

[7.4] Approx. for the Binomial Dist.

Ex: Let  $X_1, \dots, X_n$  be a RS from a Bernoulli

$$X_i \sim \text{BIN}(1, p)$$

Consider  $Y_n = \sum_{i=1}^n X_i$

Let  $p \rightarrow 0$  as  $n \rightarrow \infty$  in such a way that  $\mu = np$  is fixed.

$$M_{Y_n}(t) = (pe^t + q)^n = \left[ \frac{\mu}{n} e^t + 1 - \frac{\mu}{n} \right]^n$$

$$= \left( 1 + \frac{\mu(e^t - 1)}{n} \right)^n \rightarrow e^{\mu(e^t - 1)}$$

$$Y_n \xrightarrow{d} Y \sim \text{POI}(\mu)$$

Ex: Same description as above

$$W_n = \hat{p}_n = \frac{1}{n} Y_n = \frac{1}{n} \sum_{i=1}^n X_i$$

$$M_{W_n}(t) = E \left[ e^{\frac{tY_n}{n}} \right] = M_{Y_n} \left( \frac{t}{n} \right)$$

$$= (pe^{\frac{t}{n}} + q)^n$$

$$= \left[ p \left( 1 + \frac{t}{n} + \frac{t^2}{2n^2} + \frac{t^3}{3!n^3} + \dots \right) + 1 - p \right]^n$$

$$= \left[ 1 + \frac{pt}{n} + \underbrace{\frac{pt^2}{2n^2} + \frac{pt^3}{3!n^3} + \dots}_{\frac{d(n)}{n}} \right]^n$$

$$\rightarrow e^{pt}$$

Note:  $M_{Y_n}(t) = e^{pt}$

This is the MGF for a degenerate dist at  $y=p$ .

$$f(y) = \begin{cases} 1, & y=p \\ 0, & \text{otherwise.} \end{cases}$$

$Y \sim \text{DEGENERATE}(p)$ .

$$M_Y(t) = E[e^{tY}] = \sum_{\text{all } y} e^{ty} f(y) = e^{pt} f(p) = e^{pt}$$

So  $y$  converges stochastically to  $p$  as  $n \rightarrow \infty$ .

EX: Consider the "standardized" variable

$$Z_n = \frac{Y_n - np}{\sqrt{npq}} = \frac{Y_n - np}{\sigma_n}$$

$$M_{Z_n}(t) = E[e^{tZ_n}] = E\left[\exp\left\{t \frac{Y_n - np}{\sigma_n}\right\}\right]$$

$$= e^{-\frac{npt}{\sigma_n}} E\left[e^{\frac{t}{\sigma_n} Y_n}\right] = e^{-\frac{npt}{\sigma_n}} M_{Y_n}\left(\frac{t}{\sigma_n}\right)$$

$$= e^{-\frac{npt}{\sigma_n}} (pe^{\frac{t}{\sigma_n}} + q)^n$$

$$= \left[ e^{-\frac{pt}{\sigma_n}} (pe^{\frac{t}{\sigma_n}} + q) \right]^n$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$e^{-\frac{pt}{\sigma_n}} = 1 - \frac{p}{\sigma_n} t + \frac{1}{2} \frac{p^2}{\sigma_n^2} t^2 - \frac{1}{6} \frac{p^3}{\sigma_n^3} t^3 + \dots$$

$$e^{t/\sigma_n} = 1 + \frac{t}{\sigma_n} + \frac{1}{2\sigma_n^2} t^2 + \frac{1}{6\sigma_n^3} t^3 + \dots$$

$$\begin{aligned} pe^{t/\sigma_n} + q &= 1 - p + p + \frac{pt}{\sigma_n} + \frac{pt^2}{2\sigma_n^2} + \frac{pt^3}{6\sigma_n^3} + \dots \\ &= 1 + \frac{pt}{\sigma_n} + \frac{pt^2}{2\sigma_n^2} + \frac{pt^3}{6\sigma_n^3} + \dots \end{aligned}$$

$$e^{-\frac{pt}{\sigma_n}} (pe^{t/\sigma_n} + q)$$

$$= \left( 1 - \frac{p}{\sigma_n} t + \frac{1}{2} \frac{p^2}{\sigma_n^2} t^2 - \frac{1}{6} \frac{p^3}{\sigma_n^3} t^3 + \dots \right) \times$$

