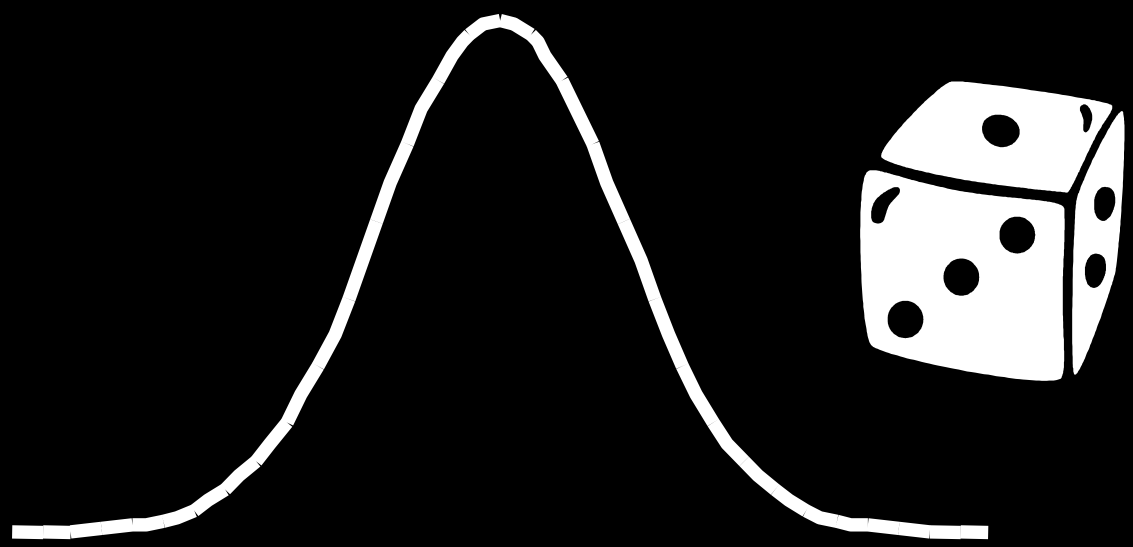
15-251

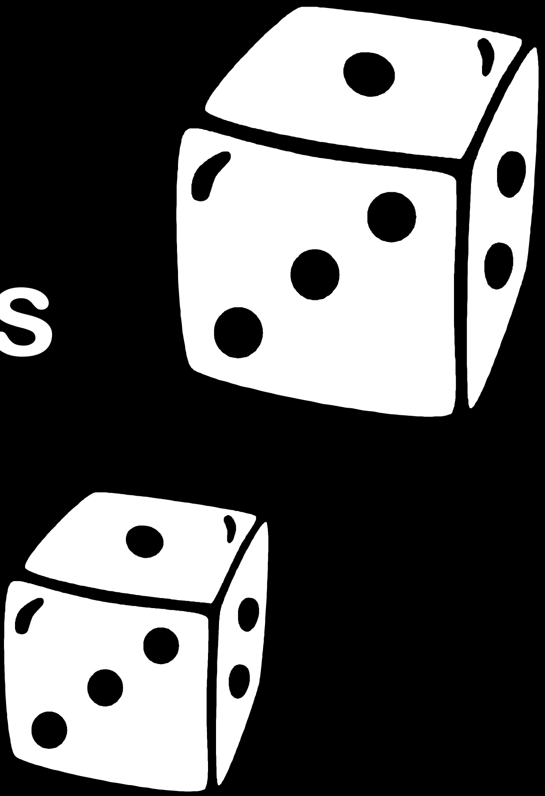
Great Theoretical Ideas in Computer Science

Probability Theory: Counting in Terms of Proportions

Lecture 10, September 25, 2008



Some Puzzles



Teams A and B are equally good

Teams A and B are equally good

In any one game, each is equally likely to win

Teams A and B are equally good

In any one game, each is equally likely to win

What is the most likely length of a "best of 7" series?

Teams A and B are equally good

In any one game, each is equally likely to win

What is the most likely length of a
“best of 7” series?

Flip coins until either 4 heads or 4 tails
Is this more likely to take 6 or 7 flips?

6 and 7 Are Equally Likely

6 and 7 Are Equally Likely

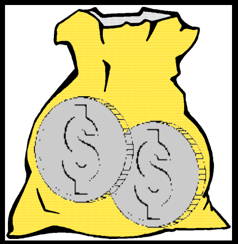
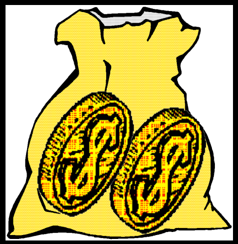
To reach either one, after 5 games, it must be 3 to 2

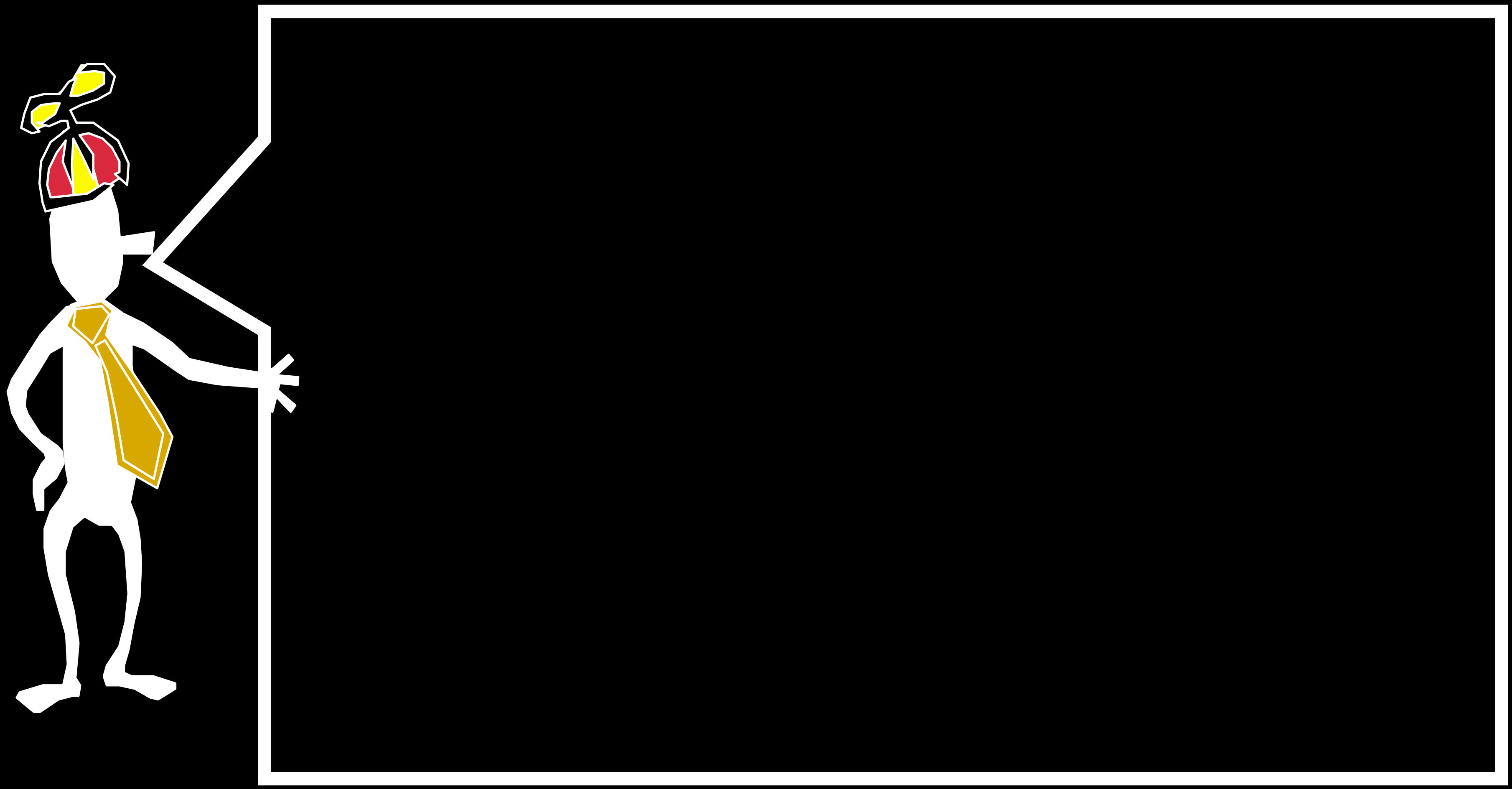
6 and 7 Are Equally Likely

To reach either one, after 5 games, it must be 3 to 2

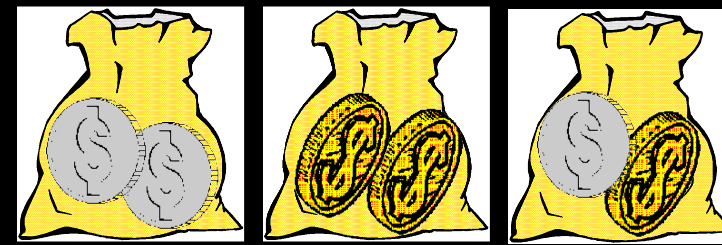
½ chance it ends 4 to 2; ½ chance it doesn't

Silver and Gold

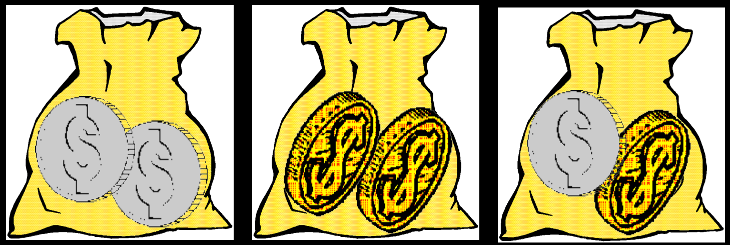


Silver and Gold



One bag has two silver coins, another has two gold coins, and the third has one of each

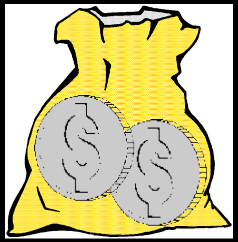
Silver and Gold

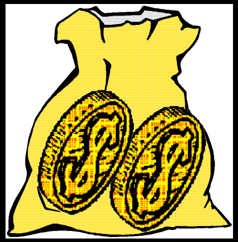


One bag has two silver coins, another has two gold coins, and the third has one of each

One bag is selected at random. One coin from it is selected at random. It turns out to be gold

Silver and Gold

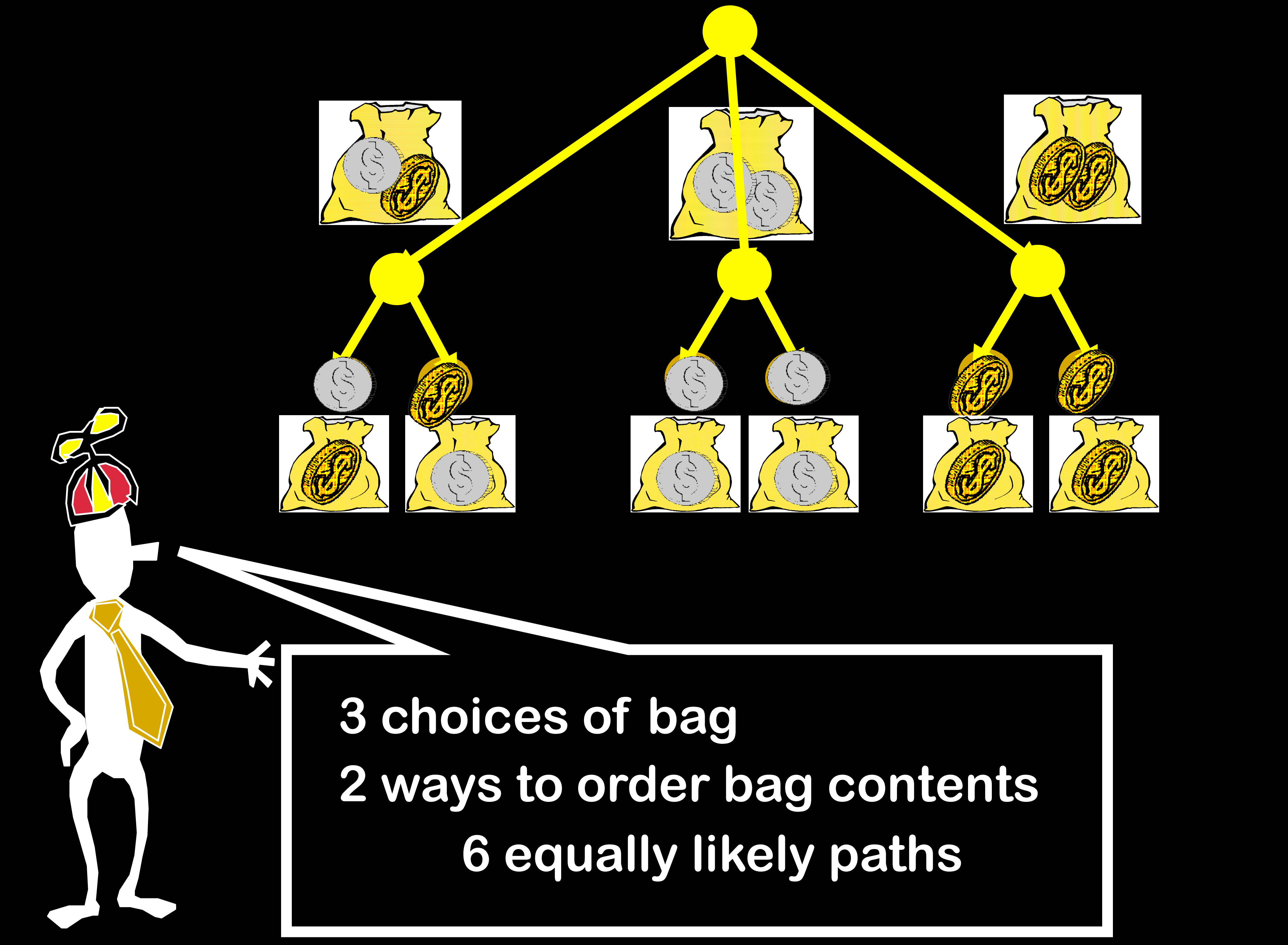


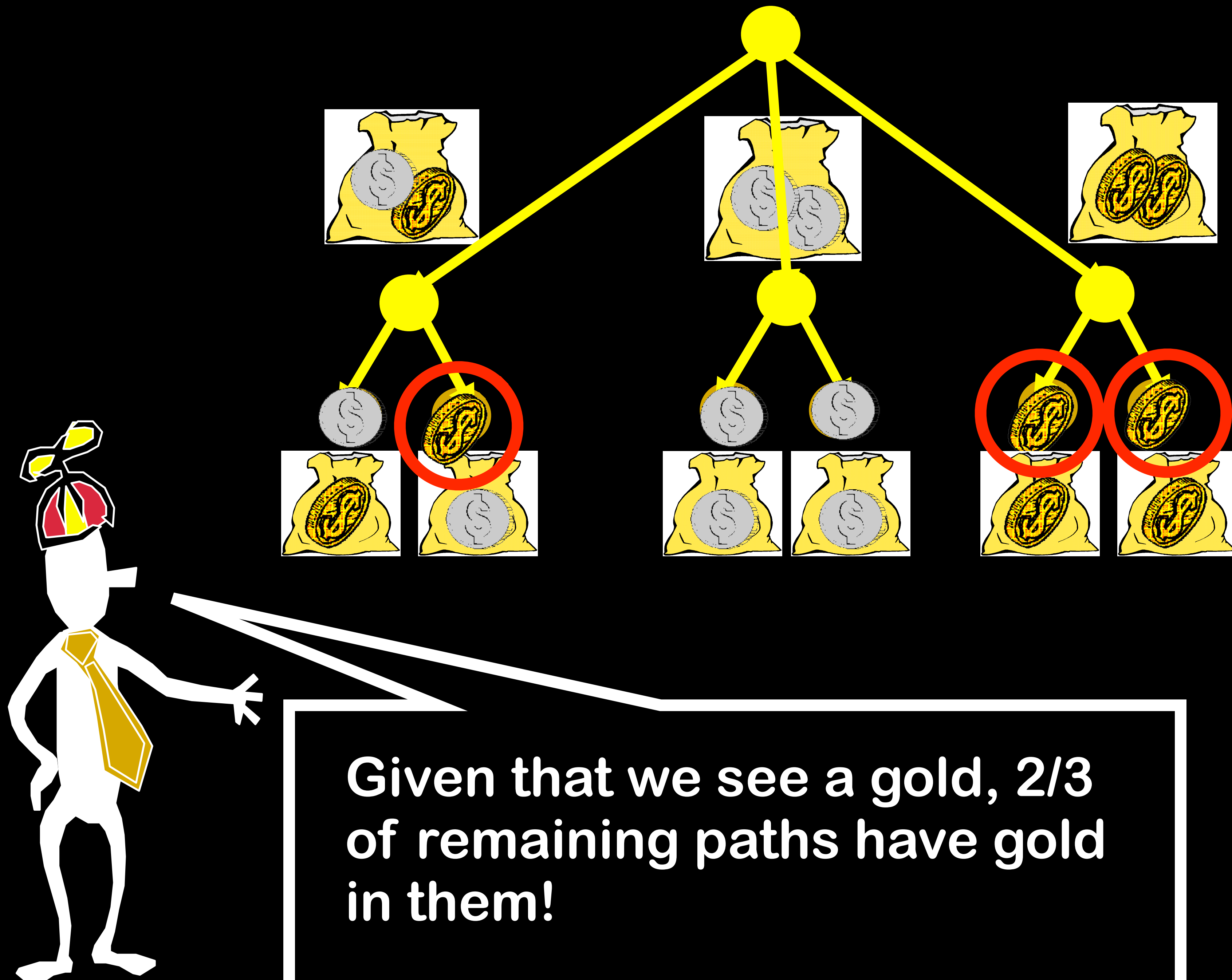


One bag has two silver coins, another has two gold coins, and the third has one of each

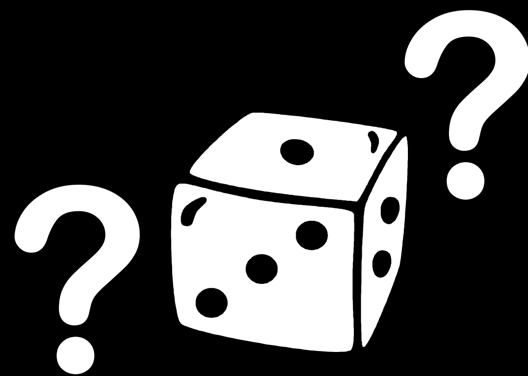
One bag is selected at random. One coin from it is selected at random. It turns out to be gold

What is the probability that the other coin is gold?


3 choices of bag
2 ways to order bag contents
6 equally likely paths

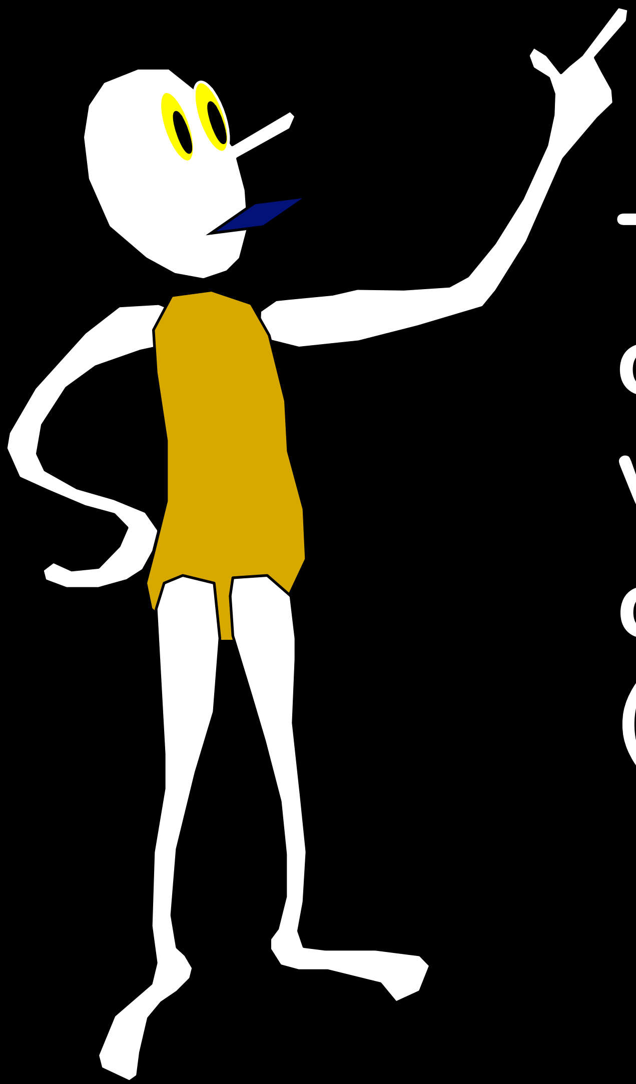


Given that we see a gold, 2/3 of remaining paths have gold in them!



So, sometimes, probabilities can be counter-intuitive

Language of Probability



The formal language of probability is a very important tool in computer science (and science)

Finite Probability Distribution

Finite Probability Distribution

A (finite) probability distribution **p** is a finite set **S** of elements, together with a non-negative real weight, or probability **p(x)** for each element **x in S**

Finite Probability Distribution

A (finite) probability distribution **p** is a finite set **S** of elements, together with a non-negative real weight, or probability **p(x)** for each element **x in S**

The weights must satisfy:

Finite Probability Distribution

A (finite) probability distribution **p** is a finite set **S** of elements, together with a non-negative real weight, or probability **p(x)** for each element **x in S**

The weights must satisfy:

$$\sum_{x \in S} p(x) = 1$$

Finite Probability Distribution

A (finite) probability distribution p is a finite set S of elements, together with a non-negative real weight, or probability p(x) for each element x in S

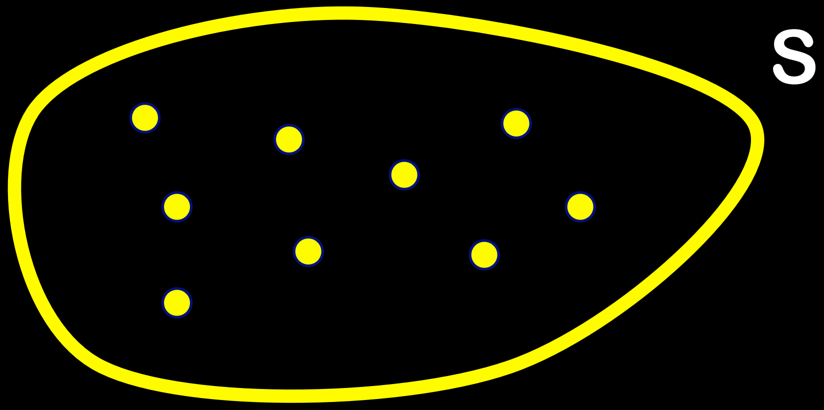
The weights must satisfy:

$$\sum_{x \in S} p(x) = 1$$

S is often called the sample space and elements x in S are called samples

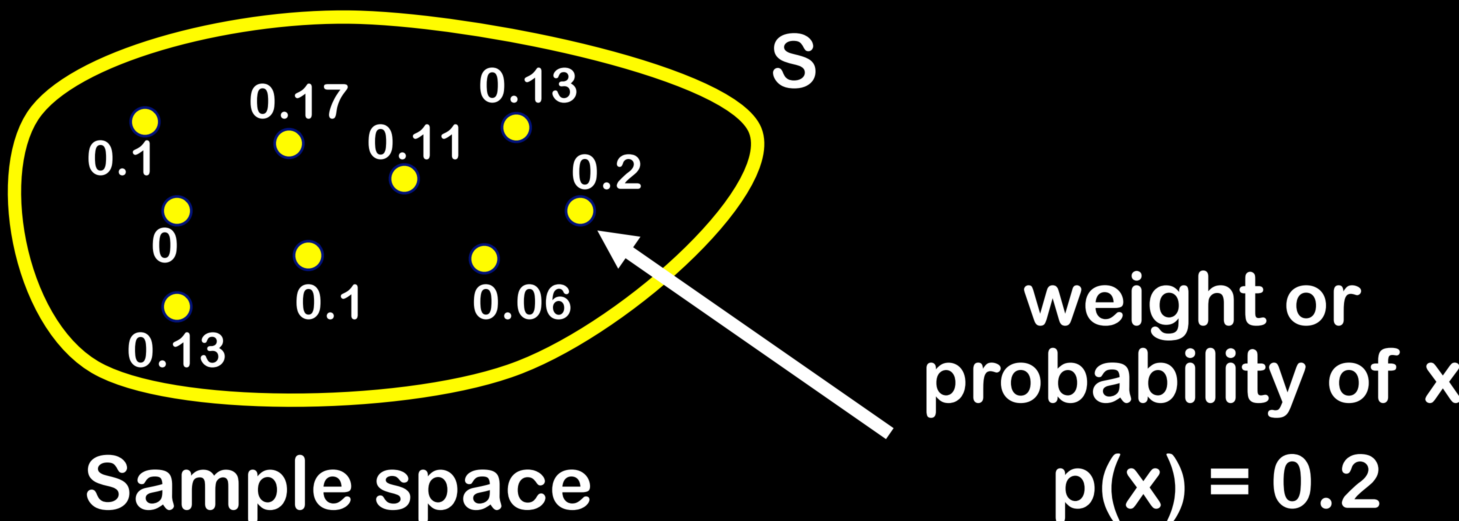
Sample Space

Sample Space



Sample space

Sample Space



Sample space

Events

Events

Any set $E \subseteq S$ is called an event

Events

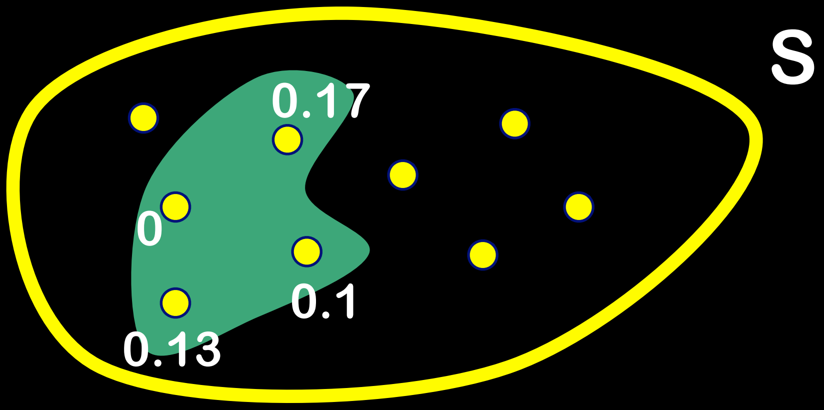
Any set $E \subseteq S$ is called an event

