Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 5

Basic Probability – Part I

How Good are Your Intutions?

"Seven Odd Questions" $\boxed{\text{Hacking, 2001}}$

“Odd Question” # 1

About as many boys as girls are born in hospitals. Many babies are born every week at City General. In Cornwall, a country town, there is a small hospital where only a few babies are born every week. A *normal* week is one where between 45% and 55% of babies are females. An *unusual* week is one where more than 55% are girls, or more than 55% are boys.

Which of the following is true:

(a) Unusual weeks occur equally often at City General and at Cornwall.

(b) Unusual weeks are more common at City General.

(c) Unusual weeks are more common at Cornwall.

“Odd Question” # 2

Pia is thirty-one years old, single, outspoken, and smart. She was a philosophy major. When a student, she was an ardent supporter of Native American rights, and she picketed a department store that had no facilities for nursing mothers.

Rank the following statements in order from most probable to least probable.

(a) Pia is an active feminist.
(b) Pia is a bank teller.
(c) Pia works in a small bookstore.
(d) Pia is a bank teller and an active feminist.
(e) Pia is a bank teller and an active feminist who takes yoga classes.
(f) Pia works in a small bookstore and is an active feminist who takes yoga classes.

“Odd Question” # 3

In Lotto 6/49, a standard government-run lottery, you choose 6 out of 49 numbers (1 through 49). You win the biggest prize–maybe millions of dollars–if these 6 are drawn. (The prize money is divided between all those who choose the lucky numbers. If no one wins, then most of the prize money is put back into next weeks lottery.) Suppose your aunt offers you, *free*, a choice between two ticket in the lottery, with numbers as shown:

I. You win if 1,2,3,4,5, and 6 are drawn.

II. You win if 39, 36, 32, 21, 14, and 3 are drawn.

Do you prefer I, II, or are you indifferent between the two?

(a) Prefer I

(b) Prefer II

(c) Indifferent

“Odd Question” # 4

To throw a total of 7 with a pair of dice, you have to get a 1 and a 6, or a 2 and a 5, or a 3 and a 4. To throw a total of 6 with a pair of dice, you have to get a 1 and a 5, or a 2 and a 4, or a 3 and another 3. With two fair dice, you would expect:

(a) To throw 7 more frequently than 6.
(b) To throw six more frequently than 7.
(c) To throw 6 and 7 equally often.

“Odd Question” # 5

You have been called to jury duty in a town where there are two taxi companies, Green Cab Ltd. and Blue Taxi Inc. Blue taxi uses cars painted blue; Green Cabs uses green cars. Green cabs dominates the market, with 85% of the taxis on the road. On a misty winter night a taxi sideswiped another car and drove off. A witness says it was a blue cab. The witness is tested under conditions like those on the night of the accident, and 80% of the time she correctly reports the color of the cab that is seen. That is, regardless of whether she is shown a blue or a green cab in misty evening light, she gets the color right 80% of the time.

You conclude, on the basis of this information:

(a) The probability that the sideswiper was blue is 0.8.

(b) It is more likely that the sideswiper was blue, but the probability is less than 0.8.

(c) It is just as probable that the sideswiper was green as that it was blue.

(d) It is more likely than not that the sideswiper was green.

“Odd Question” # 6

You are a physician. You think it is quite likely that one of your patients has strep throat, but you aren’t sure. You take some swabs from the throat and send them to a lab for testing. The test is (like nearly all lab tests) not perfect.

- If the patient has strep throat, then 70% if the time the lab says yes. But 30% of the time it says NO.
- If the patient does not have strep throat, then 90% of the time the lab says NO. But 10% of the time it says YES.

You send five successive swabs to the lab, from the same patient. and get back these results in order: YES, NO, YES, NO, YES. You conclude:

(a) These results are worthless.

(b) It is likely that the patient does not have strep throat.

(c) It is slightly more likely than not, that the patient does have strep throat.

(d) It is very much more likely than not, that the patient does have strep throat.

“Odd Question” # 7

‘Imitate” a coin. That is, write down a sequence of 100 H (for heads) and T (for tails) without tossing a coin–but a sequence that you think will fool everyone into thinking it is the reporting of tossing a fair coin.

