coincide. ■

Remark 8.2. *1. Suppose X is a random variable such that $E(|X|^k) < \infty$ for $k = 1, 2, \ldots, n$. Then,*

$$\left| \phi_X(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} E(X^k) \right| = \left| E(e^{itX}) - E\left(\sum_{k=0}^{n} \frac{(it)^k}{k!} X^k \right) \right| \leq E\left(\left| e^{itX} - \sum_{k=0}^{n} \frac{(itX)^k}{k!} \right| \right)$$
$$\leq E\left(\min \left\{ \frac{2|tX|^n}{n!}, \frac{|tX|^{n+1}}{(n+1)!} \right\} \right).$$

Note that,

$$E\left(\min \left\{ \frac{2|tX(\omega)|^n}{n!}, \frac{|tX(\omega)|^{n+1}}{(n+1)!} \right\} \right) \leq \int_{\Omega} \frac{2|t|^n}{n!} |X(\omega)|^n dP = 2\frac{|t|^n}{n!} E(|X|^n).$$

Hence, in this context there is no need to assume that $E(|X|^{n+1})$ exists, only $E(|X|^n)$.

2. In the case where $E(|X|^n)$ exist and, if for all t

$$\lim_{n\to\infty} \frac{|t|^n E(|X|^n)}{n!} = 0$$

we have $\phi_X(t) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} E(X^k)$.

3. Different bounds can be obtained for the $E\left(\min \left\{ \frac{2|tX(\omega)|^n}{n!}, \frac{|tX(\omega)|^{n+1}}{(n+1)!} \right\} \right) = E\left(\min\left\{g(\omega), h(\omega)\right\}\right)$. In particular, for any $\epsilon > 0$ and $A = \{\omega : |X(\omega)| > \epsilon\}$

$$E\left(\min \left\{ \frac{2|tX(\omega)|^n}{n!}, \frac{|tX(\omega)|^{n+1}}{(n+1)!} \right\} \right) \leq \int g(\omega) I_A dP + \int h(\omega) I_{A^c} dP$$
$$\leq 2\frac{|t|^n}{n!} \int_A |X(\omega)|^n dP + \frac{|t|^{n+1}}{(n+1)!} \epsilon \int |X(\omega)|^n dP$$

or

$$E\left(\min \left\{ \frac{2|tX(\omega)|^n}{n!}, \frac{|tX(\omega)|^{n+1}}{(n+1)!} \right\} \right) = \int g(\omega) I_A dP + \int h(\omega) I_{A^c} dP$$
$$\leq 2\frac{|t|^n}{n!} \int_A |X(\omega)|^n dP + \frac{|t|^{n+1}}{(n+1)!} \epsilon^{n+1}$$

3. If $X \sim N(\mu, \sigma^2)$ then $E(e^{itX}) = e^{i\mu t - \frac{\sigma^2}{2} t^2}$.

The characteristic function for a random vector $X \in \mathbb{R}^d$ is defined as follows.

Definition 8.2. *The characteristic function of a random vector* $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}^d, \mathcal{B}^d)$ *with distribution* F_X *is the complex valued function*

$$\phi_X(t) := E(e^{it^T X}) \text{ for } t \in \mathbb{R}^d.$$

Remark 8.3. *If* $X \sim N(\mu, \Sigma)$ *where* $\mu \in \mathbb{R}^d$ *and* Σ *is a* $d \times d$ *matrix, the characteristic function* $\phi_X(t)$ *is given by* $E(e^{it^T X}) = e^{it^T\mu - \frac{1}{2}t^T\Sigma t}$.

It follows directly from the definition of a characteristic function ϕ that if $F = G$ where F and G are distribution functions, then ϕ associated with F is identical to the ϕ associated with G. That is, if two distributions coincide, so do their characteristic functions. The next theorem establishes that if two characteristic functions are the same they are associated with the same distribution function.

Theorem 8.3. *Let F and G be two distributions with the same characteristic function. That is,*

$$\int_{\mathbb{R}} e^{itx} dF(x) = \int_{\mathbb{R}} e^{itx} dG(x) \text{ for all } t \in \mathbb{R}.$$

Then, $F = G$.

