Lec 22

3a

Since $\epsilon > 0$ is arbitrary, the "only if" part is proved.

Next consider the "if" part. Since $C_0(\mathbb{R}) \subset C_B(\mathbb{R})$, (2.9) and Theorem 9.2.2 imply that $\mu_n \longrightarrow^{v} \mu$. As noted in Remark 9.2.1, if $\{\mu_n\}_{n\geq 1}, \mu$ are probability measures then $\mu_n \longrightarrow^{v} \mu$ iff $\mu_n \longrightarrow^{d} \mu$. So the proof is complete. □

Definition 9.2.2:

Alt. Def.

(a) A sequence of probability measures $\{\mu_n\}_{n\geq 1}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called *tight* if for any $\epsilon > 0$, there exists $M = M_\epsilon \in (0,\infty)$ such that

$$\sup_{n\geq 1} \mu_n([-M, M]^c) < \epsilon. \tag{2.10}$$

*

Def for seq. of RVs

(b) A sequence of random variables $\{X_n\}_{n\geq 1}$ is called *tight* or *stochastically bounded* if the sequence of probability distributions $\{\mu_n\}_{n\geq 1}$ of $\{X_n\}_{n\geq 1}$ is tight, i.e., given any $\epsilon > 0$, there exists $M = M_\epsilon \in (0,\infty)$ such that

$$\sup_{n\geq 1} P(|X_n| > M) < \epsilon. \tag{2.11}$$

Remark 9.2.3: In Definition 9.2.2 (b), the random variables X_n, $n \geq 1$ need not be defined on a common probability space. If X_n is defined on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$, $n \geq 1$, then (2.11) needs to be replaced by

$$\sup_{n\geq 1} P_n(|X_n| > M) < \epsilon.$$

Do in class (skipped)

Example 9.2.3: Let $X_n \sim \text{Uniform}(n, n+1)$. Then, for any given $M \in (0,\infty)$,

$$P(|X_n| > M) \geq P(X_n > M) = 1 \quad \text{for all} \quad n > M.$$

Consequently, for any $M \in (0,\infty)$,

$$\sup_{n\geq 1} P(|X_n| > M) = 1$$

and the sequence $\{X_n\}_{n\geq 1}$ cannot be stochastically bounded.

not tight

Example: Let $\mu_n = \delta_{n \bmod 3}$ — point mass at $n \bmod 3$.

i.e. $\left.\begin{array}{l}\mu_1 = \delta_1 \\ \mu_2 = \delta_2 \\ \mu_3 = \delta_0\end{array}\right\}$ repeats

$\left.\begin{array}{l}\mu_4 = \delta_1 \\ \mu_5 = \delta_2 \\ \mu_6 = \delta_0\end{array}\right\}$

$\vdots$

Q. Is $\{\mu_n\}$ tight?

Yes. Take $[a,b] = [0,3]$.

Then $\forall \varepsilon > 0$ & $\forall n \in \mathbb{N}$,

$$\mu_n([0,3]) = \delta_n([0,3])_{\bmod 3}$$
$$= \mathbb{1}_{[0,3]}(n)\Big|_{n=0,1,2}$$
$$= 1 > 1-\varepsilon.$$

Lec 23 (1a) Example: For $n \geq 1$, let $X_n \sim \text{Uniform}(a_n, 2+a_n)$ where $a_n = (-1)^n$. Then

$X_1 = U(-1, 2-1=1)$ so $|X_1| \leq 1$

$X_2 = U(1, 2+1=3)$ so $|X_2| \leq 3$

$X_3 = U(-1,1)$

$X_4 = U(1,3)$ $\Rightarrow |X_n| \leq 3 \quad \forall n \in \mathbb{N}$.

$\vdots$

$$\Rightarrow \sup_{n \geq 1} P(|X_n| > 3) < \varepsilon \quad \forall \varepsilon > 0$$

$\Rightarrow$ seq. of prob. dist'ns of $\{X_n\}$ is tight

* However, $\{X_n\}_{n\geq 1}$ does NOT converge in distribution to a RV X

Properties

- Any finite collection of probability measures is tight.
- Union of 2 tight collections of prob. meas. is tight.
- Any sub-collection of a tight collection is tight.

Thm 11.1.10 : If $\{\mu_n\}$ is a tight sequence of probability measures, then there is a subsequence $\{\mu_{n_k}\}$ & a prob. measure μ s.t. $\mu_{n_k} \Rightarrow \mu$ as $n \to \infty$.

i.e. $\{\mu_{n_k}\}$ converges weakly to μ

Pf Idea : By Helly Selection Principle, $\exists$ F_{n_k} & F s.t. $F_{n_k}(x) \to F(x)$ at all continuity points of F.

Using tightness, can show that F is actually a prob. distribution function (& in particular $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$, as desired).

Cor 11.1.11 : Let $\{\mu_n\}$ be a tight seq. of prob. dist'ns on $\mathbb{R}$. Suppose that μ is the only possible weak limit of $\{\mu_n\}$, meaning $\mu_{n_k} \Rightarrow \nu$ implies that $\nu = \mu$.