$$\Pr_D[E] = \sum_{x \in E} p(x)$$

Events

Any set $E \subseteq S$ is called an event

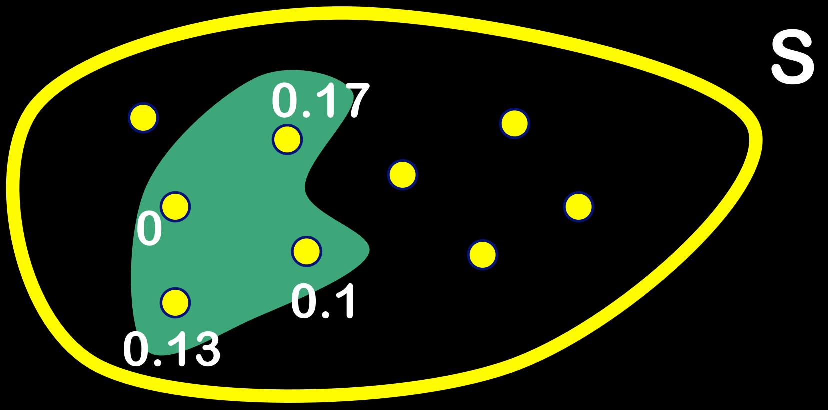
$$\Pr_D[E] = \sum_{x \in E} p(x)$$



Events

Any set $E \subseteq S$ is called an event

$$Pr_D[E] = \sum_{x \in E} p(x)$$



$$Pr_D[E] = 0.4$$

Uniform Distribution

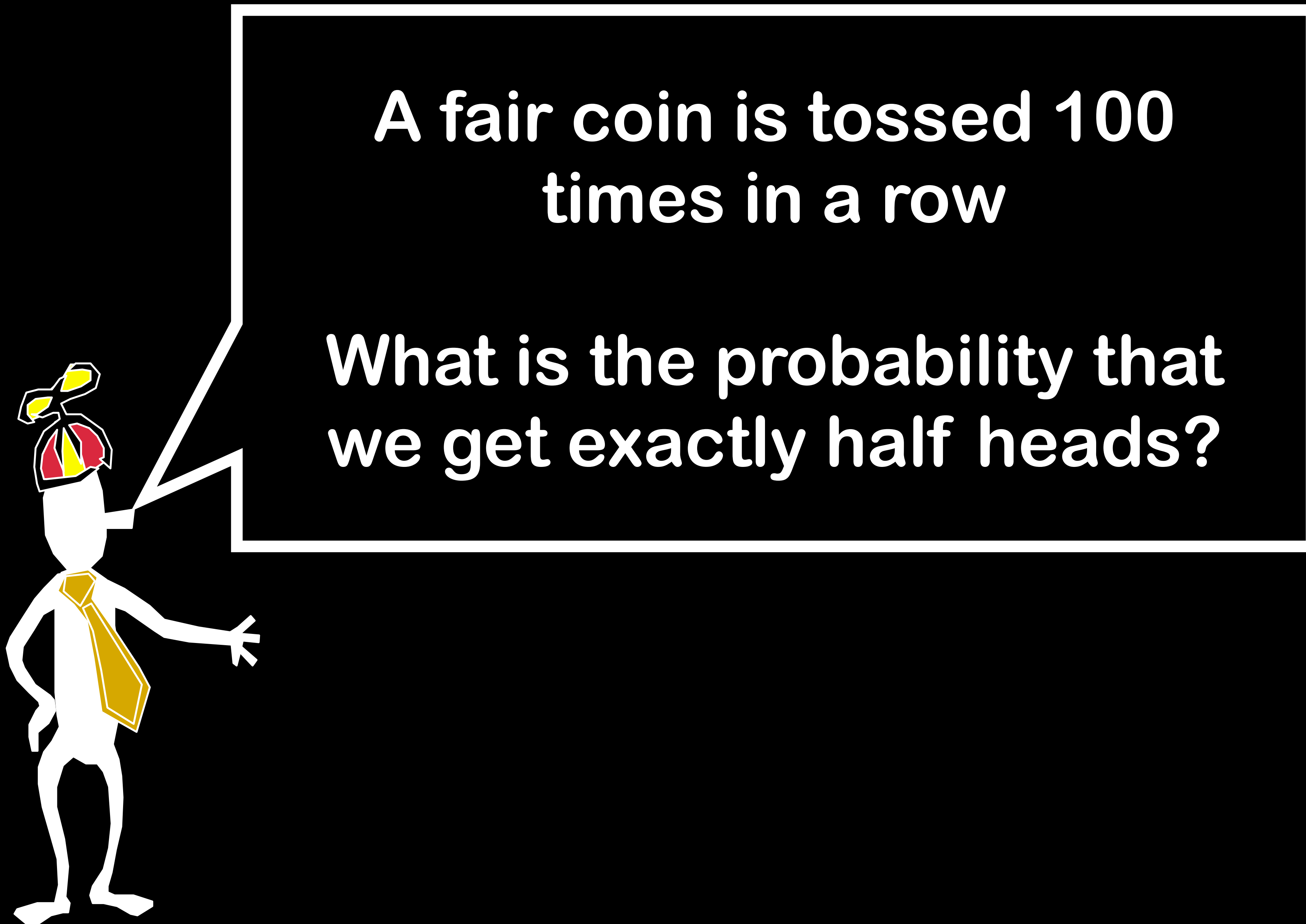
Uniform Distribution

If each element has equal probability, the distribution is said to be uniform

Uniform Distribution

If each element has equal probability, the distribution is said to be uniform

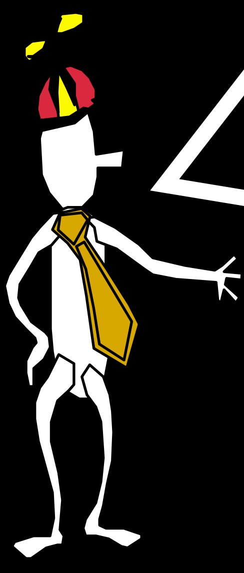
$$\Pr_D[E] = \sum_{x \in E} p(x) = \frac{|E|}{|S|}$$


A fair coin is tossed 100 times in a row
What is the probability that we get exactly half heads?

Using the Language

Using the Language

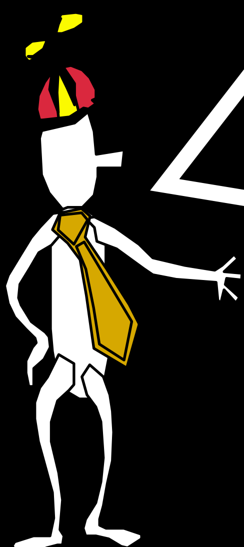
The sample space **S** is the set of all outcomes $\{H,T\}^{100}$



Using the Language

The sample space **S** is the set of all outcomes $\{H,T\}^{100}$

Each sequence in **S** is equally likely, and hence has probability $1/|S| = 1/2^{100}$

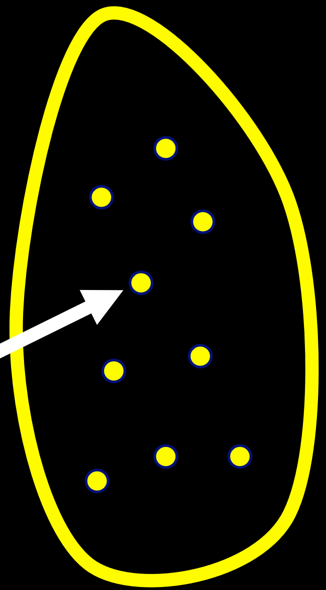


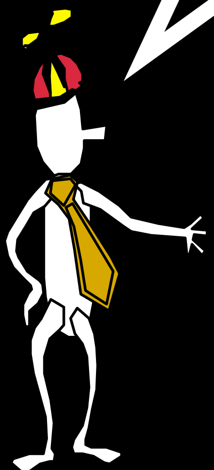
Visually

S = all sequences of 100 tosses

x = HHTTT……TH

p(x) = 1/|S|

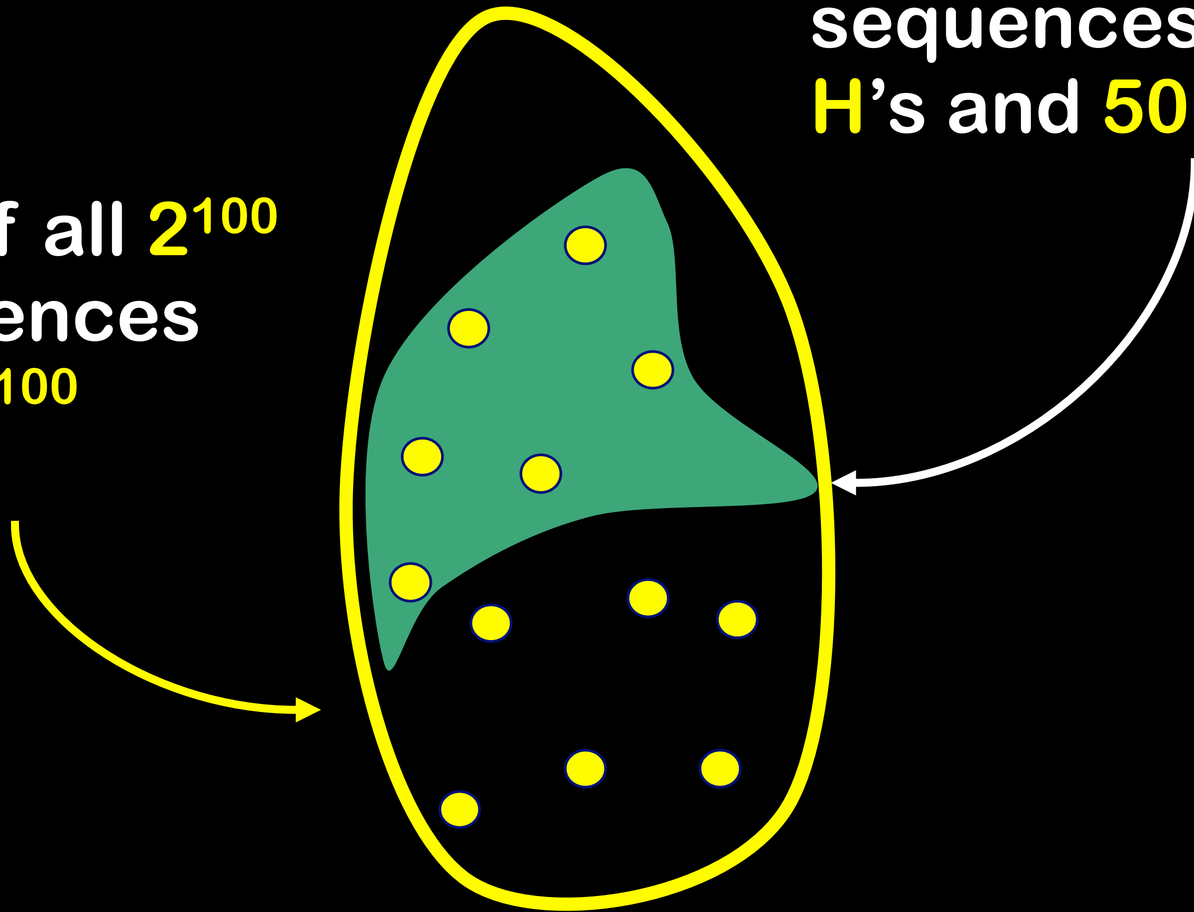




Set of all 2^{100} sequences $\{H,T\}^{100}$

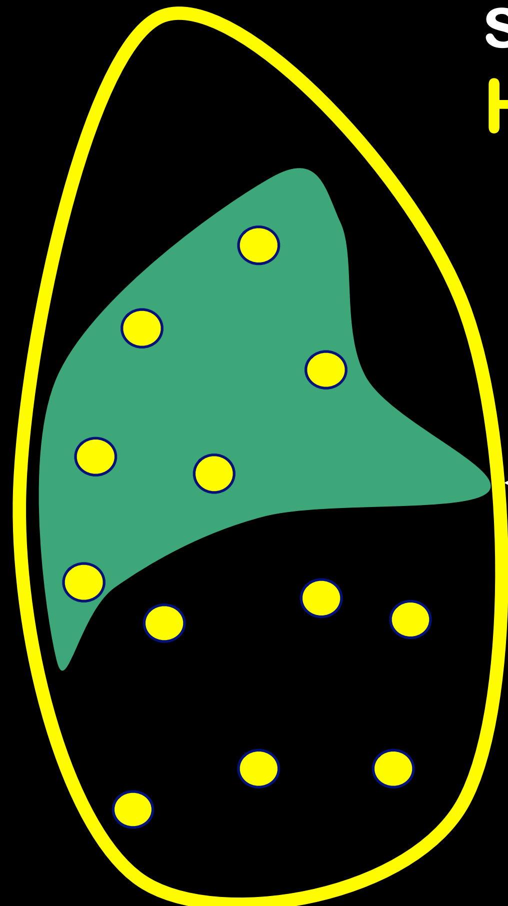
Event E = Set of sequences with 50 H’s and 50 T’s

Set of all 2^{100} sequences $\{H,T\}^{100}$



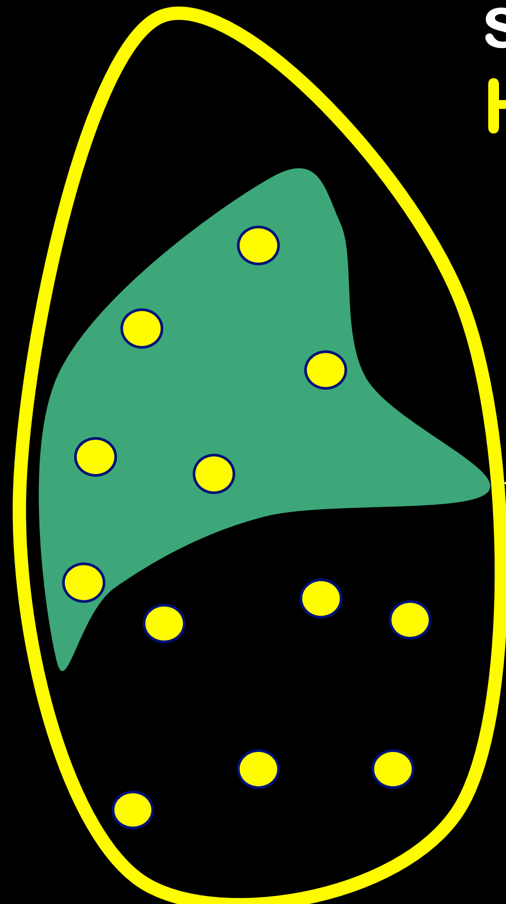
Event E = Set of sequences with 50 H’s and 50 T’s

Set of all 2^{100} sequences $\{H,T\}^{100}$



Probability of event E = proportion of E in S

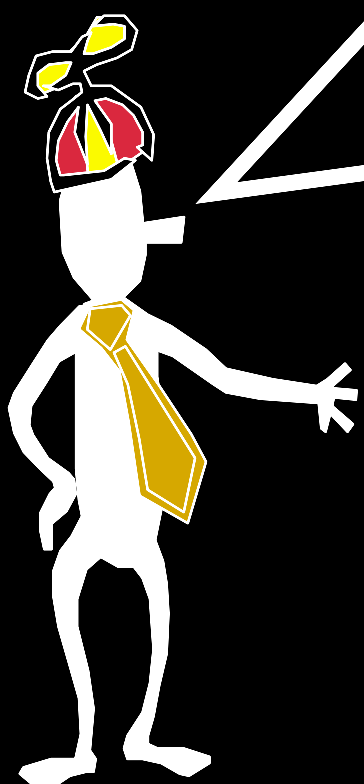
Set of all 2^{100} sequences $\{H,T\}^{100}$



Event E = Set of sequences with 50 H's and 50 T's

Probability of event E = proportion of E in S

$$\binom{100}{50} / 2^{100}$$



Suppose we roll a white die and a black die

What is the probability that sum is 7 or 11?

Same Methodology!

Same Methodology!

$$S = \{ \begin{array}{llllll} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \}$$

Same Methodology!

$$
\begin{array}{llllll}
S = \{\ (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & \boxed{(1,6)}, \\
(2,1), & (2,2), & (2,3), & (2,4), & \boxed{(2,5)}, & (2,6), \\
(3,1), & (3,2), & (3,3), & \boxed{(3,4)}, & (3,5), & (3,6), \\
(4,1), & (4,2), & \boxed{(4,3)}, & (4,4), & (4,5), & (4,6), \\
(5,1), & \boxed{(5,2)}, & (5,3), & (5,4), & (5,5), & (5,6), \\
\boxed{(6,1)}, & (6,2), & (6,3), & (6,4), & (6,5), & (6,6)\ \}
\end{array}
$$

Same Methodology!

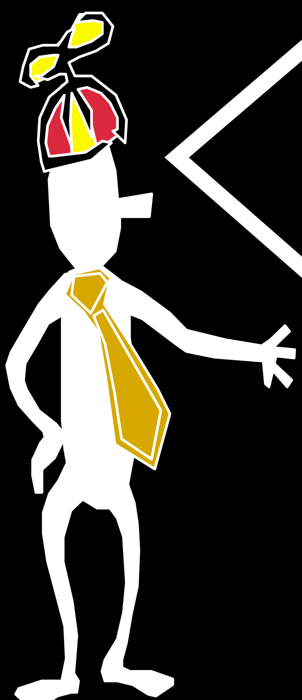
$$
\begin{array}{llllllll}
S = \{ & (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & \boxed{(1,6)}, \\
& (2,1), & (2,2), & (2,3), & (2,4), & \boxed{(2,5)}, & (2,6), \\
& (3,1), & (3,2), & (3,3), & \boxed{(3,4)}, & (3,5), & (3,6), \\
& (4,1), & (4,2), & \boxed{(4,3)}, & (4,4), & (4,5), & (4,6), \\
& (5,1), & \boxed{(5,2)}, & (5,3), & (5,4), & (5,5), & \boxed{(5,6)}, \\
& \boxed{(6,1)}, & (6,2), & (6,3), & (6,4), & \boxed{(6,5)}, & (6,6) \}
\end{array}
$$

Same Methodology!

$$
\begin{aligned}
S = \{ &(1,1), \ (1,2), \ (1,3), \ (1,4), \ (1,5), \ \boxed{(1,6)}, \\
&(2,1), \ (2,2), \ (2,3), \ (2,4), \ \boxed{(2,5)}, \ (2,6), \\
&(3,1), \ (3,2), \ (3,3), \ \boxed{(3,4)}, \ (3,5), \ (3,6), \\
&(4,1), \ (4,2), \ \boxed{(4,3)}, \ (4,4), \ (4,5), \ (4,6), \\
&(5,1), \ \boxed{(5,2)}, \ (5,3), \ (5,4), \ (5,5), \ \boxed{(5,6)}, \\
&\boxed{(6,1)}, \ (6,2), \ (6,3), \ (6,4), \ \boxed{(6,5)}, \ (6,6) \}
\end{aligned}
$$

Pr[E] = |E|/|S| = proportion of E in S = 8/36

23 people are in a room



23 people are in a room

Suppose that all possible birthdays are equally likely

23 people are in a room

Suppose that all possible birthdays are equally likely

What is the probability that two people will have the same birthday?

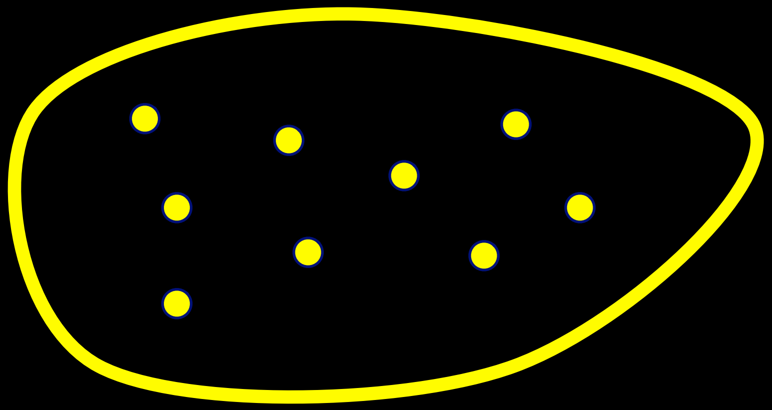
And The Same Methods Again!

And The Same Methods Again!

Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$

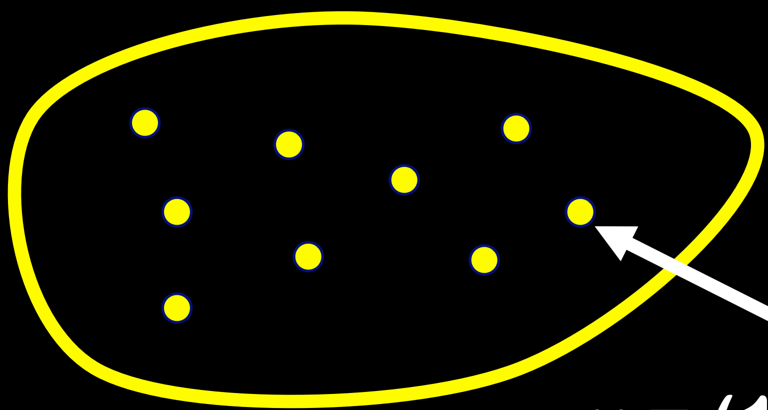
And The Same Methods Again!

Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



And The Same Methods Again!

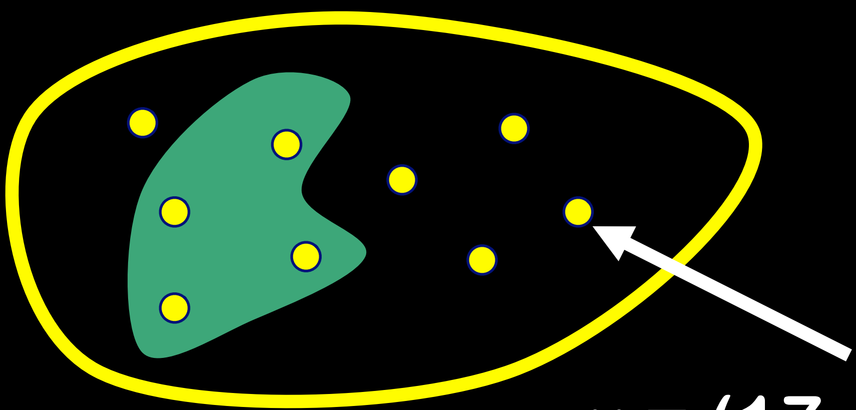
Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



$$x = \underbrace{(17,42,363,1,\ldots, 224,177)}_{\text{23 numbers}}$$

And The Same Methods Again!

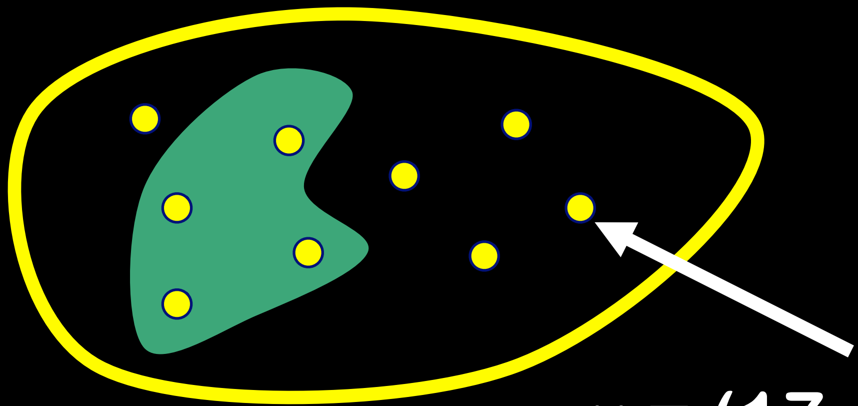
Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



$$x = \underbrace{(17,42,363,1,\ldots, 224,177)}_{\text{23 numbers}}$$

And The Same Methods Again!

Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



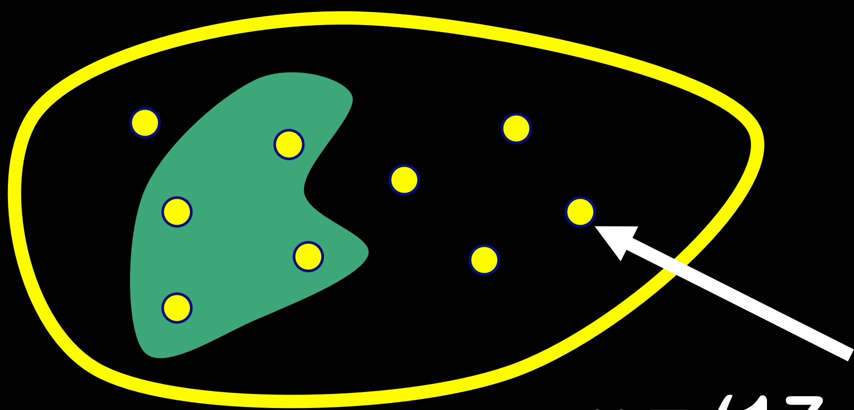
$x = (17,42,363,1,\ldots, 224,177)$

23 numbers

Event $E = \{ x \in W \mid \text{two numbers in } x \text{ are same} \}$

And The Same Methods Again!

Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



$x = (17,42,363,1,\ldots, 224,177)$

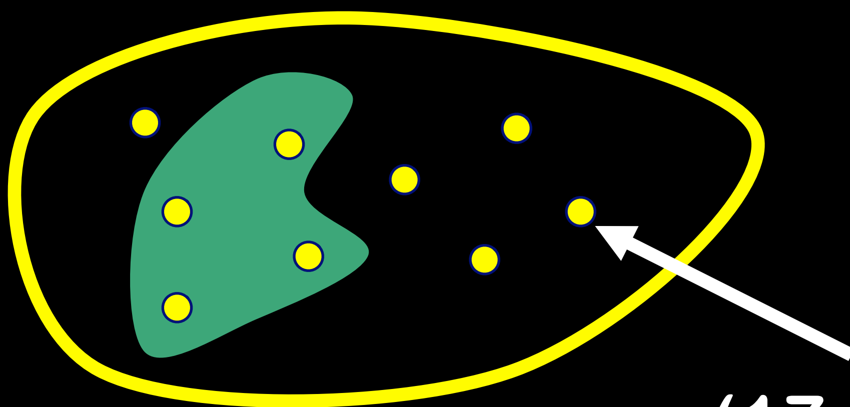
23 numbers

Event $E = \{\, x \in W \mid \text{two numbers in } x \text{ are same} \,\}$

What is $|E|$?

And The Same Methods Again!

