

## 2.4 The Binomial Distribution

## → Bernoulli Experiment

1 0

A **Bernoulli experiment** is a random experiment where the outcome can be classified as one of two mutually exclusive ways (Heads/Tails, Pass/Fail)

A sequence of Bernoulli trials occurs when a Bernoulli experiment is performed several **independent** times, and the success probability,  $p$ , remains the same.

coin  $p = 0.5$

# Bernoulli experiment

- e.g. Flipping a fair coin. If I count the event “heads” as a success, this is a Bernoulli experiment with  $p=0.5$ .
- If I toss the coin 10 times, results correspond to 10 Bernoulli trials with  $p=0.5$

# Bernoulli Distribution

$p$  prob of a success

$$X \sim \text{Bernoulli}(p)$$

"is distributed as"

If random variable,  $X$ , has a Bernoulli distribution:

$$\rightarrow f(x) = p^x (1-p)^{1-x},$$

$$x = \{0,1\}$$

if  $x=1$ ,  
 ~~$p^1(1-p)^0$~~

$$\rightarrow \square E[X] = \sum_{x=0}^1 x \underbrace{p^x (1-p)^{1-x}}_{f(x)} = 0(1-p) + 1(p) = \boxed{p}$$

$$\rightarrow \square \text{Var}[X] = \sum_{x=0}^1 (\underline{x} - \underline{p})^2 \underbrace{p^x (1-p)^{1-x}}_{f(x)} = \boxed{p(1-p)}$$

$$\rightarrow \square SD[X] = \sqrt{p(1-p)}$$

$$\text{Var}[X] = E[(x - \mu)^2]$$

$$X \sim \text{Bern}(p)$$

## Definition: Random sample

An observed sequence of  $n$  Bernoulli trials can be written as a vector of zeroes and ones, with length  $n$ . We call this a random sample of size  $n$  from a Bernoulli distribution.

- $X_i$  denotes the Bernoulli random variable associated with the  $i^{\text{th}}$  trial.

$$\begin{array}{ccccccc}
 X_1 & X_2 & X_3 & X_4 & X_5 & & \\
 \hline
 1 & 0 & 0 & 0 & 1 & & 
 \end{array}$$

$x_1 = 1$        $i = 4$        $n = 5$

(Note: In the original image,  $X_4$  is circled, and  $X_5$  has an arrow pointing to it from the label  $i=4$ . There is also a stray  $X_4$  to the right.)

1  
0

succ  
fail

$$p = 0.3$$

$$p^x (1-p)^{1-x}$$

$$0.3 (1-0.3)^{1-0} = 0.7$$

## Example: Sequence of Bernoulli Trials

Suppose 30% of all lottery tickets are winners. If five tickets are purchased, (0, 1, 0, 0, 1) is one possible observed outcome.

0.3  
0.7

X<sub>1</sub> F S F F S

X<sub>2</sub>

{0, 1}

Assuming independence, the probability of this exact outcome is  $(0.7)(0.3)(0.7)(0.7)(0.3) = (0.7)^3(0.3)^2$

$$0 = x_1 \quad x_2 = 1$$

$p$  is prob of getting a 1  
"success"

▪  $X \sim \text{Bernoulli}(p=0.3)$

$x = \{0, 1\}$



## Binomial Distribution

Often, we are interested only in the total number of successes, but not the actual order of occurrence.

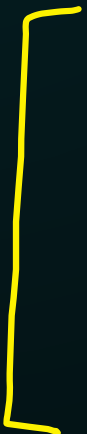
If we let  $X$  = the # of observed successes in  $n$  Bernoulli trials, then the possible values of  $X$  are  $0, 1, 2, \dots, n$ .

- For  $x$  successes, there are  $n - x$  failures.
- $X$  has a **binomial distribution**.



# Binomial Distribution

$X$  is a binomial random variable if the following are all true

- 
1. A Bernoulli (success/fail) experiment is performed a constant number of times,  $n$ . *Fixed*
  2. The random variable,  $X$ , is the number of successes in  $n$  trials.
  3. All trials are independent
  4. The success probability,  $p$ , for every trial is constant. (The failure probability,  $1 - p$ , is also constant).

# Binomial Distribution

total # success in  $n$  trials

Notation:

$$X \sim \text{Binomial}(n, p)$$

pmf

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

★

$$E[X] = np$$

★

$$Var[X] = np(1-p)$$

if  $n=1$  : Bernoulli Dist

# Lottery Ticket - Binomial Example

$$p = .3$$

Suppose 30% of lottery tickets are winning tickets.

