Def A collection $(\nu_\alpha)_{\alpha \in I}$ of probability measures on $\mathbb{R}$ is *tight* if for all $\varepsilon > 0$ there exists an $R > 0$ such that $\nu_\alpha[\mathbb{R} \setminus [-R,R]] < \varepsilon \quad \forall \alpha \in I$

Remark We require that the overwhelming majority of mass of all the members of the collection is carried by the same compact set.

This is a sort of (pre-)compactness of the collection of measures in the topology of weak convergence. (It guarantees that we can extract subsequential limits)

Exercise (ex. 5/week 2) If $(F_n)_{n\in\mathbb{N}}$ is a sequence of c.d.f.'s associated to a tight family of probability measures on $\mathbb{R}$, then there exists a subsequence $(F_{n_k})_{k\in\mathbb{N}}$ that converges pointwise to a c.d.f F at all continuity points of F.

Theorem Let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{R}$, and (χ_n) the associated characteristic functions $\chi_n(\theta) = \int_{\mathbb{R}} e^{i\theta x}\, d\nu_n(x)$.

Then (ν_n) converges weakly iff $\chi_n(\theta) \longrightarrow \chi(\theta) \quad \forall \theta \in \mathbb{R}$, where the function $\chi : \mathbb{R} \to \mathbb{C}$ is continuous at 0.

Remarks

- This proves the implications (i) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (i) of the theorem about equivalent formulations of weak convergence. In fact the statement about the latter is stronger — we do not require ~~a priori~~ explicitly that χ is a characteristic fn of a measure.
- The proof strategy of the non-trivial "if" part consists of two steps: to show the convergence of $(\nu_n)_{n\in\mathbb{N}}$
  1. Show precompactness (tightness).
     — every subsequence has a further converging subsequence
  2. Characterize explicitly any subsequential limit
     — all converging subsequences have the same limit

Proof: "only if": $x \mapsto e^{i\theta x}$ is continuous and bounded, so

$$\chi_n(\theta) = \int_{\mathbb{R}} e^{i\theta x}\, d\nu_n(x) \underset{\substack{n\to\infty \\ \uparrow \\ \text{since } \nu_n \xrightarrow{w} \nu}}{\longrightarrow} \int_{\mathbb{R}} e^{i\theta x}\, d\nu(x) =: \chi(\theta)$$

continuous at $\theta = 0$ as a characteristic fn

"if" Suppose $\chi_n(\theta) \longrightarrow \chi(\theta)$ $\forall \theta \in \mathbb{R}$
and $\lim_{\theta\to 0} \chi(\theta) = 1$.

To show tightness, use the auxiliary calculation

$$\int_{-u}^{u} (1 - e^{i\theta x})\, d\theta = 2u - 2\,\frac{\sin(ux)}{x}.$$

Divide by u, integrate against ν_n and use Fubini:

$$\frac{1}{u}\int_{-u}^{u} (1 - \chi_n(\theta))\, d\theta = 2\int_{\mathbb{R}} \underbrace{\left(1 - \frac{\sin(ux)}{ux}\right)}_{\geq 0} d\nu_n(x)$$

because $\left|\frac{\sin(\xi)}{\xi}\right| \leq 1$

$$\geq 2 \int_{\mathbb{R}\setminus(\frac{-2}{u},\frac{2}{u})} \underbrace{\left(1 - \frac{1}{ux}\right)}_{\geq \frac{1}{2}} d\nu_n(x)$$

we used $|\sin(\xi)| \leq 1$

$$\geq \nu_n\left[\mathbb{R}\setminus\left(\tfrac{-2}{u}, \tfrac{2}{u}\right)\right]$$

An upper bound for the mass of the complement of an interval written in terms of the characteristic fn.

By the assumption $\chi(\theta) \to 1$ as $\theta \to 0$ we can find u small enough so that

$$\frac{1}{u}\int_{-u}^{u} (1 - \chi(\theta))\, d\theta < \varepsilon$$

By $\chi_n(\theta) \longrightarrow \chi(\theta)$ and dominated convergence (constant function dominates)

for n large $$\frac{1}{u}\int_{-u}^{u} (1 - \chi_n(\theta))\, d\theta < 2\varepsilon.$$

We conclude $\nu_n\left[\mathbb{R}\setminus(\frac{-2}{u},\frac{2}{u})\right] < 2\varepsilon$ for n large and tightness of $(\nu_n)_{n\in\mathbb{N}}$ follows.

