intro to data science with probability & statistics

CSCI 3022

Lecture 7
Feb 5, 2018

1. More discrete RVs
2. Common distributions

HW 2 - posted tonight
Notetaker still wanted
Kyle & Sofie

Department of Computer Science
UNIVERSITY OF COLORADO BOULDER

Dan Larremore

Warmup - Discrete RVs!

- Dungeons & Dragons is a game played with special dice, with 4, 6, 8, 10, 12, and 20 sides. Say you roll the d8 and the d4, and define the random variable X to be the value of the d8 minus the value of the d4.
- **What is the PMF of X? What is the CDF of X?**

	1	2	3	4
1	0	-1	-2	-3
2	1	0	-1	-2
3	2	1	0	-1
4	3	2	1	0
5	4	3	2	1
6	5	4	3	2
7	6	5	4	3
8	7	6	5	4

$f(-3) = f(7) = 1/32$

$f(-2) = f(6) = 2/32$

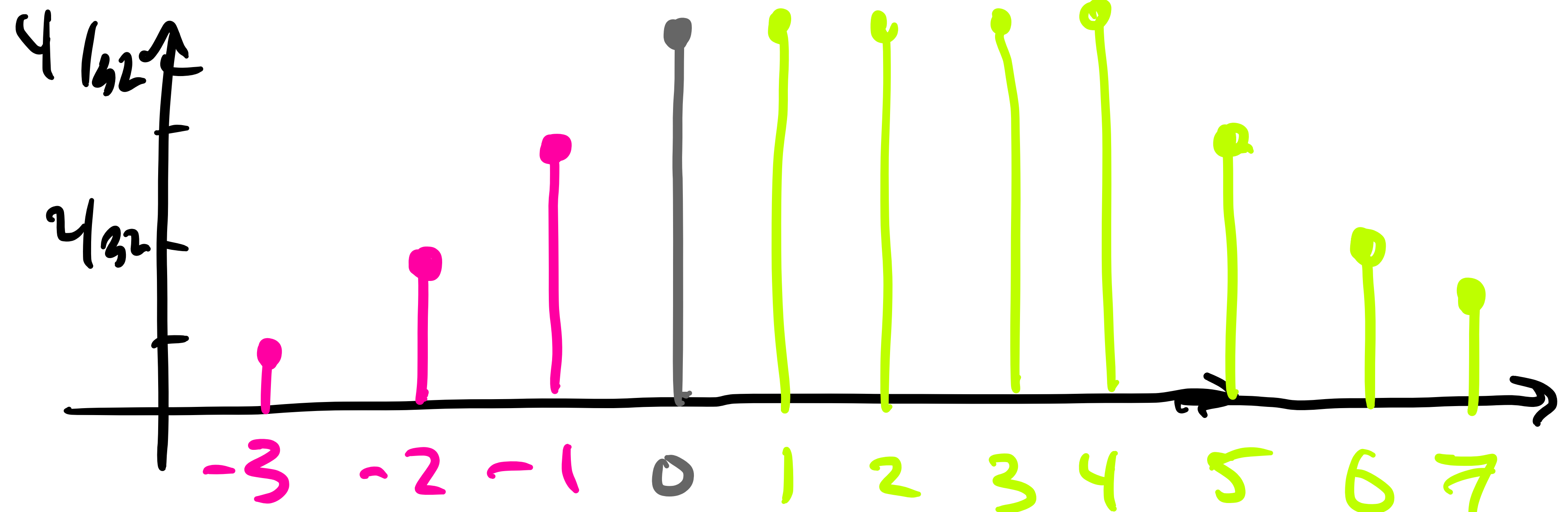
$f(-1) = f(5) = 3/32$

$f(0) = f(1) = f(2) = f(3) = f(4) = 4/32$

$F(-3) = 1/32$

$F(-2) = 3/32$

$F(-1) = 6/32, \ldots$



The Discrete Uniform Distribution

- **Definition**: the discrete *uniform* distribution assigns a probability mass of $\frac{1}{n}$ to each of n values in [a,b]. We write it as *unif(a,b).*