Proof. Let $F(x) - G(x) = D(x)$. We need to show that

$$\int_{\mathbb{R}} e^{itx} dD(x) = 0 \text{ for all } t \in \mathbb{R} \tag{8.3}$$

implies $D(x) = 0$. We first note that $D(x)$ is the difference between two distributions functions, i.e., two bounded monotone increasing functions. Hence, $D(x)$ is of bounded variation on $\mathbb{R}$.[1]. Now, equation (8.3) holds for any trigonometric polynomial

$$T(x) = \sum_{v=-n}^{n} a_v e^{i(\lambda v)x}$$

[1]See Natanson (1955, The Theory of Functions of a Real Variable, Frederick Ungar Publishing Co., New York) Theorem 5, p. 239.

for $\lambda \in \mathbb{R}$. Consequently, (8.3) also holds for any function which is the uniform limit of a trigonometric polynomial $T(x)$. Hence, by Weierstrass approximation theorem it also holds for any continuous periodic function $h(x)$.[2]

Let g be a continuous function that vanishes outside a bounded interval I, and choose $m > 0$ sufficiently large so that $I \subset (-m, m]$. Define h_m as a continuous periodic function of periods $2m$ such that $h_m(x) = g(x)$ for $-m < x \leq m$. Then, equation (8.3) holds for h_m. Since D is of bounded variation it is possible to choose m sufficiently large so that the variation of $D(x)$ for $|x| > m$ is arbitrarily small. Hence, the integral

$$\int_{\mathbb{R}} h_m(x)dD(x) \to \int_{\mathbb{R}} g(x)dD(x) \text{ as } m \to \infty.$$

Thus,

$$\int_{\mathbb{R}} g(x)dD(x) = \int_I g(x)dD(x) = 0$$

for every continuous function that is zero outside of I. By the uniform boundedness of g (continuous on a bounded interval) it follows that

$$\int_a^b g(x)dD(x) = \int_I g dD(x) = 0$$

provided that a and b are points of continuity of D and that g is continuous for $a \leq x \leq b$. But then, $D(x)$ must be a constant on its continuity points. Hence, $G(x) = F(x)$ for $x \in C(F) \cap C(G)$. But since when F and G coincide on their points of continuity they coincide everywhere, and the proof is complete. ■

The next theorem gives an explicit representation of F in terms of ϕ.

Theorem 8.4. *For $a, b \in C(F)$ such that $a < b$,*

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

[2]See Natanson (1955) Theorem 4, p. 111.

If $\int_{\mathbb{R}} \phi(t)dt < \infty$ *the distribution function* F *has a density* f *and*

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\phi(t)dt.$$

Proof. Let $\phi(t) = \int_{\mathbb{R}} e^{itx}dF(x)$. We first prove that for every a, b, $a < b$, points of continuity of F, we have $F(b) - F(a) = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi(t)dt$. Note that even though $t = 0$ belongs to $[-T, T]$ the integral is well defined because $\frac{e^{-ita}-e^{-itb}}{it} \to b - a$ as $t \to 0$.

$$\begin{aligned}
\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{2\pi it}\phi(t)dt &= \int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{2\pi it}\left(\int_{-\infty}^{\infty} e^{itx}dF(x)\right)dt \\
&= \int_{-\infty}^{\infty}\left(\int_{-T}^{T}\frac{e^{-ita+itx}-e^{-itb+itx}}{2\pi it}dt\right)dF(x),
\end{aligned}$$

where the last equality follows from the fact that we can interchange the order of integration. Now,

$$\begin{aligned}
&\int_{-T}^{T}\frac{1}{2\pi it}(e^{it(x-a)} - e^{it(x-b)})dt \\
&= \int_{-T}^{T}\frac{1}{2\pi it}(\cos(t(x-a)) + i\sin(t(x-a)) - \cos(t(x-b)) - i\sin(t(x-b)))dt \\
&= \int_{-T}^{0}[\cdot]dt + \int_{0}^{T}[\cdot]dt.
\end{aligned}$$

If $t = -\tau$,

$$\begin{aligned}
\int_{-T}^{0}[\cdot]dt &= \int_{-T}^{0} -\frac{1}{2\pi i\tau}(\cos(\tau(x-a)) + i\sin(\tau(x-a)) - \cos(\tau(x-b)) - i\sin(\tau(x-b)))(-1)d\tau \\
&= \int_{0}^{T}\frac{1}{2\pi it}(-\cos(t(x-a)) + i\sin(t(x-a)) + \cos(t(x-b)) - i\sin(t(x-b)))dt.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\int_{-T}^{T}\frac{e^{-it(x-a)}-e^{-it(x-b)}}{2\pi it}\phi(t)dt &= 2\int_{0}^{T}\frac{1}{2\pi it}(i\sin(t(x-a)) - i\sin(t(x-b)))dt \\
&= \frac{1}{\pi}\int_{0}^{T}\frac{1}{t}(\sin(t(x-a)) - \sin(t(x-b)))dt.
\end{aligned}$$

