We will use the following auxiliary result in many calculations below.

Lemma If $c_n \in \mathbb{C}$ $(n \in \mathbb{N})$, $c_n \xrightarrow[n\to\infty]{} c$, then

$$\left(1+\frac{c_n}{n}\right)^n \xrightarrow[n\to\infty]{} e^c .$$

Proof: First use triangle inequality

$$\left|\left(1+\tfrac{c_n}{n}\right)^n - e^c\right| \le \left|\left(1+\tfrac{c_n}{n}\right)^n - e^{c_n}\right| - \left|e^{c_n}-e^c\right|$$

$\longrightarrow 0$ by continuity of exp

To estimate ~~do~~ the first term, use

$$|z^n - w^n| = |z-w| \cdot \left|z^{n-1} + z^{n-2}w + z^{n-3}w^2 + \dots + w^{n-1}\right|$$
$$\le |z-w| \cdot n \cdot M^{n-1} \qquad \text{if } |z|,|w| \le M .$$

With $z = 1+\frac{c_n}{n}$ and $w = e^{c_n/n}$ and any $\gamma > |c|$

we have $|z| \le e^{\gamma/n}$, $|w| \le e^{\gamma/n}$ for all large enough n, and thus

$$\left|\left(1+\tfrac{c_n}{n}\right)^n - e^{c_n}\right| \le \left|1+\tfrac{c_n}{n} - e^{c_n/n}\right| \cdot n \cdot \left(e^{\gamma/n}\right)^{n-1}$$

Finally note that by Taylor series remainder, we have

$$\left|e^{c_n/n} - \left(1+\tfrac{c_n}{n}\right)\right| = \mathcal{O}\left(\left(\tfrac{c_n}{n}\right)^2\right) \qquad \text{and obviously } \left(e^{\gamma/n}\right)^{n-1} \le e^{\gamma}.$$

We conclude that

$$\left|\left(1+\tfrac{c_n}{n}\right)^n - e^{c_n}\right| \le \mathcal{O}\left(\frac{c_n^2}{n}\right) \xrightarrow[n\to\infty]{} 0 \qquad \square$$

I.3 WEAK CONVERGENCE OF REAL RANDOM VARIABLES

In statistical physics, we often ~~consider~~ have the following situation

- a probabilistic model (random object) is defined,
  and it depends on a size parameter N
  (the number of microscopic constituents of the system)
- the size parameter in real world systems is large,
  which is modelled by considering the thermodynamical limit $N \to \infty$

▶ In which sense do the random objects converge as $N \to \infty$?

Appropriate notion: weak convergence of random variables / probability measures

We discuss weak convergence on $\mathbb{R}$ first
(later: in complete separable metric spaces)

Let's start by examples: in which sense do the following sequences of random variables converge?

<u>Examples</u>

(1) Continuous dependence on parameters

(a) $X_n \sim \mathrm{Exp}(\lambda_n)$. $\quad \lambda_n \xrightarrow[n\to\infty]{} \lambda$

in which sense $\quad X_n \xrightarrow[n\to\infty]{?} X \sim \mathrm{Exp}(\lambda)$

(b) $X_n \sim N(\mu_n, \sigma_n^2) \quad \mu_n \longrightarrow \mu, \quad \sigma_n^2 \longrightarrow \sigma^2$

$$X_n \xrightarrow[n\to\infty]{?} X \sim N(\mu, \sigma^2)$$

(c) $X_n \sim \mathrm{Bin}(N, p_n) \quad p_n \longrightarrow p$

$$X_n \xrightarrow[n\to\infty]{?} X \sim \mathrm{Bin}(n, p)$$

(2) The Poisson approximation of binomial distribution
(let $n \longrightarrow \infty$ but $p_n \longrightarrow 0$ $\quad p_n \approx \frac{\lambda}{n}$)
suppose $(p_n)_{n\in\mathbb{N}}$ is such that $n \cdot p_n \xrightarrow[n\to\infty]{} \lambda$.

Let $X_n \sim \mathrm{Bin}(n, p_n)$.