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

One last result: sufficient condition for a sequence of measures to be tight

Lemma 11.1.13: Let $\{\mu_n\}$ be a sequence of prob. measures on $\mathbb{R}$ with characteristic functions $\phi_n(t) = \int e^{itx}\, d\mu_n(x)$. Suppose $\exists$ function g (continuous at 0) s.t. $\lim_{n\to\infty} \phi_n(t) = g(t)$ for each $|t| < t_0$ for some $t_0 > 0$. Then $\{\mu_n\}$ is tight.

* Theorem 11.1.14 [Continuity Thm]: Let $\mu, \mu_1, \mu_2, \ldots$ be prob. measures with corresp. characteristic functions $\phi, \phi_1, \phi_2, \ldots$ Then $\mu_n \Rightarrow \mu$ iff $\phi_n(t) \to \phi(t)$ $\forall\, t \in \mathbb{R}$.

In words: prob. measures converge weakly to μ iff their char. functions converge pointwise to that of μ.

Pf: First suppose that $\mu_n \Rightarrow \mu$ as $n \to \infty$. Then since $\cos(tx)$ & $\sin(tx)$ are bounded continuous functions, we have that

$$\begin{aligned}
\phi_n(t) &= \int \cos(tx)\, d\mu_n(x) + i \int \sin(tx)\, d\mu_n(x) \\
&\to \int \cos(tx)\, d\mu(x) + i \int \sin(tx)\, d\mu(x) \quad \text{by def of weak conv.} \\
&= \phi(x) \qquad \text{as } n \to \infty \text{ for each } t \in \mathbb{R}
\end{aligned}$$

Conversely, suppose that $\phi_n(t) \to \phi(n)$ for each $t \in \mathbb{R}$. Then by using $g = \phi$ in Lemma 11.1.13, the $\{\mu_n\}$ are tight. Now suppose that $\mu_{n_k} \Rightarrow \nu$ for some subseq $\{\mu_{n_k}\}$ & some measure ν. Then

$$\phi_{n_k}(t) \to \phi_\nu(t) \quad \forall t \in \mathbb{R} \quad \text{where} \quad \phi_\nu(t) = \int e^{itx}\, d\nu(x).$$

On the other hand, we know that (by assumption)

$$\phi_{n_k}(t) \to \phi(t) \quad \forall t \in \mathbb{R}.$$

Hence, $\phi_\nu = \phi$. By Fourier uniqueness (Cor 11.1.7), this implies that $\nu = \mu$.

Thus, μ is the only possible weak limit of $\{\mu_n\}$ so by Cor 11.1.11, it follows that $\mu_n \Rightarrow \mu$ as $n \to \infty$. ∎

The Central Limit Theorem

(Ref: Rosenthal § 11.2, Billingsley § 27)

$\mu_n \Rightarrow \mu$ iff $\phi_n(t) \to \phi(t)\ \forall t \in \mathbb{R}$

Last time we proved the Continuity Theorem (11.1.4), so we are now in a position to prove the classical Central Limit Thm (CLT).

First, compute the characteristic function for the standard normal distribution $N(0,1)$, i.e.

for RV $X \sim N(0,1)$, density $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ w.r.t. Lebesgue measure.

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\,dx = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$$

Turns out that

$$\phi_X(t) = M_X(it) = e^{+(it)^2/2} = \boxed{e^{-t^2/2}} \quad \forall t \in \mathbb{R}$$

(can justify this using complex analysis)

<u>Prop</u>: If $X \sim N(0,1)$, then $\phi_X(t) = e^{-t^2/2}\ \forall t \in \mathbb{R}$.

* Thm 11.2.2 [Central Limit Thm] : Let $X_1, X_2, \ldots$ be i.i.d. random variables with finite mean m and finite variance v. Set $S_n = X_1 + \cdots + X_n$. Then as $n \to \infty$,

$$\mathcal{L}\left(\frac{S_n - nm}{\sqrt{vn}}\right) \Rightarrow \mu_N \qquad \text{where } \mu_N = N(0,1)$$

weak convergence or convergence in distribution

std. Normal distribution.

[In words, when i.i.d. RVs are added, their properly normalized sum tends toward a normal distribution as n gets large.]

Pf: By replacing X_i by $\frac{X_i - m}{\sqrt{v}}$ $\forall i$, we can assume that $m = 0$ & $v = 1$.

Let $\phi_n(t) = E\left[e^{it\, S_n/\sqrt{n}}\right]$ be the characteristic function of $S_n/\sqrt{n}$ (by definition). By the Continuity Theorem & Proposition above, it suffices to show that $\lim_{n\to\infty} \phi_n(t) = \phi(t) = e^{-t^2/2}$ for each fixed $t \in \mathbb{R}$.

More details

[We want to show that the sequence $\mu_n = \mathcal{L}\left(\frac{S_n}{\sqrt{n}}\right)$ recalling that $m=0$, $v=1$ converges weakly to $\mu_N = N(0,1)$. By Continuity Thm,

$$\mu_n \Rightarrow N(0,1) \quad \text{iff} \quad \phi_n(t) \to \underset{\substack{\phi(t) \text{ by Prop.} \\ \text{for } N(0,1)}}{\overset{}{\underset{\parallel}{e^{-t^2/2}}}} \quad \text{for } t \in \mathbb{R}.$$
]

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

$$\phi_X^{(j)}(0) = i^j E[X^j]$$

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

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

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

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

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

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

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

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

since $i^2 = -1$

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

[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