Which of these is a real sequence of coin flips?

Exhibit A

H H H H T H H T T T H H T T T H T T T H T T H H T
T T T H T T T H T H T T H H H H T T T T T H H H H
H T T H T T H H T T H H H H H T H H T H T H T H T
T H H T H H T T T H T T T T T T T T T H H T T T T

Exhibit B

H H T H T T T H H T H H H T H T T T H T H H T T T
T H H T T T H H H T H T T T H T T H H T H H T H T
T T H H H T H T T H T H H T T H H H T H T T H H H
T T H H H H T H T T H H T T T H H T H H H T T H T

How could we tell which are the real coin flips?

Hacking (2001, p. 31)

> *Hardly anyone making up a sequence of 10 tosses puts in a run of 7 heads in a row. It is true that the chance of getting 7 heads in a row with a fair coin is only 1/64. But in tossing a coin 100 times, you have at least 93 chances to start tossing 7 heads in a row, because each of the first 93 tosses could begin a run of 7.*

How could we tell which are the real coin flips?

Hacking (2001, p. 31), continued...

> *It is more probable than not, in 100 tosses, that you will get 7 heads in a row. It is certainly more probable than not, that you will get at least 6 heads in a row. Yet almost no one writes down a pretend sequence, in which there are even 6 heads in a row.*

What is Randomness?

What is Randomness?

Many possible answers:

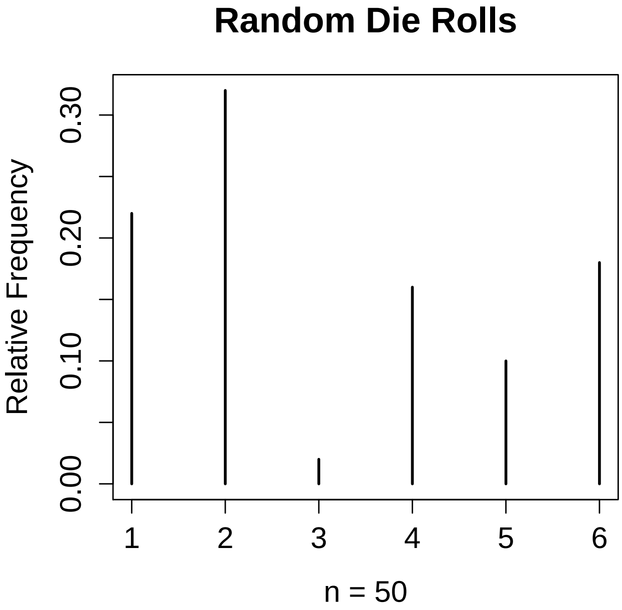
- Lack of regularity
- Complexity (length of shortest encoding)
- No memory/influence from previous trials
- Impossibility of successful gambling “system”

Our Definition of Probability for this Course

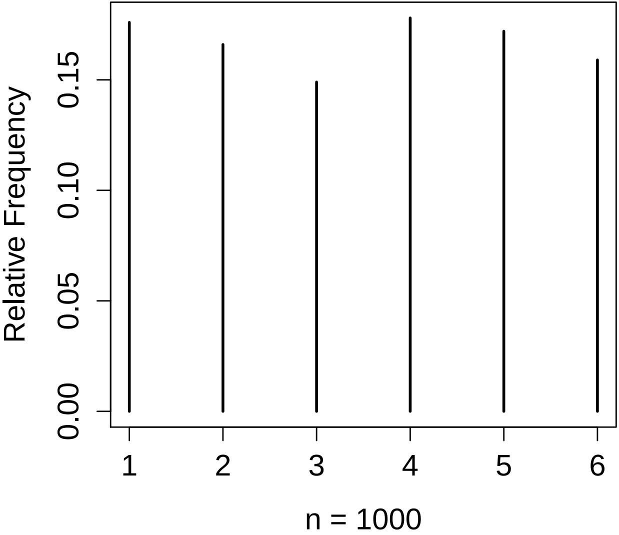
$$\text{Probability} = \text{Long-run Relative Frequency}$$

That is, relative frequencies settle down to probabilities if we carry out an experiment over, and over, and over...