Sample space $W = \{1, 2, 3, \ldots, 366\}^{23}$



$x = (17,42,363,1,\ldots, 224,177)$

$\longleftrightarrow$

23 numbers

Event $E = \{ x \in W \mid$ two numbers in x are same $\}$

What is $|E|$? **Count $|\overline{E}|$ instead!**

$\overline{E}$ = all sequences in S that have no repeated numbers

$\overline{E}$ = all sequences in S that have no repeated numbers

$|\overline{E}| = (366)(365)\ldots(344)$

$\overline{E}$ = all sequences in S that have no repeated numbers

$|\overline{E}| = (366)(365)\ldots(344)$

$|W| = 366^{23}$

$\overline{E}$ = all sequences in S that have no repeated numbers

$$|\overline{E}| = (366)(365)\ldots(344)$$

$$|W| = 366^{23}$$

$$\frac{|\overline{E}|}{|W|} = 0.494\ldots$$

$\overline{E}$ = all sequences in S that have no repeated numbers

$$|\overline{E}| = (366)(365)\ldots(344)$$

$$|W| = 366^{23}$$

$$\frac{|\overline{E}|}{|W|} = 0.494\ldots$$

$$\frac{|E|}{|W|} = 0.506\ldots$$

Sons of Adam

Sons of Adam

Adam was X inches tall

Sons of Adam

Adam was X inches tall

He had two sons:

Sons of Adam

Adam was X inches tall

He had two sons:

One was X+1 inches tall

Sons of Adam

Adam was X inches tall

He had two sons:

One was X+1 inches tall

One was X-1 inches tall

Sons of Adam

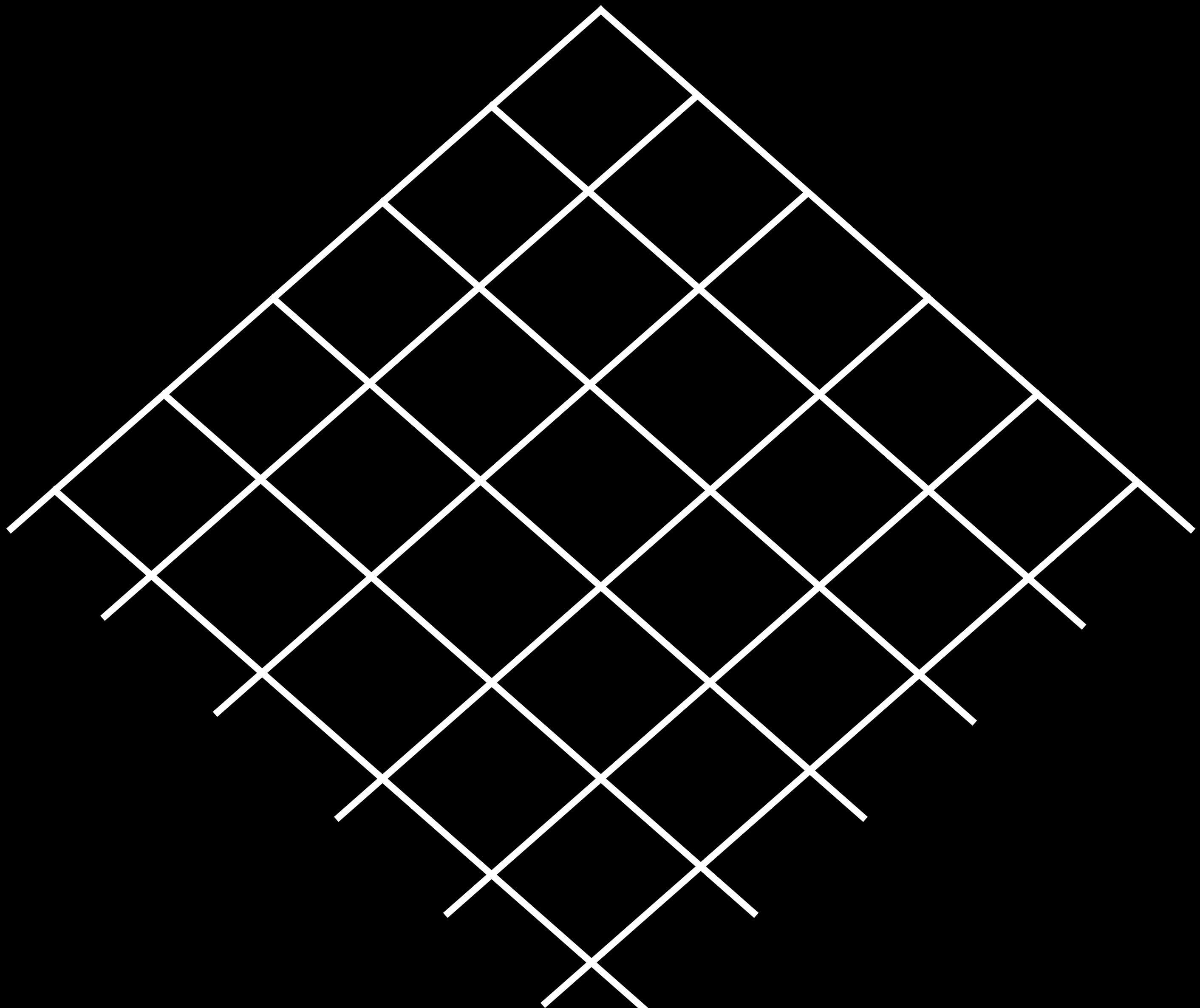
Adam was X inches tall

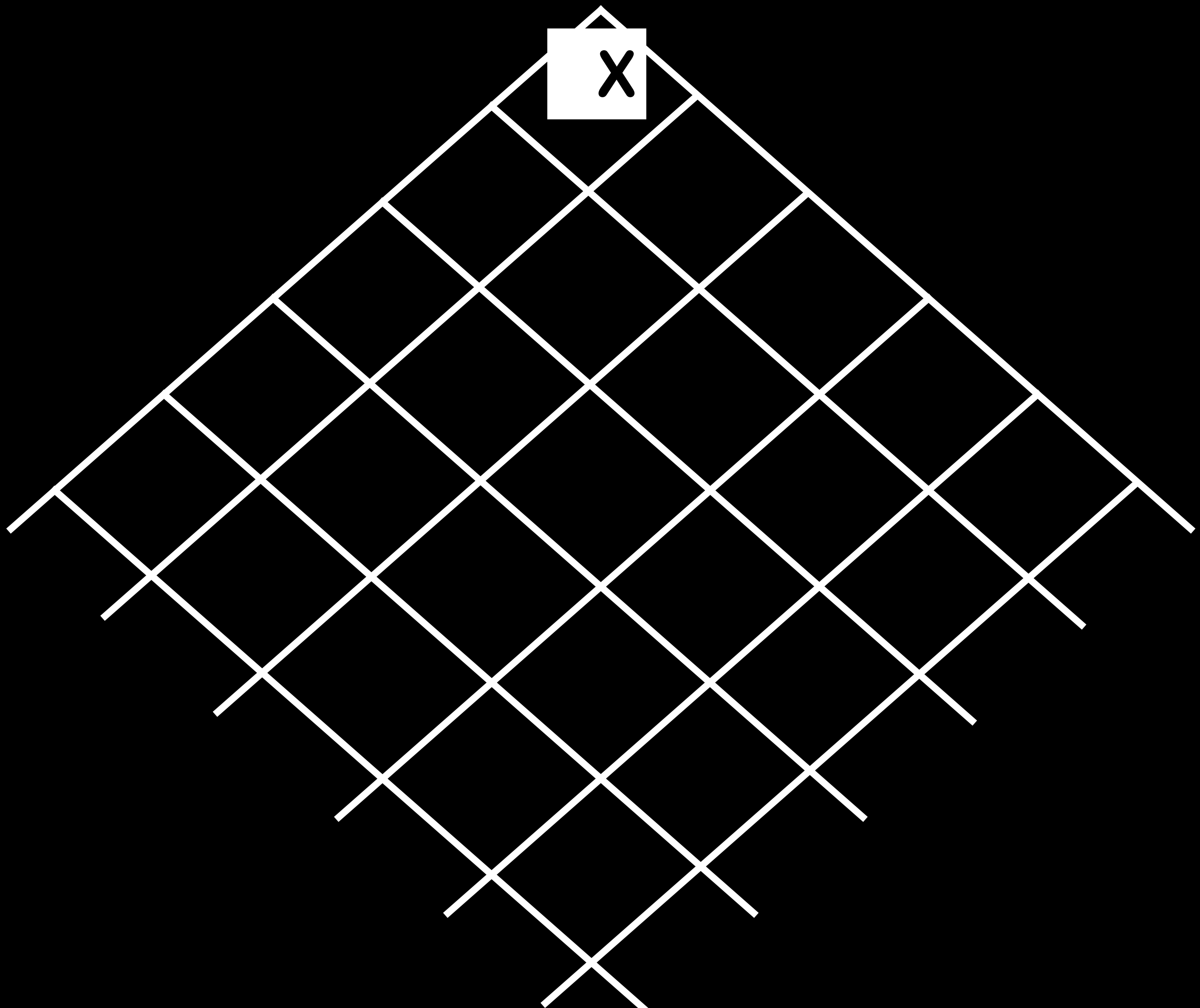
He had two sons:

