12. Characteristic Functions

An extremely useful construct in probability (and the primary ingredient in the classical proofs of many central limit theorems) is the characteristic function of a random variable, which is essentially the (inverse) Fourier transform of its distribution.

Definition. The *characteristic function* of a random variable X is defined as $\varphi(t) = E\left[e^{itX}\right]$.

When confusion may arise, we will indicate the dependence on the random variable with a subscript.

Though we have restricted our attention to real-valued random variables thus far, no new theory is required since if Z is complex valued, $E[Z] = E\left[\mathrm{Re}(Z)\right] + iE[\mathrm{Im}(Z)]$ provided that the expectations of the real and imaginary parts are well defined.
In the case of characteristic functions, Euler's formula gives $e^{itX} = \cos(tX) + i\sin(tX)$, and the sine and cosine functions are bounded and thus integrable against μ_X.
(Note that we are still assuming that the underlying random variables are real-valued.)

Several properties of characteristic functions are immediate from the definition.

(1) $\varphi(0) = E[1] = 1$

(2) $\varphi(-t) = E[\cos(-tX)] + iE[\sin(-tX)] = E[\cos(tX)] - iE[\sin(tX)] = \overline{\varphi(t)}$

(3) $|\varphi(t)| = \left|E\left[e^{itX}\right]\right| \leq E\left|e^{itX}\right| = 1$

(4) $|\varphi(t+h) - \varphi(t)| \leq E\left|e^{i(t+h)X} - e^{itX}\right| = E\left[\left|e^{itX}\right|\left|e^{ihX} - 1\right|\right] = E\left[\left|e^{ihX} - 1\right|\right]$.
Since the last term goes to zero as $h \to 0$ (by the bounded convergence theorem), φ is uniformly continuous.

(5) $\varphi_{aX+b}(t) = E\left[e^{it(aX+b)}\right] = E\left[e^{i(at)X}e^{itb}\right] = e^{itb}\varphi_X(at)$

(6) $\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}$ by 2 and 5

(7) If X_1 and X_2 are independent, then

$$\varphi_{X_1+X_2}(t) = E\left[e^{it(X_1+X_2)}\right] = E\left[e^{itX_1}e^{itX_2}\right] = E\left[e^{itX_1}\right]E\left[e^{itX_2}\right] = \varphi_{X_1}(t)\varphi_{X_2}(t).$$

We now turn to some examples.

Example 12.1 (Rademacher)**.** If $P(X = 1) = P(X = -1) = \frac{1}{2}$, then its ch.f. is given by

$$\varphi(t) = \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \cos(t).$$

.

Example 12.2 (Poisson)**.** If $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$ for $k=0,1,2,...$, then its ch.f. is given by

$$\varphi(t)=\sum_{k=0}^{\infty}e^{itk}e^{-\lambda}\frac{\lambda^k}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\left(\lambda e^{it}\right)^k}{k!}=e^{-\lambda}e^{\lambda e^{it}}=e^{\lambda(e^{it}-1)}.$$

Example 12.3 (Normal)**.** If X has density $f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, then its ch.f. is given by $\varphi(t)=e^{-\frac{t^2}{2}}$.

Naive derivation:

$$\varphi(t)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{itx}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-\frac{1}{2}[(x-it)^2+t^2]}dx=e^{-\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-\frac{(x-it)^2}{2}}dx=e^{-\frac{t^2}{2}}$$

since $\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-it)^2}{2}}$ is the the density of a normal random variable with mean it and variance 1.