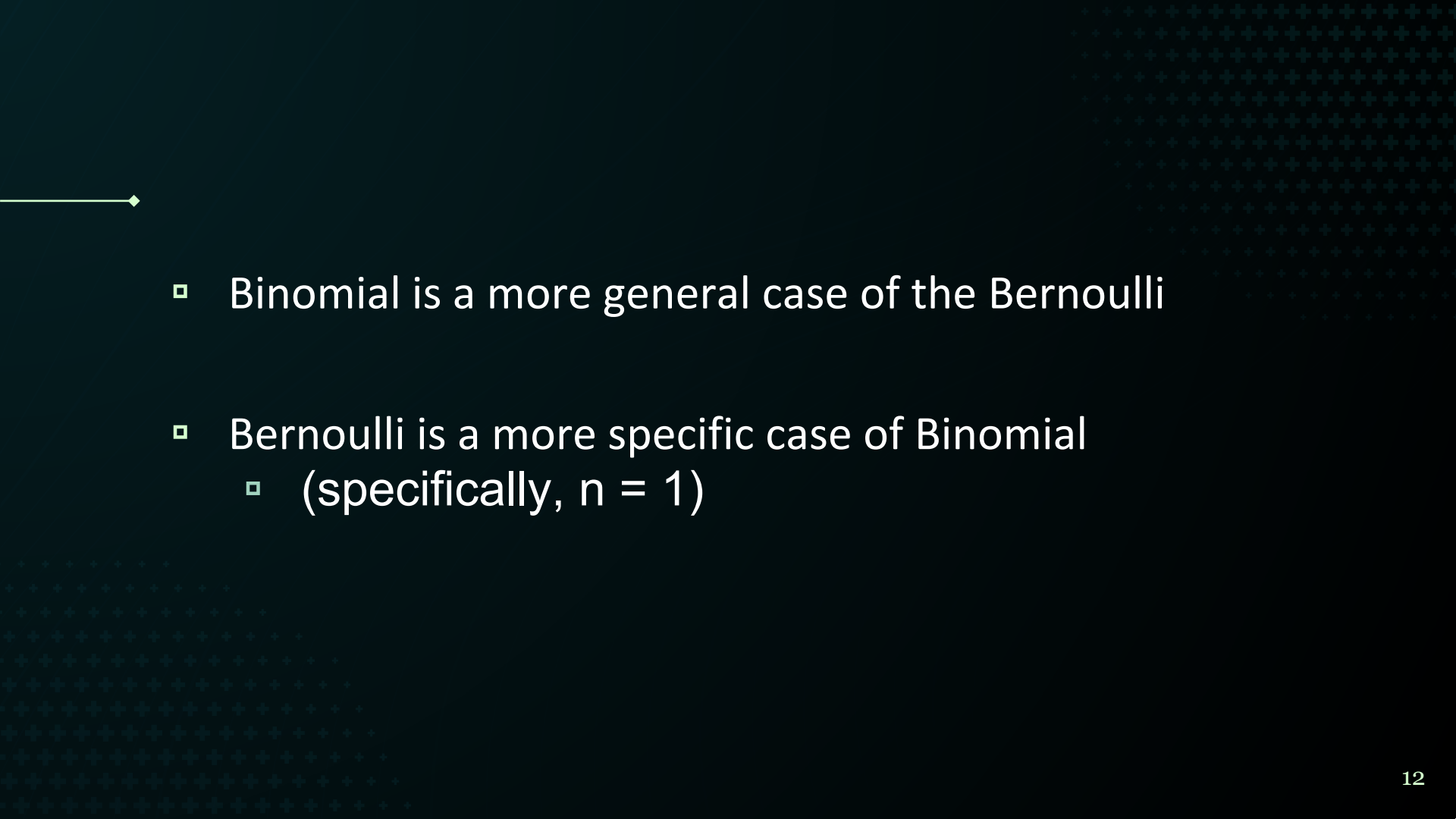
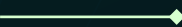
Let  $X$  = the number of winning tickets out of  $n=5$  ←  
purchased. The probability of purchasing two winning  
tickets is

$$x=2$$

$$\binom{n}{x} p^x (1-p)^{n-x}$$

$$f(2) = P(X = 2) = \underbrace{\binom{5}{2} 0.3^2 (1 - 0.3)^{5-2}}_{\text{stats}} =$$

stats ↑

- 
- 
- Binomial is a more general case of the Bernoulli
  - Bernoulli is a more specific case of Binomial
    - (specifically,  $n = 1$ )

## [ 2.5 Negative Binomial & Geometric Distribution

flip coin until 1st H

## Geometric Distribution

Say we observe a sequence of independent Bernoulli trials until the first success occurs.

