Math 639: Lecture 5

Characteristic functions, central limit theorems

Bob Hough

February 7, 2017

The binomial distribution

Let $X_1, X_2, \ldots$ be i.i.d. random variables,

$$\text{Prob}(X_1 = 1) = \frac{1}{2}, \qquad \text{Prob}(X_1 = -1) = \frac{1}{2}.$$

The sum $S_n = X_1 + X_2 + \cdots + X_n$ is the nth step of simple random walk. From the binomial theorem one obtains

$$\text{Prob}(S_{2n} = 2k) = \binom{2n}{n+k} 2^{-2n}.$$

The binomial distribution

As $n \to \infty$, the binomial distribution approximates the density of a normal distribution pointwise in the following sense.

Theorem

If $\frac{2k}{\sqrt{2n}} \to x$ *as* $n \to \infty$ *then*

$$\mathrm{Prob}(S_{2n} = 2k) \sim \frac{e^{-\frac{x^2}{2}}}{\sqrt{\pi n}}.$$

The binomial distribution

Proof.

- Stirling's formula gives the asymptotic
$$n! \sim n^n e^{-n}\sqrt{2\pi n}$$
as $n \to \infty$.
- Hence as $|n \pm k| \to \infty$,
$$\binom{2n}{n+k} = \frac{(2n)!}{(n+k)!(n-k)!}$$
$$\sim \frac{(2n)^{2n}}{(n-k)^{n-k}(n+k)^{n+k}}\sqrt{\frac{2n}{2\pi(n+k)(n-k)}}.$$

□

The binomial distribution

Proof.

- If $\frac{k}{n} \to 0$ as n increases,

$$\binom{2n}{n+k} 2^{-2n} \sim \frac{1}{\sqrt{\pi n}} \left(1 + \frac{k}{n}\right)^{-n-k} \left(1 - \frac{k}{n}\right)^{-n+k}$$

$$\sim \frac{1}{\sqrt{\pi n}} \left(1 - \frac{k^2}{n^2}\right)^{-n} \left(1 + \frac{k}{n}\right)^{-k} \left(1 - \frac{k}{n}\right)^{k}.$$

- If $k = o(n^{\frac{2}{3}})$ then this is $\sim \frac{e^{-\frac{k^2}{n}}}{\sqrt{\pi n}}$, as wanted.

□

The De Moivre-Laplace Theorem

Theorem (The De Moivre-Laplace Theorem)

If $a < b$ *then as* $m \to \infty$,

$$\mathrm{Prob}\left(a \leq \frac{S_m}{\sqrt{m}} \leq b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

The De Moivre-Laplace Theorem

Proof.

- Assume $m = 2n$ is even, as the odd case may be handled similarly.
- Calculate

$$\mathrm{Prob}\left(a \leq \frac{S_m}{\sqrt{m}} \leq b\right) = \sum_{m \in [a\sqrt{2n}, b\sqrt{2n}] \cap 2\mathbb{Z}} \mathrm{Prob}(S_{2n} = m).$$

- Inserting the asymptotic evaluation, this is

$$\sim \left(\frac{2}{n}\right)^{\frac{1}{2}} \sum_{x \in [a,b] \cap \sqrt{\frac{2}{n}}\mathbb{Z}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \sim \int_a^b \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx.$$

□

Convergence in distribution

Definition

A sequence of random variables X_n is said to *converge in distribution* to X_∞, written $X_n \Rightarrow X_\infty$, if for each interval $(a, b]$ for which a and b are points of continuity of the distribution function of X_∞,

$$\mathrm{Prob}(X_n \in (a, b]) \to \mathrm{Prob}(X_\infty \in (a, b]).$$

Geometric distribution

Example

Let X_p be the number of trials needed to get a success in a sequence of independent trials of success probability p. This has a geometric distribution, $\mathrm{Prob}(X_p \geq n) = (1-p)^{n-1}$ for $n = 1, 2, 3, \ldots$. As $p \downarrow 0$,

$$\mathrm{Prob}(pX_p > x) \to e^{-x}, \qquad x \geq 0.$$

Birthday problem

Example

Let $X_1, X_2, \ldots$ be independent and uniformly distributed on $\{1, 2, \ldots, N\}$, and let $T_N = \min\{n : X_n = X_m, \text{ some } m < n\}$. Hence

$$\mathrm{Prob}(T_N > n) = \prod_{m=2}^{n} \left(1 - \frac{m-1}{N}\right),$$

and, for $x \geq 0$,

$$\mathrm{Prob}\left(\frac{T_N}{N^{\frac{1}{2}}} > x\right) \to \exp\left(-\frac{x^2}{2}\right).$$