Formal proof:

Since $\sin(tx)e^{-\frac{x^2}{2}}$ is odd and integrable,

$$\begin{aligned}\varphi(t)&=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{itx}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\cos(tx)e^{-\frac{x^2}{2}}dx+\frac{i}{\sqrt{2\pi}}\int_{\mathbb{R}}\sin(tx)e^{-\frac{x^2}{2}}dx\\&=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\cos(tx)e^{-\frac{x^2}{2}}dx.\end{aligned}$$

Differentiating with respect to t (which can be justified using a DCT argument) gives

$$\begin{aligned}\varphi'(t)&=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\frac{d}{dt}\cos(tx)e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}-x\sin(tx)e^{-\frac{x^2}{2}}dx\\&=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\sin(tx)\left(\frac{d}{dx}e^{-\frac{x^2}{2}}\right)dx=-\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}t\cos(tx)e^{-\frac{x^2}{2}}dx=-t\varphi(t).\end{aligned}$$

It follows from the method of integrating factors that $\frac{d}{dt}\left(e^{\frac{t^2}{2}}\varphi(t)\right)=0$, hence $e^{\frac{t^2}{2}}\varphi(t)=e^{\frac{0^2}{2}}\varphi(0)=1$ for all t.

Example 12.4 (Exponential)**.** If X is absolutely continuous with density $f_X(x)=e^{-x}1_{[0,\infty)}(x)$, then its ch.f. is given by

$$\varphi(t)=\int_0^{\infty}e^{itx}e^{-x}dx=\int_0^{\infty}e^{(it-1)x}dx=\lim_{b\to\infty}\frac{1}{it-1}e^{(it-1)x}\Big|_0^b=\frac{1}{1-it}.$$

Our next task is to show that the characteristic function uniquely determines the distribution.

We first observe that

Proposition 12.1. *For all* $T>0$, $\left|\int_0^T\frac{\sin(t)}{t}dt-\frac{\pi}{2}\right|\leq\frac{T+1}{T^2}$.

Proof. (Homework)

For all $T>0$ the function $e^{-uv}\sin(u)$ is bounded in absolute value by e^{-uv}, which is integrable over $R_T=\{(u,v):0<u<T,v>0\}$, so it follows from Fubini's theorem that

$$\int_0^T \frac{\sin(u)}{u}du = \int_0^T \left(\int_0^\infty e^{-uv}\sin(u)dv\right)du = \int_0^\infty \left(\int_0^T e^{-uv}\sin(u)du\right)dv$$
$$= \int_0^\infty \left[-\frac{1}{1+v^2}e^{-uv}\left(\cos(u)+v\sin(u)\right)\Big|_{u=0}^{u=T}\right]dv$$
$$= \int_0^\infty \frac{dv}{1+v^2} - \cos(T)\int_0^\infty \frac{1}{1+v^2}e^{-Tv}dv - \sin(T)\int_0^\infty \frac{v}{1+v^2}e^{-Tv}dv.$$

Since $\int_0^\infty \frac{dv}{1+v^2} = \frac{\pi}{2}$, we have

$$\left|\int_0^T \frac{\sin(t)}{t}dt - \frac{\pi}{2}\right| = \left|\cos(T)\int_0^\infty \frac{1}{1+v^2}e^{-Tv}dv + \sin(T)\int_0^\infty \frac{v}{1+v^2}e^{-Tv}dv\right|$$
$$\leq |\cos(T)|\int_0^\infty \left|\frac{1}{1+v^2}\right|e^{-Tv}dv + |\sin(T)|\int_0^\infty \left|\frac{v}{1+v^2}\right|e^{-Tv}dv$$
$$\leq \int_0^\infty e^{-Tv}dv + \int_0^\infty ve^{-Tv}dv = \frac{T+1}{T^2}.$$

□

A little more calculus gives

Lemma 12.1. *For every* $\theta \in \mathbb{R}$, $\lim_{T\to\infty}\int_{-T}^T \frac{\sin(\theta t)}{t}dt = \pi\,\mathrm{sgn}(\theta)$ *where* $\mathrm{sgn}(\theta) = \begin{cases} -1, & \theta < 0 \\ 0, & \theta = 0 \\ 1, & \theta > 0 \end{cases}$.

Proof. (Homework)

Since $\lim_{t\to 0}\frac{\sin(\theta t)}{t} = \lim_{t\to 0}\frac{\theta\cos(\theta t)}{1} = \theta$, it is easy to see that the integral $\int_{-T}^T \frac{\sin(\theta t)}{t}dt$ exists for all $T > 0$.