Taking limits on both sides as $T \to \infty$ requires that we investigate the value of

$$\int_{0}^{T}\frac{1}{t}\sin(t(x-a))dt \text{ and } \int_{0}^{T}\frac{1}{t}\sin(t(x-b))dt \text{ as } T \to \infty.$$

Since $\int_0^T \frac{1}{t}\sin(t)dt$ is bounded for $T > 0$ and approaches $\pi/2$ as $T \to \infty$[3] we obtain

$$\int_0^\infty \frac{\sin(\alpha t)}{t} dt = \begin{cases} \pi/2, & \alpha > 0 \\ 0, & \alpha = 0 \\ -\pi/2 & \alpha < 0 \end{cases}.$$

Now, consider the following cases: (1) $x < a < b$; (2) $a < x < b$; (3) $a < b < x$; (4) $x = a < b$; (5) $x = b$. Then, we obtain

case (1): $\lim_{T\to\infty} \int_{-T}^T [\cdot] dt \to \frac{1}{\pi}(-\frac{\pi}{2}) - \frac{1}{\pi}(-\frac{\pi}{2}) = 0$,

case (2): $\lim_{T\to\infty} \int_{-T}^T [\cdot] dt \to \frac{1}{\pi}(\frac{\pi}{2}) - \frac{1}{\pi}(-\frac{\pi}{2}) = 1$,

case (3): $\lim_{T\to\infty} \int_{-T}^T [\cdot] dt \to \frac{1}{\pi}(\frac{\pi}{2}) - \frac{1}{\pi}(\frac{\pi}{2}) = 0$,

case (4): $\lim_{T\to\infty} \int_{-T}^T [\cdot] dt \to 0 - \frac{1}{\pi}(-\frac{\pi}{2}) = \frac{1}{2}$,

case (5): $\lim_{T\to\infty} \int_{-T}^T [\cdot] dt \to \frac{1}{\pi}(\frac{\pi}{2}) - 0 = \frac{1}{2}$.

Consequently,

$$\lim_{T\to\infty} \int_{-T}^T \frac{e^{-it(x-a)} - e^{-it(x-b)}}{2\pi i t} \phi(t) dt = \begin{cases} 0 & x < a \text{ or } x > b \\ \frac{1}{2} & x = a \text{ or } x = b \\ 1 & \text{if } a < x < b. \end{cases}$$

By Lebesgue's dominated convergence theorem

$$\begin{aligned} \lim_{T\to\infty} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{2\pi i t} \phi(t) dt &= \int_{-\infty}^\infty \left(\frac{1}{2} I_{\{x=a\}} + \frac{1}{2} I_{\{x=b\}} + I_{a<x<b} \right) dF(x) \\ &= \frac{1}{2}\left(F(a) - F(a-)\right) + \frac{1}{2}\left(F(b) - F(b-)\right) + \left(F(b) - F(a)\right) \\ &= F(b) - F(a), \end{aligned}$$

where the second equality comes from

$$\int_{-\infty}^\infty \frac{1}{2} I_{\{x=a\}} dF(x) = \frac{1}{2}(F(a) - F(a-))$$

and

$$\int_{-\infty}^\infty \frac{1}{2} I_{\{x=b\}} dF(x) = \frac{1}{2}(F(b) - F(b-)).$$

[3]See Apostol, T.,1974, *Mathematical Analysis.* Addison Wesley Publishing Company, p. 286.

The last equality follows because a, b are points of continuity of F.