Convergence of maxima

Example

Let $X_1, X_2, \ldots$ be independent with distribution F, and let $M_n = \max_{m \leq n} X_m$. M_n has distribution function $\mathrm{Prob}(M_n \leq x) = F(x)^n$. In particular, if X_i has an exponential distribution, so that $F(x) = 1 - e^{-x}$, then

$$\mathrm{Prob}(M_n - \log n \leq y) \to \exp(-e^{-y}), \qquad n \to \infty.$$

This is the *Gumbel distribution*.

Weak convergence

Theorem

If $F_n \Rightarrow F_\infty$ then there are random variables Y_n, $1 \leq n \leq \infty$ with distribution F_n so that $Y_n \to Y_\infty$, a.s.

Proof.

- We build the random variables on $(0, 1)$ with Borel sets and Lebesgue measure.
- Define $Y_n(x) = \sup\{y : F_n(y) \leq x\}$, and similarly Y_∞.
- Define $a_x = \sup\{y : F_\infty(y) < x\}$, $b_x = \inf\{y : F_\infty(y) > x\}$.
- Let $\Omega_0 = \{x : a_x = b_x\}$. We have $\Omega \setminus \Omega_0$ is countable, since (a_x, b_x) contains a rational number. We check that $Y_n(x) \to Y_\infty(x)$ for $x \in \Omega_0$.

□

Weak convergence

Proof.

- Recall that, for $x \in \Omega_0$, $\sup\{y : F_\infty(y) < x\} = \inf\{y : F_\infty(y) > x\}$.
- Let $y < F^{-1}(x)$ be a point of continuity. Since $x \in \Omega_0$, $F(y) < x$, and so $F_n(y) < x$ for all n sufficiently large. It follows that $F_n^{-1}(x) \geq y$ and
$$\liminf_{n\to\infty} F_n^{-1}(x) \geq F_\infty^{-1}(x).$$
- Arguing similarly, $\limsup_{n\to\infty} F_n^{-1}(x) \leq F_\infty^{-1}(x)$.

□

Weak convergence

Theorem

$X_n \Rightarrow X_\infty$ if and only if for every bounded continuous function g we have $\mathrm{E}[g(X_n)] = \mathrm{E}[g(X_\infty)]$.

Weak convergence

Proof.

- First suppose $X_n \Rightarrow X_\infty$. Choose Y_n equal in distribution to X_n and converging a.s.. Then bounded convergence gives

$$\mathrm{E}[g(X_n)] = \mathrm{E}[g(Y_n)] \to \mathrm{E}[g(Y_\infty)] = \mathrm{E}[g(X_\infty)].$$

- Now suppose $\mathrm{E}[g(X_n)] \to \mathrm{E}[g(X_\infty)]$ for all bounded continuous g. Let

$$g_{x,\epsilon}(y) = \begin{cases} 1 & y \leq x \\ 0 & y \geq x + \epsilon \\ \text{linear} & x \leq y \leq x + \epsilon \end{cases}.$$

□

Weak convergence

Proof.

- Calculate

$$\limsup_{n\to\infty} \mathrm{Prob}(X_n \leq x) \leq \limsup_{n\to\infty} \mathrm{E}[g_{x,\epsilon}(X_n)]$$
$$= \mathrm{E}[g_{x,\epsilon}(X_\infty)] \leq \mathrm{Prob}(X_\infty \leq x + \epsilon).$$

  Letting $\epsilon \downarrow 0$ gives $\limsup_{n\to\infty} \mathrm{Prob}(X_n \leq x) \leq \mathrm{Prob}(X_\infty \leq x)$.

- To obtain the other direction use $g_{x-\epsilon,\epsilon}$.