Because $\frac{\sin(\theta t)}{t}$ is even, it follows by u-substitution that

$$\int_{-T}^T \frac{\sin(\theta t)}{t}dt = 2\int_0^T \frac{\sin(\theta t)}{t}dt = 2\int_0^{\theta T}\frac{\sin(u)}{u}du = 2\mathrm{sgn}(\theta)\int_0^{|\theta|T}\frac{\sin(u)}{u}du.$$

Proposition 12.1 shows that $\lim_{T\to\infty}\int_0^{|\theta|T}\frac{\sin(u)}{u}du = \frac{\pi}{2}$ for all $\theta \neq 0$, so, since $\theta = 0$ implies that $\int_0^{|\theta|T}\frac{\sin(u)}{u}du = 0$ for all T, we have

$$\lim_{T\to\infty}\int_{-T}^T \frac{\sin(\theta t)}{t}dt = 2\,\mathrm{sgn}(\theta)\lim_{T\to\infty}\int_0^{|\theta|T}\frac{\sin(u)}{u}du \to \pi\,\mathrm{sgn}(\theta)$$

for all $\theta \in \mathbb{R}$.

□

With the previous result at our disposal, we are in a position to prove

Theorem 12.1 (Inversion Formula)**.** *Let* $\varphi(t) = \int e^{itx}d\mu(x)$ *where* μ *is a probability measure on* $(\mathbb{R}, \mathcal{B})$*. If* $a < b$*, then*

$$\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^T \frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt = \mu\left((a,b)\right) + \frac{1}{2}\mu\left(\{a,b\}\right).$$

Proof. We begin by noting that

$$(*) \qquad \left|\frac{e^{-ita}-e^{-itb}}{it}\right| = \left|\int_a^b e^{-ity}dy\right| \le \int_a^b \left|e^{-ity}\right| dy = b-a,$$

so, since $[-T,T]$ is finite and μ is a probability measure, Fubini's theorem gives

$$\begin{aligned} I_T &= \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt = \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\left(\int_{\mathbb{R}} e^{itx}d\mu(x)\right)dt \\ &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{it}dt\right)d\mu(x). \end{aligned}$$

Now

$$\frac{e^{it(x-a)}-e^{it(x-b)}}{it} = \frac{\sin(t(x-a))-\sin(t(x-b))}{t} + i\frac{\cos(t(x-b))-\cos(t(x-a))}{t},$$

and it follows from $(*)$ and the inequality $|\mathrm{Im}(z)| \le |z|$ that $\int_{-T}^{T} \frac{\cos(t(x-b))-\cos(t(x-a))}{t}dt$ exists.

Thus, since $\dfrac{\cos(t(x-b))-\cos(t(x-a))}{t}$ is an odd function, we must have

$$\begin{aligned} I_T &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{it}dt\right)d\mu(x) \\ &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{\sin(t(x-a))-\sin(t(x-b))}{t}dt + i\int_{-T}^{T}\frac{\cos(t(x-b))-\cos(t(x-a))}{t}dt\right)d\mu(x) \\ &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{\sin(t(x-a))-\sin(t(x-b))}{t}dt\right)d\mu(x) \\ &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{\sin(t(x-a))}{t}dt - \int_{-T}^{T}\frac{\sin(t(x-b))}{t}dt\right)d\mu(x). \end{aligned}$$

Lemma 12.1 shows that $\left|\int_{-T}^{T}\frac{\sin(\theta t))}{t}dt\right|$ converges to the finite constant π as $T\to\infty$, so it follows from the bounded convergence theorem and Lemma 12.1 that

$$\begin{aligned} \lim_{T\to\infty} I_T &= \lim_{T\to\infty}\int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{\sin(t(x-a))}{t}dt - \int_{-T}^{T}\frac{\sin(t(x-b))}{t}dt\right)d\mu(x) \\ &= \int_{\mathbb{R}}\lim_{T\to\infty}\left(\int_{-T}^{T}\frac{\sin(t(x-a))}{t}dt - \int_{-T}^{T}\frac{\sin(t(x-b))}{t}dt\right)d\mu(x) \\ &= \pi\int_{\mathbb{R}}\left[\mathrm{sgn}(x-a)-\mathrm{sgn}(x-b)\right]d\mu(x). \end{aligned}$$