For the last statement in the enunciation of the theorem, note that by the same argument used to establish uniform continuity of characteristic functions we have that $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\phi(t)dt$ is continuous and therefore integrable on $[a,b]$. Hence,

$$\begin{aligned}\int_a^b f(x)dx &= \int_a^b \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\phi(t)dtdx \\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\phi(t)\int_a^b e^{-itx}dxdt = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\phi(t)\int_a^b e^{-itx}dxdt \\ &= \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\phi(t)\frac{e^{-ita}-e^{-itb}}{it}dt = F(b)-F(a)\end{aligned}$$

for all $a, b \in C(F)$. Hence, f is the density of F. ■

Remark 8.4. *Gil-Pelaez (1951, Biometrika) has shown that if $x \in C(F)$*

$$F(x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{it}dt.$$

8.2 A central limit theorem for independent random variables

Theorem 8.5. *Let $\{X_j\}_{j=1,2,\cdots}$ be a sequence of IID random variables with $E(X_j) = \mu$, $V(X_j) = \sigma^2$ and $S_n = \sum_{j=1}^n X_j$.*

$$\frac{n^{-1}S_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{n^{-1}(S_n - n\mu)}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} Z \sim N(0,1).$$

Proof. Without loss of generality take $E(X_j) = 0$ and $V(X_j) = 1$ (otherwise, define $Y_j = \frac{X_j - \mu}{\sigma}$ and note that $E(Y_j) = 0$, $V(Y_j) = 1$). Then,

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{S_n}{\sqrt{n}},$$

and by the fact that $\{X_j\}_{j=1,2,\cdots}$ is IID we have

$$\phi_{\frac{S_n}{\sqrt{n}}}(t) = E\left(e^{it\frac{S_n}{\sqrt{n}}}\right) = E\left(e^{it\frac{X_1}{\sqrt{n}}}\right)\cdots E\left(e^{it\frac{X_n}{\sqrt{n}}}\right) = \left(E(e^{it\frac{X_1}{\sqrt{n}}})\right)^n = \left(\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n.$$

Since $E(X_1)=0,\ E(X_1^2)=1$,

$$\left|\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)-1-\frac{it}{\sqrt{n}}E(X_1)-\left(\frac{it}{\sqrt{n}}\right)^2\frac{1}{2!}E(X_1^2)\right|\leq E\left(\min\left\{\frac{|\frac{tX_1}{\sqrt{n}}|^3}{3!},\frac{2|\frac{tX_1}{\sqrt{n}}|^2}{2}\right\}\right)$$

$$\left|\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)-1+\frac{1}{2}\frac{t^2}{n}\right|\leq E\left(\min\left\{\frac{|tX_1|^3}{6n^{3/2}},\frac{|tX_1|^2}{n}\right\}\right)$$

$$=\frac{1}{n}E\left(\min\left\{\frac{|tX_1|^3}{6n^{1/2}},|tX_1|^2\right\}\right)$$

Now, $\min\left\{\frac{|tX_1|^3}{6n^{1/2}},|tX_1|^2\right\}\leq|tX_1|^2\in\mathcal{L}$, since $E(X_1^2)=1$. Also, $\min\left\{\frac{|tX_1|^3}{6n^{1/2}},|tX_1|^2\right\}\leq\frac{|tX_1|^3}{6n^{1/2}}\to 0$ as $n\to\infty$. Thus, by Lebesgue's Dominated Convergence Theorem,

$$E\left(\min\left\{\frac{|tX_1|^3}{6n^{1/2}},|tX_1|^2\right\}\right)\to 0,$$

and $\lim_{n\to\infty} n\left|\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)-1+\frac{1}{2}\frac{t^2}{n}\right|\to 0$.

Now, note that for $i=1,2,\ldots,n$ and $a_i,b_i\in\mathbb{C}$ with $|a_i|,|b_i|\leq 1$, $|\prod_{i=1}^n a_i-\prod_{i=1}^n b_i|\leq\sum_{i=1}^n|a_i-b_i|$. Then,

$$\left|\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n-\left(1-\frac{1}{2}\frac{t^2}{n}\right)^n\right|\leq n\left|\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)-\left(1-\frac{1}{2}\frac{t^2}{n}\right)\right|\to 0.$$

Since $\left(1-\frac{1}{2}\frac{t^2}{n}\right)^n\to e^{-\frac{1}{2}t^2}$ we see that $\phi_{\frac{S_n}{\sqrt{n}}}(t)\to\phi(t)=e^{-\frac{1}{2}t^2}$. Then, by Theorem 8.3 it must be that $Z\sim N(0,1)$. ■