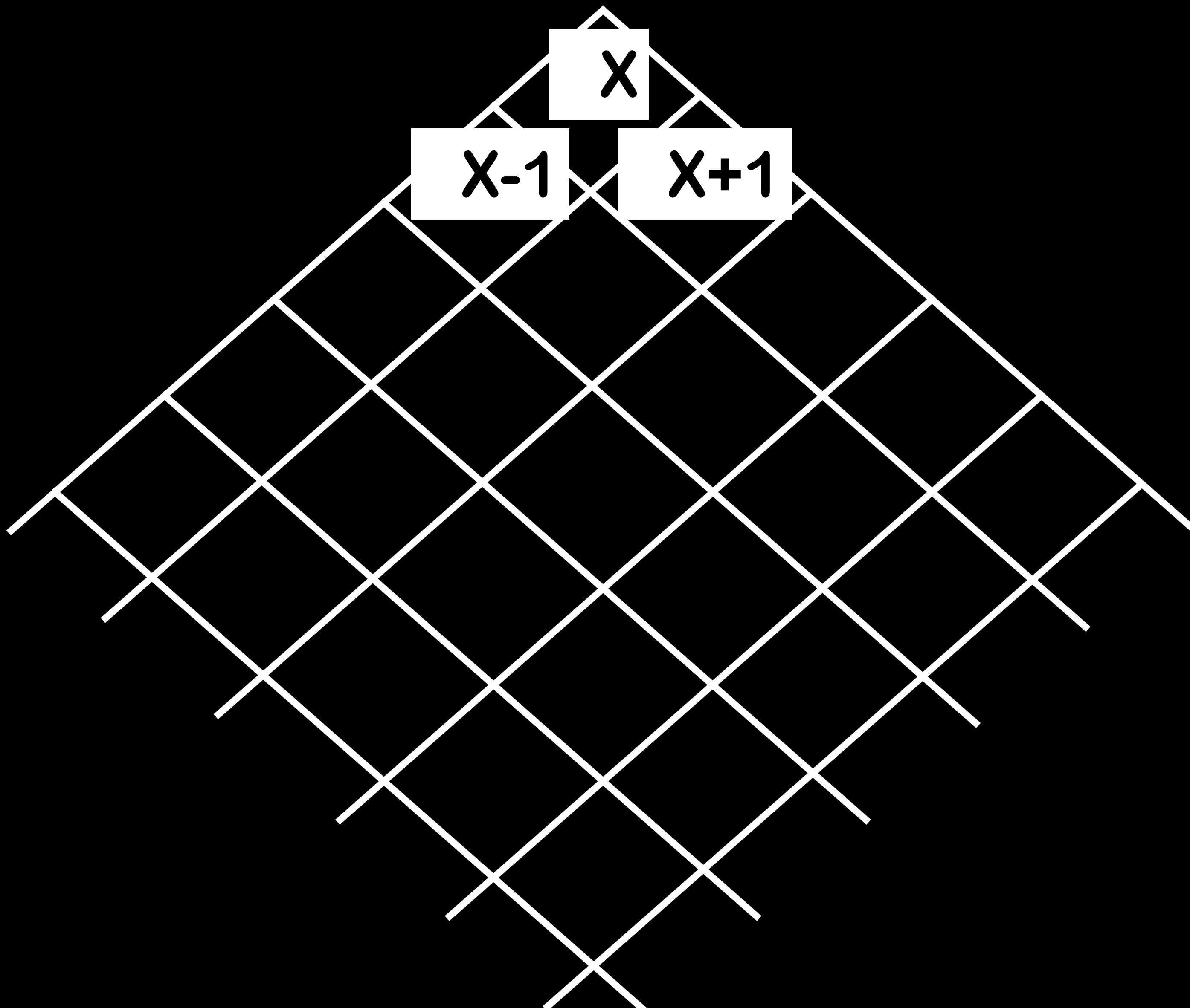
One was X+1 inches tall

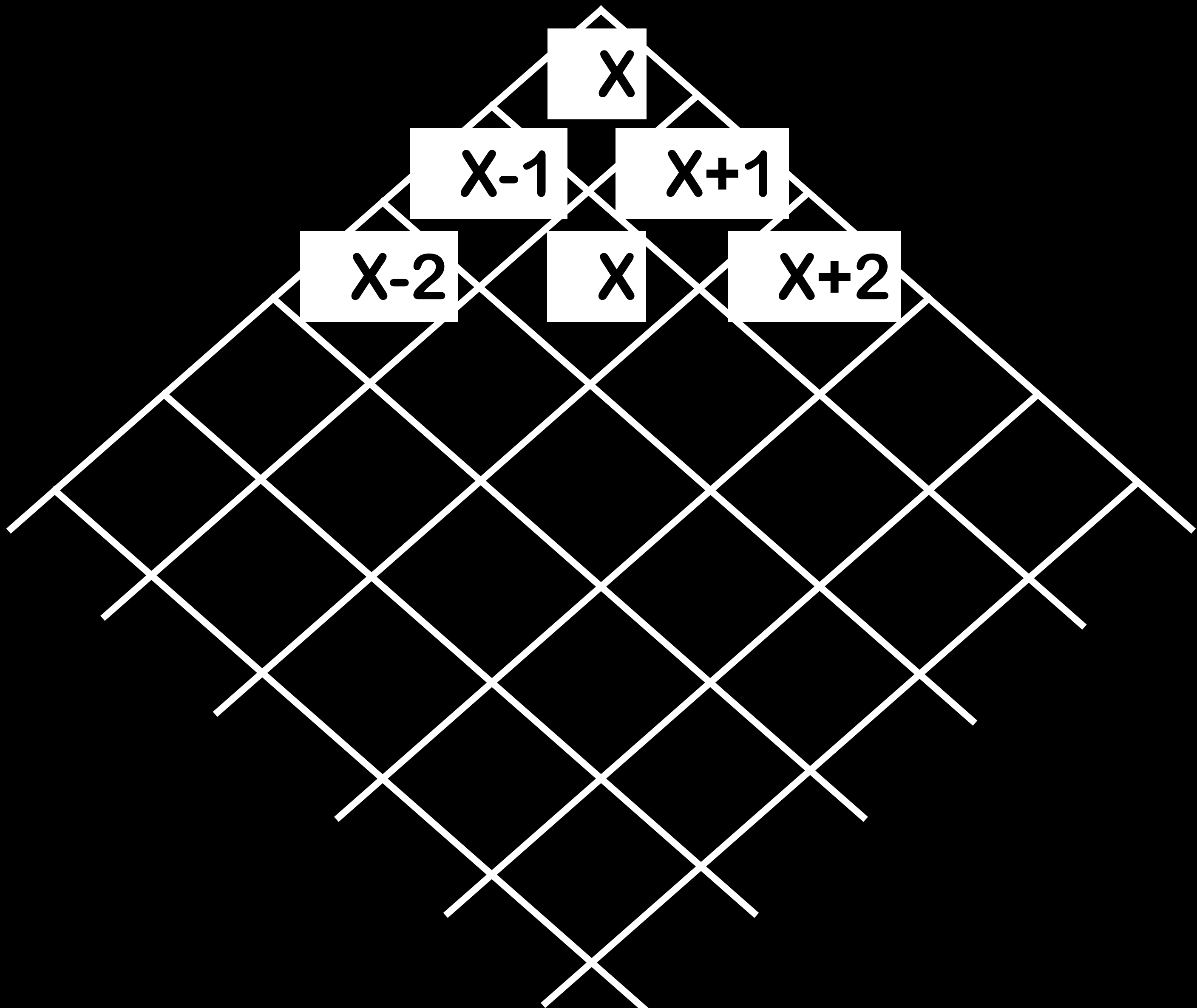
One was X-1 inches tall

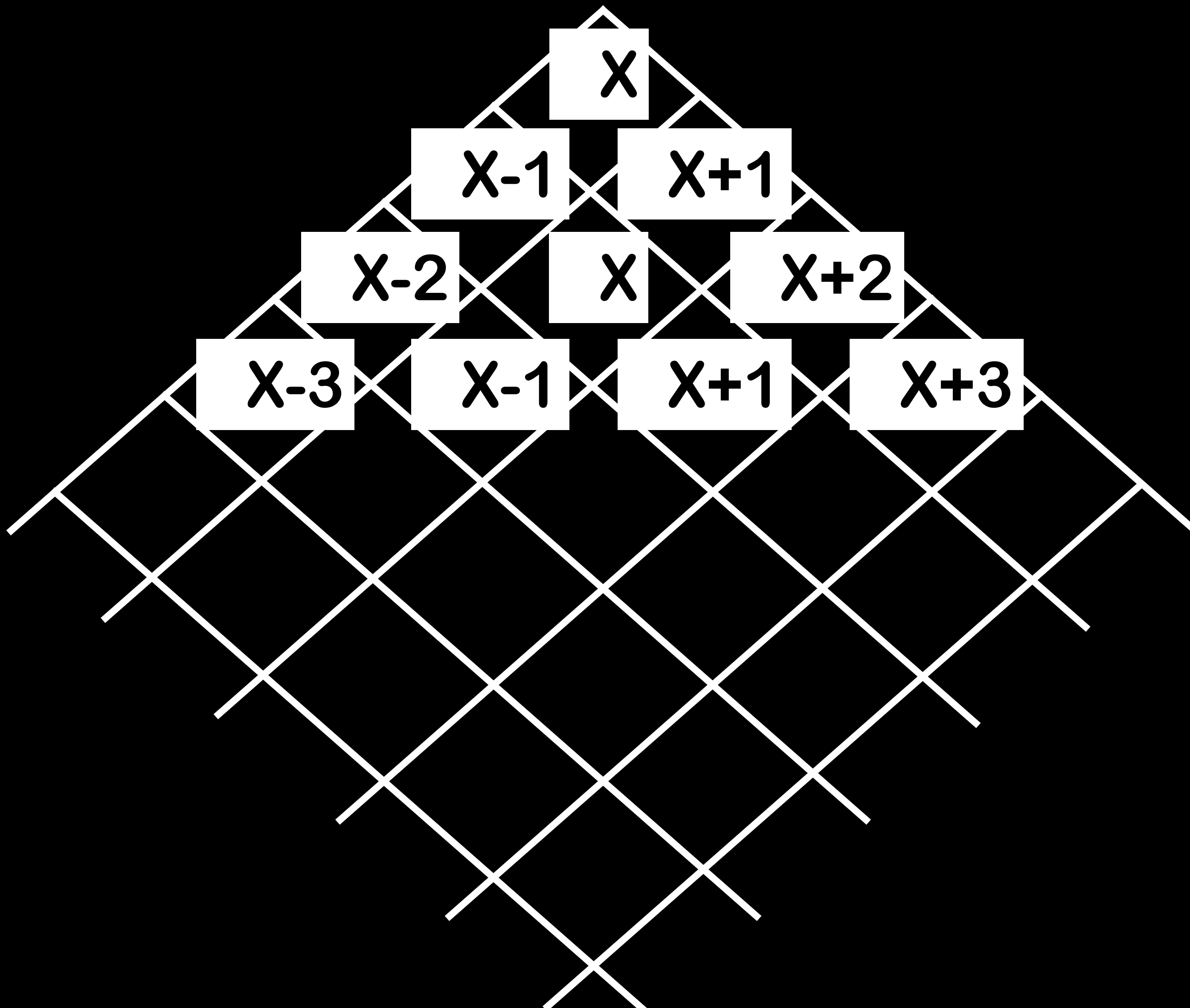
Each of his sons had two sons …

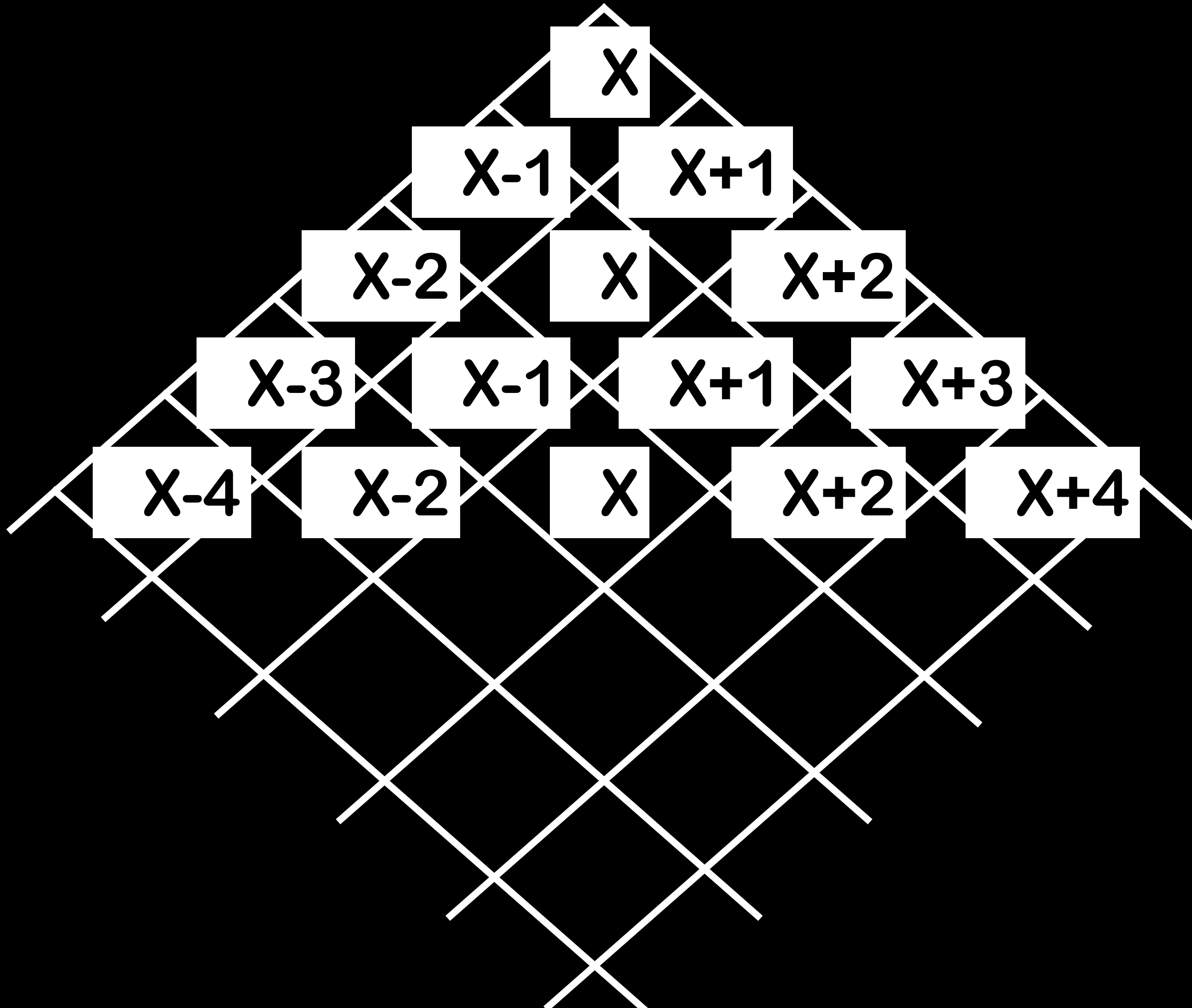


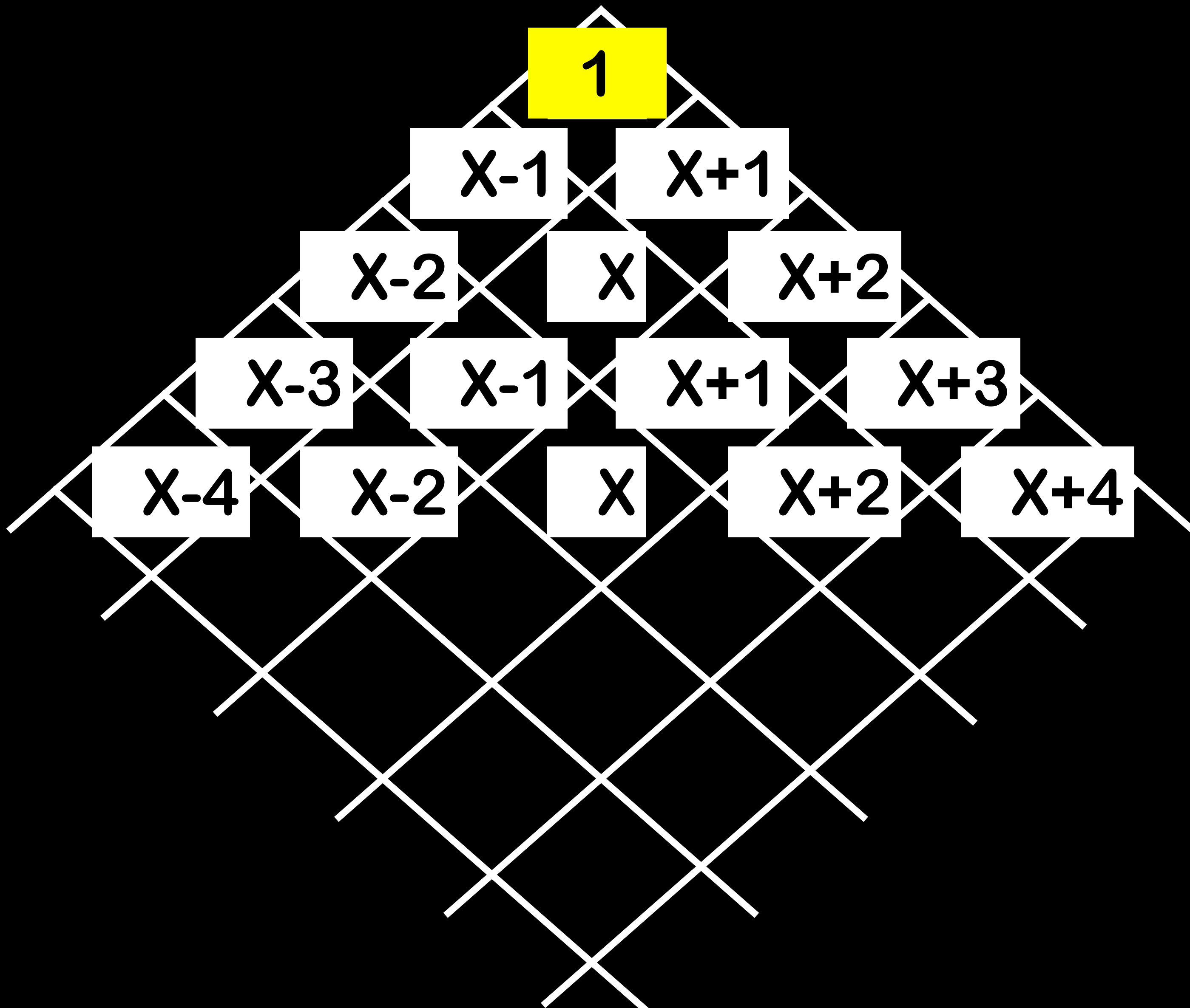

X

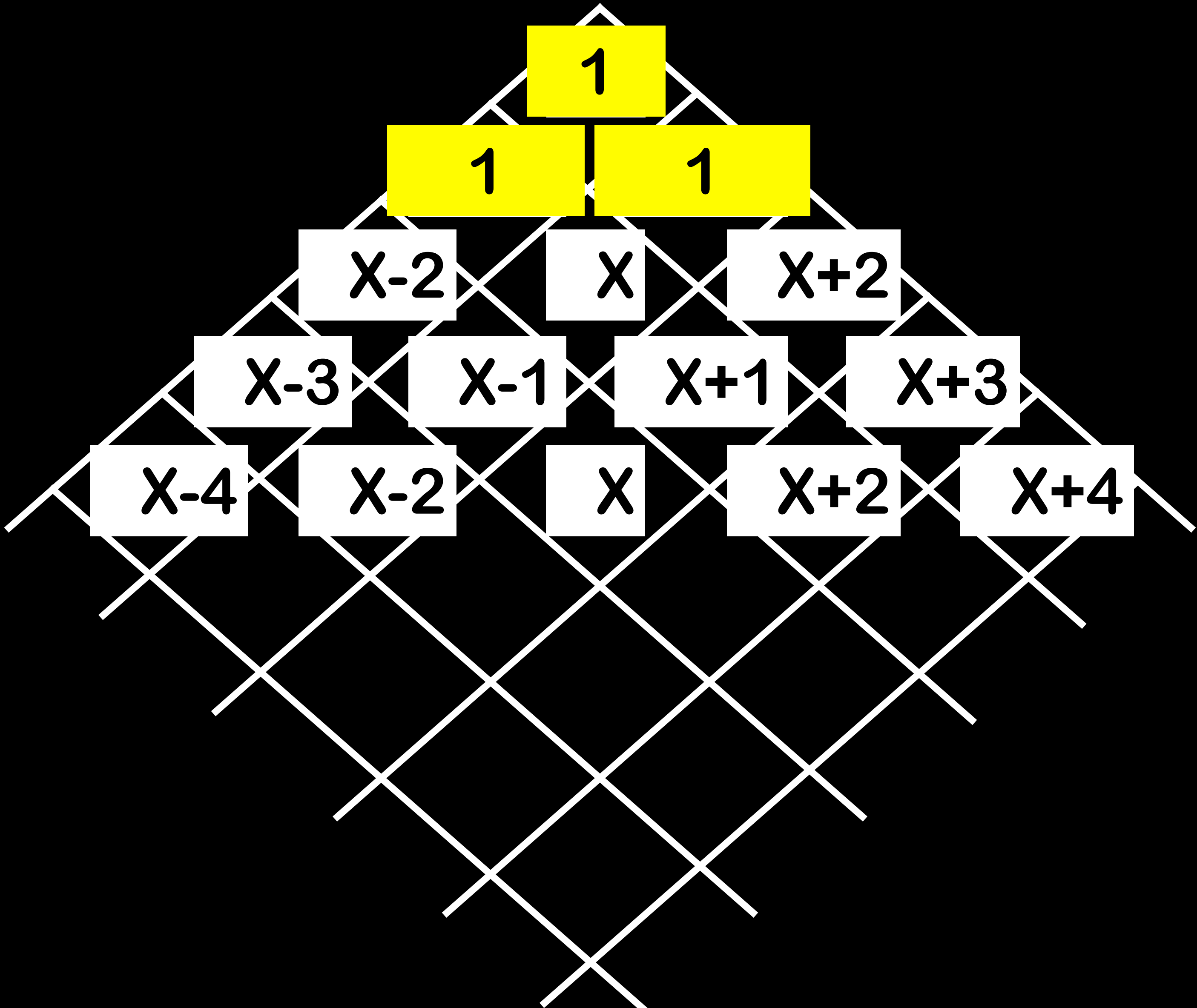

X
X-1
X+1

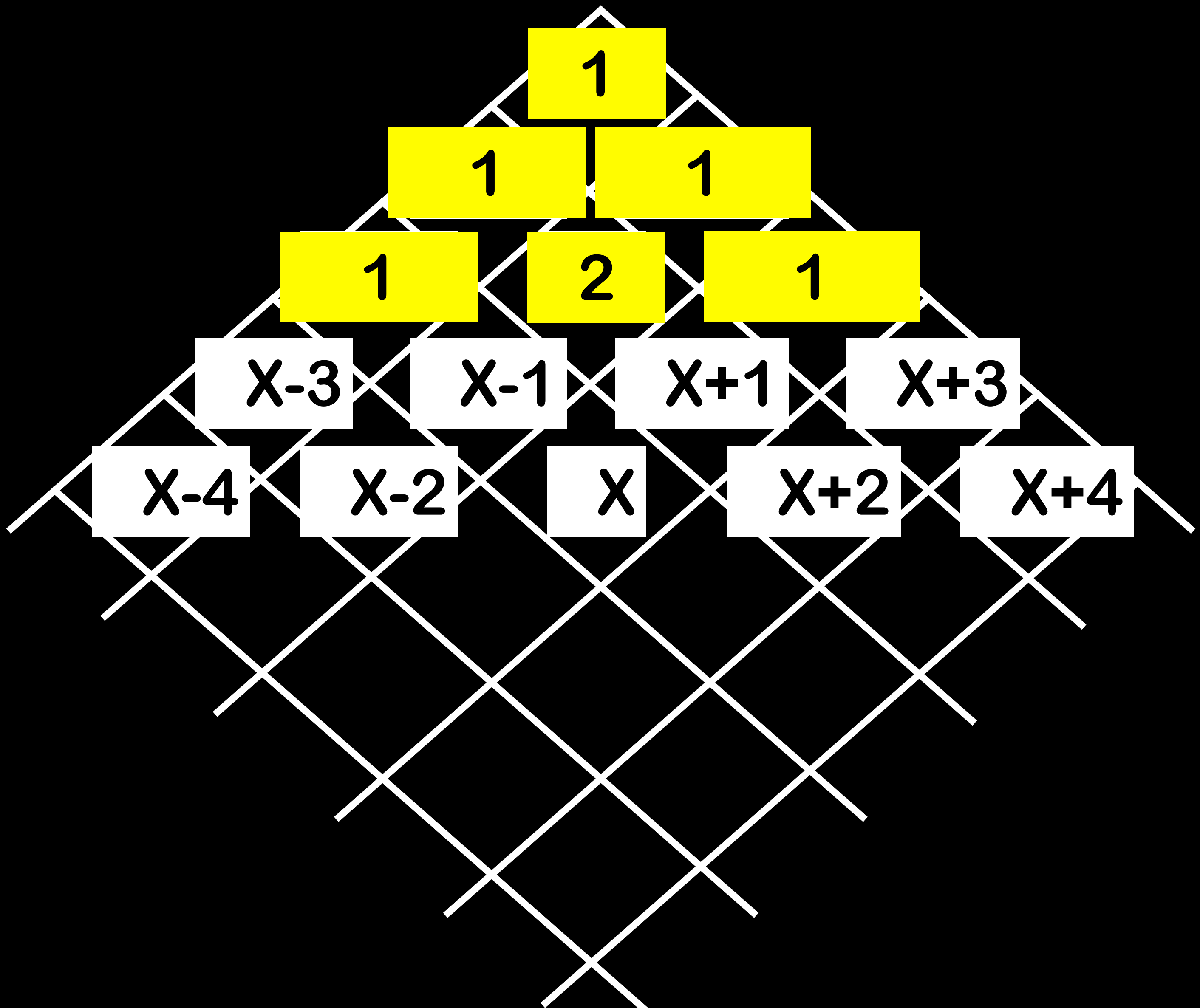

X
X-1
X+1
X-2
X
X+2

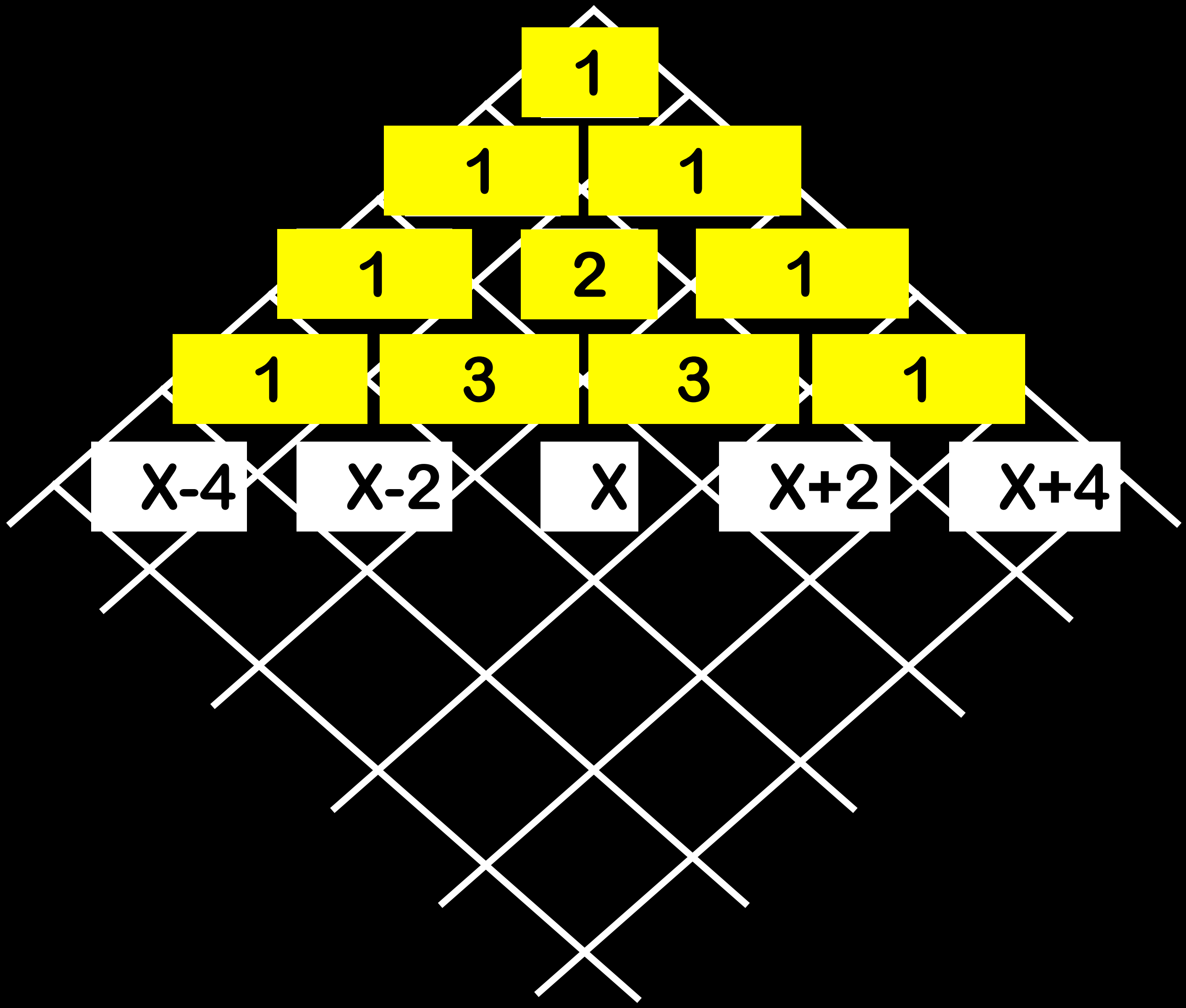

X
X-1
X+1
X-2
X
X+2
X-3
X-1
X+1
X+3

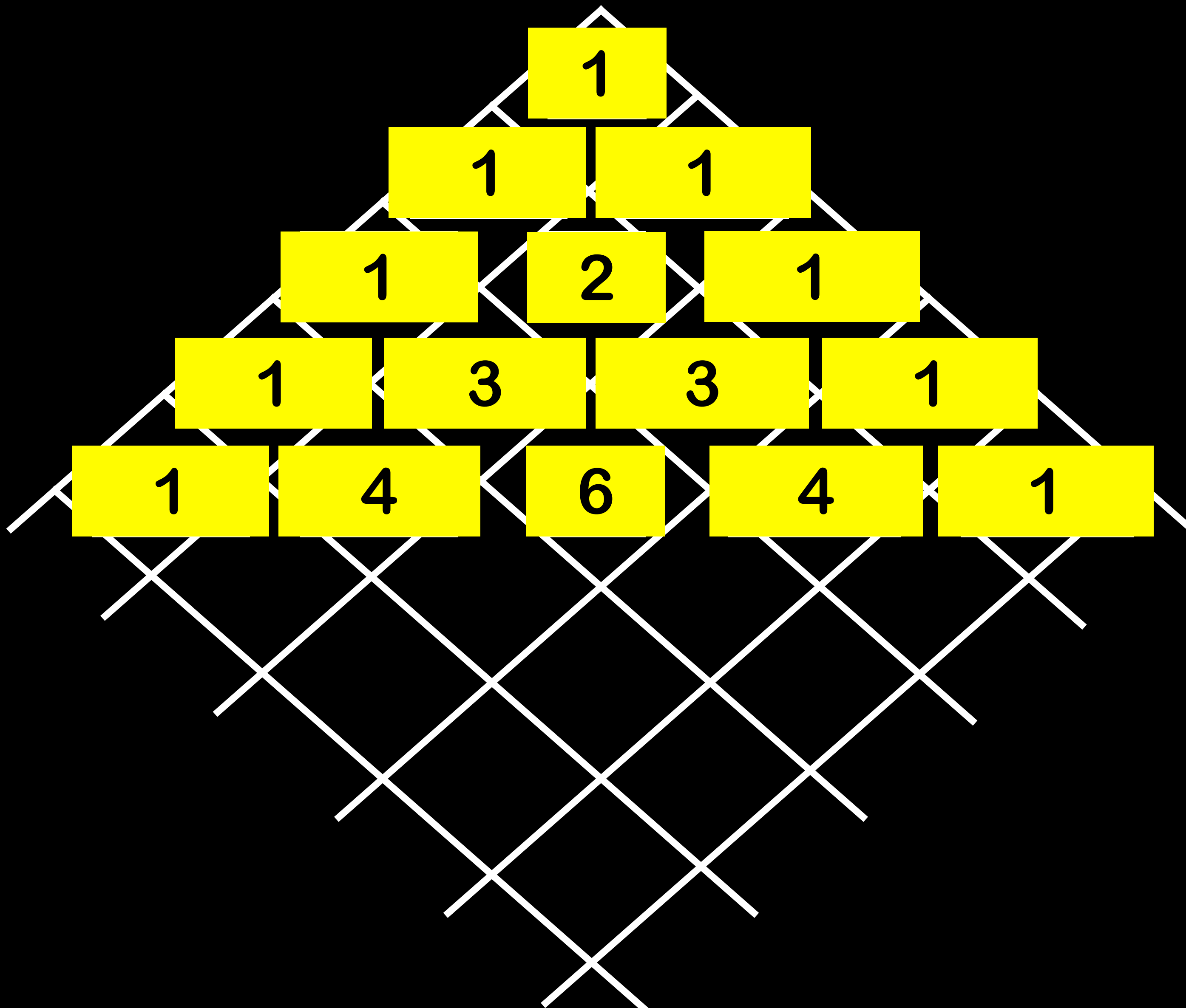

X
X-1
X+1
X-2
X
X+2
X-3
X-1
X+1
X+3
X-4
X-2
X
X+2
X+4

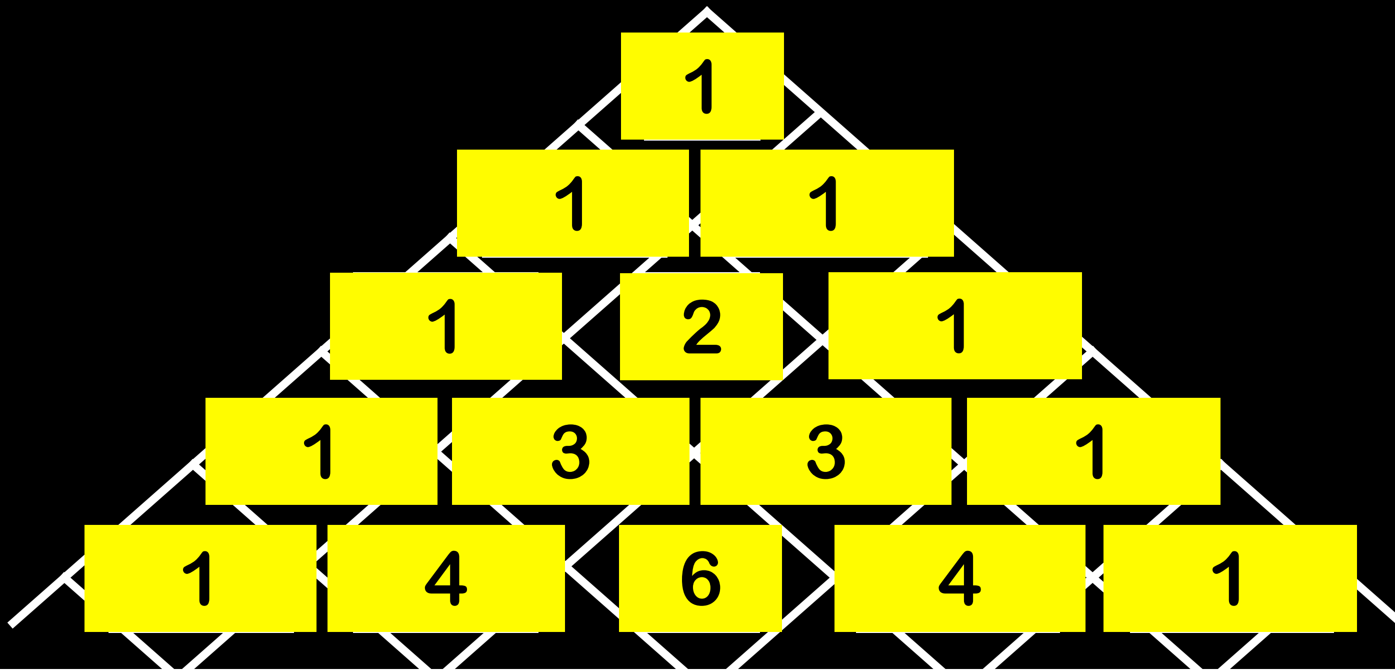

1
X-1
X+1
X-2
X
X+2
X-3
X-1
X+1
X+3
X-4
X-2
X
X+2
X+4


1
1
1
X-2
X
X+2
X-3
X-1
X+1
X+3
X-4
X-2
X
X+2
X+4


1
1
1
1
2
1
X-3
X-1
X+1
X+3
X-4
X-2
X
X+2
X+4


1
1
1
1
2
1
1
3
3
1
X-4
X-2
X
X+2
X+4


1
1 1
1 2 1
1 3 3 1
1 4 6 4 1



In the n^{th} generation there will be 2^n males, each with one of n+1 different heights:

$$h_0, h_1, \ldots, h_n$$

$h_i = (X-n+2i)$ **occurs with proportion:** $\binom{n}{i} / 2^n$

Unbiased Binomial Distribution On n+1 Elements

Unbiased Binomial Distribution On n+1 Elements

Let S be any set $\{h_0, h_1, \ldots, h_n\}$ where each element h_i has an associated probability

Unbiased Binomial Distribution On n+1 Elements

Let S be any set $\{h_0, h_1, \ldots, h_n\}$ where each element h_i has an associated probability

$$\frac{\binom{n}{i}}{2^n}$$

Unbiased Binomial Distribution On n+1 Elements

Let S be any set $\{h_0, h_1, \ldots, h_n\}$ where each element h_i has an associated probability

$$\frac{\binom{n}{i}}{2^n}$$

Any such distribution is called an Unbiased Binomial Distribution or an Unbiased Bernoulli Distribution

More Language Of Probability

More Language Of Probability

The probability of event A given event B is written Pr[A | B]

More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

More Language Of Probability

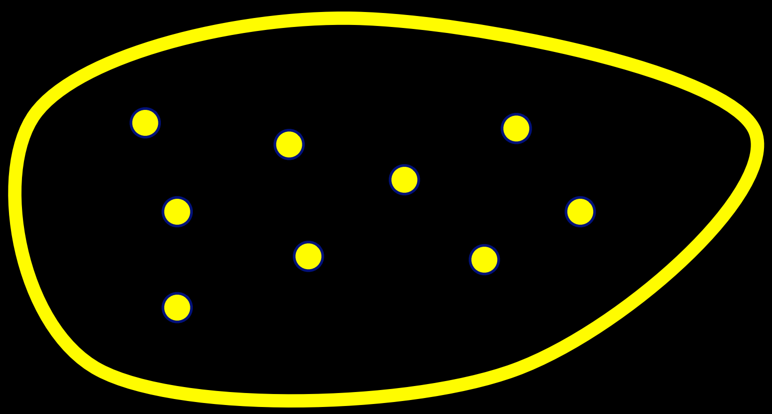
The probability of event A given event B is written Pr[A | B] and is defined to be =

$$\frac{\Pr[A \cap B]}{\Pr[B]}$$

More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

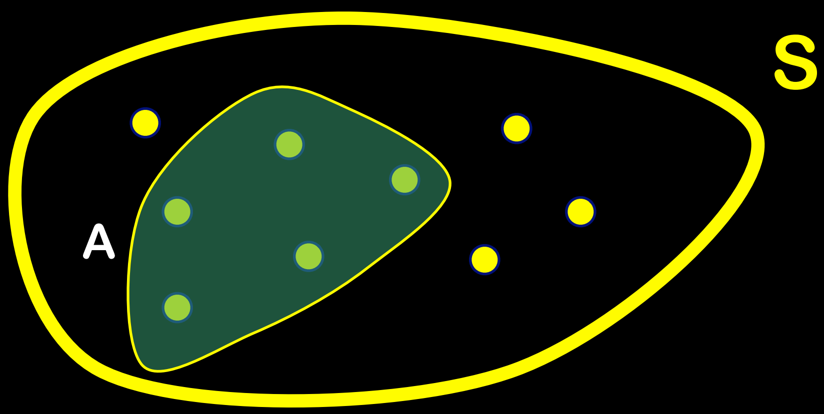
$$\frac{\Pr[A \cap B]}{\Pr[B]}$$



More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

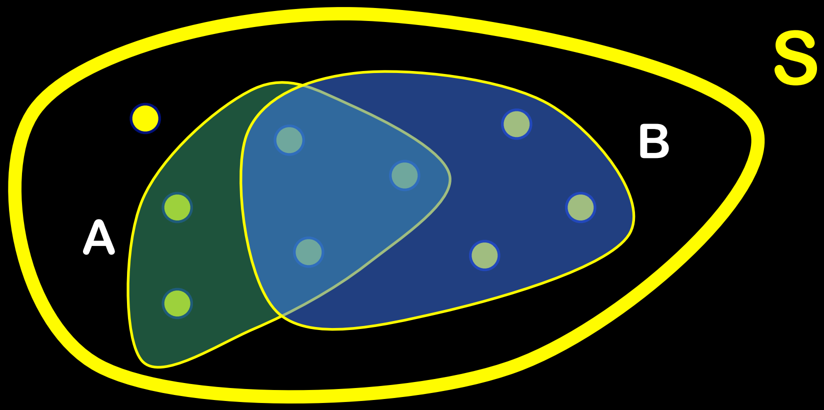
$$\frac{\Pr[A \cap B]}{\Pr[B]}$$



More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

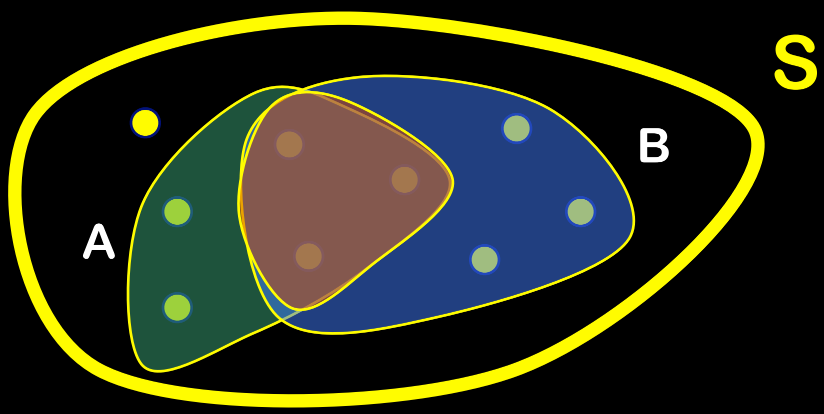
$$\frac{\Pr[\,A \cap B\,]}{\Pr[\,B\,]}$$



More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

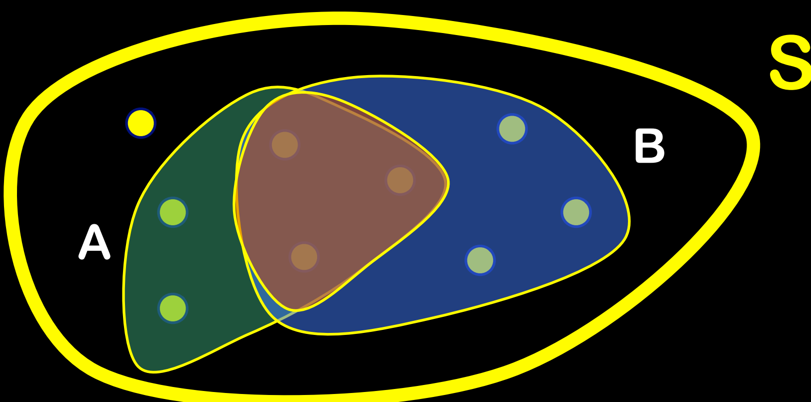
$$\frac{\Pr[A \cap B]}{\Pr[B]}$$



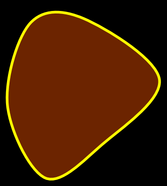
More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

$$\frac{\Pr[\, A \cap B \,]}{\Pr[\, B \,]}$$



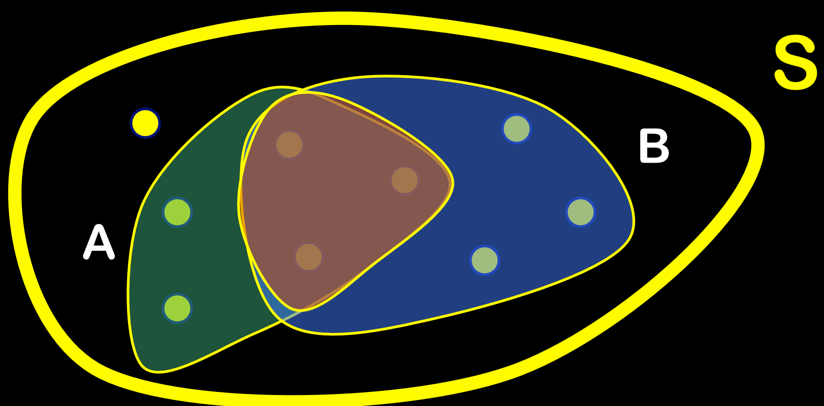
proportion of $A \cap B$



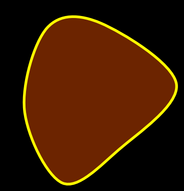
More Language Of Probability

The probability of event A given event B is written Pr[A | B] and is defined to be =

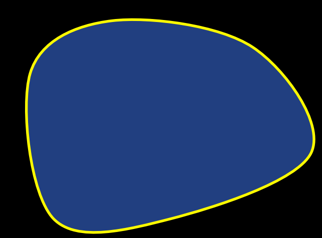
$$\frac{\Pr[\, A \cap B \,]}{\Pr[\, B \,]}$$



proportion of $A \cap B$



to B



Suppose we roll a white die and black die

Suppose we roll a white die and black die

What is the probability that the white is 1 given that the total is 7?