(Characteristic function of binomial distribution

$$\mathbb{P}[B=k] = p^k (1-p)^{n-k} \binom{n}{k}$$

$$\chi(\theta) = \mathbb{E}[e^{i\theta B}] = \sum_{k=0}^{n} p^k (1-p)^{n-k} \binom{n}{k} e^{i\theta k} = (1-p+pe^{i\theta})^n$$
)

Let χ_n be the characteristic function of X_n

$$\chi_n(\theta) = (1-p_n+p_n e^{i\theta})^n = \left(1 - \frac{\lambda(1+o(1))(1-e^{i\theta})}{n}\right)^n$$

$$\xrightarrow[n\to\infty]{} \exp\left(\lambda(e^{i\theta}-1)\right) = e^{-\lambda} e^{\lambda e^{i\theta}}$$

(Characteristic function of Poisson distribution

$$\mathbb{P}[P=k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\chi(\theta) = \mathbb{E}[e^{i\theta P}] = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{i\theta k} = e^{-\lambda} e^{\lambda e^{i\theta}}$$
)

in which sense $\quad \underset{\mathrm{Bin}(n,\, p_n)}{\underset{\wr}{X_n}} \xrightarrow[n\to\infty]{?} X \sim \mathrm{Poisson}(\lambda)$

(3) The minimum of independent uniform random variables:

Let $U_j \sim \mathrm{Unif}([0,1])$ for $j=1,2,3,\ldots$, independent ($\perp\!\!\!\perp$)

Denote $M_n = \min_{1\leq j\leq n} U_j$ and $X_n = n \cdot M_n$.

The cumulative distribution function of X_n is

$$F_n(x) = \mathbb{P}[X_n \leq x] = \mathbb{P}\left[M_n \leq \tfrac{x}{n}\right]$$
$$= 1 - \mathbb{P}\left[M_n > \tfrac{x}{n}\right] = 1 - \mathbb{P}\left[\forall j \in [\![1,n]\!] : U_j > \tfrac{x}{n}\right]$$
$$\overset{\perp\!\!\!\perp}{=} 1 - \left(1 - \tfrac{x}{n}\right)^n \xrightarrow[n\to\infty]{} 1 - e^{-x} = F_{\mathrm{Exp}(1)}(x)$$

(c.d.f. of exponential random variable)

$$X_n \xrightarrow{?} X \sim \mathrm{Exp}(1)$$

In the above examples we considered convergence of random variables in various senses:

- convergence of probabilities of events
- convergence of cumulative distribution functions
- convergence of characteristic functions

In statistical physics the natural notion of convergence is often that observable quantities converge. We idealize that an observable quantity is the expected value of ~~[illegible]~~ some well-behaved function of the state of the system. By well-behaved we ~~[illegible]~~ here mean continuous and bounded.

Def: A sequence $(X_n)_{n\in\mathbb{N}}$ of $\mathbb{R}$-valued random variables (resp. ~~[illegible]~~ $(\nu_n)_{n\in\mathbb{N}}$ of proba-measures on $\mathbb{R}$) converges weakly to a random variable X (resp. to a proba-measure ν) if for all bounded ~~[illegible]~~ continuous $f: \mathbb{R}\to\mathbb{R}$ we have

$$\mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)]$$

$$\left(\text{resp. } \int_{\mathbb{R}} f \, d\nu_n \longrightarrow \int_{\mathbb{R}} f \, d\nu\right).$$

We denote

$$X_n \xrightarrow{w} X \qquad \left(\text{resp. } \nu_n \xrightarrow{w} \nu\right)$$

The various forms of convergence are essentially equivalent.