Remark 8.5. *We observe that if the sequence of random variables $\{X_t\}_{t\in\mathbb{N}}$ is heterogeneously distributed with $\mu_t=E(X_t)$ and $V(X_t)=\sigma_t^2<\infty$, then*

$$E\left(\frac{S_n}{n}\right)=E\left(\frac{\sum_{t=1}^n X_t}{n}\right)=\frac{1}{n}\sum_{t=1}^n\mu_t,$$

$$V\left(\frac{S_n}{n}\right)=V\left(\frac{\sum_{t=1}^n X_t}{n}\right)=\frac{1}{n^2}\sum_{t=1}^n V(X_t)=\frac{1}{n^2}\sum_{t=1}^n E(X_t-\mu_t)^2=\frac{s_n^2}{n^2}.$$

Let $Y_{tn}=\frac{X_t-\mu_t}{s_n}$ and note that $E(Y_{tn})=0$ and $V(Y_{tn})=E(Y_{tn}^2)=\frac{1}{s_n^2}E(X_t-\mu_t)^2$. Then,

$$\frac{\frac{1}{n}\sum_{t=1}^n(X_t-\mu_t)}{\sqrt{\frac{s_n^2}{n^2}}}=\frac{\sum_{t=1}^n(X_t-\mu_t)}{s_n}=\sum_{t=1}^n\frac{X_t-\mu_t}{s_n}=\sum_{t=1}^n Y_{tn}.$$

Theorem 8.6. *Let $\{Y_{tn}\}_{t=1,2,\ldots,n}$ be an independent triangular array of random variables with $E(Y_{tn}) = 0$, $\sigma_{tn}^2 := V(Y_{tn}) = \frac{1}{s_n^2}E(X_t - \mu_t)^2$ with $\sum_{t=1}^n \sigma_{tn}^2 = 1$. Then, if*

$$\lim_{n\to\infty} \sum_{t=1}^n \int_{|Y_{tn}|>\epsilon} Y_{tn}^2 dP = 0$$

for all $\epsilon > 0$, we have that $S_n = \sum_{t=1}^n Y_{tn} \xrightarrow{d} N(0,1)$.

Proof. We must show that $|\phi_{S_n}(\lambda) - e^{-\frac{1}{2}\lambda^2}| = \left|\prod_{t=1}^n \phi_{Y_{tn}}(\lambda) - \prod_{t=1}^n e^{-\frac{1}{2}\lambda^2\sigma_{tn}^2}\right| \to 0$ as $n \to \infty$, since $\sum_{t=1}^n \sigma_{tn}^2 = 1$. Now,

$$\begin{aligned}
|\phi_{S_n}(\lambda) - e^{-\frac{1}{2}\lambda^2}| &= \left|\prod_{t=1}^n \phi_{Y_{tn}}(\lambda) - \prod_{t=1}^n (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2) + \prod_{t=1}^n (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2) - \prod_{t=1}^n e^{-\frac{1}{2}\lambda^2\sigma_{tn}^2}\right| \\
&\le \left|\prod_{t=1}^n \phi_{Y_{tn}}(\lambda) - \prod_{t=1}^n (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2)\right| + \left|\prod_{t=1}^n (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2) - \prod_{t=1}^n e^{-\frac{1}{2}\lambda^2\sigma_{tn}^2}\right| \\
&= T_{1n} + T_{2n}.
\end{aligned}$$

For all $z \in \mathbb{C}$ with $|z| \le 1/2$, $|e^z - 1 - z| \le |z|^2$. To see this, note that

$$|e^z - 1 - z| = \left|\sum_{j=0}^\infty \frac{z^j}{j!} - 1 - z\right| = \left|\sum_{j=2}^\infty \frac{z^j}{j!}\right| = \left|z^2 \sum_{j=0}^\infty \frac{z^j}{(j+2)!}\right| \le |z|^2 \sum_{j=0}^\infty \frac{|z|^j}{(j+2)!}.$$

But $|z| \le 1/2$, so $\sum_{j=0}^\infty \frac{|z|^j}{(j+2)!} \le \sum_{j=0}^\infty \frac{1}{2^j}\frac{1}{(j+2)!} < \frac{1}{2}\sum_{j=0}^\infty \frac{1}{2^j} = 1$.