By tightness we know that subsequential limits of $(\nu_n)_{n\in\mathbb{N}}$ exist. Suppose $(\nu_{n_k})_{k\in\mathbb{N}}$ is a convergent subseq, $\nu_{n_k} \xrightarrow[k\to\infty]{} \nu$.
Then by the same argument as in "only if" we find that the characteristic function of ν is χ.
Therefore every convergent subsequence has the same limit. $\square$

I.4 SUMS OF INDEPENDENT RANDOM VARIABLES

$(X_j)_{j\in\mathbb{N}}$ i.i.d. (abbreviation for "independent and identically distributed")

$$S_n := \sum_{j=1}^{n} X_j$$

We will discuss two limit theorems $(n\to\infty)$

(LLN) Law of large numbers : $\frac{1}{n} S_n \xrightarrow[n\to\infty]{} \mu = \mathbb{E}[X_j]$

- the average of a large number of independently repeated experiments is close to the expected value

(CLT) Central limit theorem : $\frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{w} N(0,\sigma^2)$, $\sigma^2 = \mathrm{Var}[X_j]$

- gives the random fluctuations of the average around the limit value
  - * the sum S_n fluctuates in scale $\sqrt{n}$
  - * the shape of the fluctuations is Gaussian

Laws of large numbers

Theorem (Strong law of large numbers for square integrable i.i.d. summands)
Let $(X_j)_{j\in\mathbb{N}}$ be i.i.d. and assume that
$$\mu = \mathbb{E}[X_j] \quad \text{and} \quad \sigma^2 = \mathrm{Var}[X_j] < \infty .$$
Define $S_n = \sum_{j=1}^{n} X_j$.
Then $\frac{1}{n} S_n \xrightarrow[n\to\infty]{} \mu$ almost surely.

Proof (see lecture notes or any textbook) $\square$

Remark "almost surely" = "with probability 1"
Almost sure convergence [redacted] $Y_n \xrightarrow[n\to\infty]{a.s.} Y$ means

$$\mathbb{P}\Big[\underbrace{\big\{\omega\in\Omega \;\big|\; \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\big\}}_{\text{the event that convergence happens}}\Big] = 1 .$$

Example Monte Carlo integration

$f : [0,1] \longrightarrow \mathbb{R}$ s.t. $\int_0^1 |f(x)|\,dx < \infty$

$\boxed{\int_0^1 f(x)\,dx = ?}$

(to apply the previous formulation of LLN assume also $\int_0^1 |f(x)|^2\,dx < \infty$)

$(U_j)_{j\in\mathbb{N}}$ i.i.d

$U_j \sim \mathrm{Unif}([0,1])$ $\longleftarrow$ can be sampled on a computer

Note: $(f(U_j))_{j\in\mathbb{N}}$ is i.i.d. and $\mathbb{E}[f(U_j)] = \int_0^1 f(x)\,dx$

(and $\mathrm{Var}[f(U_j)] < \infty$)

Define $I_n = \frac{1}{n}\sum_{j=1}^{n} f(U_j)$. By LLN

$$I_n \longrightarrow \int_0^1 f(x)\,dx \quad \text{(almost surely)}$$

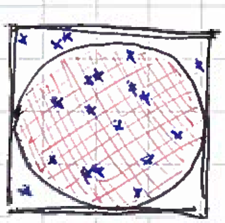
This gives a way to numerically evaluate the integral.

Generalizes to $\int_{[0,1]^d} f(x_1,x_2,\dots,x_d)\,dx_1\cdots dx_d$.

For example in high dimensional cases, this evaluation method is occasionally quite good.

Example Random approximation of π

Use the method above to calculate $\pi = \int_{[-1,1]^2} \mathbb{1}_{\{x^2+y^2\le 1\}}(x,y)\,dx\,dy$



with $n = 10\,000$ one obtains e.g.

$\pi \approx 3{,}1236$

$\pi \approx 3{,}1164$

$\pi \approx 3{,}1292$

$\pi \approx 3{,}1620$

$\vdots$

A different formulation:

Exercise (ex. 3/week 2) $(X_j)_{j\in\mathbb{N}}$ square integrable ($\mathbb{E}[X_j^2] < \infty$) uncorrelated ($\mathrm{Cov}(X_j, X_k) = 0 \ \forall j \neq k$) with $\mu = \mathbb{E}[X_j]$, $\mathrm{Var}[X_j] \le \text{const.}\ \forall j$

Then $\frac{1}{n} S_n \xrightarrow[n\to\infty]{\mathbb{P}} \mu$.

Remark: Stochastic convergence $Y_n \xrightarrow{\mathbb{P}} Y$ means $\forall \varepsilon > 0$ $\lim_{n\to\infty} \mathbb{P}[|Y_n - Y| \geq \varepsilon] \longrightarrow 0$.