Theorem Let (X_n) be $\mathbb{R}$-valued r.v.'s, (F_n) their cumulative distribution functions $F_n(x) = \mathbb{P}[X_n \leq x]$, (χ_n) their characteristic functions $\chi_n(\theta) = \mathbb{E}[e^{i\theta X_n}]$, and (ν_n) their distributions $\nu_n[B] = \mathbb{P}[X_n \in B]$.
Let also X, F, χ, ν be a r.v., its c.d.f., char. fn., and distribution. Then the following are equivalent

(1) $\forall f: \mathbb{R} \to \mathbb{R}$ continuous bounded
$$\mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)] \qquad \left(\text{i.e. } X_n \xrightarrow{w} X\right)$$

(2) For all continuity points $x \in \mathbb{R}$ of F we have
$$F_n(x) \xrightarrow[n\to\infty]{} F(x)$$

(3) For all $\theta \in \mathbb{R}$ we have $\chi_n(\theta) \xrightarrow[n\to\infty]{} \chi(\theta)$.

(4) For all $U \subset \mathbb{R}$ open we have $\liminf(\nu_n[U]) \geq \nu[U]$

(5) For all $A \subset \mathbb{R}$ closed we have $\limsup(\nu_n[A]) \leq \nu[A]$

(6) For all Borel subsets $B \subset \mathbb{R}$ such that $\nu[\partial B] = 0$ we have $\nu_n[B] \xrightarrow[n\to\infty]{} \nu[B]$.

A map of the proof

$$\begin{tikzcd}
 & (4) \arrow[dd, Leftrightarrow] \arrow[dr, Rightarrow] & & \\
 & & (6) \arrow[r, Rightarrow] & (2) \arrow[ull, Rightarrow, bend right] \\
(1) \arrow[r, Leftrightarrow] & (5) \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

We will prove the equivalence with condition (3) in the next lecture, once we have introduced "tightness".

Proof:

$\boxed{(4) \Leftrightarrow (5)}$ Just use $A = \mathbb{R} \setminus U$.

$\boxed{(4) \& (5) \Rightarrow (6)}$ Recall $\underset{\text{interior of } B}{B^\circ} \subset B \subset \underset{\text{closure of } B}{\overline{B}}$

$B^\circ = \mathbb{R} \setminus \overline{(\mathbb{R} \setminus B)}$ (open set) $\qquad$ (closed set)

and $\underset{\text{(boundary)}}{\partial B} = \overline{B} \setminus B^\circ$.

If $\nu[\partial B] = 0$ then $\nu[B^\circ] = \nu[B] = \nu[\overline{B}]$.

In this case we get, assuming (4) and (5),

$$\begin{aligned}
\liminf \nu_n[B] &\geq \liminf \nu_n[B^\circ] \\
&\overset{(4)}{\geq} \nu[B^\circ] = \nu[B] = \nu[\overline{B}] \\
&\overset{(5)}{\geq} \limsup \nu_n[\overline{B}] \geq \limsup \nu_n[B]
\end{aligned}$$

Thus the limit $\lim_{n\to\infty} \nu_n[B]$ exists and equals $\nu[B]$.

$\boxed{(6) \Rightarrow (2)}$ Note that $\partial((-\infty, x]) = \{x\}$ and $\nu[\{x\}] = 0$ iff x is a continuity point of F.

Thus by (6), we have $F_n(x) = \nu_n[(-\infty, x]] \to \nu[(-\infty, x]] = F(x)$

$\boxed{(2) \Rightarrow (4)}$ Assume first that U is an open interval $U = (a, b)$.

Note that since F has at most countably many points of discontinuity (increasing function) and open sets can be approximated from below, we can find for any $\varepsilon > 0$ continuity points a', b' such that $a < a' < b' < b$ and $\nu[(a', b')] \geq \nu[(a, b)] - \varepsilon$.

Then, assuming (2), we have

$$\begin{aligned}
\nu[(a,b)] - \varepsilon &\leq \nu[(a', b')] = F(b') - F(a') \\
&\overset{(2)}{=} \lim_{n\to\infty} \underbrace{(F_n(b') - F_n(a'))}_{= \nu_n[(a', b')]} \leq \liminf \nu_n[(a,b)].
\end{aligned}$$

This proves the claim if U is an interval.

In the general case U is an at most countable union of disjoint open intervals, $U = \bigcup_{j \in \mathbb{N}} (a_j, b_j)$.