□

Continuous mapping theorem

Theorem (Continuous mapping theorem)

Let g be a measurable function and $D_g = \{x : g \text{ discontinuous at } x\}$. If $X_n \Rightarrow X_\infty$ and $\mathrm{Prob}(X_\infty \in D_g) = 0$ then $g(X_n) \Rightarrow g(X)$. If, in addition, g is bounded, then $\mathrm{E}[g(X_n)] \to \mathrm{E}[g(X_\infty)]$.

Continuous mapping theorem

Proof.

- Let Y_n equal to X_n in distribution, with $Y_n \to Y_\infty$ a.s.
- If f is continuous, then $D_{f \circ g} \subset D_g$ so $\text{Prob}(Y_\infty \in D_{f \circ g}) = 0$ and $f(g(Y_n)) \to f(g(Y_\infty))$ a.s.
- If f is bounded, then $\mathrm{E}[f(g(Y_n))] \to \mathrm{E}[f(g(Y_\infty))]$ so $g(X_n) \Rightarrow g(X_\infty)$.
- We have $g(Y_n) \to g(Y_\infty)$ a.s., so that for bounded g, $\mathrm{E}[g(Y_n)] \to \mathrm{E}[g(Y_\infty)]$.

□

Convergence in distribution

Theorem

The following statements are equivalent:

1. $X_n \Rightarrow X_\infty$
2. *For all open sets* G, $\liminf_{n\to\infty} \mathrm{Prob}(X_n \in G) \geq \mathrm{Prob}(X_\infty \in G)$.
3. *For all closed sets* K, $\limsup_{n\to\infty} \mathrm{Prob}(X_n \in K) \leq \mathrm{Prob}(X_\infty \in K)$.
4. *For all sets* A *with* $\mathrm{Prob}(X_\infty \in \partial A) = 0$,
   $\lim_{n\to\infty} \mathrm{Prob}(X_n \in A) = \mathrm{Prob}(X_\infty \in A)$.

Convergence in distribution

Proof.

- $1 \Rightarrow 2$: Let Y_n have the same distribution as X_n and satisfy $Y_n \to Y_\infty$ a.s. Then $\liminf \mathbf{1}_G(Y_n) \geq \mathbf{1}_G(Y_\infty)$, so Fatou implies

$$\liminf_{n\to\infty} \mathrm{Prob}(Y_n \in G) \geq \mathrm{Prob}(Y_\infty \in G).$$

- $2 \Rightarrow 3$: This follows since K^c is open
- $2, 3 \Rightarrow 4$: Let $K = \overline{A}$ and $G = A^o$. Then $\partial A = K - G$ has $\mathrm{Prob}(X_\infty \in \partial A) = 0$, which implies $\mathrm{Prob}(X_\infty \in G) = \mathrm{Prob}(X_\infty \in K) = \mathrm{Prob}(X_\infty \in A)$. The claim now follows from 2 and 3.
- $4 \Rightarrow 1$: For x such that $\mathrm{Prob}(X_\infty = x) = 0$, 4 implies $\mathrm{Prob}(X_n \in (-\infty, x]) \to \mathrm{Prob}(X_\infty \in (-\infty, x])$.

□

Helly's selection theorem

Theorem (Helly's selection theorem)

For every sequence F_n of distribution functions, there is a subsequence $F_{n(k)}$ and a right continuous nondecreasing function F so that $\lim_{k\to\infty} F_{n(k)}(y) = F(y)$ at all continuity points y of F.

This convergence is called *vague*.

Helly's selection theorem

Proof.

- Let $q_1, q_2, \ldots$ be an enumeration of the rationals. By diagonalization it's possible to choose a sequence F_{n_k} such that $F_{n_k}(q) \to G(q)$ converges for each rational q.
- Define G at x, by $G(x) = \inf\{G(q) : q \in \mathbb{Q}, q > x\}$. Evidently G is right continuous.
- The convergence at points of continuity of G follows from the convergence at rational points.

□

Tight sequences

Definition

A sequence of distribution functions $\{F_n\}$ is *tight* if, for all $\epsilon > 0$ there is M_ϵ so that

$$\limsup_{n\to\infty} 1 - F_n(M_\epsilon) + F_n(-M_\epsilon) \leq \epsilon.$$

Tight sequences

Theorem

Let $\{F_n\}$ be a sequence of probability distribution functions. Every subsequential limit of $\{F_n\}$ is the distribution function of a probability measure if and only if $\{F_n\}$ is tight.