$$\left( 1 + \frac{pt}{\sigma_n} + \frac{pt^2}{2\sigma_n^2} + \frac{pt^3}{6\sigma_n^3} + \dots \right)$$

$$\sigma_n^k = (\sqrt{npq})^k = n^{k/2} (pq)^{k/2}$$

$$\sigma_n^2 = npq = O(n) \quad \sigma_n^3 = O(n^{3/2})$$

$$= 1 + \frac{pt}{\sigma_n} - \frac{pt}{\sigma_n} - \frac{p^2 t^2}{\sigma_n^2} + \frac{1}{2} \frac{pt^2}{\sigma_n^2} + \frac{p^2 t^2}{2\sigma_n^2} + \frac{1}{6} \frac{pt^3}{\sigma_n^3} + \dots$$

$$= \left[ 1 + \frac{t^2}{\sigma_n^2} \left( \frac{p}{2} - p^2 + \frac{p^2}{2} \right) + \frac{\left( \text{stuff} / \sqrt{n} p^{3/2} q^{3/2} \right)}{n} \right]^n$$

$$= \left[ 1 + \frac{t^2}{\sigma_n^2} \frac{pq}{2} + \frac{d(n)}{n} \right]^n$$

$$= \left[ 1 + \frac{t^2 pq}{n pq 2} + \frac{d(n)}{n} \right]^n$$

$$= e^{t^2/2}$$

$$Z_n \xrightarrow{d} N(0,1)$$

Note: Since  $Z_n \sim N(0,1)$

This leads to the approx

$$Y_n \sim N(np, npq)$$

This works best when the Binomial is symmetric ( $p$  is near .5)

Ex: Prob a basketball player hits at least 9 free throws in 20. Suppose  $p = .5$

$$\text{Exact } P(Y_{20} \geq 9) = \sum_{y=9}^{20} \binom{20}{y} .5^{20} = .7483$$

$$\begin{aligned} \text{Est: } P(Y_{20} \geq 9) &= 1 - P(Y_{20} \leq 8) = 1 - \Phi\left(\frac{8-10}{\sqrt{5}}\right) \\ &= 1 - \Phi(-.89) = .8133 \end{aligned}$$

## 7.4 Continued

Ex.: Prob a basket ball player hits at least 9 of 20 free throws (suppose  $p=.5$ )

Exact

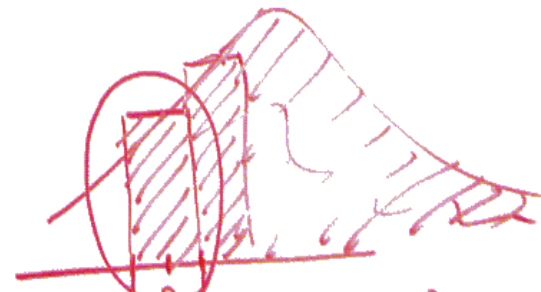
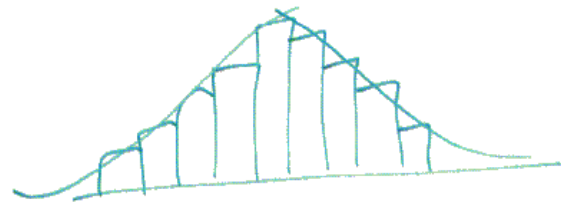
$$P(Y_{20} \geq 9) = \sum_{y=9}^{20} \binom{20}{y} .5^{20} = \boxed{.7483}$$

Est.  $Y_n \sim N(n\overset{=10}{p}, n\overset{=5}{pq})$

$$\begin{aligned} P(Y_n \geq 9) &= 1 - P(Y_n \leq 8) = 1 - \Phi\left(\frac{8-10}{\sqrt{5}}\right) \\ &= 1 - \Phi(-.89) = \boxed{.8133} \end{aligned}$$

These aren't all that close.

We're to estimate a discrete dist with a continuous one.



$$\begin{aligned} P(Y_n \geq 9) &= P(Y_n \geq 8.5) \\ &= 1 - P(Y_n \leq 8.5) = 1 - \Phi\left(\frac{8.5-10}{\sqrt{5}}\right) \\ &= \boxed{.7486} \end{aligned}$$