SNU MAE Intro. Eng. Prob

<u>Limit Thms</u> : the most important theoretical probability results.

- Law of large numbers : concerned with conditions under which the avrg of a seq of rvs converges to the expected avrg.
- Central limit thm : concerned with determining conditions under which the sum of a large number of rvs has an approximetely normal prob distribution.

<u>Prop</u> [Markov's Inequality]

X : a r.v. that takes only nonnegative values.
For any value $a>0$,

$$P\{X \geq a\} \leq \frac{E[X]}{a}.$$

<u>Pf</u> For $a>0$, let $I = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{else} \end{cases}$

Since $X \geq 0$,

$$I \leq \frac{X}{a}$$

$$\Rightarrow \underset{\substack{\parallel \\ P\{X \geq a\}}}{E[I]} \leq \frac{E[X]}{a}.$$

//

<u>Prop</u> [Chebyshev's Inequality]

X : a r.v. with finite μ & σ^2.
For any value $k>0$,

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}.$$

Pf Apply Markov's inequality to $(X-\mu)^2$ with $a=k^2$.
Then

$$P\{(X-\mu)^2 \geq k^2\} \leq \frac{E[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2}$$

$$\parallel$$

$$P\{|X-\mu| \geq k\}.$$

//

Rmk Markov's & Chebyshev's inequalities allow to derive bounds on probabilities when only mean, or mean & var are known.
(The bounds obtained are not necessarily close to the actual probability.)

Prop If $Var(X)=0$, then

$$P\{X = E[X]\} = 1.$$

i.e. the only rv having variances equal to 0 are those that are constant with prob 1.

Pf For any $n \geq 1$,

$$P\{|X-\mu| > \tfrac{1}{n}\} = 0 \quad (\because \text{Chebyshev with } k=\tfrac{1}{n})$$

$\Downarrow$ let $n \to \infty$.

Continuity property

$$0 = \lim_{n\to\infty} P\{|X-\mu| > \tfrac{1}{n}\} = P\{\lim_{n\to\infty}\{|X-\mu| > \tfrac{1}{n}\}\}$$

$$= P\{X \neq \mu\}.$$

//

Thm [The weak law of large numbers]

$X_1, X_2, \dots$: sequence of i.i.d rvs.

$E[X_i] = \mu$ finite.

Then, for any $\varepsilon > 0$,

$$P\left\{\left|\frac{X_1+\cdots+X_n}{n} - \mu\right| \geq \varepsilon\right\} \to 0 \quad \text{as } n \to \infty.$$

//

<u>Pf</u> (Assuming that the rvs have a finite σ^2.)

$$E\left[\frac{X_1+\cdots+X_n}{n}\right] = \mu$$

$$Var\left(\frac{X_1+\cdots+X_n}{n}\right) = \frac{\sigma^2}{n}$$

$$\Rightarrow P\left\{\left|\frac{X_1+\cdots+X_n}{n}-\mu\right| \geq \varepsilon\right\} \leq \frac{\sigma^2}{n\varepsilon^2} \quad (\because \text{Chebyshev})$$

//

<u>Rmk</u> Bernoulli proved the weak law of large numbers for Bernoulli rvs before Chebyshev's inequality.

<u>Central Limit Thm.</u>

(Roughly speaking) The sum of a large number of independent rvs has a distribution that is approximately normal.!

<u>Thm</u> [The central limit thm]

$X_1, X_2, \ldots$: a sequence of iid rvs, each with μ & σ^2.

Then the distribution of $\dfrac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ tends to $N(0,1)$ as $n\to\infty$.

i.e.

$$P\left\{\frac{X_1+\cdots+X_n-\mu n}{\sigma\sqrt{n}} \leq a\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{x^2}{2}}\,dx$$

as $n\to\infty$.

<u>Lemma</u> $Z_1, Z_2, \ldots$: a sequence of rvs with distribution ftn F_{Z_n} and moment gen ftn M_{Z_n} $n\geq 1$.

Z : a rv with F_Z and M_Z.

If $M_{Z_n}(t) \to M_Z(t)$ for all t,

then $F_{Z_n}(t) \to F_Z(t)$ for all t at which $F_Z(t)$ is conti.

Pf (CLT) Let $\mu=0$, $\sigma^2=1$.

Assume that the moment gen ftn of X_i, $M(t)$, exists and is finite.

Moment gen ftn of $X_i/\sqrt{n}$ is

$$E\left[\exp\left\{\frac{tX_i}{\sqrt{n}}\right\}\right] = M\left(\frac{t}{\sqrt{n}}\right)$$

$\Rightarrow$ Moment gen ftn of $\sum_{i=1}^{n}\frac{X_i}{\sqrt{n}}$ is $\left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n$.

Let $L(t) = \log M(t)$.