Thus the tightness condition rules out 'escape of mass'. For a proof, see Durrett, p. 104.

Tight sequences

Theorem

If there is a $\phi \geq 0$ *so that* $\phi(x) \to \infty$ *as* $|x| \to \infty$ *and*

$$C = \sup_n \int \phi(x) dF_n(x) < \infty$$

then F_n *is tight.*

Proof.

$1 - F_n(M) + F_n(-M) \leq \frac{C}{\inf_{|x| \geq M} \phi(x)}$. □

Metrics on distributions

Definition

The *Lévy Metric* on two distribution functions is defined by

$$\rho(F, G) = \inf\{\epsilon : F(x-\epsilon) - \epsilon \leq G(x) \leq F(x+\epsilon) + \epsilon \text{ for all } x\}.$$

One has $\rho(F_n, F) \to 0$ if and only if $F_n \Rightarrow F$.

Metrics on distributions

Definition

The *Ky Fan Metric* on two distribution functions is defined by

$$\alpha(X, Y) = \inf\{\epsilon \geq 0 : \text{Prob}(|X - Y| > \epsilon) \leq \epsilon\}.$$

Exercise

Check that the distribution functions F, G of random variables X, Y, satisfy $\rho(F, G) \leq \alpha(X, Y)$.

Convergence in distribution

Theorem

If each subsequence of $\{X_n\}$ has a sub-subsequence which converges in distribution, then $\{X_n\}$ converges in distribution.

Proof.

This follows on applying the Lévy metric. □

Uniqueness of the characteristic function

Theorem (The inversion formula)

Let $\phi(t) = \int e^{itx}\mu(dx)$ *where* μ *is a probability measure. If* $a < b$, *then*

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

Uniqueness of the characteristic function

Proof.

- Let
$$R(\theta, T) = \int_{-T}^{T} \frac{\sin \theta t}{t} dt = 2\int_0^{T\theta} \frac{\sin x}{x} dx = 2S(T\theta).$$
- As $T \to \infty$, $S(T) \to \frac{\pi}{2}$, so $R(\theta, T) \to \pi \operatorname{sgn} \theta$. Thus
$$R(x-a, T) - R(x-b, T) \to \begin{cases} 2\pi & a < x < b \\ \pi & x = a, b \\ 0 & \text{otherwise} \end{cases}.$$

□

Uniqueness of the characteristic function

Proof.

- Calculate

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi(t)dt = \frac{1}{2\pi}\int_{-T}^{T}\int\frac{e^{-ita}-e^{-itb}}{it}e^{itx}\mu(dx)dt$$

$$\frac{1}{2\pi} = \int\left[\int_{-T}^{T}\frac{\sin(t(x-a))-\sin(t(x-b))}{t}dt\right]\mu(dx)$$

$$= \frac{1}{2\pi}\int(R(x-a,T)-R(x-b,T))\mu(dx).$$

- The claim follows by bounded convergence, since $\frac{e^{-ita}-e^{-itb}}{it} = \int_a^b e^{-itx}dx$.

□

Characteristic functions

Theorem

If $\int |\phi(t)| dt < \infty$*, then* μ *has bounded continuous density*

$$f(y) = \frac{1}{2\pi} \int e^{-ity} \phi(t) dt.$$

Proof.

Check

$$\mu((a,b)) + \frac{1}{2}\mu(\{a,b\}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it}\phi(t)dt$$
$$\leq \frac{b-a}{2\pi}\int_{-\infty}^{\infty} |\phi(t)|dt.$$

Hence μ does not have atoms. □

Proof.

Calculate

$$\begin{aligned}\mu(x, x+h) &= \frac{1}{2\pi} \int \frac{e^{-itx} - e^{-it(x+h)}}{it} \phi(t) dt \\ &= \frac{1}{2\pi} \int \left(\int_x^{x+h} e^{-ity} dy \right) \phi(t) dt \\ &= \int_x^{x+h} \left(\frac{1}{2\pi} \int e^{-ity} \phi(t) dt \right) dy.\end{aligned}$$

Continuity of the integrand follows from dominated convergence. □

Method of characteristic functions

Theorem

Let μ_n, $1 \leq n \leq \infty$ be probability measures with characteristic functions ϕ_n.