Since $a<b$ by assumption, we have that $\mathrm{sgn}(x-a)-\mathrm{sgn}(x-b) = \begin{cases} 0, & x<a \text{ or } x>b \\ 1, & x=a \text{ or } x=b, \\ 2, & a<x<b \end{cases}$ thus

$$\begin{aligned} \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt &= \frac{1}{2\pi}\lim_{T\to\infty} I_T = \int_{\mathbb{R}}\frac{\mathrm{sgn}(x-a)-\mathrm{sgn}(x-b)}{2}d\mu(x) \\ &= \int_{(-\infty,a)\cup(b,\infty)} 0\,d\mu(x) + \int_{\{a,b\}}\frac{1}{2}\,d\mu(x) + \int_{(a,b)} 1\,d\mu(x) \\ &= \frac{1}{2}\mu(\{a,b\}) + \mu((a,b)). \end{aligned}$$

□

Remark. Note that the *Cauchy principal value* $\lim_{T\to\infty}\int_{-T}^{T} f(x)dx$ is not necessarily the same as

$$\int_{-\infty}^{\infty} f(x)dx := \lim_{\substack{b\to\infty \\ a\to-\infty}} \int_a^b f(x)dx.$$

For example, $\lim_{a\to\infty}\int_{-a}^{a} \frac{2x}{1+x^2}dx = 0$ since the integrand is odd, but

$$\lim_{a\to\infty}\int_{-2a}^{a} \frac{2x}{1+x^2}dx = \lim_{a\to\infty} \log\left(1+x^2\right)\Big|_{-2a}^{a} = \lim_{a\to\infty} \log\left(\frac{1+a^2}{1+4a^2}\right) = -\log(4),$$

so the improper Riemann integral $\int_{-\infty}^{\infty} \frac{2x}{1+x^2}dx$ is not defined.

Of course, since $\left|\frac{2x}{1+x^2}\right| \approx \frac{2}{|x|}$ for large $|x|$, $\int_{-\infty}^{\infty} \frac{2x}{1+x^2}dx$ is not defined as a Lebesgue integral either.

It is left as a homework exercise to imitate the proof of Theorem 12.1 to obtain

Theorem 12.2. *Under the assumptions of Theorem 12.1,*

$$\mu(\{a\}) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{-ita}\varphi(t)dt.$$

Combining Theorems 12.1 and 12.2, and noting that a Borel measure is specified by its values on open intervals (by a $\pi - \lambda$ argument), shows that probability distributions are uniquely determined by their characteristic functions.

To prove our next big result, we need the following bound on the tail probabilities of a distribution in terms of its characteristic function.

Lemma 12.2. *If φ is the characteristic function corresponding to the distribution μ, then for all $u > 0$,*

$$\mu\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \le u^{-1}\int_{-u}^{u} (1-\varphi(t))\,dt.$$

Proof. It follows from the parity of the sine and cosine functions that

$$\int_{-u}^{u} \left(1-e^{itx}\right)dt = 2u - \int_{-u}^{u} (\cos(tx) + i\sin(tx))\,dt = 2\left(u - \frac{\sin(ux)}{x}\right).$$

Dividing by u, integrating against $d\mu(x)$, and appealing to Fubini gives

$$u^{-1}\int_{-u}^{u} (1-\varphi(t))\,dt = u^{-1}\int\left(\int_{-u}^{u} 1-e^{itx}dt\right)d\mu(x) = 2\int\left(1-\frac{\sin(ux)}{ux}\right)d\mu(x).$$