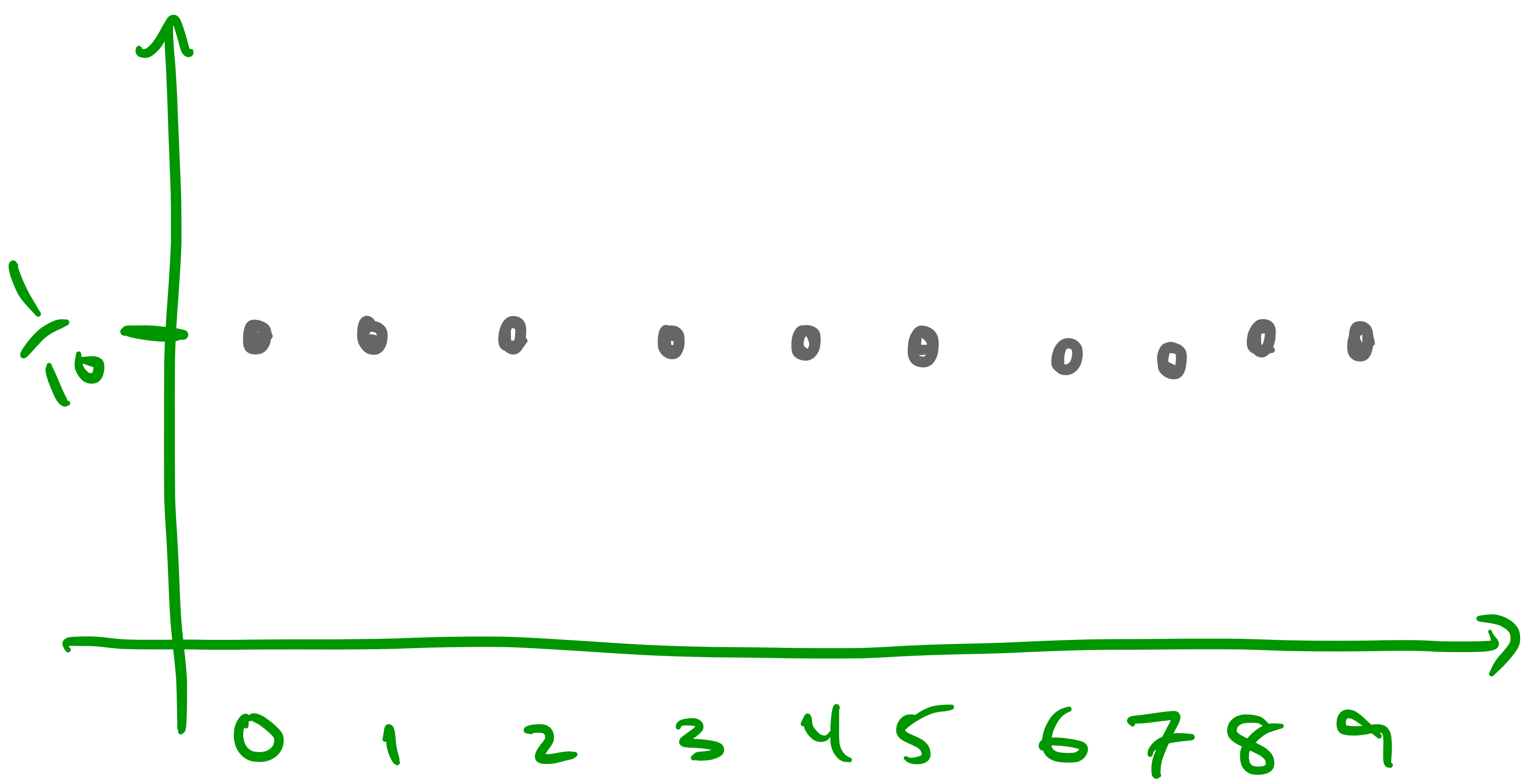
unif[0,2], n = 5

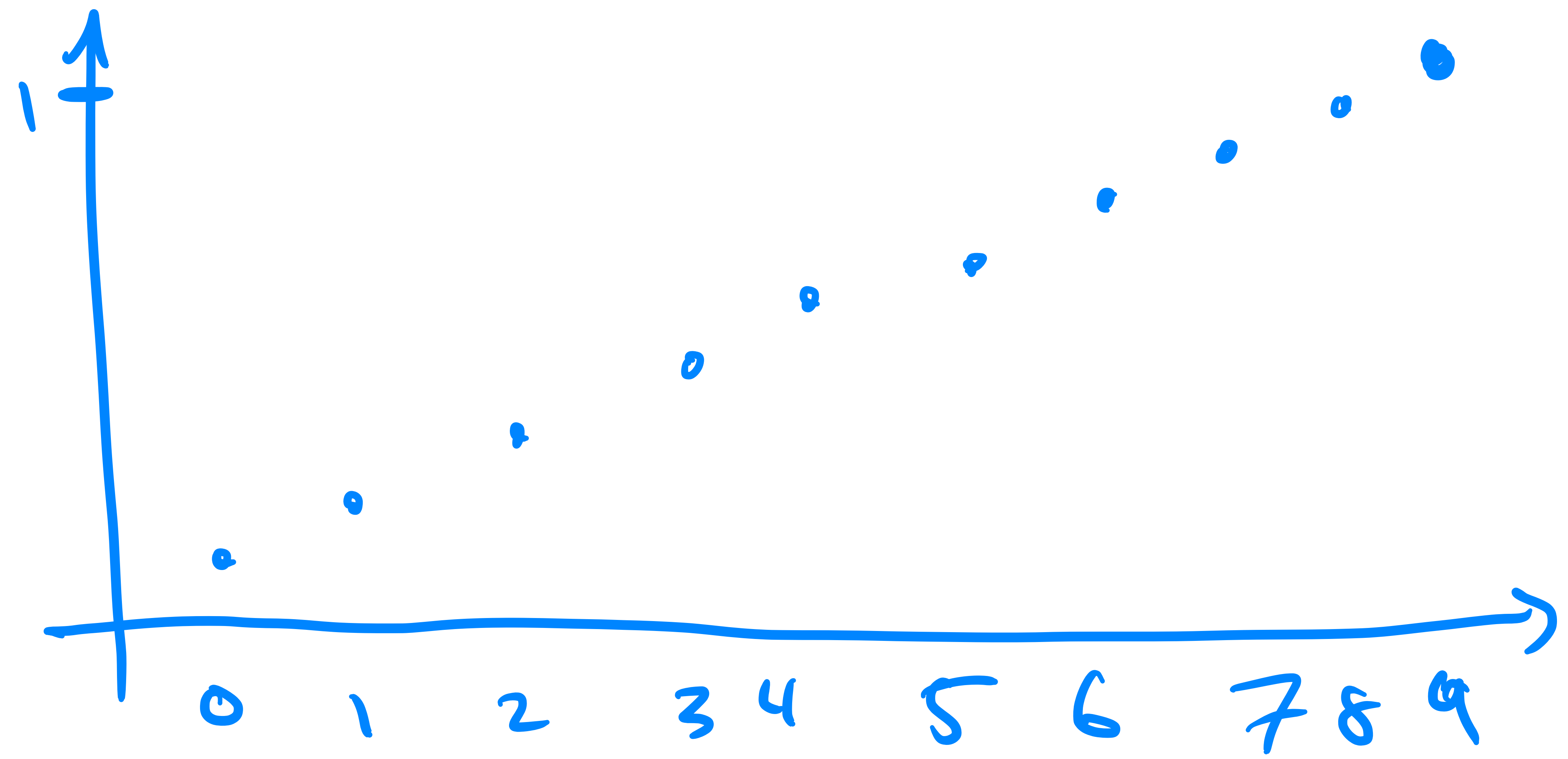
1/5	1/5	1/5	1/5	1/5
0	1/2	1	1 1/2	2

- **Ponder**: can you think of an example of a random experiment whose outcome is a discrete uniform distribution? What are *a* and *b*?

d6 ~ unif[1,6] n = 6 coin ~ unif[0,1] n = 2

- **Plot**: the probability mass function for unif[0,9], *n*=10. Then plot the CDF.

