Lecture 29: Law of Large Numbers and Central Limit Theorem

it's saying that the distribution is not changing with time

Let $X_1, X_2, \ldots$ be i.i.d. mean μ, var σ^2,
let $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ (Sample mean)

(Strong)
Law of Large Numbers: (Convergence of random variable)

$\bar{X}_n \to \mu$ as $n \to \infty$ with probability 1

(Sample mean) (True mean)

Sample mean converges to pointwise with prob. 1, to the true mean

this is an event. Either these random variables converges or they don't. And, that event has probability 1.

Example $X_j \sim \text{Bern}(p)$ (just imagine an infinite sequence of coin tosses, where the prob. of heads is p), then

$$\frac{X_1 + \cdots + X_n}{n} \to p \text{ with prob. } 1.$$

how many times did the coin land heads

n is the total number of coin flips.

Note: The coin is memoryless. The coin does not care how many failures or how many losses we had before.

Swamping

No matter how unlucky we were in the first 100 or the first million trials, that's nothing compared to ∞. So, those first 100 or the first million just get swamped out by the entire infinite future.

The way it works is not through if we're unlucky at the beginning that somehow it gets offset later by an increase in heads. The way it works is through what we call swamping.

<u>Weak Law of Large Numbers</u>:
for any $c > 0$ (some small number like 0.001),

$$P(|\overline{X}_n - \mu| > c) \to 0 \quad \text{(Convergence in probability)}$$
$$\text{as } n \to \infty.$$

This says, if n is large enough, then it's extremely ~~un~~likely that $\overline{X}_n$ and μ are extremely close to each other.

<u>Proof</u>: $P(|\overline{X}_n - \mu| > c) \leq \frac{Var(\overline{X}_n)}{c^2}$ (Chebyshev's Inequality)

$$= \frac{\frac{1}{n^2} \cdot n\sigma^2}{c^2}$$ (Variance of the sum is n times the variance of one term)

$$= \frac{\sigma^2}{nc^2} \to 0.$$ (Since, σ & c are constants and as $n \to \infty$, $\frac{\sigma^2}{nc^2} \to 0$)

$\overline{X}_n - \mu \to 0$ as $n \to \infty$ with probability of 1, but what does the distribution of $\overline{X}_n$ look like?

$$\frac{n^{1/2}(\overline{X}_n - \mu)}{\sigma} \quad \text{(Central Limit Theorem)}$$

$$\to N(0,1) \quad \text{(in distribution)}$$
$$\text{as } n \to \infty$$

CENTRAL LIMIT THEOREM : (Converge of distribution)

$$n^{1/2}\frac{(\bar{X}_n-\mu)}{\sigma} \longrightarrow N(0,1) \text{ in distribution}$$

as $n \to \infty$

(By convergence in distribution, it means the distribution of $\left[n^{1/2}\frac{(\bar{X}_n-\mu)}{\sigma}\right]$ converges to the standard normal distribution.)

Equivalently:

$$\frac{\sum_{j=1}^{n} X_j - n\mu}{\sqrt{n}\,\sigma} \to N(0,1) \text{ in distribution}$$

$\left(CDF\left[n^{1/2}\frac{(\bar{X}_n-\mu)}{\sigma}\right] \to \Phi\right)$ if $n \to \infty$

has mean $n\mu$, variance $n\sigma^2$

Proof (assume MGF $M(t)$ of X_j exists)

Note: If the MGFs converge to some other MGF, then the random variables converge in distribution.

Let's assume $\mu=0, \sigma=1$ since Consider

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{(X_i-\mu)}{\sigma}.$$

Let $S_n = \sum_{j=1}^{n} X_j$, show MGF of $\frac{S_n}{\sqrt{n}}$ goes to $N(0,1)$ MGF.

$$MGF = E\left(e^{tS_n/\sqrt{n}}\right) = E\left(e^{tX_1/\sqrt{n}}\right)\dots E\left(e^{tX_n/\sqrt{n}}\right)$$

(Independence and hence uncorrelated)

$E(XY) = E(X)E(Y)$

But, since X's are i.i.d., these are really the same thing written n times.

or, $E\left(e^{tS_n/\sqrt{n}}\right) = \left(\underbrace{MGF\left(t/\sqrt{n}\right)}\right)^n$ → MGF of X_1 evaluated at $t/\sqrt{n}$

If $n \to \infty$, $MGF(0) = 1$

$\left(MGF\left(t/\sqrt{n}\right)\right)^n \longrightarrow 1^\infty$ (indeterminate form)

Take logs.

$$\lim_{n\to\infty} n \log M\left(t/\sqrt{n}\right) = \lim_{n\to\infty} \frac{\log M\left(t/\sqrt{n}\right)}{1/n}$$

Let $y = 1/\sqrt{n}$, then y be real

$$= \lim_{y\to 0} \frac{\log M(yt)}{y^2}$$

$$= \lim_{y\to 0} \frac{\frac{1}{M(yt)} M'(yt)\cdot t}{2y} \quad \text{(L'Hospital Rule, } \tfrac{0}{0} \text{ form)}$$

$$= \lim_{y\to 0} \frac{tM'(yt)}{2y\underbrace{M(yt)}_{\to 1}}$$

$$= \frac{t}{2} \lim_{y\to 0} \frac{M'(yt)}{y}$$

$$= \frac{t^2}{2} \lim_{y\to 0} \frac{M''(yt)}{1} \quad \text{(L'Hospital Rule)}$$

$$= \frac{t^2}{2} M''(0)$$

$M(t) = E(e^{tX_1})$
$M(0) = 1$
$M'(0) = 0$ (since we assume $\mu = 0$)
$M''(0) = $ 2nd Moment $= 1$

$$= t^2/2, \text{ which is the log of } \underbrace{e^{t^2/2}}_{N(0,1)\ MGF}$$

B Normal Approximation to the Binomial
(or Binomial Approximated by Normal)

Let $X \sim Bin(n,p)$, think of $X = \sum_{j=1}^{n} X_j$, $X_j \sim Bern(p)$ i.i.d.

So, the Central Limit theorem says that if the N is large, this will be approximately normal, atleast after we have standardized it.

$$P(a \leq X \leq b) = P\left(\frac{a-np}{\sqrt{npq}} \leq \frac{X-np}{\sqrt{npq}} \leq \frac{b-np}{\sqrt{npq}}\right)$$

(standardized)

$$\approx \phi\left(\frac{b-np}{\sqrt{npq}}\right) - \phi\left(\frac{a-np}{\sqrt{npq}}\right)$$ (Normal Approximation)

Contrast with Poisson approximation

Pois n large, p small, $\lambda = np$ "moderate"
(Poisson is relevant when we are dealing with a large number of very rare unlikely things.)

Normal n large, p close to $1/2$

(Every normal distribution is symmetric, If p is far from $1/2$, then the binomial is very very skewed and in that case, it doesnot make ~~sense~~ that much sense to approximate using a normal.)

$$P(a \le X \le b) \simeq \phi\left(\frac{b-np}{\sqrt{npq}}\right) - \phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

(Here, we are ~~using c~~ approximating a discrete distribution using something continuous)

<u>Continuity Correction</u>

It's an improvement to deal with the fact that we are using something continuous to approximate something discrete.

$$P(X=a) = P\left(a-\tfrac{1}{2} < X < a+\tfrac{1}{2}\right)$$

a is integer (this improves above approximation)