Since $|\sin(y)| = \left|\int_0^y \cos(x)dx\right| \le |y|$ for all y, we see that the integrand on the right is nonnegative, so we have

$$\begin{aligned} u^{-1}\int_{-u}^{u} (1-\varphi(t))\,dt &= 2\int\left(1-\frac{\sin(ux)}{ux}\right)d\mu(x) \ge 2\int_{|x|>\frac{2}{u}}\left(1-\frac{\sin(ux)}{ux}\right)d\mu(x) \\ &\ge 2\int_{|x|>\frac{2}{u}}\left(1-\frac{1}{|ux|}\right)d\mu(x) \ge \mu\left(\left\{x : |x| > \frac{2}{u}\right\}\right). \end{aligned}$$

□

We are now able to take our next major step toward proving the central limit theorem by relating weak convergence to the convergence of the corresponding characteristic functions.

Theorem 12.3 (Continuity Theorem). *Let μ_n, $1 \leq n \leq \infty$, be probability distributions with characteristic functions $\varphi_n(t) = \int_{\mathbb{R}} e^{itx} d\mu_n(x)$.*

(i) *If $\mu_n \Rightarrow \mu_\infty$, then $\varphi_n(t) \to \varphi_\infty(t)$ for all $t \in \mathbb{R}$.*

(ii) *If $\varphi_n(t)$ converges pointwise to a limit $\varphi(t)$ that is continuous at $t = 0$, then the sequence $\{\mu_n\}$ is tight and converges weakly to the distribution μ with characteristic function $\varphi(t)$.*

Proof.

For (i), note that since e^{itx} is bounded and continuous, if $\mu_n \Rightarrow \mu_\infty$, then it follows from Theorem 11.2 that $\varphi_n(t) \to \varphi_\infty(t)$.

For (ii), we observe that

$$u^{-1} \int_{-u}^{u} (1 - \varphi(t))\, dt \leq 2 \sup\{1 - \varphi(t) : |t| \leq u\} \to 0 \text{ as } u \to 0$$

since φ is continuous at 0 and thus $\lim_{t \to 0} \varphi(t) = \varphi(0) = 1$.

It follows that for any $\varepsilon > 0$, there is a $v > 0$ such that

$$v^{-1} \int_{-v}^{v} (1 - \varphi(t))\, dt < \frac{\varepsilon}{2}.$$

Because $|1 - \varphi(t)| \leq 2$ and $\varphi_n(t) \to \varphi(t)$ for all t, the bounded convergence theorem shows that there is an $N \in \mathbb{N}$ such that

$$\left| v^{-1} \int_{-v}^{v} (1 - \varphi_n(t))\, dt - v^{-1} \int_{-v}^{v} (1 - \varphi(t))\, dt \right| < \frac{\varepsilon}{2}$$

whenever $n \geq N$.

The last two observations and Lemma 12.2 show that

$$\mu_n\left(\left\{x : |x| > \frac{2}{v}\right\}\right) \leq v^{-1} \int_{-v}^{v} (1 - \varphi_n(t))\, dt < \varepsilon$$

for all $n \geq N$, so, since ε was arbitrary, it follows that $\{\mu_n\}_{n=1}^{\infty}$ is tight.

Now let $\{\mu_{n_m}\}_{m=1}^{\infty}$ be any subsequence. Tightness and Theorems 11.5 and 11.6 imply that there is a further subsequence which converges weakly to some probability measure μ_∞.
It then follows from part (i) that the corresponding characteristic functions converge pointwise to the characteristic function of μ_∞.
Because $\varphi_n(t) \to \varphi(t)$ for all t and characteristic functions uniquely characterize distributions, it must be the case that $\mu_\infty = \mu$.
Therefore, every subsequence of $\{\mu_n\}_{n=1}^{\infty}$ has a further subsequence which converges weakly - that is, in the weak-$*$ topology - to μ, so Lemma 8.2 shows that $\mu_n \Rightarrow \mu$. □

The crux of the proof of the nontrivial part of Theorem 12.3 was establishing tightness of the sequence $\{\mu_n\}$, and this is where we used the assumption that the limiting characteristic function is continuous at 0.

