Statistical Foundations of Data Science

# Statistical Foundations of Data Science Combinatorics Part 2

University of the Witwatersrand

2022

Statistical Foundations of Data Science

• Prove by induction that  $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$ 

Statistical Foundations of Data Science

Review Question

- Review Question
- Combinations and binomial coefficients

- Review Question
- Combinations and binomial coefficients
- Combinatorial identities

- Review Question
- Combinations and binomial coefficients
- Combinatorial identities
- Derangements

Statistical Foundations of Data Science

• Prove by induction that  $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$ 

- Prove by induction that  $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$
- Base case n = 1 both sides are equal to 1.

- Prove by induction that  $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$
- Base case n = 1 both sides are equal to 1.
- Assume true for n = k

- Prove by induction that  $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$
- Base case n = 1 both sides are equal to 1.
- Assume true for n = k
- $1^3 + 2^3 + 3^3 + ... + k^3 + (k+1)^3 = \frac{k^2(k+1)^2}{4} + (k+1)^3$

- Prove by induction that  $1^3 + 2^3 + 3^3 + .. + n^3 = \frac{n^2(n+1)^2}{4}$
- Base case n = 1 both sides are equal to 1.
- Assume true for n = k
- $1^3 + 2^3 + 3^3 + ... + k^3 + (k+1)^3 = \frac{k^2(k+1)^2}{4} + (k+1)^3$
- $\bullet = \frac{(k+2)^2(k+1)^2}{4}$

- Prove by induction that  $1^3+2^3+3^3+..+n^3=\frac{n^2(n+1)^2}{4}$
- Base case n = 1 both sides are equal to 1.
- Assume true for n = k
- $1^3 + 2^3 + 3^3 + ... + k^3 + (k+1)^3 = \frac{k^2(k+1)^2}{4} + (k+1)^3$
- $\circ = \frac{(k+2)^2(k+1)^2}{4}$
- So true by induction!

## Combinations

Statistical Foundations of Data Science

• We have n prisoners and are selecting k for execution! How many ways can we select these k.

#### Combinations

- We have n prisoners and are selecting k for execution! How many ways can we select these k.
- Alternatively we are selecting a team of k players from n who tried out.

#### **Combinations**

- We have n prisoners and are selecting k for execution! How many ways can we select these k.
- Alternatively we are selecting a team of k players from n who tried out.
- In either case we'll call the answer  $\binom{n}{k}$ , it's also occasionally called nCk.

# Combinations - Properties

Statistical Foundations of Data Science

> Before deriving a formula let's talk about some useful identities.

# Combinations - Properties

- Before deriving a formula let's talk about some useful identities.
- $\binom{n}{1} = n$

- Before deriving a formula let's talk about some useful identities.
- $\binom{n}{1} = n$
- $\binom{n}{0} = 1$

- Before deriving a formula let's talk about some useful identities.
- $\binom{n}{1} = n$
- $\binom{n}{0} = 1$

- Before deriving a formula let's talk about some useful identities.
- $\binom{n}{1} = n$
- $\binom{n}{0} = 1$
- $\binom{n}{k} = \binom{n}{n-k}$

## Combinations - Formula

Statistical Foundations of Data Science

• Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.

#### Combinations - Formula

- Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.
- Pick first *k*!.

#### Combinations - Formula

- Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.
- Pick first *k*!.
- Now we can reorder the n-k not picked in (n-k)! ways to get  $nPk = \frac{n!}{(n-k)!}$

- Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.
- Pick first *k*!.
- Now we can reorder the n-k not picked in (n-k)! ways to get  $nPk = \frac{n!}{(n-k)!}$
- As we don't care about the order we execute them in  $nCk = \binom{n}{k} = \frac{n!}{(n-k)!k!}$ .

- Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.
- Pick first *k*!.
- Now we can reorder the n-k not picked in (n-k)! ways to get  $nPk = \frac{n!}{(n-k)!}$
- As we don't care about the order we execute them in  $nCk = \binom{n}{k} = \frac{n!}{(n-k)!k!}$ .
- Properties hold  $\binom{n}{k} = \binom{n}{n-k}$  is obvious from the formula.

- Line up the *n* prisoners, in order of how much you feel like executing them. *n*! ways.
- Pick first *k*!.
- Now we can reorder the n-k not picked in (n-k)! ways to get  $nPk = \frac{n!}{(n-k)!}$
- As we don't care about the order we execute them in  $nCk = \binom{n}{k} = \frac{n!}{(n-k)!k!}$ .
- Properties hold  $\binom{n}{k} = \binom{n}{n-k}$  is obvious from the formula.
- $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$  is algebra.