If  $X$  is the number of trials needed to observe the 1<sup>st</sup> success, then  $X$  follows a **Geometric Distribution** with parameter,  $p$ .

$p = .2$  1st success on 8th trial?

$$f(8) = \underline{0.2} (1 - 0.2)^7$$

Geometric Distribution

$f(x; p)$

$$\rightarrow f(x) = p(1-p)^{x-1},$$

$$x = 1, 2, 3, \dots$$

$$X \sim \text{Geom}(p)$$

$$\begin{cases} \square \underline{E[X]} = 1/p \\ \square \text{Var}[X] = \frac{1-p}{p^2} \end{cases}$$

$$\sum_{x=1}^{\infty} x \cdot f(x) = 1/p$$

$$p = 0.5$$

$$\frac{1}{0.5} = 2$$

$$\left. \vphantom{\frac{1}{0.5}} \right\} p = 0.1$$

$$\frac{1}{0.1} = 10$$

$$P\{X \leq x\} = \sum_{x=1}^x x \cdot p(1-p)^{x-1}$$

# Pop Quiz

Can you show that the geometric distribution is a valid pmf?     *How?*



# Negative Binomial Distribution

More generally, suppose we observe a sequence of independent Bernoulli trials until the  $r^{\text{th}}$  success occurs.

If  $X$  is the number of trials needed to observe the  $r^{\text{th}}$  success, then  $X$  follows a **Negative Binomial** distribution with parameters  $r, p$ .

success prob = 0.4

How many trials until I get 3rd success?

