03/28

Limit Laws

Law of large Numbers: "the proportion of heads after n flips of a fair coin converges to $\frac{1}{2}$ as $n \to \infty$" — with prob. 1.

Thm (Weak LLN.) Let $X_1, X_2, \ldots$ be a sequence of ~~indep, mean zero r.v's~~ ~~indep. r.v's~~. i.i.d. r.v's. with mean μ. and finite variance σ^2.

Then
$$\frac{X_1 + \cdots + X_n}{n} \to \mu \quad \text{as } n \to \infty$$
in probability.

Def $Z_n \to Z$ (r.v's) in prob. if $\forall \varepsilon > 0$,
$$\mathbb{P}\{|Z_n - Z| > \varepsilon\} \to 0 \quad \text{as } n \to \infty.$$

$$\mathrm{Var}(Z_n) = \frac{1}{n^2} \sum_1^n \underbrace{\mathrm{Var}(X_i)}_{\sigma} \quad \text{by independence}$$
$$= \frac{\sigma^2}{n}.$$

Chebychev's ineq: $\mathbb{P}\{|Z - \mu| > \varepsilon\} \le \frac{\mathrm{Var}(Z)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \to 0$ as $n \to \infty$.

Remarks

WLLN Can be strengthened:

(a) Variance $< \infty$ is not needed

(b) ~~imp~~ convergence "in prob" $\longmapsto$ "a.s.";

Thm (Strong LLN) Let $X_1, X_2, \dots$ be a sequence of i.i.d. r.v's with mean μ. Then

$$\frac{X_1 + \dots + X_n}{n} \to \mu \quad \text{almost surely} \quad (\text{"a.s"})$$

Def $Z_n \to Z$ a.s. if $\mathbb{P}\{\lim_{n\to\infty} Z_n = Z\} = 1$

Generally, event E happens a.s if $\mathbb{P}(E) = 1$

W/o proof, But:

Prop A.s. Convergence in probability implies convergence in probability. (not obvious!)

Fix $\varepsilon > 0$.

$$\mathbb{P}\{|Z_n - Z| > \varepsilon\} \le \mathbb{P}\{\underbrace{\sup_{k\ge n} |Z_n - Z| > \varepsilon}_{=: E_n}\}$$

Events E_n ~~is a sequence of decreasing~~ form a decreasing sequence $(E_1 \supset E_2 \supset \dots)$

"Continuity of Probability" (See lec 1-2) implies

$$\mathbb{P}(E_n) \to \mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right)$$

$$= \mathbb{P}\{\forall n \in \mathbb{N}\ \exists k \ge n : |Z_n - Z| > \varepsilon\}$$

$$\le \mathbb{P}\{Z_n \not\to Z\} = 0. \quad (\text{by a.s. convergence})$$

Hence $\mathbb{P}\{|Z_n - Z| > \varepsilon\} \to 0$ as $n \to \infty$ $\Rightarrow$ converges in prob.

Remark Not vice versa: Convergence in prob $\not\Rightarrow$ convergence a.s. (Ex.)

Central Limit Thm gives more info;
states that the limiting distribution of # heads is normal.

Normalization of the sum: Recall if X has mean μ, var σ^2 $\Rightarrow$

$$Z := \frac{X-\mu}{\sqrt{\sigma}}$$

has mean 0, var. 1 ("z-score")

$S_n := X_1 + \dots + X_n$ sum of iid r.v's with mean μ, var σ^2

$\mathbb{E}[S_n] = n\mu$; $Var(S_n) = n\sigma^2$.

Thm (CLT) Let $X_1, X_2, \dots$ be a sequence of iid r.v's with mean μ, var. σ^2.
Denote $S_n := X_1 + \dots + X_n$. Then

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \to N(0,1) \quad \text{in distribution}$$

Def $Z_n \to Z$ in distribution if $\forall$ pt a of continuity of CDF F_Z,
$F_{Z_n}(a) \to F_Z(a)$ as $n \to \infty$.
i.e. CDF's converge pointwise (on the set of continuity of F_Z).

Rem Conv. in distr. $\Rightarrow$ $\mathbb{P}\{a < Z_n \le b\} \to \mathbb{P}\{a < Z \le b\}$ as $n \to \infty$ (Ex!)

CLT $\Rightarrow$

CDF of $N(0,1)$

$$\mathbb{P}\left\{a < \frac{S_n - n\mu}{\sigma\sqrt{n}} < b\right\} \to \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,dx.$$

~~Rem A.s. convergence ⇓ Convergence in prob ⇓ Convergence in distr.~~

~~Proof of CLT~~ Towards proof of CLT : Via MGF.

(assume $M(t) = \mathbb{E}e^{tX_i} < \infty \ \forall t$).

- We already mentioned a uniqueness theorem : "MGF determine the distribution uniquely", i.e. if $M_X(t) = M_Y(t)\ \forall t \in \mathbb{R}$ then X has the same distr. as Y.

- Now, stronger statement is also true - in the limit :

Lem (Levy's Continuity Lemma) Let $Z_1, Z_2, \ldots$ be a sequence of r.v's and Z be a r.v. If
$$M_{Z_n}(t) \to M_Z(t) \quad \forall t \in \mathbb{R}$$
then $Z_n \to Z$ in distribution.

w/o proof.

Proof of CLT WLOG, $\mu = 0$, $\sigma = 1$ (Consider $\frac{X_i - \mu}{\sigma}$ instead of X_i).