# Pascal's Triangle

```
1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
```

# Pascal's Triangle

$$1 = \binom{0}{0}$$

$$1 = \binom{1}{0} \quad 1 = \binom{1}{1}$$

$$1 = \binom{2}{0} \quad 2 = \binom{2}{1} \quad 1 = \binom{2}{2}$$

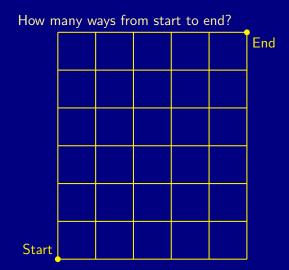
$$1 = \binom{3}{0} \quad 3 = \binom{3}{1} \quad 3 = \binom{3}{2} \quad 1 = \binom{3}{3}$$

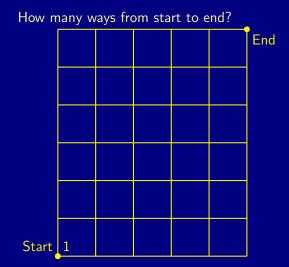
$$1 = \binom{4}{0} \quad 4 = \binom{4}{1} \quad 6 = \binom{4}{2} \quad 4 = \binom{4}{3} \quad 1 = \binom{4}{4}$$

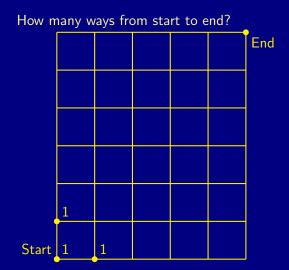
$$1 = \binom{5}{0} \quad 5 = \binom{5}{1} \quad 10 = \binom{5}{2} \quad 10 = \binom{5}{3} \quad 5 = \binom{5}{4} \quad 1 = \binom{5}{5}$$

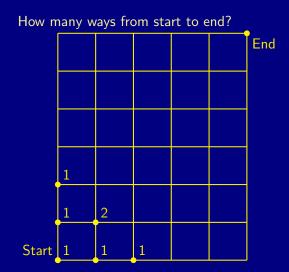
# Pascal's Triangle

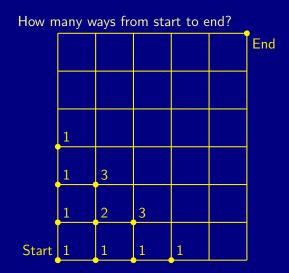
```
 \begin{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \begin{pmatrix} 2 \\ 2 \end{pmatrix} & \begin{pmatrix} 3 \\ 2 \end{pmatrix} & \begin{pmatrix} 3 \\ 1 \end{pmatrix} & \begin{pmatrix} 2 \\ 2 \end{pmatrix} & \begin{pmatrix} 3 \\ 2 \end{pmatrix} & \begin{pmatrix} 3 \\ 3 \end{pmatrix} & \begin{pmatrix} 3 \\ 4 \end{pmatrix} & \begin{pmatrix} 4 \\ 1 \end{pmatrix} & \begin{pmatrix} 4 \\ 2 \end{pmatrix} & \begin{pmatrix} 4 \\ 3 \end{pmatrix} & \begin{pmatrix} 4 \\ 4 \end{pmatrix} & \begin{pmatrix} 5 \\ 5 \end{pmatrix} & \begin{pmatrix} 5 \\ 0 \end{pmatrix} & \begin{pmatrix} 5 \\ 1 \end{pmatrix} & \begin{pmatrix} 5 \\ 2 \end{pmatrix} & \begin{pmatrix} 5 \\ 2 \end{pmatrix} & \begin{pmatrix} 5 \\ 3 \end{pmatrix} & \begin{pmatrix} 5 \\ 4 \end{pmatrix} & \begin{pmatrix} 5 \\ 5 \end{pmatrix} & \begin{pmatrix} 6 \\ 5 \end{pmatrix} & \begin{pmatrix} 7 \\ 6 \end{pmatrix} & \begin{pmatrix} 7 \\ 7 \end{pmatrix} & \begin{pmatrix} 7 \\ 0 \end{pmatrix} & \begin{pmatrix} 7 \\ 1 \end{pmatrix} & \begin{pmatrix} 7 \\ 2 \end{pmatrix} & \begin{pmatrix} 7 \\ 3 \end{pmatrix} & \begin{pmatrix} 7 \\ 4 \end{pmatrix} & \begin{pmatrix} 8 \\ 5 \end{pmatrix} & \begin{pmatrix} 7 \\ 6 \end{pmatrix} & \begin{pmatrix} 7 \\ 7 \end{pmatrix} & \begin{pmatrix} 7 \\ 7 \end{pmatrix} & \begin{pmatrix} 7 \\ 4 \end{pmatrix} & \begin{pmatrix} 8 \\ 5 \end{pmatrix} & \begin{pmatrix} 7 \\ 6 \end{pmatrix} & \begin{pmatrix} 7 \\ 7 \end{pmatrix} & \begin{pmatrix}
```

