[7.4] Approximations to the Binomial Dist.

In the last section, we learned about three examples

First, we let $X_i \sim BIN(1,p)$. Then $Y_n = \sum_{i=1}^{n} X_i$. $Y_n \sim BIN(n,p)$

- We found the dist for $Y_n \xrightarrow{d} POI(\mu)$ when $n \to \infty$, $p \to 0$ and $\mu = np$.

  or $Y_n \dot{\sim} POI(\mu)$

- We also learned $$W_n = \hat{p}_n = \frac{Y_n}{n} \xrightarrow{d} DEG(p)$$

  or that $\hat{p}_n$ converges stochastically to p.

- The last example taught $$Z_n = \frac{Y_n - np}{\sqrt{npq}} \xrightarrow{d} Z \approx N(0,1)$$

This followed by either finding the limit of the MGFs of Z_n or using the Central Limit Theorem itself

We can also write $Y_n \dot{\sim} N(np, npq)$ $\dot{\sim}$ = approximately dist.

EX: The prob a basketball player hits a shot is 0.50.
If he takes 20 shots, what is the prob. he hits at least 9?

Exact : $P[Y_{20} \geqslant 9] = 1 - P[Y_{20} \leqslant 8]$

$$= 1 - \sum_{y=0}^{8} \binom{20}{y} 0.5^y (1 - .5)^{20-y} = \boxed{0.7483}$$

Estimated - We can use the Normal dist as an approx to this! We assume $Y \mathrel{\dot\sim} N(np, npq)$ or $\dot\sim N(10, 5)$

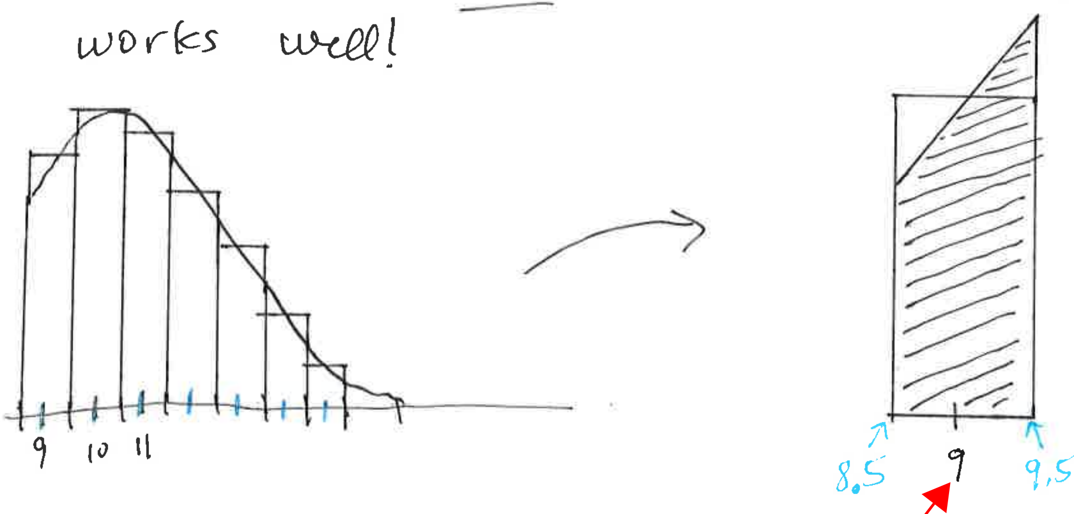
Thus, $P[Y_{20} \geqslant 9] = 1 - P[Y_{20} \leqslant 8] \doteq 1 - \Phi\left(\frac{8 - 10}{\sqrt{5}}\right)$

$$= 1 - \Phi(-0.89) = 1 - 0.1867 = \boxed{0.8133}$$

Note that .8133 is not all that close to .7483.

That can be fixed! Remember that Binomial is a discrete dist and the Normal is continuous

If we consider that the exact number on the bars is in the middle, we can make an adjustment that works well!



Treat the 9th bar as if 9 were in the middle of 8.5 and 9.5

So $P[X=9] \doteq P\left[\frac{8.5-10}{\sqrt{5}} < Z < \frac{9.5-10}{\sqrt{5}}\right] = P[-.67 < Z < -.22] = .4129 - .2514$

$= .1615$

Compare that to the exact answer $\binom{20}{9}.5^{20} = .1602$

close!

So, applying the concept with the previous statement gives

$$P[Y_{20} \geq 9] = P\left[Z > \frac{8.5 - 10}{\sqrt{5}}\right] = P[Z > -.67] = .7486$$

Exact $= .7483$

Big Improvement

<u>Ex</u>: Suppose $Y_n \sim POI(n)$. This arises from $X_i \sim POI(1)$ and $Y_n = \sum X_i \sim POI(n)$. Thus, by the CLT, $Y_n \overset{\cdot}{\sim} N(n, n)$. Suppose we want to find $P[10 \leq Y_{20} \leq 30]$?

Exact Answer: $P[10 \leqslant Y_{20} \leqslant 30] = \sum_{y=10}^{30} \frac{e^{-20} 20^y}{y!} = \boxed{0.982}$

WOW!

Approx: $P[10 \leqslant Y_{20} \leqslant 30] \doteq P\left[\frac{9.5 - 20}{\sqrt{20}} \leqslant Z \leqslant \frac{30.5 - 20}{\sqrt{20}}\right]$

$$= P[-2.35 \leqslant Z \leqslant 2.35] = \boxed{0.9811}$$