- Let's compute MGF of
$$Z_n := \frac{S_n}{\sqrt{n}} = \frac{X_1}{\sqrt{n}} + \cdots + \frac{X_n}{\sqrt{n}}.$$

$$M_{Z_n}(t) = M_{\frac{X_1}{\sqrt{n}}}(t) \cdots M_{\frac{X_n}{\sqrt{n}}}(t) = \left(M_{\frac{X_1}{\sqrt{n}}}(t)\right)^n \quad \text{(by i.i.d. assumption)}.$$

- Asymptotic analysis of
$$M_{\frac{X_1}{\sqrt{n}}}(t) = \mathbb{E}\left[\exp\left(\frac{tX_1}{\sqrt{n}}\right)\right].$$

Taylor expansion

$$\exp\left(\frac{tX_1}{\sqrt{n}}\right) = 1 + \frac{tX_1}{\sqrt{n}} + \frac{t^2X_1^2}{2n} + \bar{o}\left(\frac{t^2X_1^2}{2n}\right)$$

$$\Rightarrow \quad \mathbb{E}\left[\exp\left(\frac{tX_1}{\sqrt{n}}\right)\right] = 1 + \frac{t}{\sqrt{n}}\underbrace{\mathbb{E}[X_1]}_{=0} + \frac{t^2}{2n}\underbrace{\mathbb{E}[X_1^2]}_{=1} + \underbrace{\bar{o}\left(\frac{\overbrace{t^2}^{\text{fixed}}\overbrace{\mathbb{E}[X_1^2]}^{=1}}{2n}\right)}_{=\bar{o}\left(\frac{1}{n}\right)}$$

(interchanged limit and $\mathbb{E}$ non-rigorously)

$$\| $$

$$M_{\frac{X_1}{\sqrt{n}}}(t) = 1 + \frac{t^2}{2n} + \bar{o}\left(\frac{1}{n}\right)$$

$$\Rightarrow M_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \bar{o}\left(\frac{1}{n}\right)\right]^n = e^{\frac{t^2}{2} + \bar{o}(1)} \qquad \text{(why?)}$$

$$\to e^{t^2/2} \quad \text{as } t \to \infty$$

$$= M_Z(t) \quad \text{for} \quad Z \sim N(0,1).$$

~~Hence~~ By Levy's Continuity Lemma,

$Z_n \to Z$ in distribution. $\square$

<u>Remark</u>: CLT says, intuitively, that

$$S_n = X_1 + \dots + X_n \underset{\text{approximately}}{\sim} N(\mu n, \sigma^2 n)$$

<u>Remarks</u>: ~~Uniform convergence in CLT~~ 1. Compare with Poisson limit thm (rare successes)

CLT (frequent successes) ↑ const prob.

2. Using characteristic functions instead of MGF

Prop Convergence in probability implies convergence in distribution

Thus, conv. a.s. $\Longrightarrow$ conv. in prob. $\Longrightarrow$ conv. in distr

(WLLN) $\not\Leftarrow$ (SLLN) $\not\Leftarrow$ (CLT)

Assume $Z_n \to Z$ in prob

We will show that $\forall$ pt. a of continuity of $F(a)$,

$$\text{(i)} \quad \limsup_n \underbrace{F_n(a)}_{\substack{\| \\ P\{Z_n \le a\}}} \le \underbrace{F(a)}_{\substack{\| \\ P\{Z \le a\}}}$$

and
$$\text{(ii)} \quad \liminf_n F_n(a) \ge F(a).$$

(Together, (i) & (ii) would imply that $\lim F_n(a) = F(a)$, i.e. $Z_n \to Z$ in distr.)

Proof of (i). Let $\varepsilon > 0$. Conv. in probability yields that

the events $E_n := \{ |Z_n - Z| \le \varepsilon \}$ are likely: $P(E_n^c) \to 0$.

$$F_n(a) = P\{Z_n \le a\} = P\{Z_n \le a, \text{ and } E_n\} + P\{Z_n \le a \text{ and } E_n^c\}$$
(disjoint)

$$\le P\{Z \le a + \varepsilon\} + P(E_n^c) \qquad (\Delta \text{ ineq.})$$

$$= F(a+\varepsilon) + \underbrace{P(E_n^c)}_{\downarrow \atop 0}$$

Hence $\limsup_n F_n(a) \le F(a+\varepsilon)$.

This happens $\forall \varepsilon > 0$. Let $\varepsilon \to 0_+$, use continuity of F $\Rightarrow$

$$\limsup_n F_n(a) \le F(a)$$

~~as dis~~ So we proved (i). The proof of (ii) is similar (Ex)

$\square$

Ex Suppose that, whenever invited to a party, a person attends alone with prob. 1/3, attends with a guest with prob. 1/3, and does not attend with prob. 1/3.
A company invited all 300 of its employees and their guests to a party.
What is the prob. that at least 320 will attend?

X_i = # of people employee i will bring, including him/her self.

X_i are independent r.v's with pmf

$$P\{X_i=0\} = P\{X_i=1\} = P\{X_i=2\} = 1/3.$$

Thus $\mu = E[X_i] = 1$, $\sigma^2 = Var(X_i) = 2/3$ (check!)

# of people attending the party is

$$S_n = X_1 + \dots + X_n, \quad n = 300.$$

$$P\{S_n \geq 320\} = P\Big\{ \frac{S_n - n\mu}{\sigma\sqrt{n}} \geq \underbrace{\frac{320 - n\mu}{\sigma\sqrt{n}}}_{\frac{320 - 300\cdot 1}{\sqrt{\frac{2}{3}\cdot 300}} = 1.41} \Big\} \overset{CLT}{=} 1 - \underset{\text{tabulated}}{\Phi}(1.41) = 1 - 0.921 = 0.079 = 8\%.$$

Remark: $S_n \approx N(\mu n, \sigma^2 n) = N(300, 200)$.