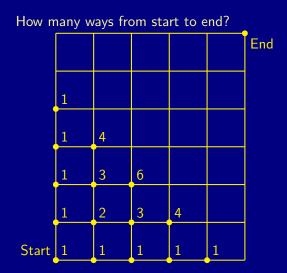


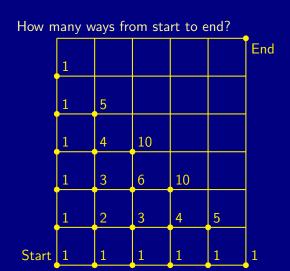


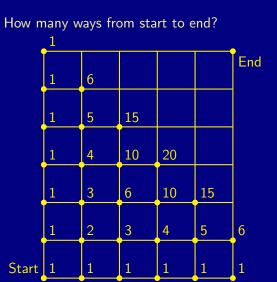


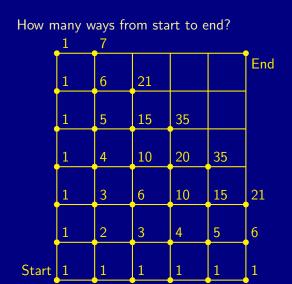


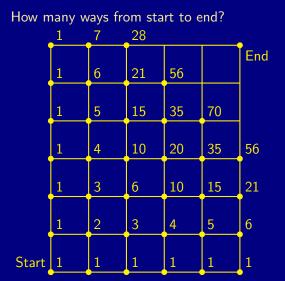


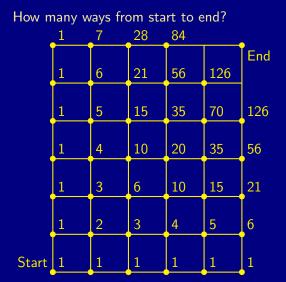


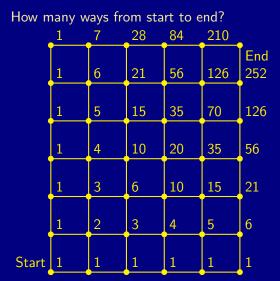


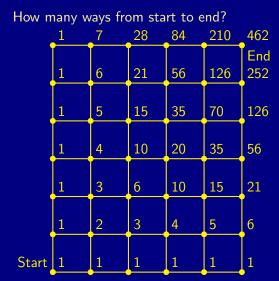












Statistical Foundations of Data Science

• Alternative solution. We'll make 11 steps. 6 up and 5 right. Choose 6 of the 11 steps to be "up steps.  $\binom{11}{6}$ 

- Alternative solution. We'll make 11 steps. 6 up and 5 right. Choose 6 of the 11 steps to be "up steps.  $\binom{11}{6}$
- $\binom{11}{6} = \frac{11!}{5!6!}$

- Alternative solution. We'll make 11 steps. 6 up and 5 right. Choose 6 of the 11 steps to be "up steps.  $\binom{11}{6}$
- $\binom{11}{6} = \frac{11!}{5!6!}$
- $\bullet = \frac{11 \times 10 \times 9 \times 8 \times 7}{5 \times 4 \times 3 \times 2 \times 1}$

- Alternative solution. We'll make 11 steps. 6 up and 5 right. Choose 6 of the 11 steps to be "up steps.  $\binom{11}{6}$
- $\binom{11}{6} = \frac{11!}{5!6!}$
- $0 = \frac{11 \times 10 \times 9 \times 8 \times 7}{5 \times 4 \times 3 \times 2 \times 1}$
- $\bullet = \frac{11 \times 3 \times 2 \times 7}{1} = 462$

# Binomial Theorem and it's relation to choosing

• 
$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$$

## Binomial Theorem and it's relation to choosing

Statistical Foundations of Data Science

• 
$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$$

• To see this write out (x + y)(x + y)...(x + y).