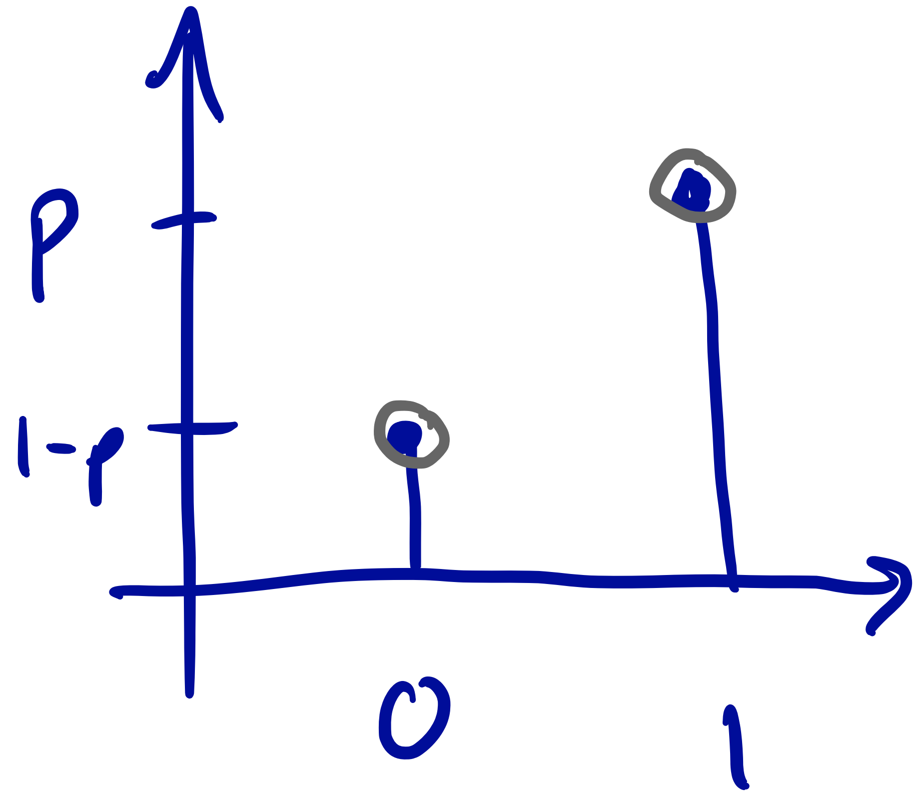




The Bernoulli Distribution

- **Definition**: a discrete RV *X* has a *Bernoulli distribution* with parameter *p*, where $0 \leq p \leq 1$, if its probability mass function (PMF) is given by:

$$f(1) = P(X=1) = p \quad \text{and} \quad f(0) = P(X=0) = 1\text{-}p$$



- We denote this distribution by *Ber*(p)

$X \sim Ber(p)$

- Look closely—what does this PMF remind you of?

biased coin

- What is the CDF of the Bernoulli distribution?

$F(0) = 1-p$

$F(1) = 1$

Counting

- Let's set aside the Bernoulli distribution for a moment and turn to counting.
- Believe it or not, *counting* comes up all over the place in probability, and therefore in CS, math, physics, engineering, data science, etc.
- Some counting is easy: how many integers are there in the interval [0,9]?
- We're interested in counting that requires math though: Ancsa, Brendan, Caterina, David, and Elly are standing in line. How many different orders could they stand in?
- If there are 10 problems on an exam and you need 7 or more correct to pass, how many different ways are there to pass the test?

Counting

- We'll talk about two key kinds of counting problems today:

- Counting **permutations** means counting the number of ways a set of objects can be ordered (or, more precisely, permuted).

Ancsa, Brendan, Caterina, David, and Elly are standing in line. How many different orders could they stand in?

- Counting **combinations** means counting the number of ways that a set of objects can combined into subsets.

If there are 10 problems on an exam and you need 7 or more correct to pass, how many different ways are there to pass the test?

Counting - Permutations

- How many ways are there of ordering 1 object? 1
- How many ways are there of ordering 2 objects? A, B or B, A 2
- What about 3 objects?

