Large Population (250,000)

proportion of "support"   support 50,000

not 200000

$p = 0.2$

Draw one: $X_1 = \begin{cases} 1 & \text{support} \\ 0 & \text{not} \end{cases}$

$Pr(X_1 = 1) = 0.2$

Draw another (without replacement)

$Pr(X_2 = 1) = \frac{49,999}{249,999}$ or $\frac{50,000}{249,999}$

≈ 0.2

...

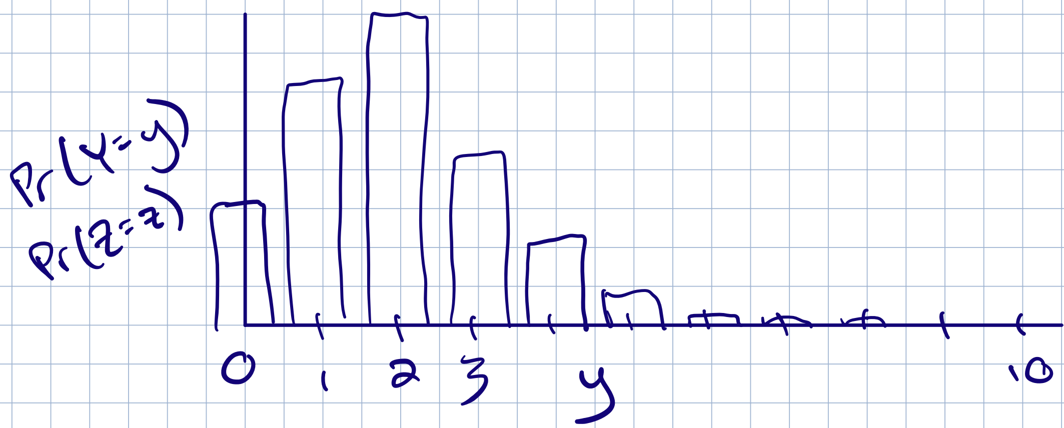
$Pr(X_{10} = 1) \approx 0.2$

What probability model can we use for the total number of "support" in a sample of size 10?

$$Y = X_1 + \ldots + X_{10} = \sum_{i=1}^{10} X_i$$

$$Y \overset{\text{approx}}{\sim} \text{Binomial}(p = 0.2, n = 10)$$

$$Pr(Y = y) = p^y (1-p)^{n-y}$$



$$0 \quad \frac{1}{10} \quad \frac{2}{10} \quad \frac{3}{10} \quad z \quad \frac{9}{10} \quad 1$$

What is the probability model for the sample proportion?

$$\hat{p} = Z = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n} Y$$

$$\Pr(Z = 1/n) = \Pr(Y = 1)$$

$$\Pr(Z = 2/n) = \Pr(Y = 2)$$

$$\boxed{\Pr(Z = z) = p^{nz}(1-p)^{n-nz}}$$

$\Rightarrow$ the sampling distribution of the sample proportion is a <u>rescaled binomial distribution</u>

$\Rightarrow$ Recall from Chapter 4:
(4.3.2)

When n is large and
$np \geq 10$
$n(1-p) \geq 10$

Then $\text{Binomial}(n,p)$
$\approx \text{Normal}(np, \sqrt{np(1-p)})$

$\Rightarrow$ Now, under the same conditions:

$$\hat{p} \overset{\text{approx}}{\sim} (p, \sqrt{\frac{p(1-p)}{n}})$$

(we will talk about large N in lecture 16.
For now, focus on small N)