$p = 0.4$

$r = 3$

$$X \sim NB(3, 0.4)$$

## Negative Binomial Distribution

~~$\binom{8}{3}$~~   $\binom{7}{2}$

$$X \sim NB(r, p)$$

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \dots$$

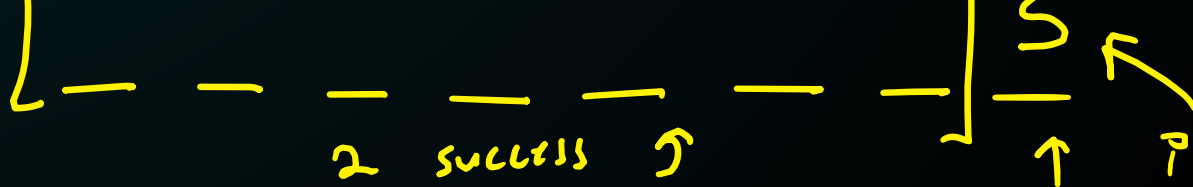
$$E[X] = r/p$$

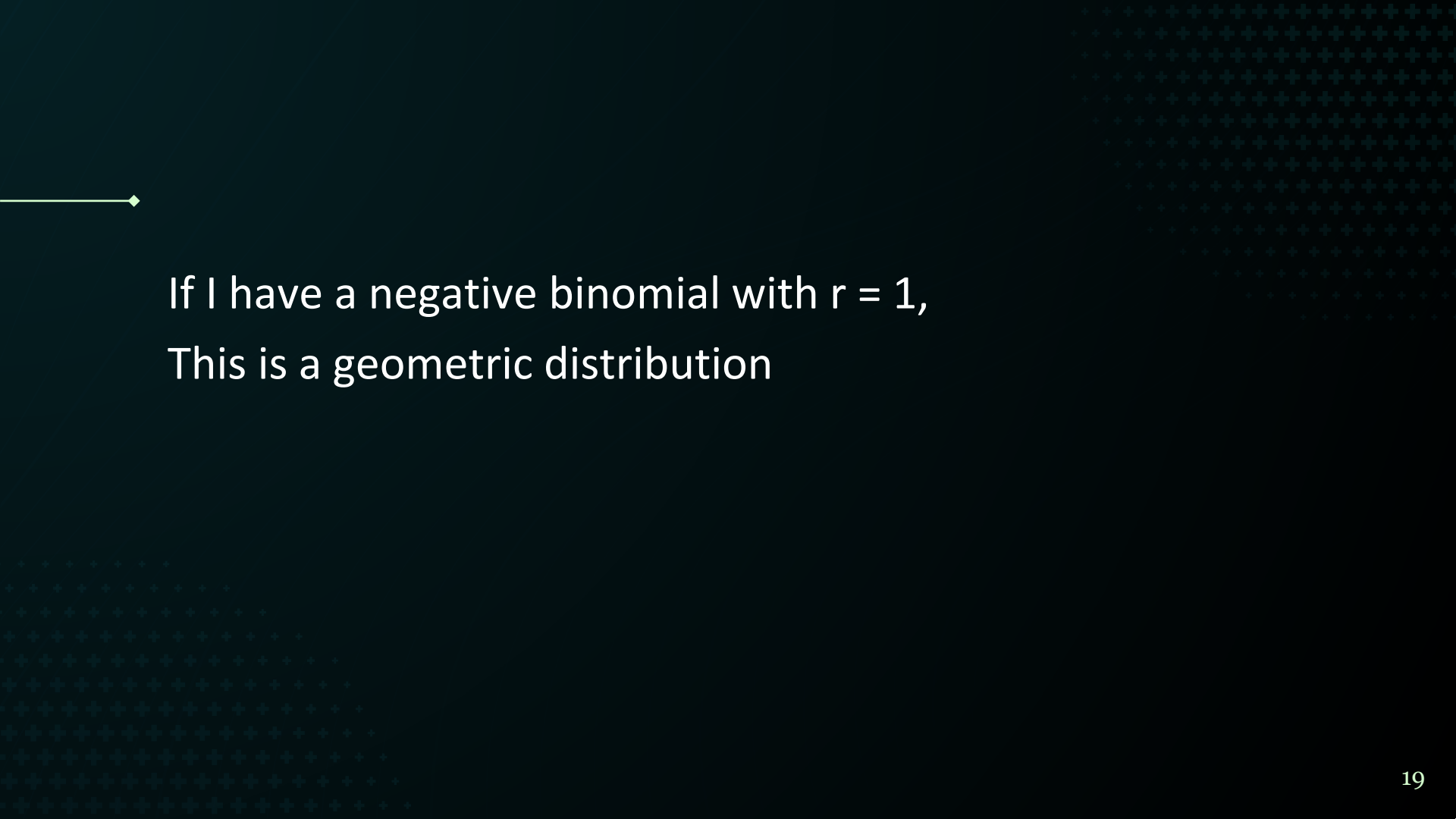
$$Var[X] = \frac{r(1-p)}{p^2}$$

$$f(x) = \binom{x-1}{3-1} 0.4^3 (0.6)^{x-3}$$

$r=3$

$f(8)$





If I have a negative binomial with  $r = 1$ ,  
This is a geometric distribution

# Examples

2.4 - 2.5

- 1 A magical beer machine vending machine gives a random beer to the customer. It gives you a stout 30% of the time and an IPA 70% of the time. Let  $X$  be the number of stouts you get out of 20 beers.  $p = 0.3$

$X \sim \text{Binom}$

$n = 20$

- What is the distribution of  $X$ ? What is its pmf?

$$\underline{X \sim \text{Binomial}(20, 0.3)} \quad f(x) = \binom{20}{x} 0.3^x (0.7)^{20-x}$$

$$P[X < 7]$$

- What is the probability of getting fewer than 7 stouts?

$$\underline{P[X \leq 6]} = \underbrace{P[X=0] + P[X=1] + \dots + P[X=6]}_{= .608}$$

$$\begin{aligned} \underline{d\text{binom}()} &\leftarrow \text{pmf} \\ \underline{p\text{binom}()} &\leftarrow \text{cdf} \end{aligned}$$

`pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)`

2 A magical beer machine vending machine randomly gives you a stout 30% of the time and an IPA 70% of the time. Thor gently smashes the machine until it gives him a stout. Let  $X$  represent the number of trials required for Thor to get his first stout.

$p$   $(1-p) \leftarrow \text{fail}$  OR  $q \leftarrow \text{fail}$   $p = 0.3$

What is the distribution of  $X$ ? What is its pmf?

$$X \sim \text{Geo}(0.3) \rightarrow f(x) = 0.3 (0.7)^{x-1}$$

What is the probability of getting a stout on the 5<sup>th</sup> trial? ✓

$$\overset{\text{pmf}}{\hookrightarrow} f(5) = 0.3 (0.7)^4 = \underline{0.07203} \quad \text{dgeom}(4, 0.3)$$

What is the probability of getting a stout within the first 5 trials?

$$P[X \leq 5] \quad \text{cdf} \quad \begin{array}{cc} p_{\text{geom}}( ) & \text{cdf} \\ d_{\text{geom}}( ) & \text{pmf} \end{array}$$

.8319

3 A random variable  $X$  has a binomial distribution

$$E[X] = np$$

$$Var[X] = np(1-p)$$

with  $\mu = 6$ ,  $\sigma^2 = 3.6$ .