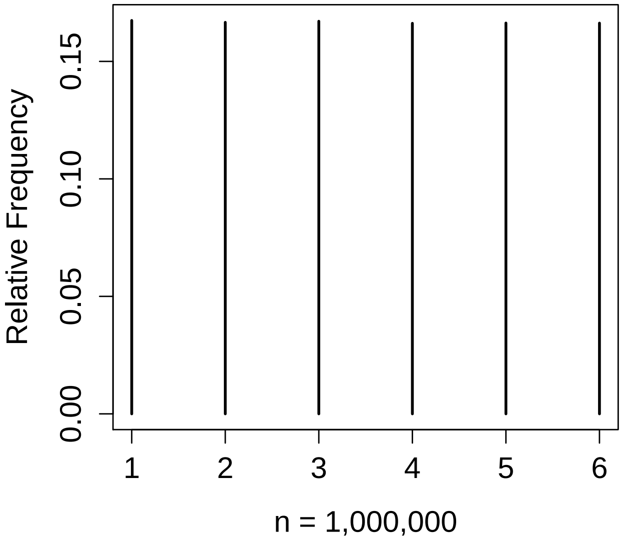

Random Die Rolls
Relative Frequency
0.00
0.10
0.20
0.30
1
2
4
5
6
n = 10


Random Die Rolls
Relative Frequency
0.00
0.10
0.20
0.30
1
2
3
4
5
6
n = 50

Random Die Rolls



Random Die Rolls



What do you think of this argument?

- The probability of flipping heads is $1/2$: if we flip a coin many times, about half of the time it will come up heads.
- The last ten throws in a row the coin has come up heads.
- The coin is bound to come up tails next time – it would be very rare to get 11 heads in a row.

(a) Agree

(b) Disagree

The Gambler's Fallacy

Relative frequencies settle down to probabilities, but this does not mean that the trials are dependent.

Dependent = "Memory" of Prev. Trials

Independent = No "Memory" of Prev. Trials

Another Argument

Lucie visits Albert. As she enters, he rolls four dice and shouts "Hooray!" for he has just rolled four sixes. Lucie: "I bet you've been rolling the dice for a long time to get that result!" Now, Lucie may have many reasons for saying this – perhaps Albert is a lunatic dice-roller. But simply on the evidence that he has just rolled four sixes, is her conclusion reasonable?

(a) Yes

(b) No

The *Inverse* Gambler's Fallacy

This is true:

Albert is more likely to get four sixes if he rolls many times than if he rolls only once.

The *Inverse* Gambler's Fallacy

This is true:

Albert is more likely to get four sixes if he rolls many times than if he rolls only once.

However:

Even if he has been rolling many times, it is *still* extremely unlikely to get four sixes on the particular roll that Lucie observes.

The *Inverse* Gambler's Fallacy

This is true:

Albert is more likely to get four sixes if he rolls many times than if he rolls only once.

However:

Even if he has been rolling many times, it is *still* extremely unlikely to get four sixes on the particular roll that Lucie observes.

The outcome of that roll doesn't tell us anything about whether he has rolled the dice before, just like six heads in a row doesn't mean we're "due" for a tails.

The Mathematics of Probability

Terminology

Terminology

Random Experiment

An experiment whose outcomes are random.

Terminology

Random Experiment

An experiment whose outcomes are random.

Basic Outcomes

Possible outcomes (mutually exclusive) of random experiment.

Terminology

Random Experiment

An experiment whose outcomes are random.

Basic Outcomes

Possible outcomes (mutually exclusive) of random experiment.

Sample Space: S

Set of all basic outcomes of a random experiment.

Terminology

Random Experiment

An experiment whose outcomes are random.

Basic Outcomes

Possible outcomes (mutually exclusive) of random experiment.