## Binomial Theorem and it's relation to choosing

- $(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$
- To see this write out (x + y)(x + y)...(x + y).
- coefficent of  $x^i y^{n-i}$ , well choose i of the n to have the x and the other n-i to have the y

## Choosing problems

Statistical Foundations of Data Science

> You are given a 5 card poker hand. How many ways can you get a full house?

### Choosing problems

- You are given a 5 card poker hand. How many ways can you get a full house?
- $\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2}=3744$

- You are given a 5 card poker hand. How many ways can you get a full house?
- $\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2} = 3744$
- The probability of drawing such a hand?  $3744/\binom{52}{5} = 0.00144057623$

#### Multinomial Coefficients

Statistical Foundations of Data Science

> Suppose I have n prisoners and I want to choose s for a soccer team and g to guillotine!

- Suppose I have n prisoners and I want to choose s for a soccer team and g to guillotine!
- $\binom{n}{s,g}\binom{n}{s}\binom{n-s}{g} = \frac{n!}{s!(n-s)!} \times \frac{(n-s)!}{g!(n-s-g)!} = \frac{n!}{s!g!(n-s-g)!}$

- Suppose I have n prisoners and I want to choose s for a soccer team and g to guillotine!
- This can also be done combinatorially!

Statistical Foundations of Data Science

• Recall that there are n! ways to arrange the first n integers.

- Recall that there are n! ways to arrange the first n integers.
- Example there are 5! = 120 ways to arrange the five numbers 1 through 5. For example. 2, 4, 5, 1, 3 and 1, 2, 3, 4, 5.

- Recall that there are n! ways to arrange the first n integers.
- Example there are 5! = 120 ways to arrange the five numbers 1 through 5. For example. 2, 4, 5, 1, 3 and 1, 2, 3, 4, 5.
- How many ways are there to arrange these numbers if no number can remain in it's natural position?

- Recall that there are n! ways to arrange the first n integers.
- Example there are 5! = 120 ways to arrange the five numbers 1 through 5. For example. 2, 4, 5, 1, 3 and 1, 2, 3, 4, 5.
- How many ways are there to arrange these numbers if no number can remain in it's natural position?
- Example 1, 4, 3, 5, 2 wouldn't count because of the 1.

- Recall that there are n! ways to arrange the first n integers.
- Example there are 5! = 120 ways to arrange the five numbers 1 through 5. For example. 2, 4, 5, 1, 3 and 1, 2, 3, 4, 5.
- How many ways are there to arrange these numbers if no number can remain in it's natural position?
- Example 1, 4, 3, 5, 2 wouldn't count because of the 1.
- Example 5, 3, 4, 1, 2 would count.

- Recall that there are n! ways to arrange the first n integers.
- Example there are 5! = 120 ways to arrange the five numbers 1 through 5. For example. 2, 4, 5, 1, 3 and 1, 2, 3, 4, 5.
- How many ways are there to arrange these numbers if no number can remain in it's natural position?
- Example 1, 4, 3, 5, 2 wouldn't count because of the 1.
- Example 5, 3, 4, 1, 2 would count.
- Example 3, **2**, 4, 1, **5** wouldn't coutn because of the 2 and the 5.

Statistical Foundations of Data Science

Inclusion-exclusion!

- Inclusion-exclusion!
- There are n! arrangements, we subtract off the  $n \times (n-1)!$  that have each particular fixed point.

- Inclusion-exclusion!
- There are n! arrangements, we subtract off the  $n \times (n-1)!$  that have each particular fixed point.
- We then add back the ones that have each pair of 2 fixed points! There are  $\binom{n}{2} \times (n-2)!$  of these.

- Inclusion-exclusion!
- There are n! arrangements, we subtract off the  $n \times (n-1)!$  that have each particular fixed point.
- We then add back the ones that have each pair of 2 fixed points! There are  $\binom{n}{2} \times (n-2)!$  of these.
- Again we need to subtract off those taken 3 at a time and add back those taken 4 at a time and so on.

- Inclusion-exclusion!
- There are n! arrangements, we subtract off the  $n \times (n-1)!$  that have each particular fixed point.
- We then add back the ones that have each pair of 2 fixed points! There are  $\binom{n}{2} \times (n-2)!$  of these.
- Again we need to subtract off those taken 3 at a time and add back those taken 4 at a time and so on.
- Finally we get  $n!(1-1+\frac{1}{2}-\frac{1}{6}+\frac{1}{24}-...+\frac{(-1)^n}{n!}) \approx \frac{n!}{e}$