As an illustration of how weak convergence may fail without the continuity assumption, consider the case $\mu_n = N(0,n)$. Then μ_n has ch.f.

$$\varphi_n(t) = e^{-\frac{nt^2}{2}} \to \begin{cases} 0, & t \neq 0 \\ 1, & t = 0 \end{cases},$$

which is discontinuous at 0. To see that μ_n has no weak limit, observe that for any $x \in \mathbb{R}$,

$$\mu_n\left((-\infty, x]\right) = \frac{1}{\sqrt{2\pi n}} \int_{-\infty}^{x} e^{-\frac{t^2}{2n}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{\sqrt{n}}} e^{-\frac{s^2}{2}} ds \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{s^2}{2}} ds = \frac{1}{2}.$$

We turn next to the problem of representing characteristic functions as power series with explicit remainders. To start, we have

Theorem 12.4. *If $E\left[|X|^n\right] < \infty$, then the characteristic function φ of X has a continuous derivative of order n given by*

$$\varphi^{(n)}(t) = \int (ix)^n e^{itx} d\mu(x).$$

Proof. (Homework)

Argue by induction and justify differentiating under the integral with the DCT or some applicable corollary thereof. □

It follows from Theorem 12.4 that if $E\left[|X|^n\right] < \infty$, then $\varphi^{(n)}(0) = \int (ix)^n d\mu(x) = i^n E\left[X^n\right]$.

The above observation combined with Taylor's theorem shows that if X has finite absolute nth moment, then

$$\varphi(t) = \sum_{k=0}^{n} \frac{\varphi^{(k)}(0)}{k!} t^k + r_n(t)t^n = \sum_{k=0}^{n} \frac{(it)^k}{k!} E\left[X^k\right] + r_n(t)t^n.$$

where the *Peano remainder* $r_n(t) \to 0$ as $t \to 0$.

In particular, we have

Corollary 12.1. *If X has mean 0 and finite variance σ^2, then*

$$\varphi(t) = 1 + itE[X] - \frac{t^2}{2} E[X^2] + r_2(t)t^2 = 1 - \frac{1}{2}\sigma^2 t^2 + o(t^2)$$

where $o(t^2)$ denotes a quantity which, when divided by t^2, tends to 0 as $t \to 0$.

Remark. To verify the statement of Taylor's theorem given above, set

$$P_n(t) = \sum_{k=0}^{n} \frac{\varphi^{(k)}(0)}{k!} t^k = \varphi(0) + \varphi'(0)t + ... + \frac{\varphi^{(n)}(0)}{n!} t^n, \quad r_n(t) = \begin{cases} \frac{\varphi(t) - P_n(t)}{t^n}, & t \neq 0 \\ 0, & t = 0 \end{cases}.$$

Then one needs only to prove that $\lim_{t \to 0} r_n(t) = 0$.

Applying L'Hospital's theorem $n-1$ times gives

$$\begin{aligned}\lim_{t\to 0} r_n(t) &= \lim_{t\to 0}\frac{\varphi(t)-P_n(t)}{t^k} = \lim_{t\to 0}\frac{\frac{d}{dt}\left[\varphi(t)-P_n(t)\right]}{\frac{d}{dt}t^k} \\ &= \ldots = \lim_{t\to 0}\frac{\frac{d^{n-1}}{dt^{n-1}}\left[\varphi(t)-P_n(t)\right]}{\frac{d^{n-1}}{dt^{n-1}}t^n} = \lim_{t\to 0}\frac{\varphi^{(n-1)}(t)-\varphi^{(n-1)}(0)-t\varphi^{(n)}(0)}{n!t} \\ &= \frac{1}{n!}\left[\lim_{t\to 0}\frac{\varphi^{(n-1)}(t)-\varphi^{(n-1)}(0)}{t}-\varphi^{(n)}(0)\right] = \frac{1}{n!}\left[\varphi^{(n)}(0)-\varphi^{(n)}(0)\right] = 0.\end{aligned}$$