Sample Space: S

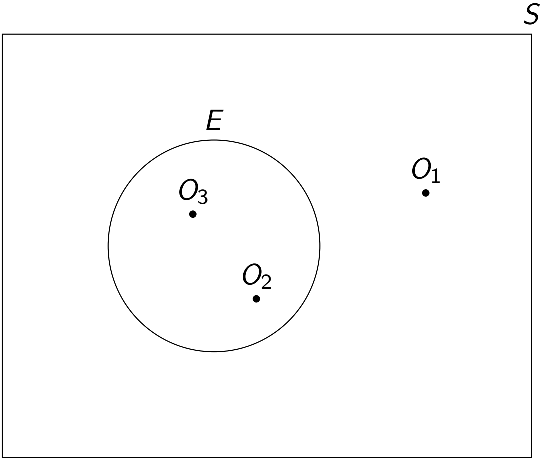
Set of all basic outcomes of a random experiment.

Event: E

A subset of the Sample Space (i.e. a collection of basic outcomes).

In set notation we write $E \subseteq S$.

Visual Representation



Example

Example

Random Experiment

Tossing a pair of dice.

Example

Random Experiment

Tossing a pair of dice.

Basic Outcome

An ordered pair (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$, e.g. $(2, 5)$

Example

Random Experiment

Tossing a pair of dice.

Basic Outcome

An ordered pair (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$, e.g. $(2, 5)$

Sample Space: S

All ordered pairs (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$

Example

Random Experiment

Tossing a pair of dice.

Basic Outcome

An ordered pair (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$, e.g. $(2, 5)$

Sample Space: S

All ordered pairs (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$

Event: $E = \{\text{Sum of two dice is less than 4}\}$

$\{(1, 1), (1, 2), (2, 1)\}$

Probability is Defined on *Sets*, and Events are Sets

Complement of an Event: $A^c = \text{not } A$

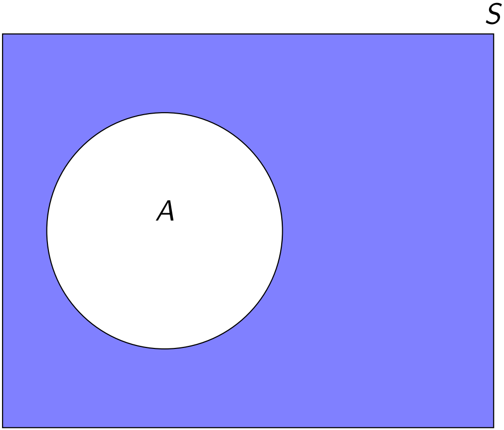


Figure : The complement A^c of an event $A \subseteq S$ is the collection of all basic outcomes from S not contained in A.

Intersection of Events: $A \cap B = A$ and B

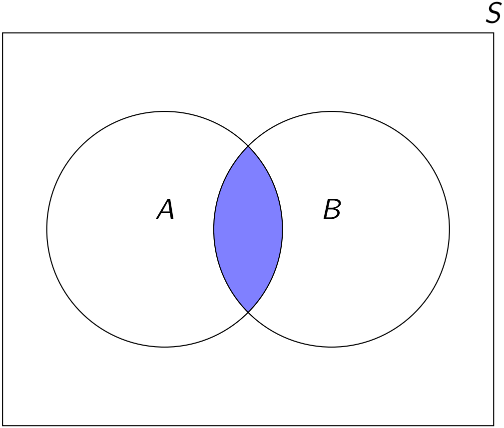


Figure : The intersection $A \cap B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in both A and B

Union of Events: $A \cup B = A$ or B

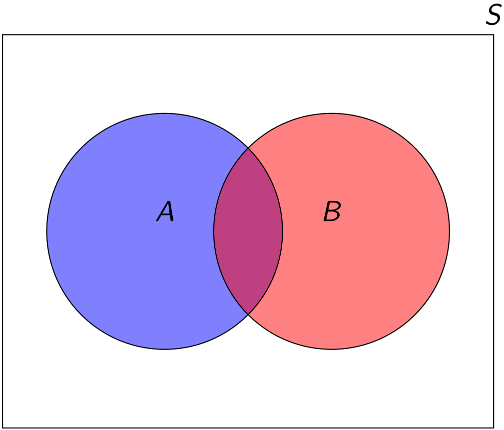


Figure : The union $A \cup B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in A, B or both.