Suppose we roll a white die and black die

What is the probability that the white is 1 given that the total is 7?

event A = {white die = 1}

Suppose we roll a white die and black die

What is the probability that the white is 1 given that the total is 7?

event A = {white die = 1}

event B = {total = 7}

S = { (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) }

event A = {white die = 1}

event B = {total = 7}

$$S = \{ \boxed{(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ (1,6),}$$

$$(2,1),\ (2,2),\ (2,3),\ (2,4),\ (2,5),\ (2,6),$$

$$(3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6),$$

$$(4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6),$$

$$(5,1),\ (5,2),\ (5,3),\ (5,4),\ (5,5),\ (5,6),$$

$$(6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6) \}$$

event A = {white die = 1}

event B = {total = 7}

S = { (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) }

event A = {white die = 1}

event B = {total = 7}

$$S = \{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),$$
$$(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),$$
$$(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),$$
$$(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),$$
$$(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),$$
$$(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \}$$

$$\Pr[A \mid B] =$$

event A = {white die = 1}

event B = {total = 7}

$$
\begin{aligned}
S = \{\, & \boxed{(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ \boxed{(1,6)},} \\
& (2,1),\ (2,2),\ (2,3),\ (2,4),\ \boxed{(2,5)},\ (2,6), \\
& (3,1),\ (3,2),\ (3,3),\ \boxed{(3,4)},\ (3,5),\ (3,6), \\
& (4,1),\ (4,2),\ \boxed{(4,3)},\ (4,4),\ (4,5),\ (4,6), \\
& (5,1),\ \boxed{(5,2)},\ (5,3),\ (5,4),\ (5,5),\ (5,6), \\
& \boxed{(6,1)},\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6) \,\}
\end{aligned}
$$

$$
\Pr[\, A \mid B \,] = \frac{\Pr[\, A \cap B \,]}{\Pr[B]}
$$

event A = {white die = 1}

event B = {total = 7}

$$
\begin{aligned}
S = \{ & \boxed{(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ \boxed{(1,6)},} \\
& (2,1),\ (2,2),\ (2,3),\ (2,4),\ \boxed{(2,5)},\ (2,6), \\
& (3,1),\ (3,2),\ (3,3),\ \boxed{(3,4)},\ (3,5),\ (3,6), \\
& (4,1),\ (4,2),\ \boxed{(4,3)},\ (4,4),\ (4,5),\ (4,6), \\
& (5,1),\ \boxed{(5,2)},\ (5,3),\ (5,4),\ (5,5),\ (5,6), \\
& \boxed{(6,1)},\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6) \}
\end{aligned}
$$

$$
\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{|A \cap B|}{|B|}
$$

event A = {white die = 1}

event B = {total = 7}

$$S = \{ \boxed{(1,1), \quad (1,2), \quad (1,3), \quad (1,4), \quad (1,5), \quad \boxed{(1,6)}},$$
$$(2,1), \quad (2,2), \quad (2,3), \quad (2,4), \quad \boxed{(2,5)}, \quad (2,6),$$
$$(3,1), \quad (3,2), \quad (3,3), \quad \boxed{(3,4)}, \quad (3,5), \quad (3,6),$$
$$(4,1), \quad (4,2), \quad \boxed{(4,3)}, \quad (4,4), \quad (4,5), \quad (4,6),$$
$$(5,1), \quad \boxed{(5,2)}, \quad (5,3), \quad (5,4), \quad (5,5), \quad (5,6),$$
$$\boxed{(6,1)}, \quad (6,2), \quad (6,3), \quad (6,4), \quad (6,5), \quad (6,6) \}$$

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{|A \cap B|}{|B|} = \frac{1}{6}$$

event A = {white die = 1}

event B = {total = 7}

Independence!

Independence!

A and B are independent events if

Independence!

A and B are independent events if

$$\Pr[\,A \mid B\,] = \Pr[\,A\,]$$

Independence!

A and B are independent events if

$$\Pr[\, A \mid B \,] = \Pr[\, A \,]$$

$$\Leftrightarrow$$

$$\Pr[\, A \cap B \,] = \Pr[\, A \,] \Pr[\, B \,]$$

$$\Leftrightarrow$$

$$\Pr[\, B \mid A \,] = \Pr[\, B \,]$$

Declaration of Independence

A_1, A_2, …, A_k are independent events if knowing if some of them occurred does not change the probability of any of the others occurring

Declaration of Independence

A_1, A_2, …, A_k are **independent events** if knowing if some of them occurred does not change the probability of any of the others occurring

E.g., $\{A_1, A_2, A_3\}$ are independent events if:

Declaration of Independence

$A_1, A_2, \ldots, A_k$ are independent events if knowing if some of them occurred does not change the probability of any of the others occurring

E.g., $\{A_1, A_2, A_3\}$ are independent events if:

$$\Pr[A_1 \mid A_2] = \Pr[A_1] \quad \Pr[A_1 \mid A_3] = \Pr[A_1]$$
$$\Pr[A_2 \mid A_1] = \Pr[A_2] \quad \Pr[A_2 \mid A_3] = \Pr[A_2]$$
$$\Pr[A_3 \mid A_1] = \Pr[A_3] \quad \Pr[A_3 \mid A_2] = \Pr[A_3]$$

Declaration of Independence

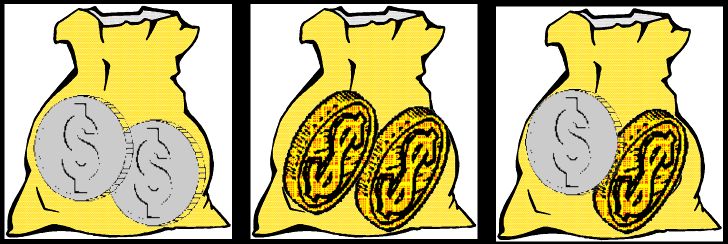
$A_1, A_2, \ldots, A_k$ are **independent events** if knowing if some of them occurred does not change the probability of any of the others occurring

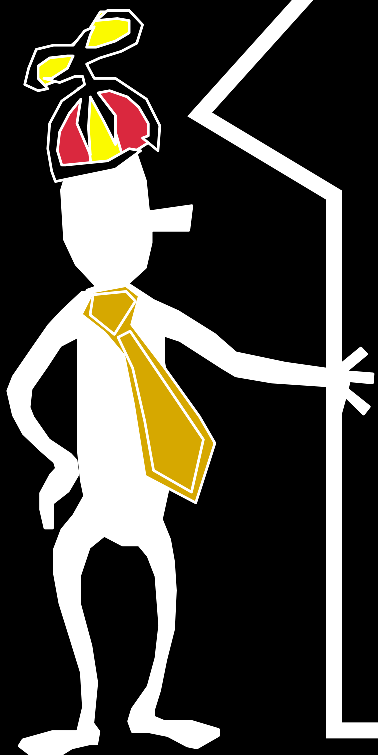
E.g., $\{A_1, A_2, A_3\}$ are independent events if:

$$Pr[A_1 \mid A_2 \cap A_3] = Pr[A_1]$$
$$Pr[A_2 \mid A_1 \cap A_3] = Pr[A_2]$$
$$Pr[A_3 \mid A_1 \cap A_2] = Pr[A_3]$$

$$Pr[A_1 \mid A_2] = Pr[A_1] \quad Pr[A_1 \mid A_3] = Pr[A_1]$$
$$Pr[A_2 \mid A_1] = Pr[A_2] \quad Pr[A_2 \mid A_3] = Pr[A_2]$$
$$Pr[A_3 \mid A_1] = Pr[A_3] \quad Pr[A_3 \mid A_2] = Pr[A_3]$$

Silver and Gold





One bag has two silver coins, another has two gold coins, and the third has one of each

One bag is selected at random. One coin from it is selected at random. It turns out to be gold

What is the probability that the other coin is gold?

Let G_1 be the event that the first coin is gold

Let G_1 be the event that the first coin is gold

$Pr[G_1] = 1/2$

Let G_1 be the event that the first coin is gold

$Pr[G_1] = 1/2$

Let G_2 be the event that the second coin is gold

Let G_1 be the event that the first coin is gold

$Pr[G_1] = 1/2$

Let G_2 be the event that the second coin is gold

$Pr[G_2 \mid G_1] = Pr[G_1 \text{ and } G_2] / Pr[G_1]$

Let G_1 be the event that the first coin is gold

$\Pr[G_1] = 1/2$

Let G_2 be the event that the second coin is gold

$$\Pr[G_2 \mid G_1] = \Pr[G_1 \text{ and } G_2] / \Pr[G_1]$$

$$= (1/3) / (1/2)$$

Let G_1 be the event that the first coin is gold

$Pr[G_1] = 1/2$

Let G_2 be the event that the second coin is gold

$$Pr[G_2 \mid G_1] = Pr[G_1 \text{ and } G_2] / Pr[G_1]$$

$$= (1/3) / (1/2)$$

$$= 2/3$$

Let G_1 be the event that the first coin is gold

$\Pr[G_1] = 1/2$

Let G_2 be the event that the second coin is gold

$$\Pr[G_2 \mid G_1] = \Pr[G_1 \text{ and } G_2] / \Pr[G_1]$$

$$= (1/3) / (1/2)$$

$$= 2/3$$

Note: G_1 and G_2 are not independent

Monty Hall Problem

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

You select some door

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

You select some door

Announcer opens one of others with no prize

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

You select some door

Announcer opens one of others with no prize

You can decide to keep or switch

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

You select some door

Announcer opens one of others with no prize

You can decide to keep or switch

What to do?

Monty Hall Problem

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Each has probability 1/3

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Each has probability 1/3

Staying

we win if we chose the correct door

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Each has probability 1/3

Staying	Switching
we win if we chose the correct door	we win if we chose the incorrect door

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Each has probability 1/3

Staying

we win if we chose the correct door

Pr[choosing correct door] = 1/3

Switching

we win if we chose the incorrect door

Monty Hall Problem

Sample space = { prize behind door 1, prize behind door 2, prize behind door 3 }

Each has probability 1/3

Staying

we win if we chose the correct door

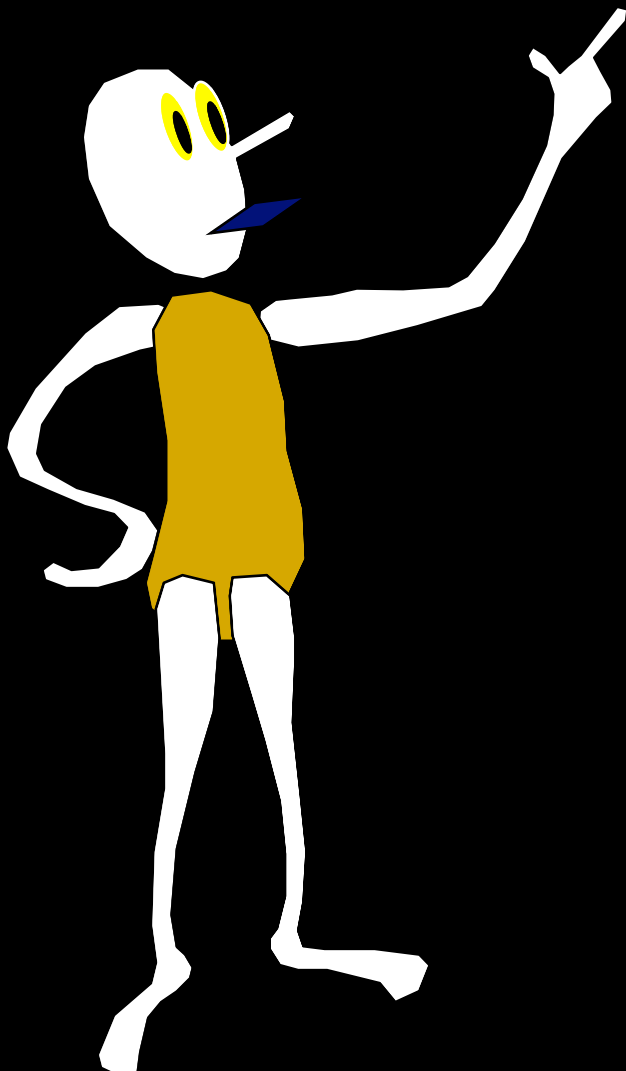
Pr[choosing correct door] = 1/3

Switching

we win if we chose the incorrect door

Pr[choosing incorrect door] = 2/3

Why Was This Tricky?

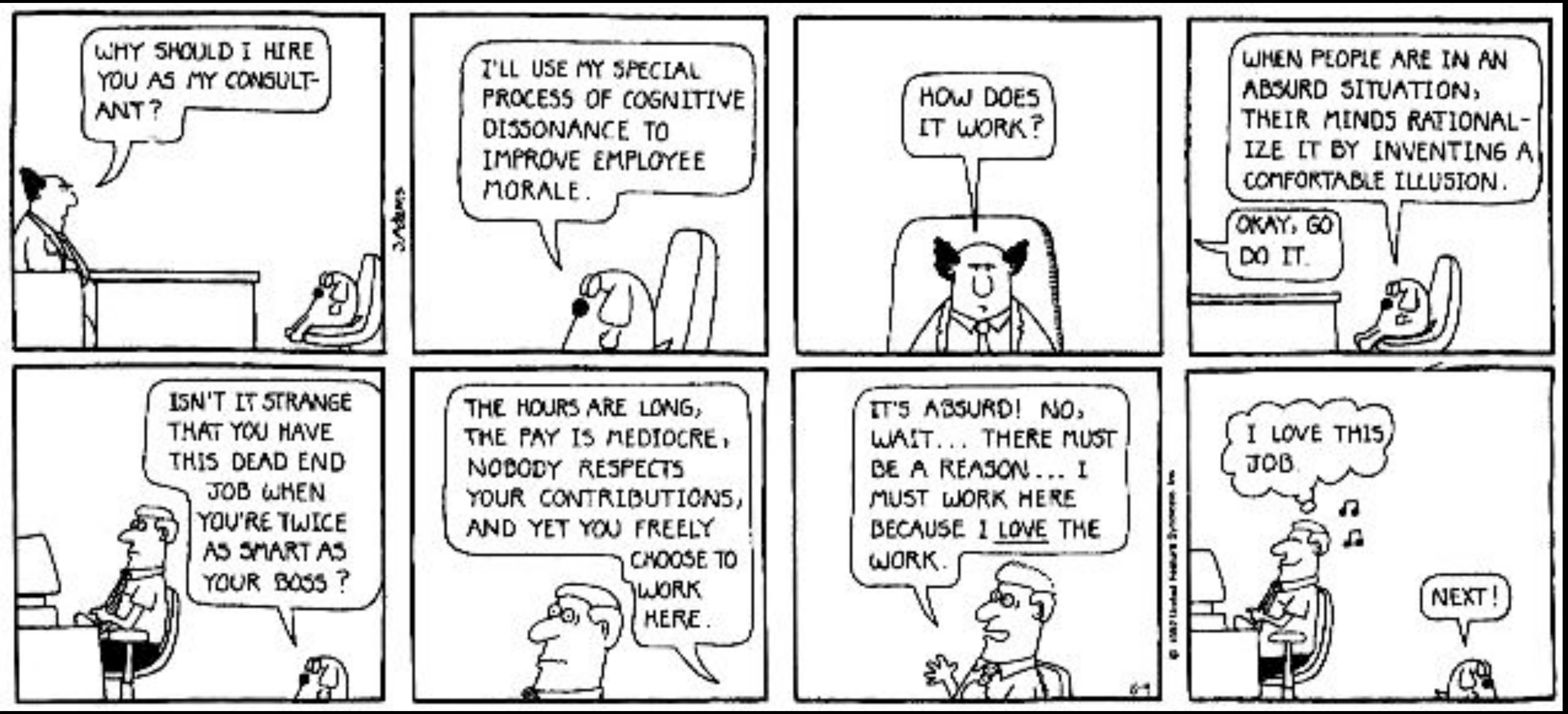


We are inclined to think:

"After one door is opened, others are equally likely…"

But his action is not independent of yours!

Cognitive Dissonance



Monty Meets Monkeys

(from article by John Tierney)

Experiment: Psychologists first observe that a monkey seeks out red, blue, and green M&Ms about equally

Monty Meets Monkeys

(from article by John Tierney)

Experiment: Psychologists first observe that a monkey seeks out red, blue, and green M&Ms about equally

The monkey is given a choice of red or blue candy. It chooses red.

Monty Meets Monkeys

(from article by John Tierney)

Experiment: Psychologists first observe that a monkey seeks out red, blue, and green M&Ms about equally

The monkey is given a choice of red or blue candy. It chooses red.

If the monkey is then given a choice of blue or green, it is more likely to choose green.

Monty Meets Monkeys

Psychological explanation:
Monkey rationalizes its initial rejection of blue by telling itself it doesn't really like blue. (Cognitive dissonance)

Monty Meets Monkeys

Psychological explanation: Monkey rationalizes its initial rejection of blue by telling itself it doesn't really like blue. (Cognitive dissonance)

Probabilistic explanation: If the monkey slightly prefers red over blue, only three ways it can rank green. It prefers green in two of the three.

Monty Meets Monkeys

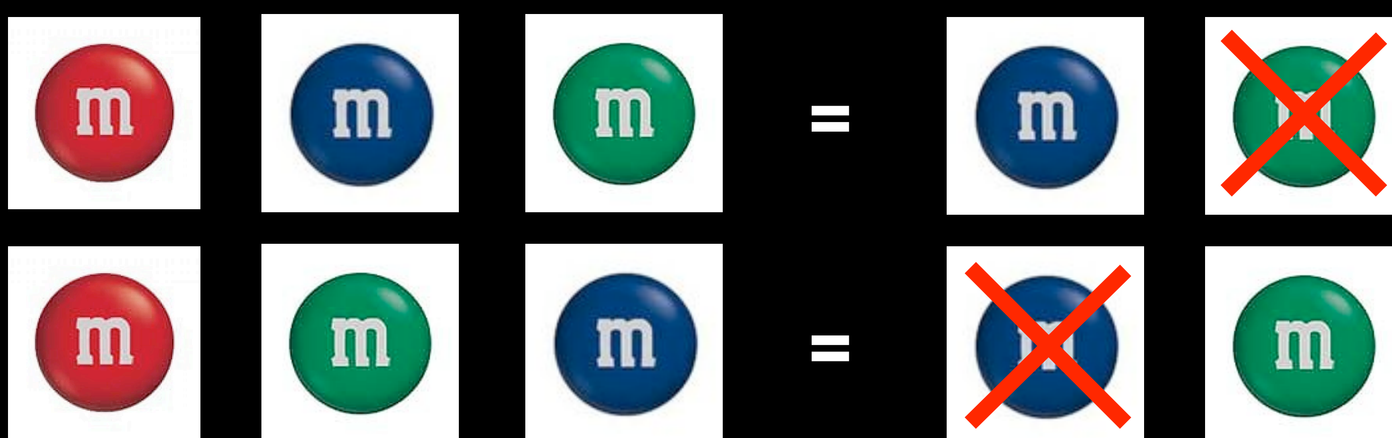
Psychological explanation: Monkey rationalizes its initial rejection of blue by telling itself it doesn't really like blue. (Cognitive dissonance)

Probabilistic explanation: If the monkey slightly prefers red over blue, only three ways it can rank green. It prefers green in two of the three.

Monty Meets Monkeys

Psychological explanation: Monkey rationalizes its initial rejection of blue by telling itself it doesn't really like blue. (Cognitive dissonance)

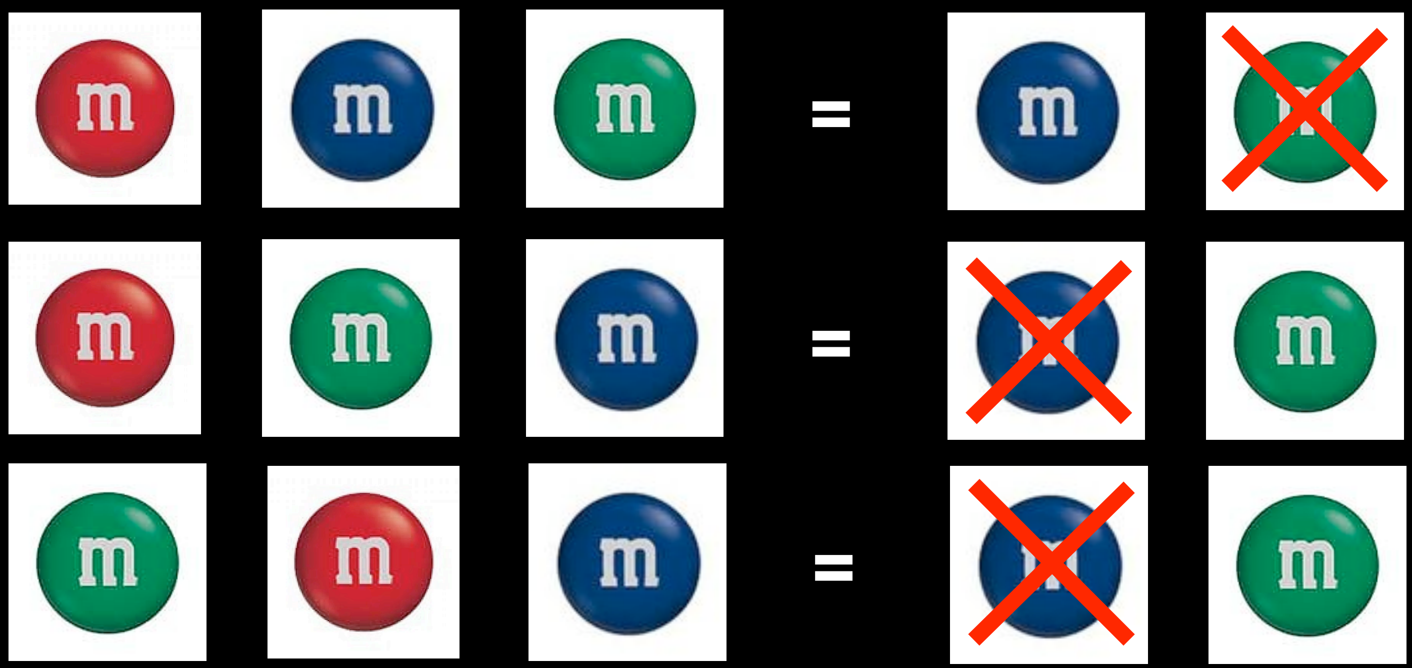
Probabilistic explanation: If the monkey slightly prefers red over blue, only three ways it can rank green. It prefers green in two of the three.



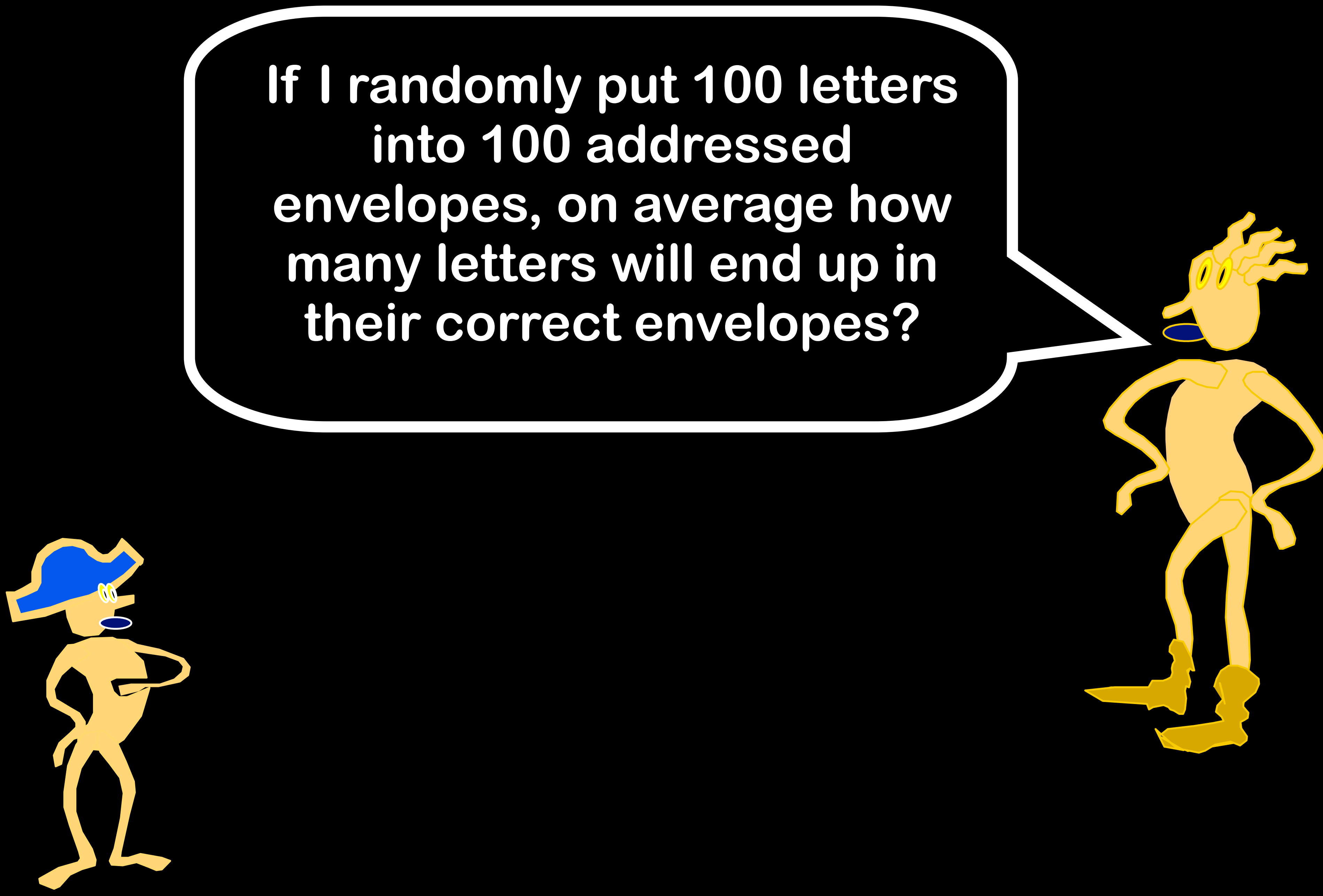
Monty Meets Monkeys

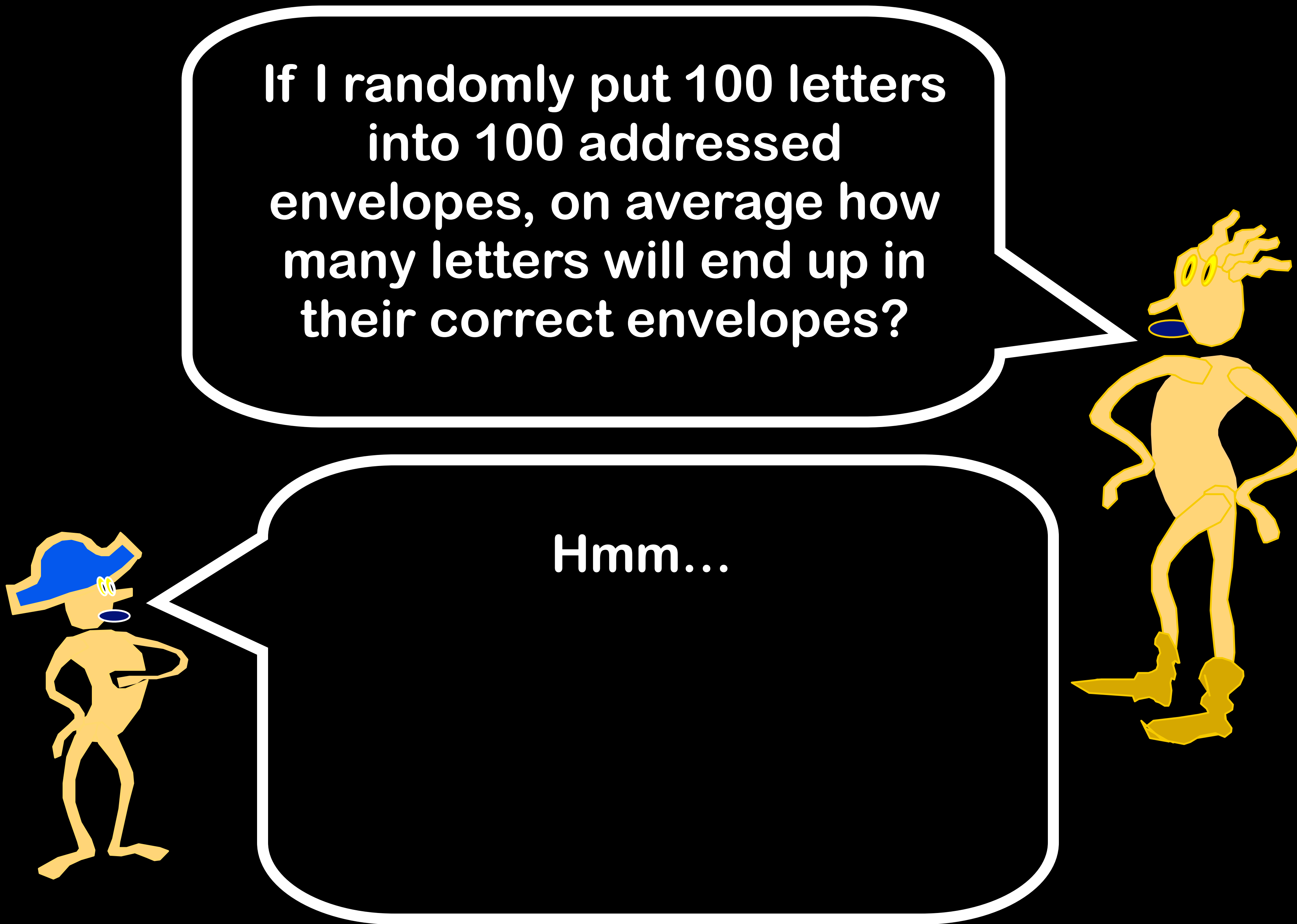
Psychological explanation: Monkey rationalizes its initial rejection of blue by telling itself it doesn't really like blue. (Cognitive dissonance)