1. *If $\mu_n \Rightarrow \mu$ then $\phi_n(t) \to \phi(t)$ for all t.*
2. *If $\phi_n(t)$ converges pointwise to a limit $\phi(t)$ that is continuous at 0, then the associated sequence of measures is tight, and converges weakly to the measure μ with characteristic function ϕ.*

Method of characteristic functions

Proof.

- Item 1 is immediate.
- $\int_{-u}^{u} 1 - e^{itx} dt = 2u - \frac{2 \sin ux}{x}$.
- Hence

$$u^{-1} \int_{-u}^{u} (1 - \phi_n(t)) dt = 2 \int \left(1 - \frac{\sin ux}{ux}\right) \mu_n(dx)$$
$$\geq 2 \int_{|x| \geq \frac{2}{u}} \left(1 - \frac{1}{|ux|}\right) \mu_n(dx)$$
$$\geq \mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right).$$

- Since $\phi(0) = 1$ and ϕ is continuous at 0, the corresponding integral against ϕ tends to 0 as $u \to 0$.

□

Method of characteristic functions

Proof.

- Given $\epsilon > 0$, let u sufficiently small so that

$$\frac{1}{u}\int_{-u}^{u}(1-\phi(t))dt < \epsilon.$$

  By monotone convergence, the same bound, but replacing ϵ with 2ϵ, holds for ϕ_n for all n sufficiently large. Hence $\{\mu_n\}$ is tight.
- By tightness, any subsequence of $\{\mu_n\}$ has a further subsequence which is convergent in distribution. Hence this subsequence has characteristic function converging to ϕ, which is the characteristic function of its limiting measure μ.
- The convergence in general now follows from the Lévy metric.

□

Method of characteristic functions

Theorem

If $\int |x|^n \mu(dx) < \infty$, then the characteristic function ϕ has n continuous derivatives, and

$$\phi^{(n)}(t) = \int (ix)^n e^{itx} \mu(dx).$$

Proof.

Exercise. □

Method of characteristic functions

The following estimate is obtained from Taylor's theorem with remainder.

Lemma

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

Method of characteristic functions

Theorem

If $\mathrm{E}\left[|X|^2\right] < \infty$, *then*

$$\phi(t) = 1 + it\,\mathrm{E}[X] - \frac{t^2}{2}\mathrm{E}\left[X^2\right] + o(t^2).$$

Proof.

The error term is $\leq t^2\,\mathrm{E}\left[|t||X|^3 \wedge 2|X|^2\right]$. This tends to 0 as $t \to 0$ by dominated convergence. □

Method of characteristic functions

Theorem

If $\limsup_{h\downarrow 0} \frac{\phi(h)-2\phi(0)+\phi(-h)}{h^2} > -\infty$, *then* $\mathrm{E}\left[|X|^2\right] < \infty$.

Method of characteristic functions

Proof.

We have $(e^{ihx} - 2 + e^{-ihx})/h^2 = -2(1 - \cos hx)/h^2 \leq 0$ and $2(1 - \cos hx)/h^2 \to x^2$ as $h \to 0$. By Fatou and Fubini,

$$\int x^2 dF(x) \leq 2 \liminf_{h \to 0} \int \frac{1 - \cos hx}{h^2} dF(x)$$
$$= -\limsup_{h \to 0} \frac{\phi(h) - 2\phi(0) + \phi(-h)}{h^2} < \infty.$$

□

Polya's criteria

Theorem (Polya's criteria)

Let $\phi(t)$ be real non-negative and have $\phi(0) = 1$, $\phi(t) = \phi(-t)$ and ϕ is decreasing and convex on $(0, \infty)$ with

$$\lim_{t \downarrow 0} \phi(t) = 1, \qquad \lim_{t \uparrow \infty} \phi(t) = 0.$$

Then there is a probability measure ν on $(0, \infty)$, so that

$$\phi(t) = \int_0^\infty \left(1 - \left|\frac{t}{s}\right|\right)^+ \nu(ds).$$

This exhibits ϕ as the convex combination of characteristic functions of probability measures, hence as the characteristic function of a probability measure.