Also, note that by Lindeberg's condition

$$\sigma_{tn}^2 = E(I_{\{Y_{tn}\le\epsilon\}}Y_{tn}^2) + E(I_{\{Y_{tn}>\epsilon\}}Y_{tn}^2) \le \epsilon^2 + E(I_{\{Y_{tn}>\epsilon\}}Y_{tn}^2) \to \epsilon^2$$

as $n \to \infty$. Since ϵ can be made arbitrarily small $\lim\limits_{n\to\infty} \max\limits_{1\le t\le n} \sigma_{tn}^2 = 0$.

Letting $z = -\frac{1}{2}\lambda^2\sigma_{tn}^2$ and taking n to be sufficiently large we can make $|z| \le 1/2$. Hence, $T_{1n} \le \sum_{t=1}^n |\phi_{Y_{tn}}(\lambda) - (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2)|$. Using item 3 in Remark 8.2, for $n = 2$, we have

$$\begin{aligned}
\left|\phi_{Y_{nt}}(\lambda) - \left(1 - \frac{1}{2}\lambda^2\sigma_{tn}^2\right)\right| &\le E\left(\min\left\{\frac{|\lambda Y_{tn}|^3}{3!}, |\lambda Y_{tn}|^2\right\}\right) \\
&\le \lambda^2 E(Y_{tn}^2 I_{\{|Y_{tn}|>\epsilon\}}) + \frac{1}{6}|\lambda|^3\epsilon E(Y_{tn}^2).
\end{aligned}$$

Thus,

$$\sum_{t=1}^{n} |\phi_{Y_{tn}}(\lambda) - (1 - \frac{1}{2}\lambda^2\sigma_{tn}^2)| \leq \lambda^2 \sum_{t=1}^{n} E(Y_{tn}^2 I_{\{|Y_{tn}|>\epsilon\}}) + \frac{1}{6}|\lambda|^3\epsilon \sum_{t=1}^{n} \sigma_{tn}^2$$
$$= \lambda^2 \sum_{t=1}^{n} E(|Y_{tn}|^2 I_{\{|Y_{tn}|>\epsilon\}}) + \frac{1}{6}|\lambda|^3\epsilon \to \epsilon\frac{1}{6}|\lambda|^3, \text{ as } n \to \infty,$$

since $\sum_{t=1}^{n} E\left(|Y_{tn}|^2 I_{\{|Y_{tn}|>\epsilon\}}\right) \to 0$ by Lindeberg's condition. Now, for T_{2n} we have

$$T_{2n} = \left|\prod_{t=1}^{n} e^{-\frac{1}{2}\lambda^2\sigma_{tn}^2} - \prod_{t=1}^{n}(1 - \frac{1}{2}\lambda^2\sigma_{tn}^2)\right| \leq \sum_{t=1}^{n} |-\frac{1}{2}\lambda^2\sigma_{tn}^2|^2 = \frac{1}{4}\lambda^4 \sum_{t=1}^{n}(\sigma_{tn}^2)^2$$
$$\leq \frac{1}{4}\lambda^4 \left(\max_{1\leq t\leq n}\sigma_{tn}^2\right) \underbrace{\sum_{t=1}^{n}\sigma_{tn}^2}_{=1} = \frac{1}{4}\lambda^4 \left(\max_{1\leq t\leq n}\sigma_{tn}^2\right) \to 0,$$

completing the proof. ■

Remark 8.6. *We observe that*

$$\lim_{n\to\infty} \sum_{t=1}^{n} E(|Y_{tn}|^{2+\delta}) = 0 \text{ for some } \delta > 0 \implies \lim_{n\to\infty} \sum_{t=1}^{n} \int_{|Y_{tn}|>\epsilon} Y_{tn}^2 dP = 0,$$

for all $\epsilon > 0$. This is easily verified by noting that

$$E|Y_{tn}|^{2+\delta} \geq E\left(I_{|Y_{tn}|>\epsilon}|Y_{tn}|^{2+\delta}\right) \text{ for all } \epsilon > 0$$
$$\geq \epsilon^\delta E(I_{|Y_{tn}|>\epsilon}|Y_{tn}|^2).$$

Hence, $\sum_{t=1}^{n} E(|Y_{tn}|^{2+\delta}) \geq \epsilon^\delta \sum_{t=1}^{n} E(I_{|Y_{tn}|>\epsilon}|Y_{tn}|^2)$. Letting $n \to \infty$, we have, for fixed ϵ,

$$\lim_{n\to\infty} \sum_{t=1}^{n} E(|Y_{tn}|^{2+\delta}) = 0 \implies \lim_{n\to\infty} \sum_{t=1}^{n} E(I_{|Y_{tn}|>\epsilon}|Y_{tn}|^2) = 0.$$

