§27 Ross L1

Weierstrass Approximation Thm By Bernstein ~ 1940 ?

Polynomials:

- $B_n(f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \longrightarrow f(x)$

$f_n(f) \longrightarrow f$ uniformly for $x \in [0,1]$

Ideas leading to this —

- Binomial Thm: $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$

$\therefore a = x,\ b = 1-x \Rightarrow \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = 1$

- Recall, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$$k\binom{n}{k} = \frac{n!}{(k-1)(n-k)!} = n\frac{(n-1)!}{(k-1)!(n-1-(k-1))!}$$

$$= n\binom{n-1}{k-1}$$

So $\quad k(k-1)\binom{n}{k} = n(k-1)\binom{n-1}{k-1} = n(n-1)\binom{n-2}{k-2}$

- Main insight:

$$\underbrace{\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}}\ = 1$$

<u>Claim:</u> only the k st $\frac{k}{n}$ is near x contribute to this sum. when n is large

To prove this: assume $\left|\frac{k}{n} - x\right| > \delta$

$\Leftrightarrow \quad (k-nx)^2 > n^2\delta^2 \Leftrightarrow \frac{(k-nx)^2}{n^2\delta^2} > 1$

Let $A_\delta = \left\{k : \frac{(k-nx)^2}{n^2\delta^2} > 1\right\}$ = "the k so $\left|\frac{k}{n} - x\right| > \delta$"

<u>Then</u>: $\displaystyle \sum_{k\in A_\delta} \underbrace{\binom{n}{k} x^k (1-x)^{n-k}}_{\text{all pos terms}} \leq \underbrace{\sum_{k\in A} \frac{(k-nx)^2}{n^2\delta^2}\binom{n}{k} x^k (1-x)^{n-k}}_{\text{" Estimate for the sum over all } k \text{ not near } \frac{k}{n}=x\text{"}}$

$$\sum_{k \in A_\delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{n^2\delta^2} \sum_{k\in A_\delta} \underbrace{(k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k}}_{\text{pos}}$$

$$\leq \frac{1}{n^2\delta^2} \sum_{k=0}^{n} (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k}$$

$$= \frac{1}{n^2\delta^2} \left\{ \sum_{k=0}^{n} k^2 \binom{n}{k} x^k (1-x)^{n-k} - \sum_{k=0}^{n} 2knx \binom{n}{k} x^k (1-x)^{n-k} + \underbrace{\sum_{k=0}^{n} n^2x^2 \binom{n}{k} x^k (1-x)^{n-k}}_{n^2x^2} \right\}$$

$$\overset{*}{=} \frac{1}{n^2\delta^2} \left\{ \sum_{k=0}^{n} k^2 \binom{n}{k} x^k (1-x)^{n-k} - 2nx \sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} + n^2x^2 \right\}$$

$$\sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} = \sum_{k=1}^{n} n \binom{n-1}{k-1} x^k (1-x)^{n-k}$$

$$= nx \sum_{k=1}^{n} \binom{n-1}{k-1} x^{k-1} (1-x)^{(n-1)-(k-1)} \underset{j=k-1}{=} nx \underbrace{\sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j}}_{1}$$

$$= nx$$

$$\sum_{k=0}^{n} k(k-1)\binom{n}{k}x^k(1-x)^{n-k} =$$

$$\sum_{k=2}^{n}\binom{n-2}{k-2}x^2x^{k-2}(1-x)^{n-k}$$

$j = k-2$

$$= n(n-1)x^2 \underbrace{\sum_{j=0}^{n-2}\binom{n-2}{j}x^j(1-x)^{n-2-j}}_{1}$$

$$= n(n-1)x^2$$

$$\sum_{k=0}^{n} k(k-1)\binom{n}{k}x^k(1-x)^{n-k} = \sum_{k=0}^{n} k^2\binom{n}{k}x^k(1-x)^{n-k} - nx$$

$$\therefore \sum_{k=0}^{n} k^2\binom{n}{k}x^k(1-x)^{n-k} = n(n-1)x^2 + nx$$

Conclude by (*) $\cancel{\sum_{k\in A}\binom{n}{k}x^k(1-x)^{n-k} \le \frac{1}{n^2\delta}\left\{n(n-1)x^2 + nx - 2n^2x + n^2\right\}}$