Mutually Exclusive and Collectively Exhaustive

Mutually Exclusive Events

We call a collection of events $E_1, E_2, E_3, \ldots$ *mutually exclusive* if they are *pairwise disjoint*, that is if the intersection of *any two different events* in the collection is empty (formally $E_i \cap E_j = \emptyset$ for any $i \neq j$).

Mutually Exclusive and Collectively Exhaustive

Mutually Exclusive Events

We call a collection of events $E_1, E_2, E_3, \ldots$ *mutually exclusive* if they are *pairwise disjoint*, that is if the intersection of *any two different events* in the collection is empty (formally $E_i \cap E_j = \emptyset$ for any $i \neq j$).

Collectively Exhaustive Events

We call a collection of events $E_1, E_2, E_3, \ldots$ *collectively exhaustive* if they, collectively, contain *all of the basic outcomes in* S. Another way of saying this is that the union $E_1 \cup E_2 \cup E_3 \cup \cdots = S$, the sample space.

Mutually Exclusive but *not Collectively Exhaustive*

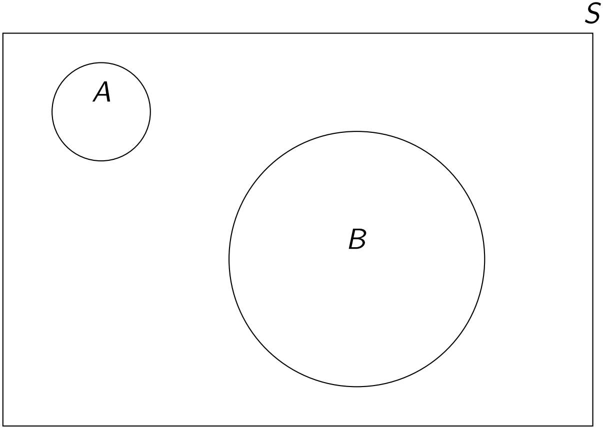


Figure : Although $A \cap B = \emptyset$, $A \cup B \neq S$

Collectively Exhaustive but *not Mutually Exclusive*

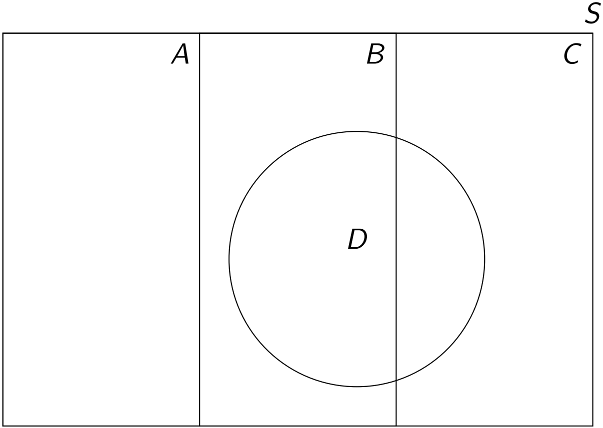


Figure : Although $A \cup B \cup C \cup D = S$, $B \cap D \neq \emptyset$ and $C \cap D \neq \emptyset$.

Collectively Exhaustive *and* Mutually Exclusive

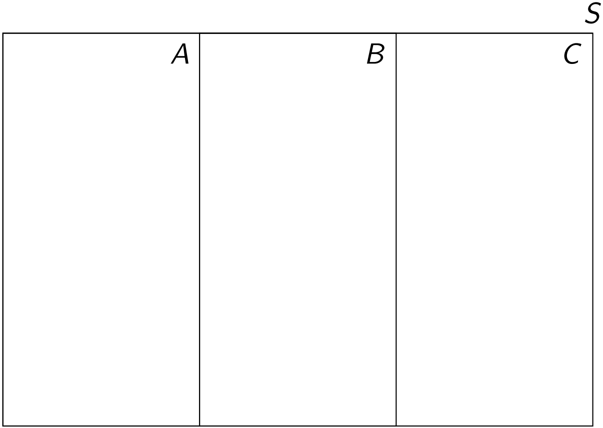


Figure : $A \cup B \cup C = S$, and $A \cap B = B \cap C = A \cap C = \emptyset$.