The requirement that $\lim_{n\to\infty} \sum_{t=1}^{n} E(|Y_{tn}|^{2+\delta}) = 0$ is called Lyapounov's condition. Note that

$$\sum_{t=1}^{n} E(|Y_{tn}|^{2+\delta}) = \sum_{t=1}^{n} E\left|\frac{X_t - \mu_t}{s_n}\right|^{2+\delta} = \sum_{t=1}^{n} \frac{E|X_t - \mu_t|^{2+\delta}}{s_n^{2+\delta}} = \frac{1}{s_n^{2+\delta}} \sum_{t=1}^{n} E|X_t - \mu_t|^{2+\delta}.$$

and $E|X_t + (-\mu_t)|^{2+\delta} \leq 2^{1+\delta}(E|X_t|^{2+\delta} + |-\mu_t|^{2+\delta})$. *This inequality is a special case of Loéve's c_r-Inequality, which states that for m finite, $r > 0$*

$$E(|\sum_{t=1}^{m} X_t|^r) \leq c_r \sum_{t=1}^{m} E|X_t|^r, \text{ where } c_r = \begin{cases} 1 & \text{if } r \leq 1 \\ m^{r-1} & \text{if } r > 1 \end{cases}$$

So, $E|X_t - \mu_t|^{2+\delta} \leq 2^{1+\delta}E|X_t|^{2+\delta} + 2^{1+\delta}|\mu_t|^{2+\delta}$. *If* $E|X_t|^{2+\delta}$ *and* $|\mu_t|^{2+\delta} < C$ *uniformly in* t, *then* $\sum_{t=1}^{n} E|\frac{X_t-\mu_t}{s_n}|^{2+\delta} \leq \frac{nC}{s_n^{2+\delta}} = \frac{C}{\frac{s_n^2}{n}s_n^\delta} < C' < \infty$, *if* $\inf_n \frac{s_n^2}{n} > 0$.
Consequently, we have that if $\frac{s_n^2}{n} > 0$ *uniformly in* n *and* $E|X_t|^{2+\delta}$, $|\mu_t|^{2+\delta} < \infty$ *uniformly in* t, *Liapounov's condition holds. By consequence, Lindeberg's condition holds.*

Theorem 8.7. *(Lévy's Continuity Theorem) Let* $\{F_n\}_{n\in\mathbb{N}}$ *be a sequence of distribution functions in* $\mathbb{R}$ *with* $F_n \Longrightarrow F$ *(*F_n *converges pointwise to* F *for every point of continuity of* F*), where* F *is any non-negative, bounded, non-decreasing, right-continuous function. Let* $\{\phi_n\}_{n\in\mathbb{N}}$ *be the sequence of characteristic functions for* F_n. *If*

$$\phi_n(t) \to \phi(t) \text{ where } \phi(t) \text{ is continuous at } t = 0,$$

F *is a distribution function and* ϕ *is its characteristic function.*

Proof. See Billingsley (1986, Probability and Measure, Chapter 5). ■

The following theorem allows the use of the central limit theorems we studied to obtain the asymptotic distribution of random vectors. It os known in Statistics as the Cramér-Wold device.

Theorem 8.8. *Let* $\{X_n\}_{n=0,1,2,\cdots}$ *be a sequence of random vectors taking values in* $\mathbb{R}^K$. *Then, for any* $\lambda \in \mathbb{R}^K$

$$\lambda^T X_n \xrightarrow{d} \lambda^T X_0 \Leftrightarrow X_n \xrightarrow{d} X_0.$$

Proof. If $X_n \xrightarrow{d} X_0$

$$\phi_{t^T X_n}(x) = E(e^{ixt^T X_n}) = \phi_{X_n}(xt) \to \phi_{X_0}(xt) = \phi_{t^T X_0}(x)$$

which shows that $\lambda^T X_n \xrightarrow{d} \lambda^T X_0$.

If $\lambda^T X_n \xrightarrow{d} \lambda^T X_0$ then

$$\phi_{X_n}(x) = E(e^{ix^T X_n}) = \phi_{x^T X_n}(1) \to \phi_{x^T X_0}(1) = \phi_{X_0}(x)$$

which shows that $X_n \xrightarrow{d} X_0$. ■