$\cancel{= \frac{1}{n^2\delta}\{}$

By (*):

$$\sum_{k\in A_\delta} \binom{n}{k} x^k (1-x)^{n-k} \le \frac{1}{n^2\delta^2}\left\{n(n-1)x^2 + nx - 2n^2x^2 + n^2x^2\right\}$$

$$= \frac{1}{n^2\delta^2}\Big\{\underbrace{\cancel{n^2x^2} - nx^2 + nx - \cancel{2n^2x^2} + \cancel{n^2x^2}}_{\text{The } n^2 \text{ term cancels !}}\Big\}$$

$$= \frac{x(1-x)}{n\delta^2} \xrightarrow[n\to\infty]{} 0$$

So: only the terms near $\frac{k}{n} = x$ contribute

Thus: $B_n(f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$

$$= \underbrace{\sum_{A_\delta} f\left(\tfrac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k}}_{\text{smaller than } \frac{x(1-x)}{n\delta^2} \text{ !}} + f(x)\underbrace{\sum_{A_\delta^c} (\,)\, x^k(1-x)^{n-k}}_{\simeq \text{ whole sum } \simeq 1}$$

<u>Thm</u>: $B_n(f) \to f$ uniformly on $[0,1]$.

Pf. We show: $\forall \varepsilon \; \exists N$ st $\forall_{\substack{n>N \\ x\in[0,1]}} \; |B_n(f)(x) - f(x)| < \varepsilon$

So fix ε. We find N. But

$$|B_n(f)(x) - f(x)| = \Big| \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-k)^k$$

$$- f(x) \underbrace{\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}}_{=1}$$

$$= \Big| \sum_{k=0}^{n} \left[f\left(\tfrac{k}{n}\right) - f(x) \right] \binom{n}{k} x^k (1-x)^{n-k} \Big|$$

$$\leq \sum_{k=0}^{n} \left| f\left(\tfrac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k}$$

$$= \Big| \sum_{k\in A_\delta} \Big| + \Big| \sum_{k\notin A_\delta} \Big|$$

choose δ so that $\left|f\left(\tfrac{k}{n}\right) - f(x)\right| < \frac{\varepsilon}{2}$ when $\left|\tfrac{k}{n} - x\right| < \delta$.

Then: $\left|\sum_{k \notin A_\delta}\right| \le \frac{\varepsilon}{2} \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} < \frac{\varepsilon}{2}$

putting in all the terms makes it bigger

for all n!

<u>But</u>: we know

$$\left|\sum_{k \in A_\delta}\right| \le 2M \sum_{A_\delta} \binom{n}{k} x^k (1-x)^{n-k} \le 2M \frac{(x(1-x))}{n\delta^2}$$

$|f(\frac{k}{n}) - f(x)| \le 2M$
$M = \max f$ on $[0,1]$

$x(1-x) \le \frac{1}{4}$

Choose N st $\quad 2M \frac{\cancel{x(x-1)} \to 1/4}{N\delta^2} < \frac{\varepsilon}{2}$

$$N > \frac{4M}{4\delta^2}$$

Then $n > N \Rightarrow \left|\sum_{k \in A_\delta}\right| \le \frac{\varepsilon}{2}$ & $\left|\sum_{k=0}^{n}\right| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ ✓

☒ Bernstein Polynomials via Binomial Dist.

☒ Recall our proof that Bernstein polynomials converge uniformly to cont fn f on $[0,1]$:

$$B_n(f) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$

The idea:

$$\sum_{A_\delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{x(1-x)}{n\delta^2} \xrightarrow[n\to\infty]{} 0$$

$$A_\delta = \left\{k : \left|\tfrac{k}{n} - x\right| > \delta\right\}$$

In fact: What we have proven is the law of large numbers for the binomial distribution in probability theory.

Since we're already done the work, we might as well make the connection.

Assume a coin tossing experiment in which:

Prob of heads $= p \equiv x$

Prob of tails $= q \equiv 1-x$

- Assume n tosses of the coin.

