To do this, ~~set~~ take $\phi_{X_1}(t) = E[e^{itX_1}]$. Then as $n \to \infty$, use a Taylor expansion and $\underbrace{\text{Prop 11.0.1}}_{\phi_X^{(j)}(0) = i^j E[X^j]}$ to get

$$\phi_n(t) = E\left[e^{it(X_1 + \cdots + X_n)/\sqrt{n}}\right]$$

$$= E\left[e^{i(t/\sqrt{n})X_1} e^{i(t/\sqrt{n})X_2} \cdots e^{i(t/\sqrt{n})X_n}\right]$$

$$= E\left[e^{i(t/\sqrt{n})X_1}\right]^n \quad \text{since } X_i\text{'s are i.i.d}$$

$$= \left(\phi_{X_1}(t/\sqrt{n})\right)^n$$

$$= \Big(1 + \frac{it}{\sqrt{n}} \underbrace{E[X_1]}_{= 0 \text{ by assumption}} + \frac{1}{2!}\left(\frac{it}{\sqrt{n}}\right)^2 \overbrace{E[X_1^2]}^{= 1 \text{ by assumption}} + o\left(\frac{1}{n}\right)\Big)^n$$

by Taylor expansion of $\phi_{X_1}(t/\sqrt{n})$ & Prop 11.0.1 (j^{th} deriv. at $t=0$)

$$= \left(1 + \frac{i^2 t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$= \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

since $i^2 = -1$

$$\to e^{-t^2/2} \text{ as } n \to \infty, \text{ as claimed.}$$

[Note that $o\left(\frac{1}{n}\right)$ means a quantity q_n s.t. $\frac{q_n}{1/n} \to 0$ as $n \to \infty$]

"little o" notation

4/26/18
More Details

Taylor Series of $f(x)$

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$$= f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(x)}{2!}(x-a)^2 + \underbrace{\frac{f'''(a)}{3!}(x-a)^3 + \cdots}_{\text{higher order terms}}$$

So apply this to $\phi_{X_1}(t/\sqrt{n}) = E\left[e^{it/\sqrt{n}\, X_1}\right]$

$t/\sqrt{n}$ is "x" in above expression, evaluate at $a=0$

$$\phi_{X_1}(t/\sqrt{n}) = \underbrace{\phi_{X_1}(0)}_{=\,1} + \frac{\phi_{X_1}'(0)}{1!}(t/\sqrt{n}) + \frac{\phi_{X_1}''(0)}{2!}(t/\sqrt{n})^2 + \underbrace{\text{h.o.t.}}_{o(\frac{1}{n})}$$

Recall that $\phi_{X_1}^{(j)}(0) = i^j E[X_1^j]$

$$\Rightarrow \phi_{X_1}(t/\sqrt{n}) = 1 + i\overbrace{E[X_1]}^{=\,0}(t/\sqrt{n}) + \frac{i^2}{2!}\overbrace{E[X_1^2]}^{=\,1}\, t^2/n + o\left(\tfrac{1}{n}\right)$$

$$= 1 + it/\sqrt{n}\,(0) - t^2/2n\,(1) + o\left(\tfrac{1}{n}\right)$$

why $o(\frac{1}{n})$?
(of order smaller than the order of the last term in the expansion, here this is $c\cdot\frac{1}{n}$)

This means a function $g(n)$ s.t. $\frac{1}{n}$ grows much faster than $g(n)$

i.e. $\dfrac{g(n)}{\frac{1}{n}} \to 0$ as $n\to\infty$

For example, $g(n) = c\cdot n^{-3/2}$ or $c\cdot n^{-2}$ or $c\cdot n^{-\alpha}$ where $\alpha \geq 3/2$

$$\left[\lim_{n\to\infty} \frac{c \cdot n^{-3/2}}{\frac{1}{n}} = \lim_{n\to\infty} \frac{c\cdot n^{-3/2}}{n^{-1}} = c \cdot \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0 \right]$$

Thus, $\phi_{X_1}(t/\sqrt{n}) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right) \to e^{-t^2/2}$ as $n\to\infty$

why?

Taylor expand

$$\Big(\underbrace{e^{-t^2/2n}}_{=f(t)}\Big)^n = \left[\underbrace{e^0}_{=1} + \cancel{\frac{e^{-t^2/2n}}{1!}\left(-\frac{t}{n}\right)\Big|_{t=0}}_{0} (t-0) + \frac{\left(-\frac{1}{n}\right)}{2!}(t-0)^2 + o\left(\frac{t^2}{n}\right)\right]^n$$

$$= \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right]^n$$

$$f'(t) = -\frac{t}{n} e^{-t^2/2n} \Rightarrow f'(0) = 0$$

$$f''(t) = -\frac{t}{n}\left(-\frac{t}{n}\right)e^{-t^2/2n} + e^{-t^2/2n}\left(-\frac{1}{n}\right)$$

$$= \frac{t^2}{n^2} e^{-t^2/2n} - \frac{1}{n} e^{-t^2/2n}$$

$$= e^{-t^2/2n}\left(\frac{t^2}{n^2} - \frac{1}{n}\right) \Rightarrow f''(0) = e^0\left(0 - \frac{1}{n}\right) = -\frac{1}{n}$$