A B C	B A C	C A B
A C B	B C A	C B A

6

- What is the formula for the number of possible permutations of *n* objects? $n!$

$n = 5$

$$\underbrace{\quad}_{5 \text{ possib.}}^{\#1} \; \underbrace{\quad}_{4 \text{ poss}}^{\#2} \; \underbrace{\quad}_{3 \text{ p.}}^{\#3} \; \underbrace{\quad}_{2 \text{ p}}^{\#4} \; \underbrace{\quad}_{1 \text{ p}}^{\#5}$$

$$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$$

$$n! = n \cdot (n-1) \cdot (n-2) \cdots (2)(1)$$

$$0! \overset{\text{defn}}{=} 1$$

Counting - Permutations 2

in Korea.

- Say there are 10 people in a race. How many ways are there of awarding the gold, silver, and bronze?

$10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 10!$

places 1 2 3 | 4 5 6 7 8 9 10

G S B

1 $10 \times 9 \times 8 = 720$

2 10! total permutations, but 7! that I don't care about for each possible 1,2,3 order.

$$\frac{10!}{7!} = \frac{10 \cdot 9 \cdot 8 \cdot \cancel{7} \cdot \cancel{6} \cdot \cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot \cancel{1}}{\cancel{7} \cdot \cancel{6} \cdot \cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot \cancel{1}}$$

Counting - Combinations

- ~~We could approach the previous problem another way…~~

10 people…. ask: how many ways could I choose a subset of 3?

For each subset → 3! orderings.

call this [10 choose 3]

\# possible G, S, B $\quad [10 \text{ choose } 3] \cdot 3! = \dfrac{10!}{7!}$

$$[10 \text{ choose } 3] = \frac{10!}{7!\,3!}$$

$$\boxed{n \text{ choose } k = \frac{n!}{k!\,(n-k)!}}$$

Counting - Combinations

- The previous slide shows us how to use permutations to get to combinations.
- If I have n objects and I choose k of them, I can do this in n choose k ways.
- Various notations: $\binom{n}{k}$ $\quad n\cancel{C}k \quad$ $\frac{n!}{k!(n-k)!}$ $\quad C_{n,k}$

{n \choose k}

formula

book

Counting - Combinations

- The previous slide shows us how to use permutations to get to combinations.
- If I have *n* objects and I choose *k* of them, I can do this in *n* choose *k* ways.
- Various notations: $\binom{n}{k}$ $\quad nCk \quad$ $\frac{n!}{k!(n-k)!}$ $\quad C_{n,k}$
- So, how many ways are there to choose 3 runners from a total of 10?

$$\binom{10}{3} = \frac{10!}{3!(10-3)!} = \frac{10!}{3!\,7!} = \frac{10\cdot \cancel{9}^{3}\cdot \cancel{8}^{4}}{\cancel{3}\cdot\cancel{2}\cdot 1} = 10\cdot 12 = 120$$

Counting - Combinations

- Let's connect some dots.
- Let's compute: $C_{1,1}$ $\quad \frac{1!}{1!\,0!} = 1$
- $C_{2,1}$, $C_{2,2}$ $\quad \frac{2!}{1!\,1!} = 2 \qquad \frac{2!}{2!\,0!} = 1$
- $C_{3,1}$, $C_{3,2}$, $C_{3,3}$
- Now $C_{1,0}$, $C_{2,0}$, $C_{3,0}$. And $C_{0,0}$.
- What do you notice?

Pascal's Δ, Binomial Coeffs.

$$
\begin{array}{ccccccccc}
 & & & & C_{0,0} & & & & \\
 & & & C_{1,0} & & C_{1,1} & & & \\
 & & C_{2,0} & & C_{2,1} & & C_{2,2} & & \\
 & C_{3,0} & & C_{3,1} & & C_{3,2} & & C_{3,3} &
\end{array}
$$

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

Combinations applied!

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

- If there are 10 problems on an exam and you need 7 or more correct to pass, how many different ways are there to pass the test?

Pass: 7 8 9 10

$$\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}$$

$$120 + 45 + 10 + 1$$

$$165 + 11$$

$$\boxed{176}$$