Axioms of Probability

Let S be the sample space. With each event $A \subseteq S$ we associate a real number $P(A)$ called the probability of A, satisfying the following conditions:

Axioms of Probability

Let S be the sample space. With each event $A \subseteq S$ we associate a real number $P(A)$ called the probability of A, satisfying the following conditions:

Axiom 1 $0 \leq P(A) \leq 1$

Axioms of Probability

Let S be the sample space. With each event $A \subseteq S$ we associate a real number $P(A)$ called the probability of A, satisfying the following conditions:

Axiom 1 $0 \leq P(A) \leq 1$

Axiom 2 $P(S) = 1$

Axioms of Probability

Let S be the sample space. With each event $A \subseteq S$ we associate a real number $P(A)$ called the probability of A, satisfying the following conditions:

Axiom 1 $0 \leq P(A) \leq 1$

Axiom 2 $P(S) = 1$

Axiom 3 If $A_1, A_2, A_3, \dots$ are mutually exclusive events, then
$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

“Classical” Probability

When all of the basic outcomes are equally likely, calculating the probability of an event is simply a matter of counting – count up all the basic outcomes that make up the event, and divide by the total number of basic outcomes.

Review of Rules for Counting

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, ..., n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Review of Rules for Counting

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, $\ldots$, n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Corollary – Number of Possible Orderings

$k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1 = k!$

Review of Rules for Counting

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, $\ldots$, n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Corollary – Number of Possible Orderings

$k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1 = k!$

Permutations – Order n people in k slots

$P_k^n = \frac{n!}{(n-k)!}$ (Order Matters)

Review of Rules for Counting

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, $\ldots$, n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Corollary – Number of Possible Orderings

$k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1 = k!$

Permutations – Order n people in k slots

$P_k^n = \frac{n!}{(n-k)!}$ (Order Matters)

Combinations – Choose committee of k from group of n

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$, where $0! = 1$ (Order Doesn't Matter)

A Fairly Ridiculous Example

Roger Federer and Novak Djokovic have agreed to play in a tennis tournament against six Penn professors. Each player in the tournament is randomly allocated to one of the eight rungs in the ladder (next slide). Federer always beats Djokovic and, naturally, either of the two pros always beats any of the professors. What is the probability that Djokovic gets second place in the tournament?

Rung 1

Quarterfinalist 1

Rung 2

Semifinalist 1

Rung 3

Quarterfinalist 2

Rung 4

Winner

Rung 5

Quarterfinalist 3

Rung 6

Semifinalist 2

Rung 7

Quarterfinalist 4

Rung 8

- There are 8! basic outcomes, the number of ways to arrange players on the tournament ladder where order matters.

- There are $8!$ basic outcomes, the number of ways to arrange players on the tournament ladder where order matters.
- There are 8 rungs we could put Federer on.

- There are 8! basic outcomes, the number of ways to arrange players on the tournament ladder where order matters.
- There are 8 rungs we could put Federer on.
- For each rung Federer could be on, there are only 4 rungs that Djokovic can occupy to avoid meeting Federer before the final.

- There are 8! basic outcomes, the number of ways to arrange players on the tournament ladder where order matters.
- There are 8 rungs we could put Federer on.
- For each rung Federer could be on, there are only 4 rungs that Djokovic can occupy to avoid meeting Federer before the final.
- Now six rungs remain: there are 6! ways to put the professors on these rungs, since order matters.

- There are 8! basic outcomes, the number of ways to arrange players on the tournament ladder where order matters.
- There are 8 rungs we could put Federer on.
- For each rung Federer could be on, there are only 4 rungs that Djokovic can occupy to avoid meeting Federer before the final.
- Now six rungs remain: there are 6! ways to put the professors on these rungs, since order matters.

$$\frac{8 \times 4 \times 6!}{8!} = \frac{8 \times 4}{7 \times 8} = 4/7 \approx 0.57$$