An <u>outcome</u> is a sequence of n H's & T's

$$\begin{matrix} H & T & T & H & H & H & \cdots & T \\ 1 & 2 & 3 & 4 & 5 & 6 & \cdots & n \end{matrix}$$

There are 2^n possible sequences

- The probability of a given sequence with k H's & $n-k$ T's is $p^k q^{n-k}$

There are $\binom{n}{k}$ sequences with k H's & $n-k$ T's

$\therefore \binom{n}{k} p^k q^{n-k} = \text{Prob}\{X_n = k\}$

X_n = total number of H's after n trials

- Note: $\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = (1+1)^n = 2^n$

so all sequences are accounted for ✓

- $$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

$\Rightarrow$ total probability is $= 1$ ✓

- X_n = total # of heads is a R.V. taking k values $1, \dots, n$ with

$$\text{Prob}\{X_n = k\} = \binom{n}{k} p^k q^{n-k}$$

<u>Conclude</u>: $\underset{\text{(mean)}}{E(X_n)} = \sum_{k=0}^{n} \underset{\substack{\uparrow \\ \text{value} \\ \text{of outcome}}}{k} \underbrace{\binom{n}{k} p^k q^{n-k}}_{\text{prob of that outcome}} \overset{\text{last time}}{=} nx$

- (We used $k\binom{n}{k} = n\binom{n-1}{k-1}$)

- $Var(X_n) = \sigma_x^2 = \sum (k - nx)^2 \underbrace{\binom{n}{k} p^k q^{n-k}}_{\text{prob of the outcome}}$

value of outcome — mean

last time

$= nx(1-x) \equiv npq$

"measures the spread of probable outcomes from the mean"

$\sigma_x = SD = \sqrt{npq}$

- Our Main Lemma Proved –

$$\sum_{k \in A_\delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{x(1-x)}{n\delta^2}$$

$$A_\delta = \left\{ k : \left| \frac{k}{n} - x \right| > \delta \right\}$$

Prob sum/n after n-trials is further than δ from x.

This says —

$$\text{Prob}\left\{\left|\frac{1}{n}\bar{X}_n - x\right| > \delta\right\} = \frac{x(1-x)}{n\delta^2}$$

or

$$\text{Prob}\left\{\left|\frac{1}{n}\bar{X}_n - p\right| > \delta\right\} = \frac{p(1-p)}{n\delta^2}$$

Weak Law of Large Numbers —

$$\lim_{n\to\infty} \text{Prob}\left\{\left|\underbrace{\frac{1}{n}\bar{X}_n}_{\text{Ave}} - \underbrace{\mu}_{\text{mean}}\right| > \delta\right\} = 0$$

$$\bar{X}_n = X_1 + \cdots + X_n \qquad \text{iid RV's}$$

$X_k \equiv$ RV. for kth toss
$= \begin{cases} 1 & \text{if } H \text{ prob } p \\ 0 & \text{if } T \text{ prob } 1-p=q \end{cases}$

We have proven the weak law of large numbers with an estimate.
for binomial distribution

Central Limit Theorem:

$$N(\mu,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

"prob dist with mean μ & variance σ^2"

$X_1, \dots, X_n$ all iid RV's mean μ, var σ^2

$$\Rightarrow \quad \overline{X}_n = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$X_i \equiv i\text{th trial} \equiv \begin{cases} 1 & \text{prob } p \\ 0 & \text{prob } 1-p \end{cases}$$

$$E(X_i) = \text{mean} = \mu = p$$

$$\mathrm{Var}(X_i) = \underbrace{(0-p)^2}_{\substack{\text{outcome} \\ \text{minus mean}}} \cdot \underbrace{(1-p)}_{\substack{\text{prob} \\ \text{of outcome}}} + (1-p)^2 p = p^2(1-p) + (1-p)^2 p$$

$$= p(1-p)(p+1-p)$$

$$= p(1-p)$$

$$\therefore \overline{X}_n \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right) \text{ in limit } n \to \infty.$$

For us —

$$\text{Prob}\left\{\frac{1}{n}\overline{X}_n = \frac{k}{n}\right\} = \frac{1}{n}\binom{n}{k} p^k (1-p)^{n-k}$$

CLT $\Rightarrow$

$$\text{Prob}\left\{\frac{1}{n}\overline{X}_n \leq y\right\} = \frac{1}{n}\sum_{\frac{k}{n} \leq y} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\sim \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\, p} e^{-\frac{(x-p)^2}{2\frac{p(1-p)}{n}}}\, dx$$

(Pf. in prob class —)