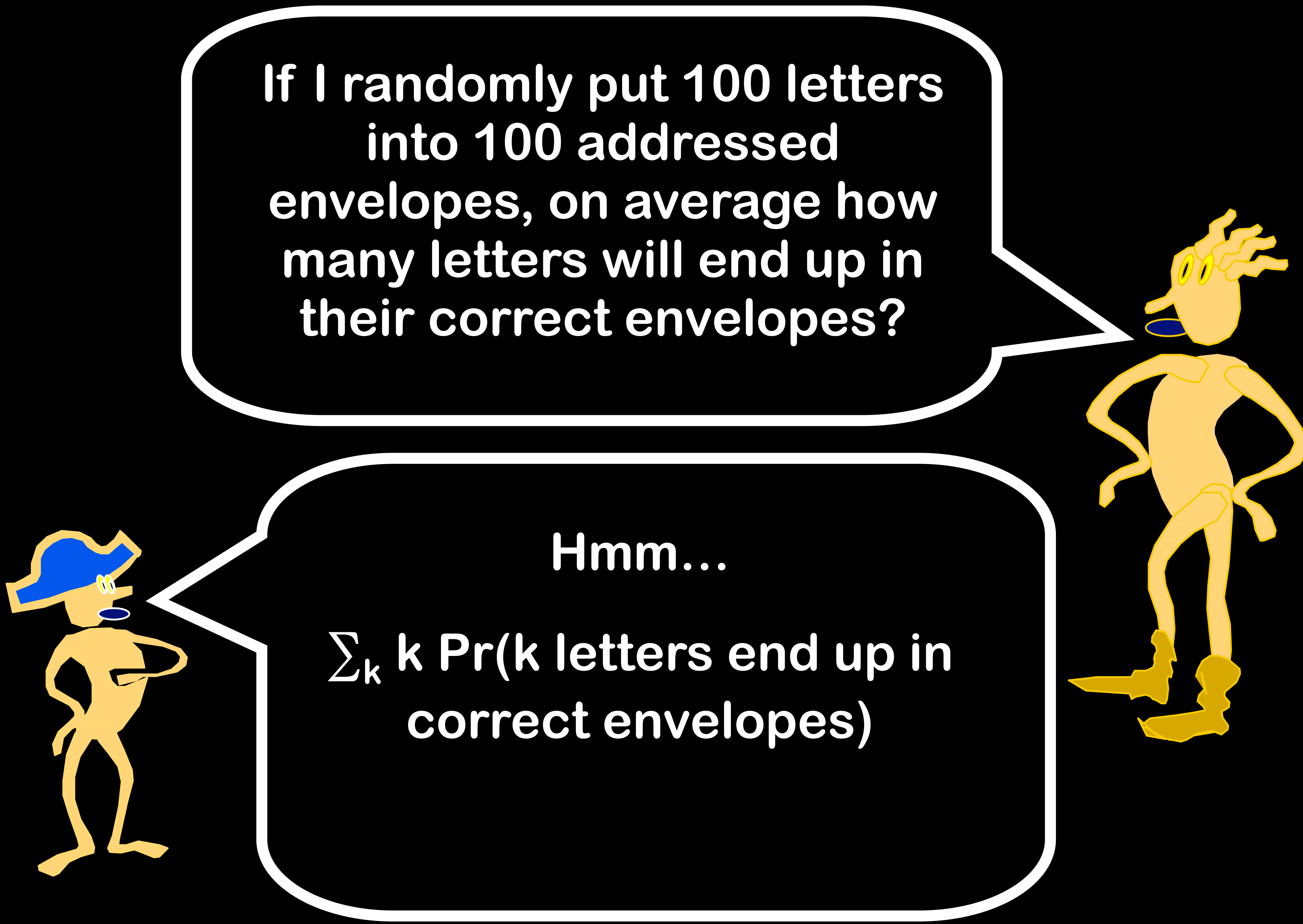
Probabilistic explanation: If the monkey slightly prefers red over blue, only three ways it can rank green. It prefers green in two of the three.



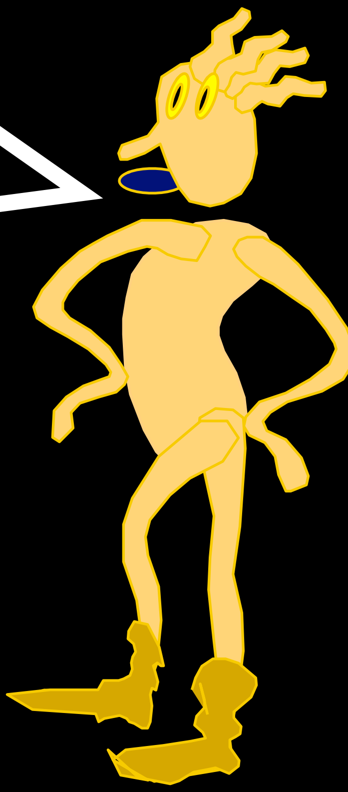
Next, we’ll learn about a formidable tool in probability that will allow us to solve problems that seem really really messy…


If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?


If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?
Hmm…


If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?
Hmm…
$\sum_k$ k Pr(k letters end up in correct envelopes)

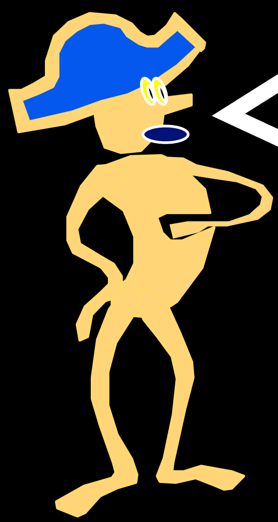
If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

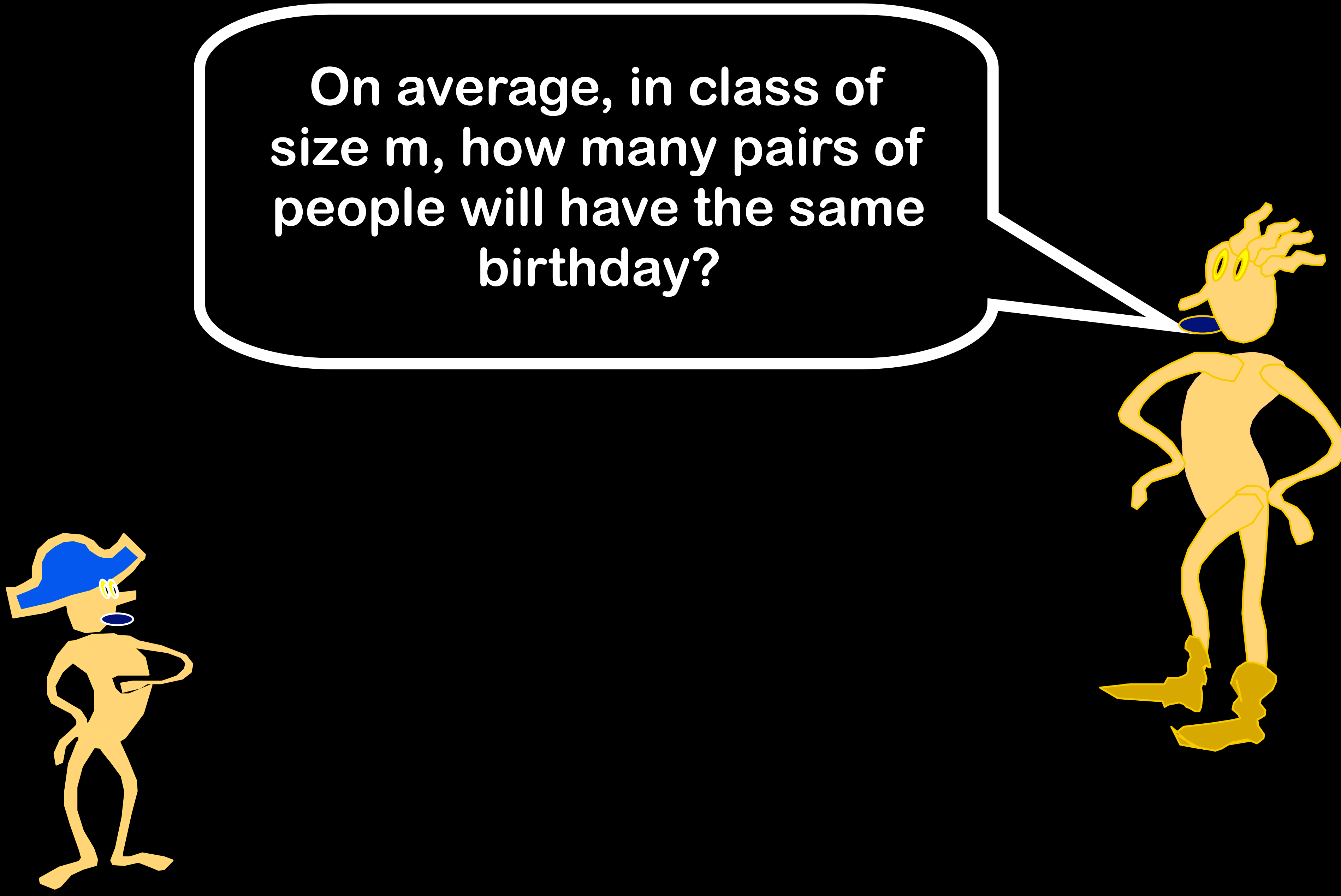


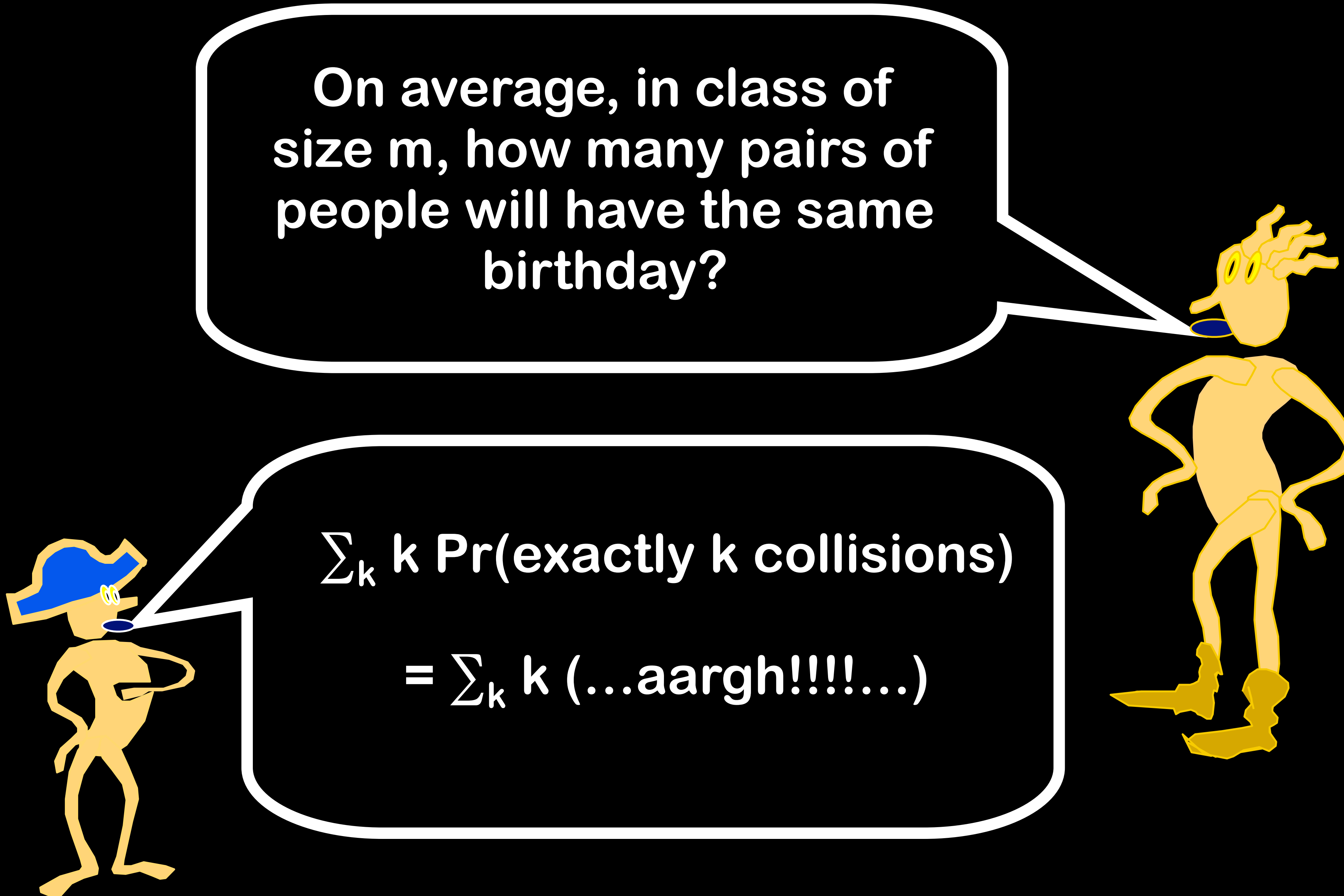
Hmm…

$\sum_k$ k Pr(k letters end up in correct envelopes)

= $\sum_k$ k (…aargh!!…)




On average, in class of size m, how many pairs of people will have the same birthday?


On average, in class of size m, how many pairs of people will have the same birthday?
$\sum_k$ k Pr(exactly k collisions)
= $\sum_k$ k (…aargh!!!!…)

The new tool is called "Linearity of Expectation"

Random Variable

To use this new tool, we will also need to understand the concepts of

Random Variable

and

Expectations

Random Variable

To use this new tool, we will also need to understand the concepts of

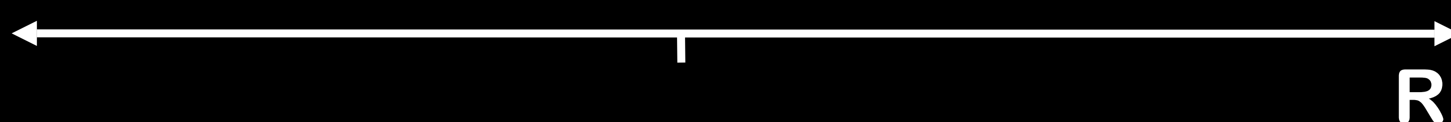
Random Variable

and

Expectations

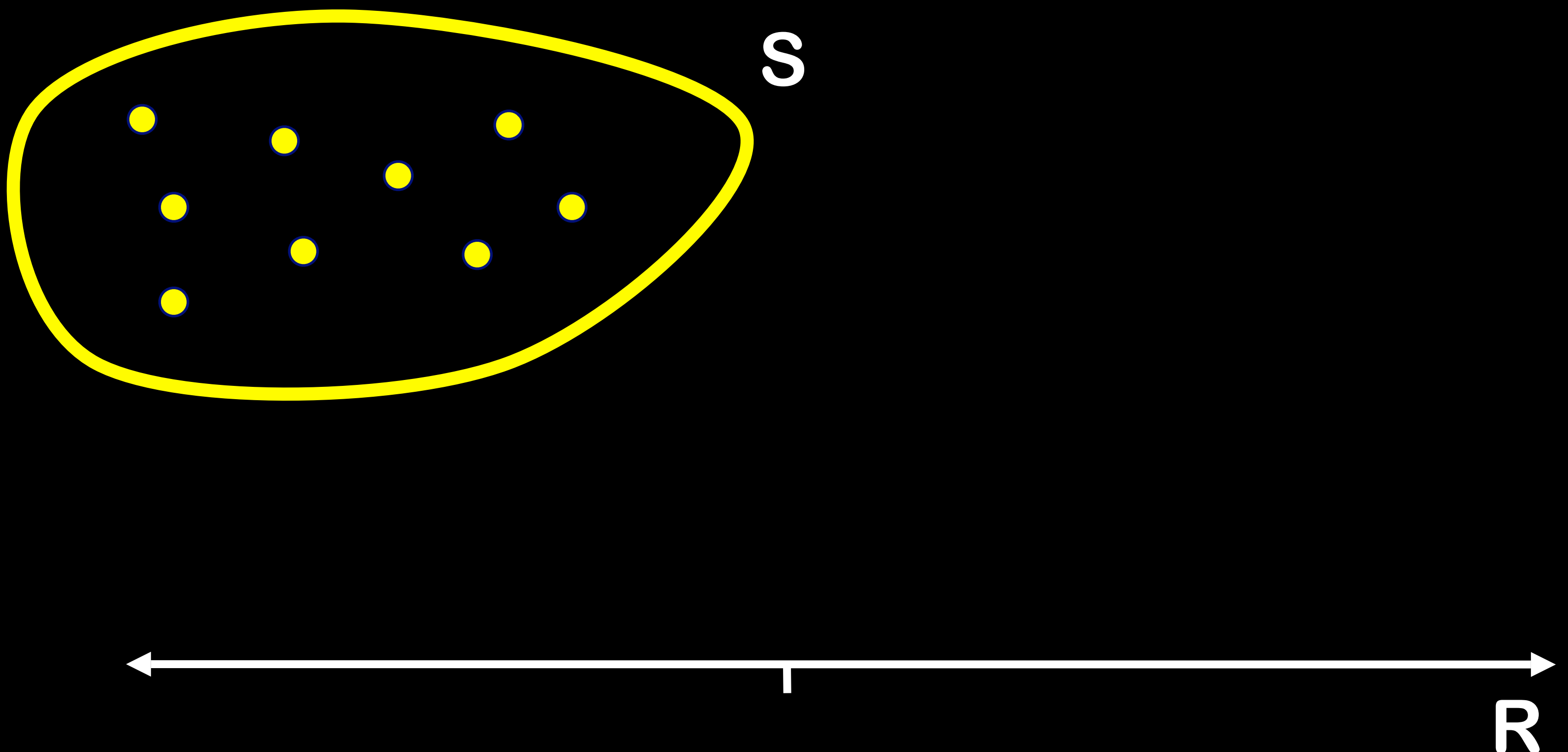
Basic, but need to understand it well

Random Variable



Random Variable

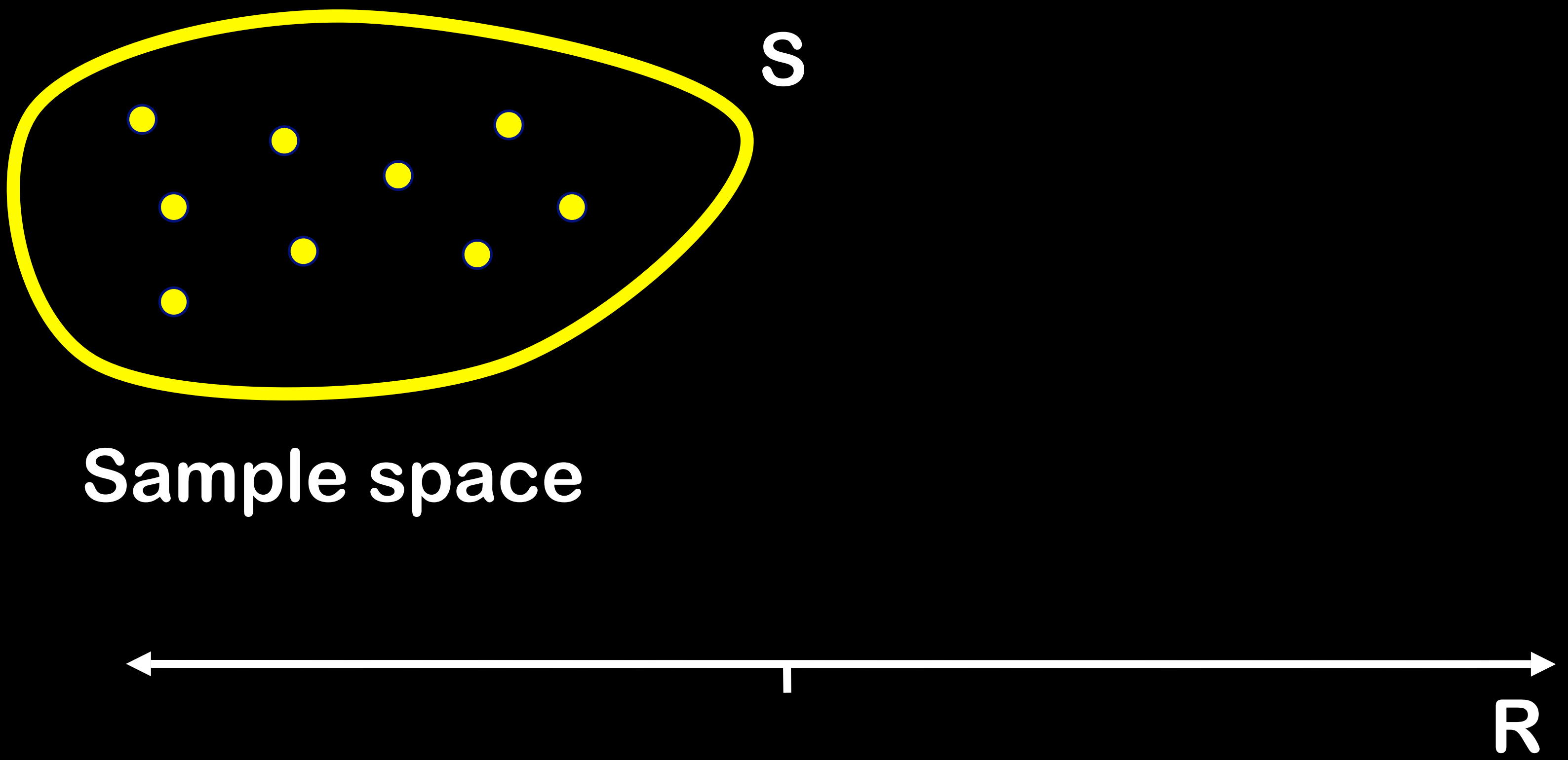
Let S be sample space in a probability distribution



Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S



Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

X(3,4) = 3,   X(1,6) = 1

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

X(3,4) = 3, X(1,6) = 1

Y = sum of values of the two dice

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

$X(3,4) = 3$, $X(1,6) = 1$

Y = sum of values of the two dice

$Y(3,4) = 7$, $Y(1,6) = 7$

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

$X(3,4) = 3, \qquad X(1,6) = 1$

Y = sum of values of the two dice

$Y(3,4) = 7, \qquad Y(1,6) = 7$

W = (value of white die)$^{\text{value of black die}}$

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a real-valued function on S

Examples:

X = value of white die in a two-dice roll

$X(3,4) = 3, \qquad X(1,6) = 1$

Y = sum of values of the two dice

$Y(3,4) = 7, \qquad Y(1,6) = 7$

W = (value of white die)$^{\text{value of black die}}$

$W(3,4) = 3^4, \qquad Y(1,6) = 1^6$

Tossing a Fair Coin n Times

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (1/2)^n$ for all x in S

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (½)^n$ for all x in S

Random Variables (say $n = 10$)

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (½)^n$ for all x in S

Random Variables (say n = 10)

X = # of heads

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (½)^n$ for all x in S

Random Variables (say $n = 10$)

X = # of heads

$X(HHHTTHTHTT) = 5$

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (½)^n$ for all x in S

Random Variables (say $n = 10$)

X = # of heads

$X(HHHTTHTHTT) = 5$

Y = (1 if #heads = #tails, 0 otherwise)

Tossing a Fair Coin n Times

S = all sequences of $\{H, T\}^n$

p = uniform distribution on S

$\Rightarrow p(x) = (½)^n$ for all x in S

Random Variables (say n = 10)

X = # of heads

X(HHHTTHTHTT) = 5

Y = (1 if #heads = #tails, 0 otherwise)