__Note__: $e^{-t^2/2} = e^{-\frac{t^2 n}{2n}} = \left(e^{-\frac{t^2}{2n}}\right)^n$

Taylor expand this!

$$= 1 - \frac{t^2}{2} + o(t^2)$$

Also, $e^{-t^2/2} = \underbrace{e^0}_{1} + \underbrace{\cancel{\frac{e^{-t^2/2}}{1!}(-t)}\Big|_{t=0}}_{0} (t-0) + \frac{(-1)}{2!}(t-0)^2 + o(t^2)$

Lec 24 (1a)

See errata in Rosenthal

[~~Formally, the limit holds since for any $\varepsilon > 0$, for sufficiently large n we have $q_n \geq -\varepsilon/n$ and $q_n \leq \varepsilon/n$, so that $\liminf_n \phi_n(t) \geq e^{-t^2/2} - \varepsilon$ and $\limsup_n \phi_n(t) \leq e^{-t^2/2} + \varepsilon$.~~ $\blacksquare$]

The classical CLT immediately implies the following result (simpler version, perhaps).

<u>Cor 11.2.3</u>: Let $X_1, X_2, \ldots$ be i.i.d. RVs with finite mean m & finite variance v. Let $S_n = X_1 + \cdots + X_n$. Then for each fixed $x \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left(\frac{S_n - nm}{\sqrt{nv}} \leq x\right) = \Phi(x).$$

CDF for $N(0,1)$

i.e. $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$

<u>Note</u>: This can also be written as

$$P(S_n \leq nm + x\sqrt{nv}) \to \Phi(x) \quad \text{as } n\to\infty.$$

In words, $\underbrace{X_1 + \cdots + X_n}_{S_n}$ is approx. equal to nm with deviations from this value of order $\sqrt{n}$.

Std deviation of S_n is $\sqrt{vn}$

<u>Remark</u>: It is not essential in the CLT to divide by $\sqrt{v}$. Without doing so, the theorem asserts instead that

$$\mathcal{L}\left(\frac{S_n - nm}{\sqrt{n}}\right) \Rightarrow N(0, \underline{v})$$

weak convergence to Normal dist'n with mean 0, variance v

<u>Remark</u>: The CLT $\Rightarrow$ WLLN.

[For details, see Rosenthal / Billingsley]

Generalizations of CLT

Classical CLT (Thm 11.2.2 & Cor 11.2.3) is very useful in many areas of science, however it has some limitations:

- No quantitative bounds on the convergence in Cor 11.2.3.
- Condition that RVs are i.i.d. is sometimes too restrictive.

- The <u>Berry-Esseen Theorem</u> solves the first problem above.
  ↳ $X_1, X_2, \ldots$ i.i.d. with finite mean m, positive finite variance v, (i.e. not $= 0$) and $E[|X_i - m|^3] = \rho < \infty$. Then

$$\left| P\left(\frac{S_n - nm}{\sqrt{nv}} \le x\right) - \Phi(x) \right| \le \frac{3\rho}{\sqrt{nv^3}}.$$

This theorem provides a bound on the convergence in Cor 11.2.3 that depends only on the third moment.

- For the 2^{nd} problem above, we will discuss 2 of many results addressing the i.i.d. restriction.

Consider triangular arrays – collections $\{Z_{nk} : n \geq 1, 1 \leq k \leq r_n\}$ of random variables s.t. for each row, $\{Z_{nk}\}_{1 \leq k \leq r_n}$ is mutually indep. (If $r_n = n$ they form an actual triangle).

Note:
- Not assumed that RVs in each row are identically distributed.
- Not assumed that different rows are indep.

Assume (for simplicity) that $E[Z_{nk}] = 0$ for $\forall n, k$.

Set $\sigma_{nk}^2 = E[Z_{nk}^2] < \infty$ (assumed finite),

$S_n = Z_{n1} + \cdots + Z_{nr_n}$, and

$s_n^2 = Var(S_n) = \sigma_{n1}^2 + \cdots + \sigma_{nr_n}^2$.

① For such a triangular array, the Lindeberg CLT states that $\mathcal{L}(S_n/s_n) \Rightarrow N(0,1)$ provided that for each $\varepsilon > 0$,

$$\lim_{n \to \infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} E\left[Z_{nk}^2 \mathbb{1}_{|Z_{nk}| \geq \varepsilon s_n} \right] = 0.$$

In words, as $n \to \infty$, the tails of the Z_{nk} contribute less & less to the variance of S_n.