$M(0)=1$, $M'(0)=\mu=0$.
$M''(0)=E[X^2]$

$$\Rightarrow L(0) = 0$$

$$L'(0) = \frac{M'(0)}{M(0)} = \mu = 0$$

$$L''(0) = \frac{M(0)\,M''(0) - [M'(0)]^2}{[M(0)]^2} = E[X^2] = 1$$

Want to show $\left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n \to e^{\frac{t^2}{2}}$

i.e. $n\,L\left(\frac{t}{\sqrt{n}}\right) \to \frac{t^2}{2}$.

$$\lim_{n\to\infty}\frac{L\left(\frac{t}{\sqrt{n}}\right)}{n^{-1}} = \lim_{n\to\infty}\frac{L'\left(\frac{t}{\sqrt{n}}\right)\left(-\frac{1}{2}n^{-\frac{3}{2}}t\right)}{-n^{-2}}$$ L'Hopital's rule

$$= \lim_{n\to\infty}\frac{L'(t/\sqrt{n})\,t}{-2n^{-\frac{1}{2}}}$$

$$= \lim_{n\to\infty}\frac{L''\left(\frac{t}{\sqrt{n}}\right)\left(-\frac{1}{2}\cancel{n^{-\frac{3}{2}}}t^2\right)}{-\cancel{n^{-\frac{3}{2}}}}$$ L'Hopital's rule

$$= \frac{t^2}{2}. \quad //$$

The gen. case can be proved by considering $X_i^* = \frac{X_i-\mu}{\sigma}$, since $E[X_i^*]=0$, $Var(X_i^*)=1$. //

Ex Rolling 10 fair dice. prob that the sum is btwn 30 & 40 ?
[30, 40]

Sol X_i : value of the ith die. $i=1,\cdots,10$

$$E[X_i] = \frac{7}{2}$$

$$\begin{aligned} Var[X_i] &= E[X_i^2] - (E[X_i])^2 \\ &= \frac{1}{6}(1+4+9+16+25+36) - \left(\frac{7}{2}\right)^2 \\ &= \frac{91}{6} - \frac{49}{4} = \frac{35}{12}. \end{aligned}$$

$$P\{29.5 \le X \le 40.5\} = P\left\{\frac{29.5-35}{\sqrt{\frac{35}{12}}} \le \frac{X-35}{\sqrt{}} \le \frac{40.5-35}{\sqrt{}}\right\}$$

$$\approx 2\Phi(1.0184) - 1$$

$$\approx .692$$

Thm [CLT for indept rvs]

$X_1, X_2, \cdots$: a sequence of indept rvs
$\mu_i = E[X_i]$, $\sigma_i^2 = Var(X_i)$.

If (a) X_i are uniformly bounded, i.e. for some M, $P\{|X_i| < M\} = 1$ for all i, and

(b) $\sum_{i=1}^{\infty} \sigma_i^2 = \infty$, then

$$P\left\{\frac{\sum_{i=1}^{n}(X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}} \le a\right\} \to \Phi(a) \quad \text{as } n \to \infty.$$

(Borel for Bernoulli rvs)
(Kolmogorov)

Thm The strong law of large numbers

$X_1, X_2, \ldots$: a sequence of iid rvs, each with $\mu = E[X_i]$ finite.

Then, with prob 1,

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

(i.e. $P\{\lim_{n\to\infty} \frac{(X_1+\cdots+X_n)}{n} = \mu\} = 1$)

Pf (assuming that $E[X_i^4] = K$ is finite)

Assume $\mu = 0$. $S_n := \sum_{i=1}^{n} X_i$.

$$E[S_n^4] = E[(X_1+\cdots+X_n)(X_1+\cdots+X_n)(X_1+\cdots+X_n)(X_1+\cdots+X_n)]$$

X_i^4, $X_i^3 X_j$, $X_i^2 X_j^2$, $X_i^2 X_j X_k$, $X_i X_j X_k X_l$ i, j, k, l all diff.

$E[\cdot] = 0$ $(\because E[X_i] = 0)$

For a given pair i & j,

$\binom{4}{2} = 6$ terms that will equal $X_i^2 X_j^2$

i i j j   j j i i
i j j i   j i i j
j i j i   i j i j

$$E[S_n^4] = nE[X_i^4] + 6\underbrace{\binom{n}{2}}_{\text{choosing a pair}} E[X_i^2 X_j^2]$$

$$= nK + 3n(n-1) E[X_i^2]E[X_j^2]$$

$$0 \le Var(X_i^2) = E[X_i^4] - (E[X_i^2])^2$$

$$\therefore (E[X_i^2])^2 \le K$$

$\Rightarrow E[S_n^4] \le nK + 3n(n-1)K$

$\Rightarrow E\left[\frac{S_n^4}{n^4}\right] \le \frac{K}{n^3} + \frac{3K}{n^2}$

$\Rightarrow E\left[\sum_{n=1}^{\infty} \frac{S_n^4}{n^4}\right] = \sum_{n=1}^{\infty} E\left[\frac{S_n^4}{n^4}\right] < \infty$

$\Rightarrow$ With prob 1, $\sum_{n=1}^{\infty} \frac{S_n^4}{n^4} < \infty$.

Convergence of a series $\Rightarrow$ nth term $\to 0$

$\therefore \frac{S_n^4}{n^4} \to 0$

$\Rightarrow \frac{S_n}{n} \to 0$.

$\therefore$ With prob 1, $\frac{S_n}{n} \to 0$ as $n \to \infty$.

When $\mu \neq 0$, apply the same argument to $X_i - \mu$.

Then with prob 1, $\lim_{n\to\infty} \sum_{i=1}^{n} \frac{X_i - \mu}{n} = 0 \Rightarrow \sum_{i=1}^{n} \frac{X_i}{n} \to \mu$

//

Rmk • Wk law of large numbers:

for any specified large n^*,

$\frac{X_1 + \cdots + X_{n^*}}{n^*}$ is likely to be near μ.

$\approx$ It leaves open the possibility that large $\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right|$ can occur infinitely often for values of n larger than n^*.

• Strong law says this cannot occur.

In particular, with prob 1, for any $\varepsilon > 0$,

$\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right|$ will be greater than ε

only a finite number of times.