Y(HHHTTHTHTT) = 1, Y(THHHHTTTTT) = 0

Notational Conventions

Notational Conventions

Use letters like A, B, E for events

Notational Conventions

Use letters like A, B, E for events

Use letters like X, Y, f, g for R.V.'s

Notational Conventions

Use letters like A, B, E for events

Use letters like X, Y, f, g for R.V.'s

R.V. = random variable

Two Views of Random Variables

Two Views of Random Variables

Think of a R.V. as

Two Views of Random Variables

Think of a R.V. as

A function from S to the reals R

Two Views of Random Variables

Think of a R.V. as

A function from S to the reals R

Or think of the induced distribution on R

Two Views of Random Variables

Think of a R.V. as

Input to the function is random

A function from S to the reals R

Or think of the induced distribution on R

Two Views of Random Variables

Input to the function is random

Think of a R.V. as

A function from S to the reals R

Or think of the induced distribution on R

Randomness is “pushed” to the values of the function

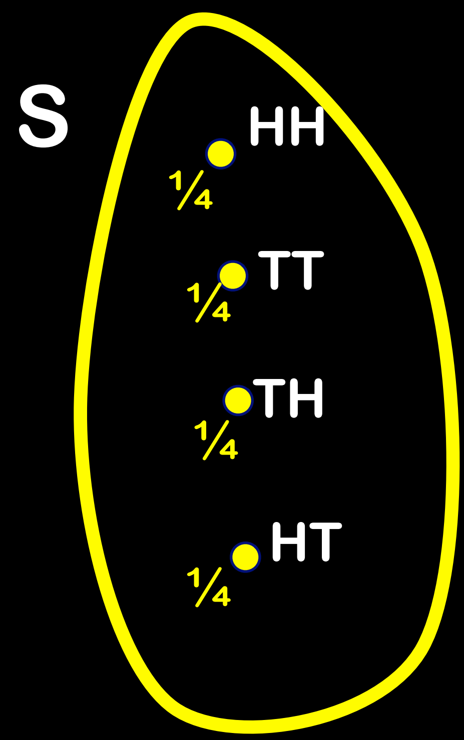
Two Coins Tossed

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads

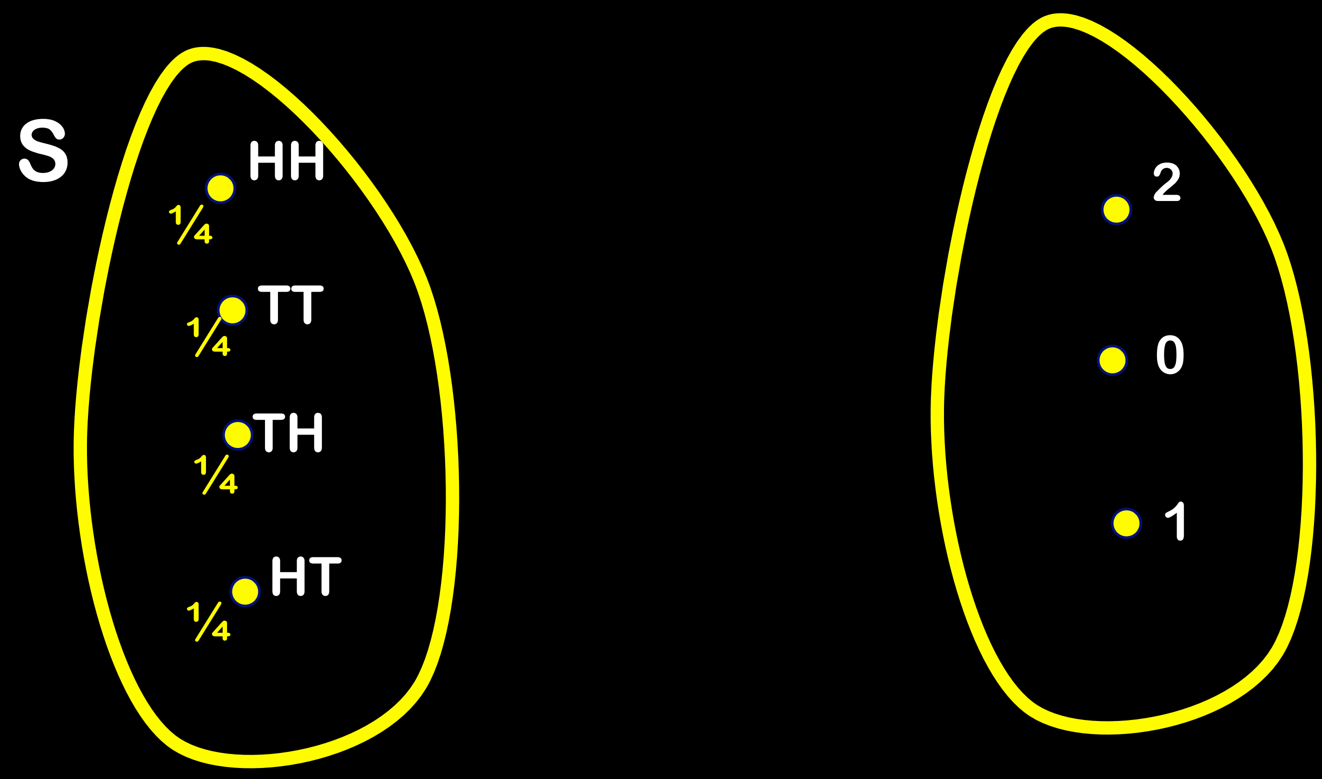
Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads



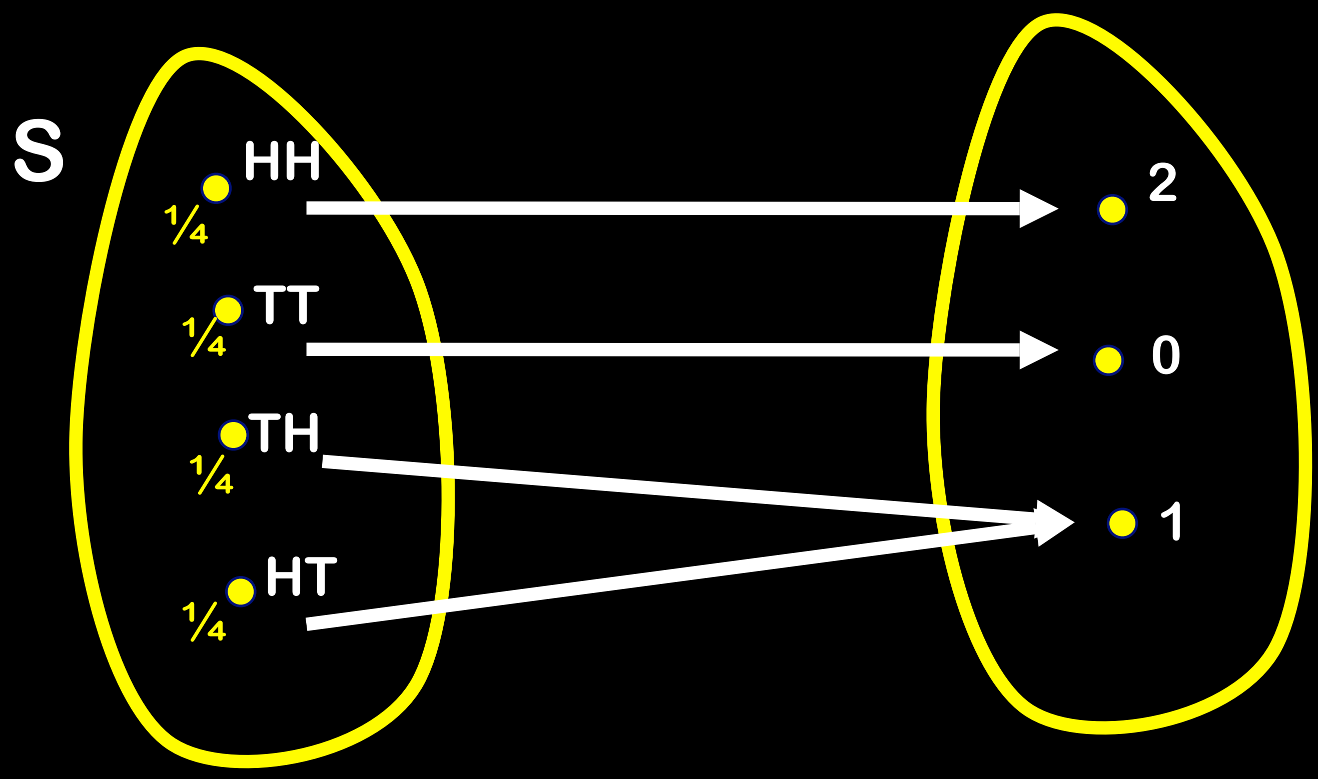
Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads



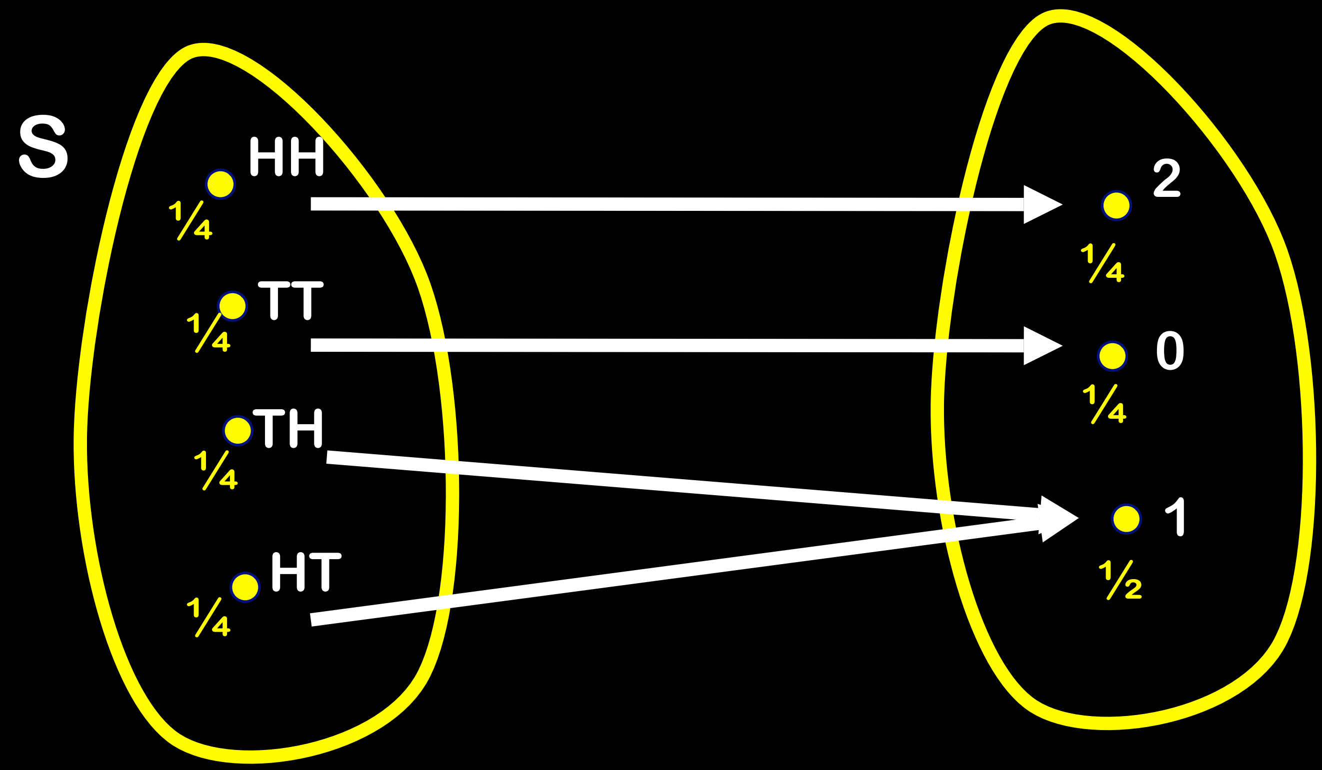
Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads



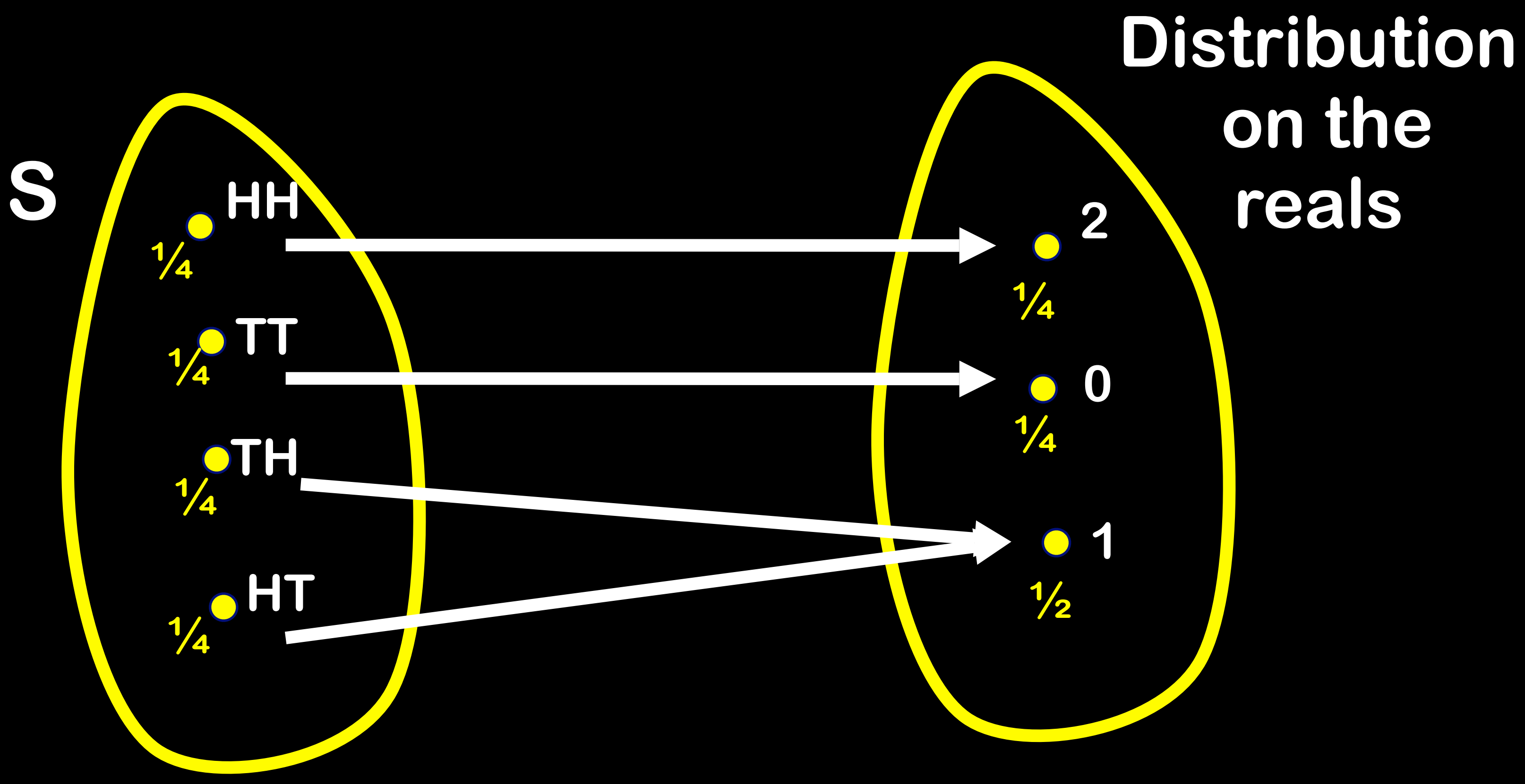
Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads

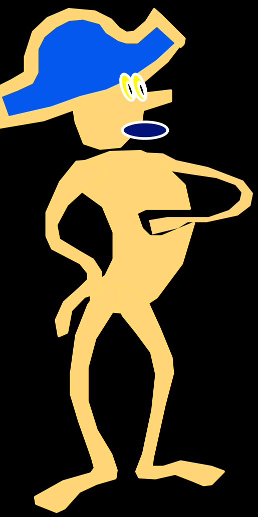


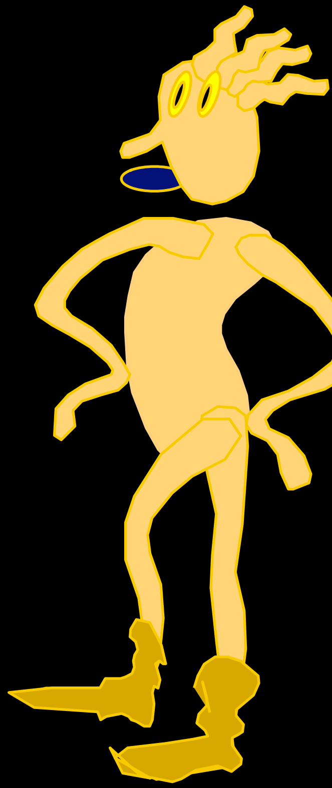
Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads

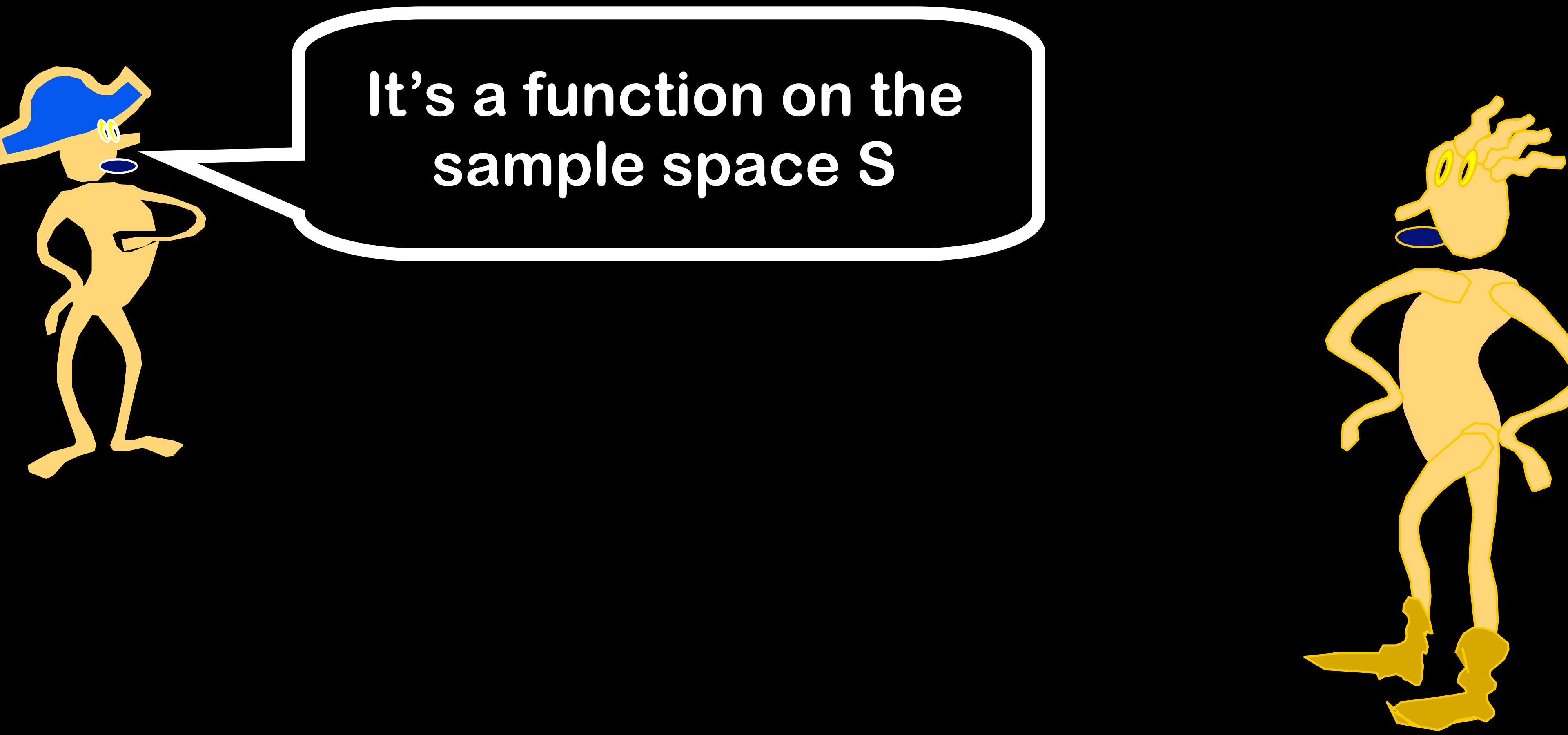


It’s a Floor Wax And a Dessert Topping

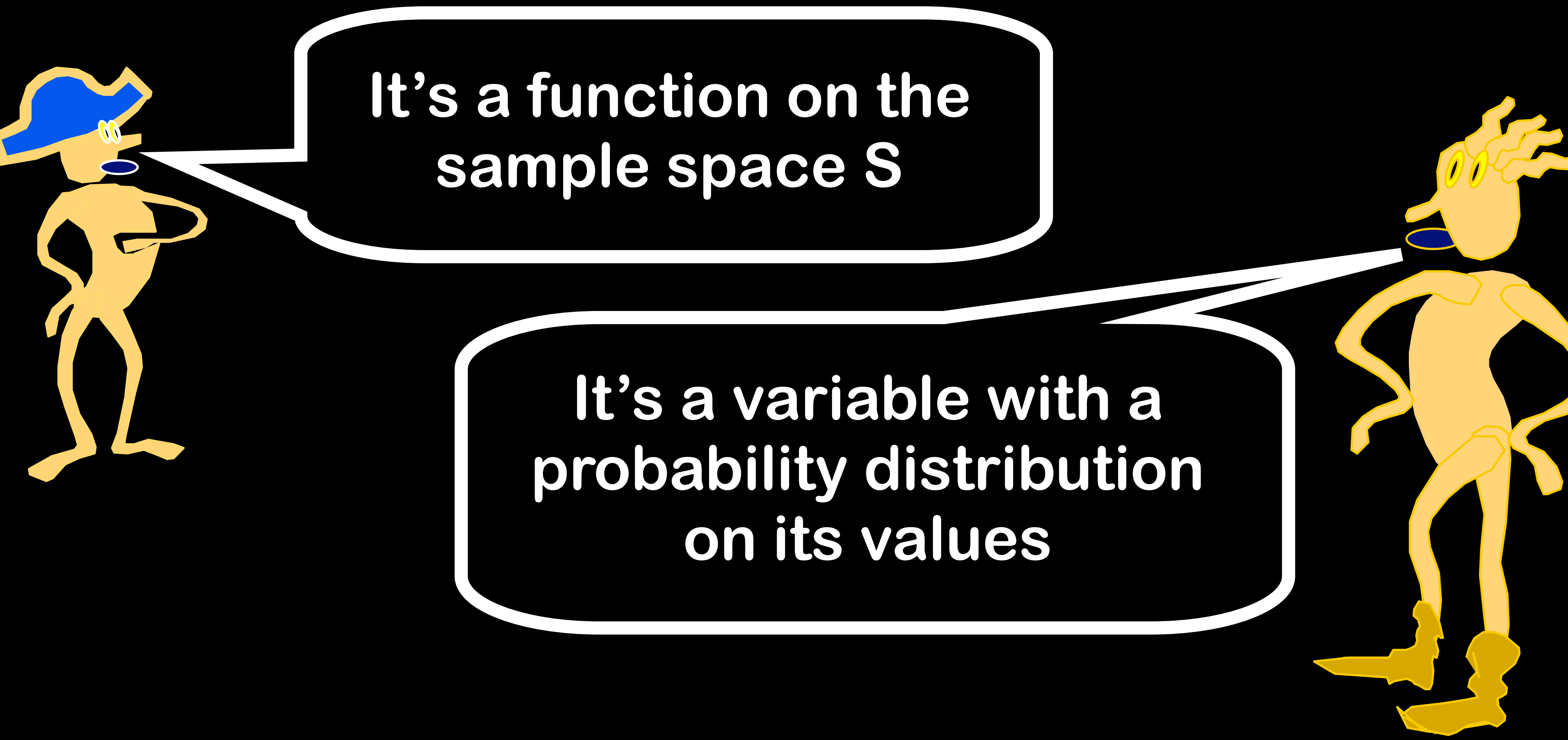




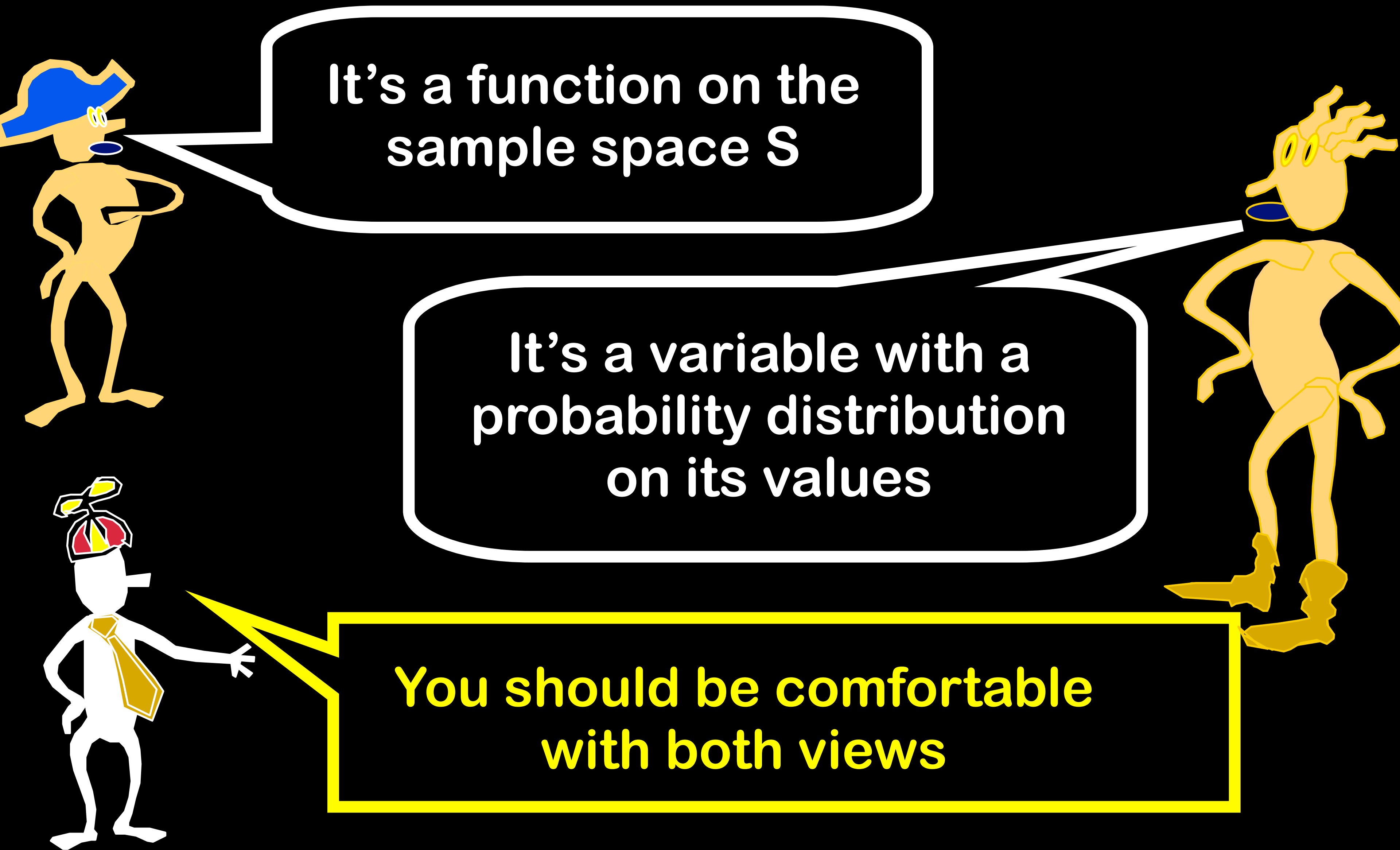
It’s a Floor Wax And a Dessert Topping



It’s a Floor Wax And a Dessert Topping



It’s a Floor Wax And a Dessert Topping



From Random Variables to Events

From Random Variables to Events

For any random variable X and value a, we can define the event A that "$X = a$"