Q. If Lindeberg condition is NOT satisfied, then what other limiting distributions may arise?

def: A distribution μ is a <u>possible limit</u> if $\exists$ a triangular array (as defined above) with $\sup_n s_n^2 < \infty$ and

$\lim_{n\to\infty} \max_{1\le k\le r_n} \sigma_{nk}^2 = 0$ ← s.t. no 1 term dominates the contribution to $Var(S_n)$

such that $\mathcal{L}(S_n) \Rightarrow \mu$. (By assuming $s_n^2 = 1$ so $\frac{S_n}{s_n^2} = S_n$ — See B p. 372)

Q. What distributions are possible limits?

A. Normal distributions $N(0,v)$, others?
(when Lindeberg cond. is satisfied & $s_n^2 \to v$)

② Yes, infinitely divisible distributions having mean 0 & finite variance

def: A distribution μ is <u>infinitely divisible</u> if $\forall\, n \in \mathbb{N}$, $\exists$ a distribution ν_n s.t. the n-fold convolution of ν_n equals μ (i.e. $\nu_n * \nu_n * \cdots * \nu_n = \mu$).

Recall: this means that if $X_1, \ldots, X_n \sim \nu_n$ are indep., then $X_1 + \cdots + X_n \sim \mu$.

If μ is infinitely divisible, then we can take $r_n = n$ and $\mathcal{L}(X_{nk}) = \nu_n$ in the triangular array to get that $\mathcal{L}(S_n) \Rightarrow \mu$.

Proof of converse, see Billingsley.

Note: A stronger condition than Lindeberg's that is often easier to check is the Lyapounov condition:

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{r_n} E\left[|Z_{nk}|^{2+\delta}\right] = 0 \quad \text{for some } \delta > 0.$$

$\left(|Z_{nk}|^{2+\delta} \text{ integrable}\right)$

Lemma: Lyapounov's condition $\Rightarrow$ Lindeberg's condition.

Pf: See Billingsley.

---

* Recall def of convolution:

indep. continuous distributions w/ densities f & g

$$(f*g)(t) = \int_{-\infty}^{\infty} f(z)\, g(t-z)\, dz = \int_{-\infty}^{\infty} f(t-z)\, g(z)\, dz$$

More on Infinitely Divisible Distributions

(Ref: § 28 Billingsley)

Examples:

1. μ is a mass of v (recall: variance) at the origin, characteristic function $\phi(t) = e^{-vt^2/2}$.
   (i.e. char. fct. of a centered normal distribution $N(0,v)$).

2. Poisson distribution

3. Gamma & exponential distributions

4. Cauchy dist'n

Thm: Suppose that

$$\phi(t) = \exp\left(\int_{\mathbb{R}} (e^{itx} - 1 - itx) \frac{1}{x^2} \, d\mu(x) \right)$$

where μ is a finite measure. Then ϕ is the characteristic function of an infinitely divisible distribution with mean 0 and variance $\mu(\mathbb{R})$.

skip if needed

Example: Suppose $Z_\lambda \sim \text{Poisson}(\lambda)$ and that

$X_{n1}, \dots, X_{nn}$ are independent s.t.

$$P(X_{nk} = 1) = \frac{\lambda}{n}, \quad P(X_{nk} = 0) = 1 - \frac{\lambda}{n} \quad \text{for } 1 \le k \le n.$$

(Poisson approx to binomial (here $p = \frac{\lambda}{n}$) example in Billingsley 25.2)

$$X_{n1} + \cdots + X_{nn} \Rightarrow \underbrace{Z_\lambda}_{\text{follows Poisson dist'n}}$$

This contrasts with CLT in which limit law is Normal.

Poisson dist'n is a "possible limit law" for indep. triangular arrays.

Weak Convergence & Method of Moments (Ref: §11.4 R, §30 B)

* There is another way of proving weak convergence of probability measures which does not explicitly use characteristic functions or the continuity theorem.

  → Instead it uses moments.

Note: This "method of moments" is not to be confused with the statistical estimation procedure with the same name!

def: A distribution μ is determined by its moments if all its moments are finite & if no other distribution has identical moments. i.e. $\left(\int |x^k| d\mu(x) < \infty \ \forall k \in \mathbb{N} \text{ and } \int x^k d\mu(x) = \int x^k d\nu(x) \ \forall k \in \mathbb{N} \Rightarrow \mu = \nu \right)$

Thm 11.4.1: Suppose that μ is determined by its moments. Let $\{\mu_n\}$ be a sequence of distributions s.t.

$$\int x^k d\mu_n(x) < \infty \quad \forall n, k \in \mathbb{N} \quad \text{and s.t.}$$

$$\lim_{n \to \infty} \int x^k d\mu_n(x) = \int x^k d\mu(x) \quad \text{for each } k \in \mathbb{N}.$$

Then $\mu_n \Rightarrow \mu$ as $n \to \infty$.

[In words, for distributions determined by their moments, converge of moments $\Rightarrow$ weak convergence of dist'ns.]