By the above argument we get $\nu[(a_j, b_j)] - \varepsilon \cdot 2^{-j} \leq \liminf \nu_n[(a_j, b_j)]$

By Fatou's lemma then $\nu[U] - \varepsilon \leq \liminf \nu_n[U]$.

$\boxed{(1)\Rightarrow(5)}$ Let $A\subset\mathbb{R}$ be closed and assume $A\neq\emptyset$.

$x\longmapsto d(x,A):=\inf_{y\in A}|x-y|$ is continuous $\mathbb{R}\longrightarrow[0,\infty)$

For $\varepsilon>0$ set
$$f_\varepsilon(x)=\begin{cases}1-\frac{d(x,A)}{\varepsilon} & \text{if } d(x,A)\le\varepsilon\\ 0 & \text{if } d(x,A)\ge\varepsilon\end{cases}.$$

Then $f_\varepsilon:\mathbb{R}\to[0,1]$ is bounded and continuous, and we have
$$\mathbb{1}_A\le f_\varepsilon\le\mathbb{1}_{A_\varepsilon}$$
where $A_\varepsilon=\{x\in\mathbb{R}\mid d(x,A)\le\varepsilon\}$ is the ε-thickening of A.

Thus
$$\nu[A]=\int_{\mathbb{R}}\mathbb{1}_A\,d\nu\le\int_{\mathbb{R}}f_\varepsilon\,d\nu\le\int_{\mathbb{R}}\mathbb{1}_{A_\varepsilon}\,d\nu=\nu[A_\varepsilon].$$

Similarly
$$\limsup\nu_n[A]\le\limsup\int_{\mathbb{R}}f_\varepsilon\,d\nu_n\overset{(1)}{=}\int_{\mathbb{R}}f_\varepsilon\,d\nu\le\nu[A_\varepsilon].$$

But A_ε approximates the closed set A from above, $A_\varepsilon\downarrow A$ as $\varepsilon\downarrow0$.

Thus $\nu[A_\varepsilon]\longrightarrow\nu[A]$ as $\varepsilon\to0$ and (5) follows.

$\boxed{(5)\Rightarrow(1)}$ In (1) we may without loss of generality consider $0<f<1$.
(add a constant to f, and multiply f by constant, if necessary)

Fix temporarily $N\in\mathbb{N}$ and set
$$A_i=\{x\in\mathbb{R}\mid f(x)\ge\tfrac{i}{N}\}\qquad(i=0,1,2,\dots,N)$$
closed set, because f continuous.

Approximate f from above and below by elementary functions
$$\sum_{i=1}^N\frac{i-1}{N}\mathbb{1}_{A_{i-1}\setminus A_i}\le f\le\sum_{i=1}^N\frac{i}{N}\mathbb{1}_{A_{i-1}\setminus A_i}$$

Integrate against ν to get
$$\underbrace{\sum_{i=1}^N\frac{i-1}{N}\nu[A_{i-1}\setminus A_i]}_{=\frac{1}{N}\sum_{i=1}^N\nu[A_i]}\le\int_{\mathbb{R}}f\,d\nu\le\underbrace{\sum_{i=1}^N\frac{i}{N}\nu[A_{i-1}\setminus A_i]}_{=\frac{1}{N}+\frac{1}{N}\sum_{i=1}^N\nu[A_i]}$$

By (5) then (using a similar estimate for ν_n)
$$-\frac{1}{N}+\limsup\int_{\mathbb{R}}f\,d\nu_n\le\frac{1}{N}\sum_{i=1}^N\limsup\nu_n[A_i]\overset{(5)}{\le}\frac{1}{N}\sum_{i=1}^N\nu[A_i]\le\int_{\mathbb{R}}f\,d\nu.$$
Take $N\to\infty$.

We get $\limsup\int_{\mathbb{R}}f\,d\nu_n\le\int_{\mathbb{R}}f\,d\nu$. Consider $-f$ to get $\liminf\int_{\mathbb{R}}f\,d\nu_n\ge\int_{\mathbb{R}}f\,d\nu$. $\square$