Polya's criteria

Proof.

- Since ϕ is convex, it's right derivative

$$\phi'(t) = \lim_{h\downarrow 0} \frac{\phi(t+h)-\phi(t)}{h}$$

  exists and is right continuous and increasing.

- Let μ be the measure $\mu(a,b] = \phi'(b) - \phi'(a)$ for all $0 \leq a < b < \infty$. Define ν by $\frac{d\nu}{d\mu}(s) = s$.

- $\phi'(t) \to 0$ as $t \to \infty$, so

$$-\phi'(s) = \int_s^\infty \frac{\nu(dr)}{r}.$$

□

Polya's criteria

Proof.

- By Fubini's theorem

$$\phi(t) = \int_t^\infty \int_s^\infty \frac{\nu(dr)}{r} ds = \int_t^\infty r^{-1} \int_t^r ds \nu(dr)$$
$$= \int_t^\infty \left(1 - \frac{t}{r}\right) \nu(dr) = \int_0^\infty \left(1 - \frac{t}{r}\right)^+ \nu(dr).$$

- The result follows on using $\phi(-t) = \phi(t)$.

□

The Moment problem

- Suppose $\int x^k dF_n(x)$ has limit μ_k for each k.
- This implies that $\{F_n\}$ is tight, and every subsequential limit has moments μ_k
- If there is a unique distribution function F with moments μ_k, then it follows that $F_n \Rightarrow F$.
- The *moment problem* asks under which conditions the moments of a measure are unique.

The Moment problem

The *lognormal density* is

$$f_0(x) = \frac{\exp\left(-\frac{(\log x)^2}{2}\right)}{x\sqrt{2\pi}}, \qquad x \geq 0.$$

Define in $-1 \leq a \leq 1$,

$$f_a(x) = f_0(x)[1 + a\sin(2\pi \log x)].$$

Theorem

The densities f_a, $-1 \leq a \leq 1$ have the same moments.

The Moment problem

Proof.

- It suffices to check

$$\int_0^\infty x^r f_0(x) \sin(2\pi \log x) dx = 0$$

for $r = 0, 1, 2, \ldots$.

- Make the change of variables $s = \log x - r$, $ds = \frac{dx}{x}$ to write the integral as

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp(rs + r^2) \exp\left(-\frac{(s+r)^2}{2}\right) \sin(2\pi(r+s)) ds$$

$$= \frac{\exp\left(\frac{r^2}{2}\right)}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{s^2}{2}\right) \sin(2\pi s) ds = 0.$$

□

Carleman's condition

Theorem

If $\limsup_{k\to\infty} \frac{\mu_{2k}^{\frac{1}{2k}}}{2k} = r < \infty$, *then there is at most one density function* F *with* $\mu_k = \int x^k dF(x)$ *for all positive integers* k.

Carleman's condition is only slightly weaker,

$$\sum_{k=1}^{\infty} \frac{1}{\mu_{2k}^{\frac{1}{2k}}} = \infty.$$

Carleman's condition

Proof.

- Let $\nu_k = \int |x|^k dF(x)$. Then $\nu_{2k+1}^2 \leq \mu_{2k}\mu_{2k+2}$, so

$$\limsup_{k\to\infty} \frac{\nu_k^{\frac{1}{k}}}{k} = r < \infty.$$

- By Taylor's theorem

$$\left| e^{i\theta X}\left(e^{itX} - \sum_{m=0}^{n-1} \frac{(itX)^m}{m!}\right)\right| \leq \frac{|tX|^n}{n!}.$$

□

Carleman's condition

Proof.

- The characteristic function satisfies

$$\left|\phi(\theta + t) - \phi(\theta) - t\phi'(\theta) - \ldots - \frac{t^{n-1}}{(n-1)!}\phi^{(n-1)}(\theta)\right| \le \frac{|t|^n}{n!}\nu_n.$$

- Since $\nu_k \le (r+\epsilon)^k k^k$ for all k sufficiently large, and $e^k \ge \frac{k^k}{k!}$, we obtain

$$\phi(\theta + t) = \phi(\theta) + \sum_{m=1}^{\infty} \frac{t^m}{m!}\phi^{(m)}(\theta), \qquad |t| < \frac{1}{er}.$$

- The uniqueness now follows from the fact that a distribution is determined by its characteristic function.