What is the distribution of  $X$ ?

$$X \sim \text{Bin}(n=15, p=0.4)$$

$$\begin{aligned} 3.6 &= np(1-p) \\ \div 6 &= np \end{aligned}$$

$$\Rightarrow 0.6 = 1-p \Rightarrow p = 0.4$$

$$\begin{aligned} 6 &= (n)(0.4) \\ \Rightarrow n &= 15 \end{aligned}$$

Find  $P(X = 4)$ .

$$f(4) = \binom{15}{4} 0.4^4 (0.6)^{11} \leftarrow$$

Find  $F(2)$ . cdf  $P[X \leq 2] = \underline{f(0) + f(1) + f(2)}$

$$\begin{aligned} &dbinom(0, 15, 0.4) + dbinom(1, 15, 0.4) + dbinom(2, 15, 0.4) = .027 \\ &pbinom(2, 15, 0.4) = .027 \end{aligned}$$

4 Jacqueline hits her free throws with  $p = 0.9$ . *assume independence between ft.*

What is the probability that she has her first miss on the 7<sup>th</sup> free throw?

$$X \sim \text{Geom}(0.1)$$

$$f(7) = 0.1 (0.9)^6$$

$$p(\text{miss}) = 0.1 \quad p = 0.9$$

(hit all first 11)

What is the probability that she has her first miss on the 12<sup>th</sup> attempt or later?

$$P[X \geq 12] = 1 - P[X \leq 11]$$

$$\frac{\binom{29}{2}}{0.1} \frac{0.9^{27}}{0.1}$$

$$0.9^{11}$$

$$x$$

What is the probability that she has her 3<sup>rd</sup> miss on the 30<sup>th</sup> free throw?

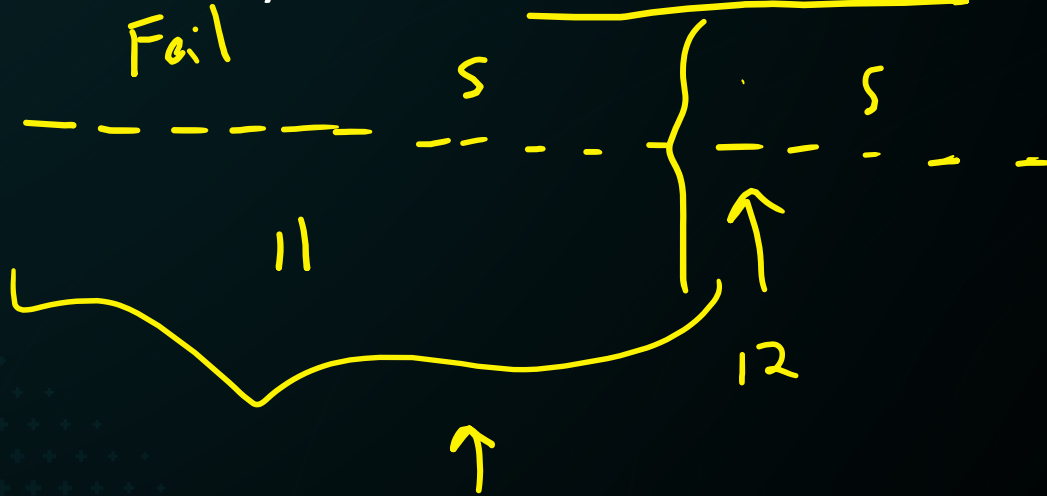
$$X \sim \text{NB}(3, 0.1)$$

$$f(30) = \left[ \binom{29}{2} \right] \underbrace{0.1}_2 \underbrace{0.9^{27}}_{27} \cdot 0.1$$

$$\text{or } \binom{29}{2} 0.1^3 0.9^{27}$$



- 
- Probability that her first ~~miss~~ <sup>success</sup> is on 12 or later?



$$p=0.1$$

$$f(x) = 0.1(0.9)^{x-1}$$

$$f(x) = p(1-p)^{x-1}$$

$$\frac{a}{1-r}$$

$$r=1-p$$

→ Jay Quellin

- Probability that first miss (success) occurs on 12 or later?

$$P[X \geq 12] = \sum_{x=12}^{\infty} p(1-p)^{x-1} = \frac{p(1-p)}{1-(1-p)}$$

$$= \frac{\cancel{p}(1-p)^{11}}{\cancel{p}}$$

$$= (1-p)^{11} = (0.9)^{11}$$

$$= 0.9^{11} = 0.313$$

$$1 - P[X \leq 11] = 1 - p_{\text{geom}}(10, 0.1)$$