Corollary 12.1 is enough to get the classical central limit theorem for i.i.d. sequences, but when we consider the Lindeberg-Feller CLT for triangular arrays, we will need a little better control on the error term. With this end in mind, we prove

Lemma 12.3. $\left|e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\right| \le \min\left\{\frac{|x|^{n+1}}{(n+1)!},\frac{2|x|^n}{n!}\right\}$

Proof. Integrating by parts gives

$$\int_0^x (x-s)^n e^{is}ds = \int_0^x \left[\frac{d}{ds}\left(-\frac{(x-s)^{n+1}}{n+1}\right)\right]e^{is}ds = \frac{x^{n+1}}{n+1}+\frac{i}{n+1}\int_0^x (x-s)^{n+1}e^{is}ds.$$

Taking $n=0$ shows that

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

and thus

$$e^{ix} = 1+ix+i^2\int_0^x (x-s)e^{is}ds.$$

If we assume that

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

then we get

$$\begin{aligned}e^{ix} &= \sum_{k=0}^{n}\frac{(ix)^k}{k!}+\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} \\ &= \sum_{k=0}^{n}\frac{(ix)^k}{k!}+\frac{i^{n+1}}{n!}\left(\frac{x^{n+1}}{n+1}+\frac{i}{n+1}\int_0^x (x-s)^{n+1}e^{is}ds\right) \\ &= \sum_{k=0}^{n+1}\frac{(ix)^k}{k!}+\frac{i^{(n+1)+1}}{(n+1)!}\int_0^x (x-s)^{n+1}e^{is}ds,\end{aligned}$$

so it follows from the principle of induction that

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

for all $n = 0, 1, 2, ...$

We will be done if we can show that the modulus of the right hand side is bounded above by both $\frac{|x|^{n+1}}{(n+1)!}$ and $\frac{2|x|^n}{n!}$.

In the first case we have

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

For the second case, note that

$$\begin{aligned}\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}ds &= \frac{i^n}{(n-1)!}\int_0^x \frac{(x-s)^n}{n}\left[\frac{d}{ds}e^{is}\right] ds \\ &= \frac{i^n}{(n-1)!}\left[-\frac{x^n}{n}+\int_0^x (x-s)^{n-1}e^{is}ds\right] \\ &= \frac{i^n}{(n-1)!}\left[-\int_0^x (x-s)^{n-1}ds+\int_0^x (x-s)^{n-1}e^{is}ds\right] \\ &= \frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}\left(e^{is}-1\right) ds,\end{aligned}$$

hence

$$\begin{aligned}\left|\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}ds\right| &\le \frac{1}{(n-1)!}\int_0^{|x|} |x-s|^{n-1}\left(\left|e^{is}\right|+1\right) ds \\ &= \frac{2}{(n-1)!}\int_0^{|x|} s^{n-1}ds = \frac{2\,|x|^n}{n!}.\end{aligned}$$

□

Observe that the upper bound $\frac{|x|^{n+1}}{(n+1)!}$ is better for small values of $|x|$, while the bound $\frac{2\,|x|^n}{n!}$ is better for $|x| > 2(n+1)$.

Applying Lemma 12.3 to $x = tX$ and taking expected values gives

Corollary 12.2.

$$\begin{aligned}\left|\varphi(t)-\sum_{k=0}^n \frac{(it)^k}{k!}E[X^k]\right| &= \left|E\left[e^{itX}-\sum_{k=0}^n \frac{(itX)^k}{k!}\right]\right| \\ &\le E\left|e^{itX}-\sum_{k=0}^n \frac{(itX)^k}{k!}\right| \le E\left[\min\left\{\frac{|tX|^{n+1}}{(n+1)!}, \frac{2\,|tX|^n}{n!}\right\}\right].\end{aligned}$$