□

The central limit theorem

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d.* $\mathrm{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2 \in (0, \infty)$. *If* $S_n = X_1 + \cdots + X_n$ *then*

$$\frac{S_n - n\mu}{\sigma n^{\frac{1}{2}}} \Rightarrow \eta$$

where η *is the standard normal distribution.*

The central limit theorem

Proof.

- By subtracting the mean, we can assume $\mu = 0$.
- We have

$$\phi(t) = \mathrm{E}\left[e^{itX_1}\right] = 1 - \frac{\sigma^2 t^2}{2} + o(t^2).$$

- For each t,

$$\mathrm{E}\left[\exp\left(\frac{itS_n}{\sigma n^{\frac{1}{2}}}\right)\right] = \left(1 - \frac{t^2}{2n} + o(n^{-1})\right)^n \to e^{-\frac{t^2}{2}}.$$

□

The Lindeberg-Feller Theorem

Theorem (The Lindeberg-Feller Theorem)

For each n, *let* $X_{n,m}$, $1 \leq m \leq n$, *be independent random variables with* $\mathrm{E}[X_{n,m}] = 0$. *Suppose*

1. $\sum_{m=1}^{n} \mathrm{E}[X_{n,m}^2] \to \sigma^2 > 0$
2. *For all* $\epsilon > 0$, $\lim_{n\to\infty} \sum_{m=1}^{n} \mathrm{E}\left[|X_{n,m}|^2 \mathbf{1}(|X_{n,m}| > \epsilon)\right] = 0.$

Then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\eta$ *as* $n \to \infty$.

The Lindeberg-Feller Theorem

Proof.

- Let $\phi_{m,n}(t) = \mathrm{E}\left[e^{itX_{n,m}}\right]$, $\sigma_{n,m}^2 = \mathrm{E}\left[X_{n,m}^2\right]$.
- We have, by Taylor expansion

$$\begin{aligned}\left|\phi_{n,m}(t) - \left(1 - \frac{t^2\sigma_{n,m}^2}{2}\right)\right| &\leq \mathrm{E}\left[|tX_{n,m}|^3 \wedge 2|tX_{n,m}|^2\right] \\ &\leq \mathrm{E}\left[|tX_{n,m}|^3\mathbf{1}(|X_{n,m}| \leq \epsilon)\right] + \mathrm{E}\left[2|tX_{n,m}|^2\mathbf{1}(|X_{n,m}| > \epsilon)\right] \\ &\leq \epsilon t^3\,\mathrm{E}\left[|X_{n,m}|^2\mathbf{1}(|X_{n,m}| \leq \epsilon)\right] + 2t^2\,\mathrm{E}\left[|X_{n,m}|^2\mathbf{1}(|X_{n,m}| > \epsilon)\right].\end{aligned}$$

- Using the second condition, we have

$$\limsup_{n\to\infty}\sum_{m=1}^n\left|\phi_{n,m}(t) - \frac{1 - t^2\sigma_{n,m}^2}{2}\right| \leq \epsilon|t|^3\sigma^2.$$

□

The Lindeberg-Feller Theorem

Proof.

- Since $\epsilon > 0$ was arbitrary,

$$\left| \prod_{m=1}^{n} \phi_{n,m}(t) - \prod_{m=1}^{n} \left(1 - \frac{t^2 \sigma_{n,m}^2}{2}\right) \right| \to 0$$

  as $n \to \infty$.

- Since $\sup_m \sigma_{n,m}^2 \to 0$ as $n \to \infty$,

$$\prod_{m=1}^{n} \left(1 - \frac{t^2 \sigma_{n,m}^2}{2}\right) \to \exp\left(-\frac{\sigma^2 t^2}{2}\right)$$

  as $n \to \infty$, so $\prod_{m=1}^{n} \phi_{n,m}(t) \to \exp\left(-\frac{\sigma^2 t^2}{2}\right)$ as $n \to \infty$, which proves the convergence.

□