Proposition If $X : \Omega \to \mathbb{R}$ is a r.v. with $\mathbb{E}[X^2] < \infty$ then as $\theta \to 0$ we have $\left(\chi(\theta) = \mathbb{E}[e^{i\theta X}] \right)$

$$\chi(\theta) = 1 + i\theta\, \mathbb{E}[X] - \frac{\theta^2}{2} \mathbb{E}[X^2] + o(\theta^2)$$

Proof: The idea is to Taylor expand $e^{i\theta X}$, which is achieved by

$$\begin{aligned} e^{i\theta x} &= 1 + i\theta \int_0^x e^{i\theta u}\, du && \leftarrow \text{easy } \cancel{\text{elementary}} \text{ integral} \\ &= 1 + i\theta \int_0^x du \left(1 + i\theta \int_0^u dv\, e^{i\theta v} \right) && \leftarrow \text{repeat the same} \\ &= 1 + i\theta x - \frac{\theta^2 x^2}{2} - \theta^2 \int_0^x du \int_0^u dv\, (e^{i\theta v} - 1) \end{aligned}$$

Then we are ready to look at the claim :

$$\mathbb{E}\left[e^{i\theta X} - \left(1 + i\theta X - \frac{\theta^2}{2} X^2 \right) \right] = -\theta^2\, \mathbb{E}\left[\int_0^X du \int_0^u dv\, (e^{i\theta v} - 1) \right]$$

Take absolute values and use "triangle inequality for integrals"

$$\left| \chi(\theta) - \left(1 + i\theta\, \mathbb{E}[X] - \frac{\theta^2}{2} \mathbb{E}[X^2] \right) \right| = |\theta|^2 \cdot \left| \mathbb{E}\left[\int_0^X du \int_0^u dv\, (e^{i\theta v} - 1) \right] \right|$$

$$\leq |\theta|^2\, \mathbb{E}\left[\int_0^{|X|} du \int_0^u dv\, |e^{i\theta v} - 1| \right].$$

Denote $R(\theta, x) = \int_0^x du \int_0^u dv\, \underbrace{|e^{i\theta v} - 1|}_{= 2\sin(\frac{\theta v}{2}) \leq \min\{2, |\theta v|\}} \leq \min\left\{ x^2, \frac{1}{6} |\theta|\, |x|^3 \right\}.$

so $\left| \chi(\theta) - \left(1 + i\theta\, \mathbb{E}[X] - \frac{\theta^2}{2} \mathbb{E}[X^2] \right) \right| \leq |\theta|^2\, \mathbb{E}[R(\theta, |X|)].$

We have $R(\theta, |X|) \leq |X|^2$ (integrable) so dominated convergence applies in

$$\lim_{\theta \to 0} \mathbb{E}[R(\theta, |X|)] \overset{\text{DCT}}{=} \mathbb{E}\Big[\underbrace{\lim_{\theta \to 0} R(\theta, |X|)}_{\substack{=0 \\ \text{since } R(\theta, x) \leq \frac{1}{6} |\theta|\, |x|^3}} \Big] = 0 \qquad \square$$

Theorem (Central limit theorem)
$(X_j)_{j\in\mathbb{N}}$ i.i.d. square integrable
$$\mathbb{E}[X_j] = \mu \ , \quad \mathrm{Var}[X_j] = \sigma^2$$

Then $$\frac{S_n - n\mu}{\sqrt{n}} \xrightarrow[n\to\infty]{w} N(0,\sigma^2).$$

Proof Without loss of generality assume $\mu = 0$ (otherwise consider $\tilde{X}_j = X_j - \mu$).
Let $$\chi(\theta) = \mathbb{E}[e^{i\theta X_j}] \underset{\uparrow}{=} 1 - \frac{\theta^2\sigma^2}{2} + o(\theta^2).$$
by previous Proposition

Then calculate the characteristic function of $\frac{S_n}{\sqrt{n}}$ using independence of (X_j)

$$\chi_{\frac{S_n}{\sqrt{n}}}(\theta) = \mathbb{E}\Big[\exp\Big(i\theta\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j\Big)\Big]$$
$$= \prod_{j=1}^{n} \mathbb{E}[e^{i\theta X_j/\sqrt{n}}] = \Big(\chi\Big(\frac{\theta}{\sqrt{n}}\Big)\Big)^n$$
$$= \Big(1 - \frac{\theta^2\sigma^2}{2n} + o\Big(\frac{\theta^2}{n}\Big)\Big)^n \xrightarrow[n\to\infty]{\text{Lemma before}} e^{-\frac{\sigma^2}{2}\theta^2}.$$

This is the characteristic function of $N(0,\sigma^2)$, see Exercise.
By the theorem about weak convergence and characteristic functions we have obtained
$$\frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{w} N(0,\sigma^2) \qquad \square$$