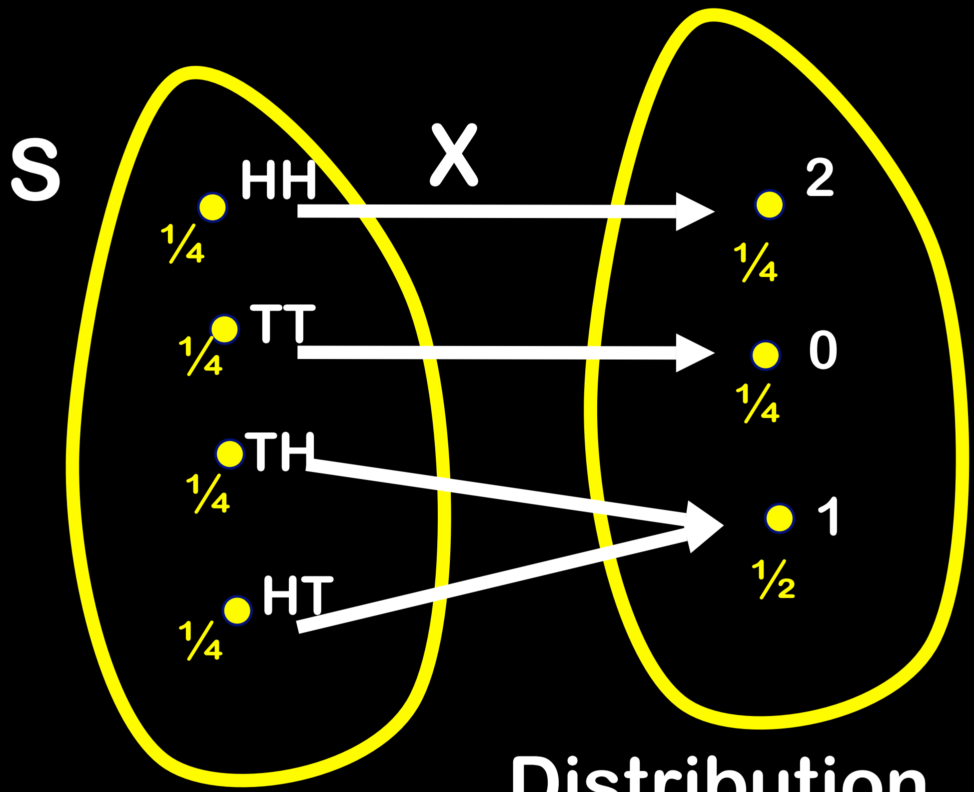
From Random Variables to Events

For any random variable X and value a, we can define the event A that "$X = a$"

$$\Pr(A) = \Pr(X=a) = \Pr(\{x \in S \mid X(x)=a\})$$

Two Coins Tossed

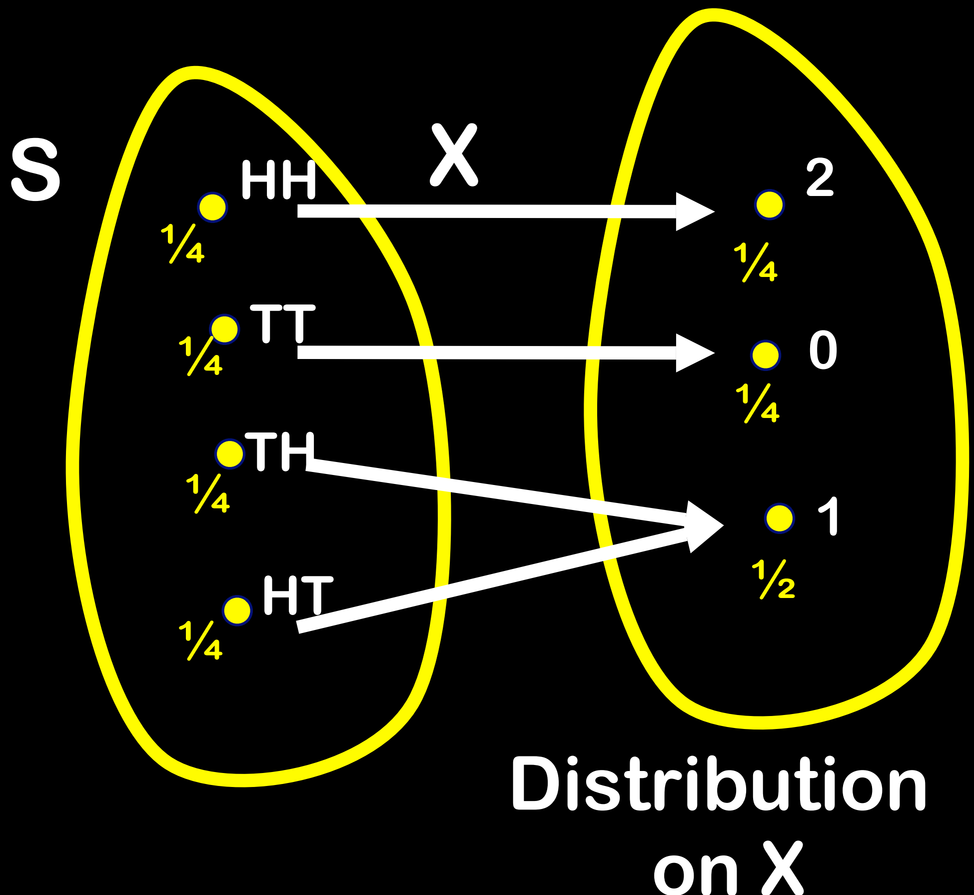
X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads



Distribution on X

Two Coins Tossed

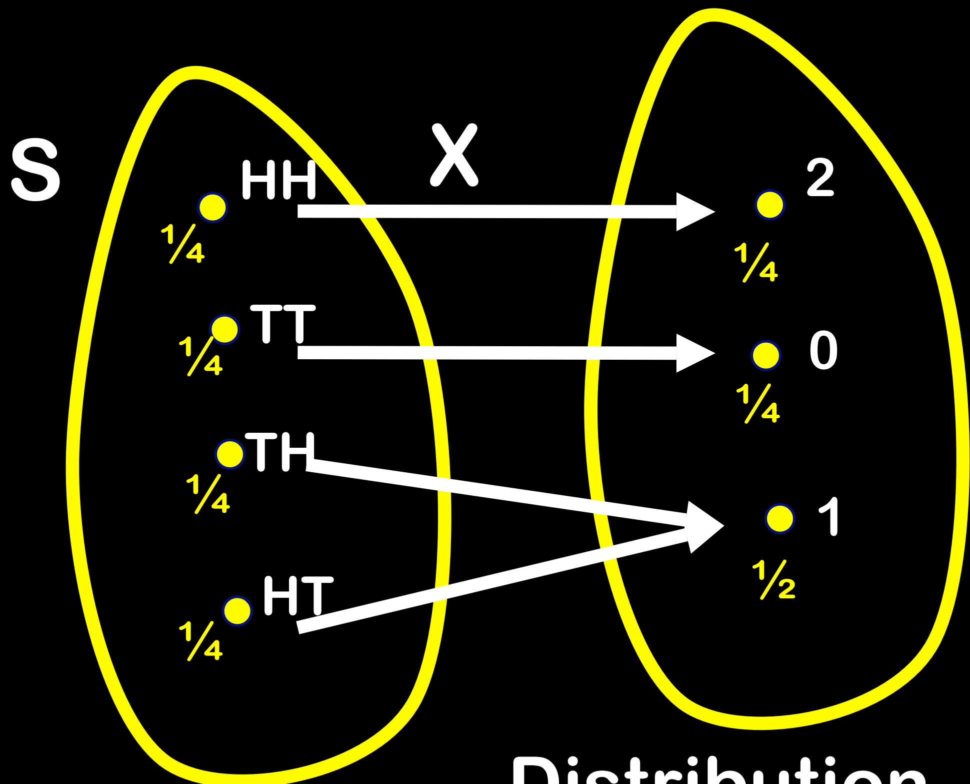
X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads



$$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$$

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads



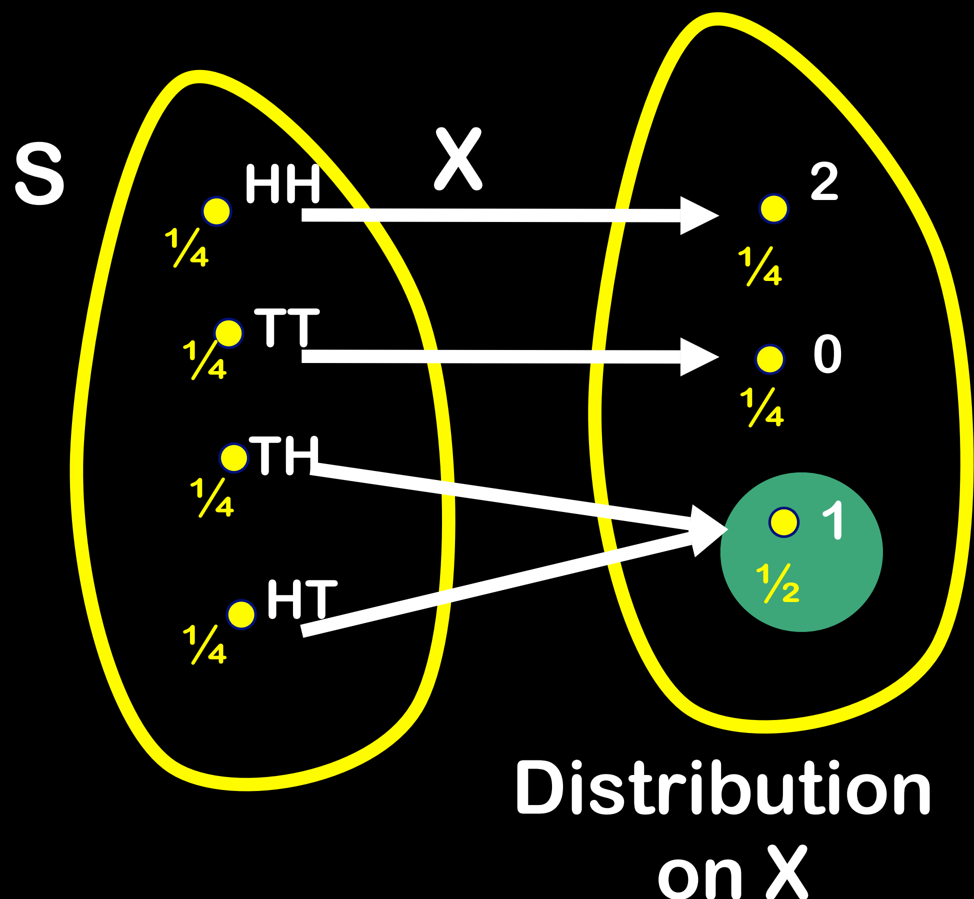
Distribution on X

$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$

$\Pr(X = 1)$

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads

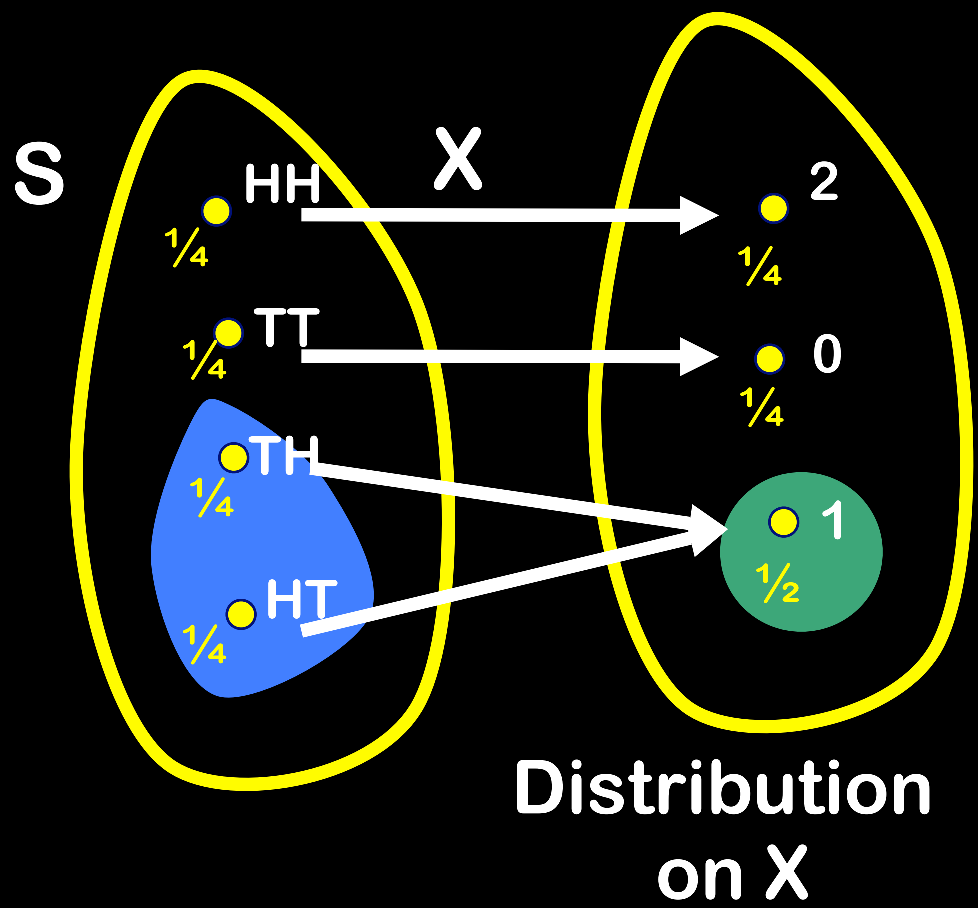


$$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$$

$\Pr(X = 1)$

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads

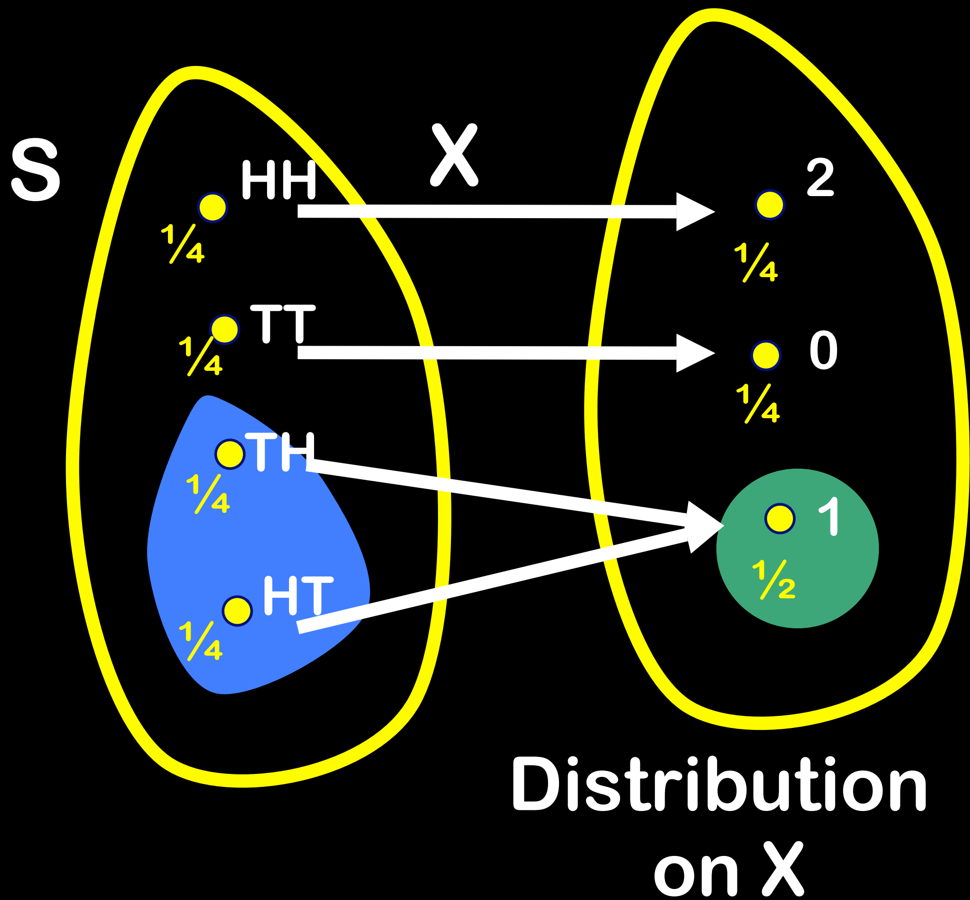


$$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$$

$\Pr(X = 1)$

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads



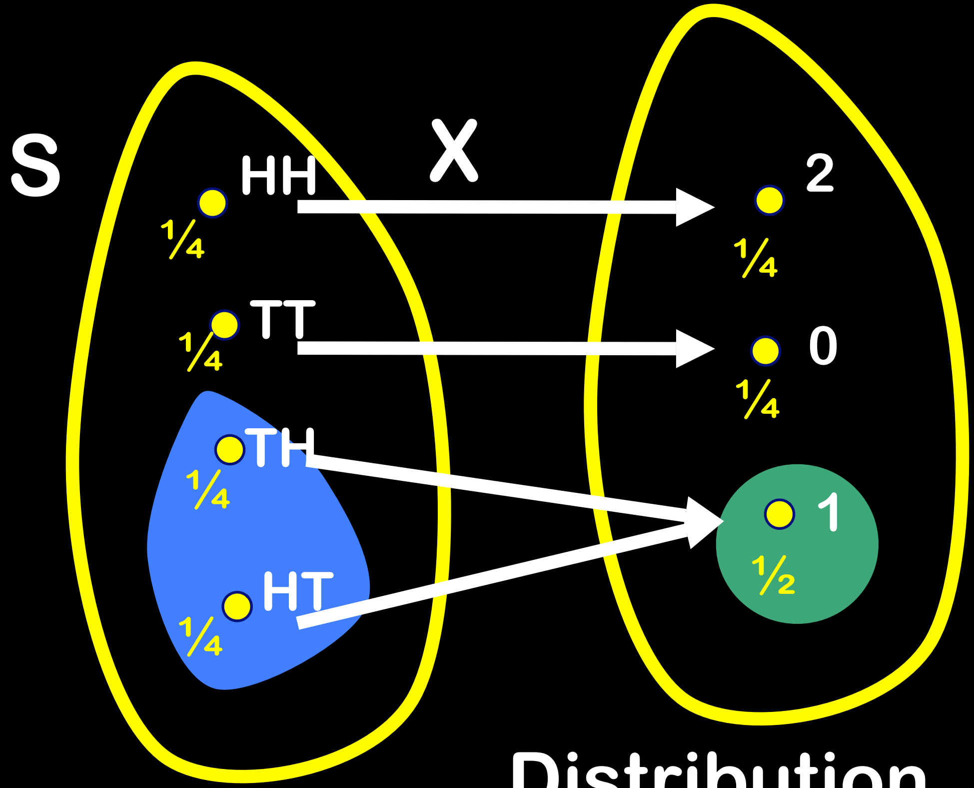
$$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$$

$\Pr(X = 1)$

$= \Pr(\{x \in S \mid X(x) = 1\})$

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts # of heads



Distribution on X

$\Pr(X = a) = \Pr(\{x \in S \mid X(x) = a\})$

$\Pr(X = 1)$

$= \Pr(\{x \in S \mid X(x) = 1\})$

$= \Pr(\{TH, HT\}) = \frac{1}{2}$

From Events to Random Variables

From Events to Random Variables

For any event A, can define the indicator random variable for A:

From Events to Random Variables

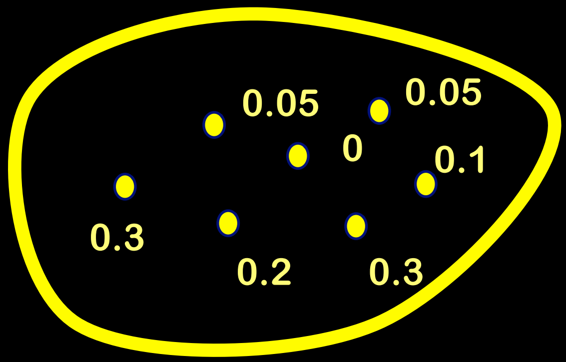
For any event A, can define the indicator random variable for A:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

From Events to Random Variables

For any event A, can define the indicator random variable for A:

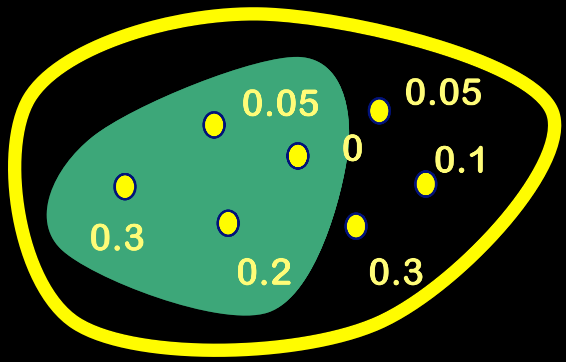
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event **A**, can define the indicator random variable for **A**:

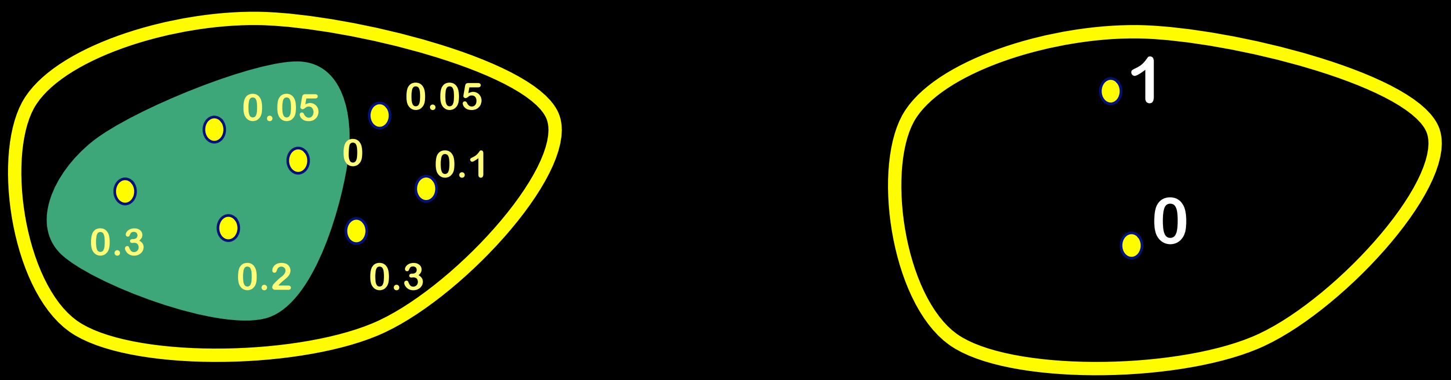
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event A, can define the indicator random variable for A:

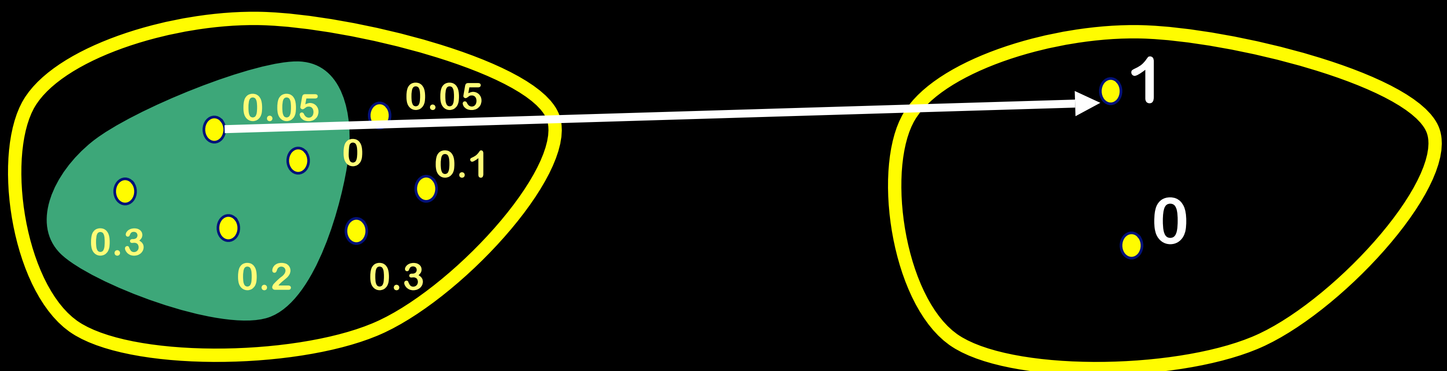
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event A, can define the indicator random variable for A:

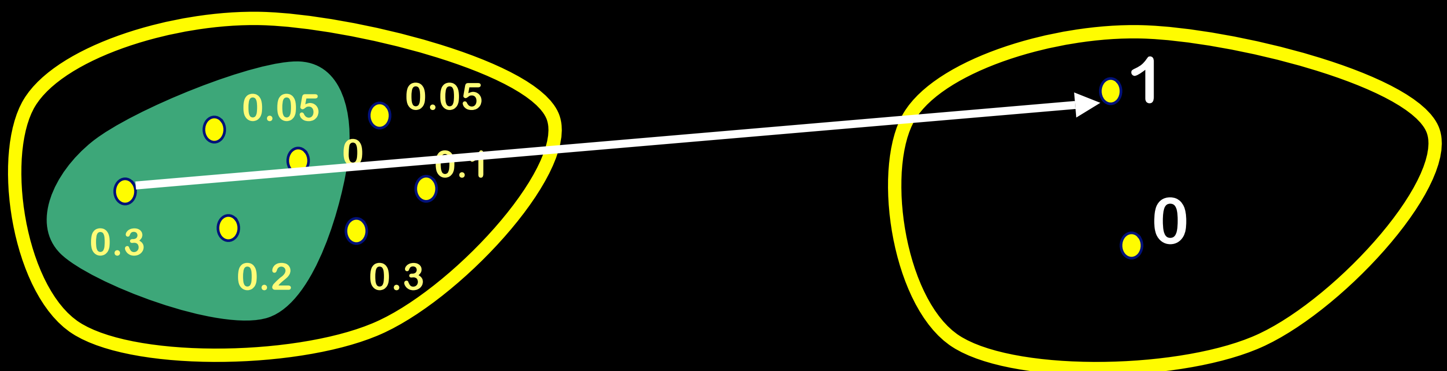
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event A, can define the indicator random variable for A:

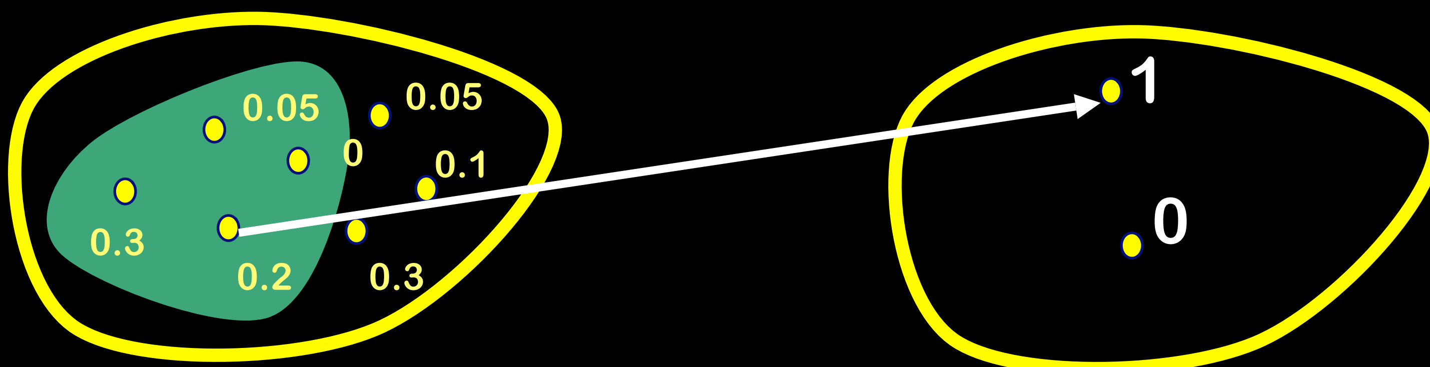
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event A, can define the indicator random variable for A:

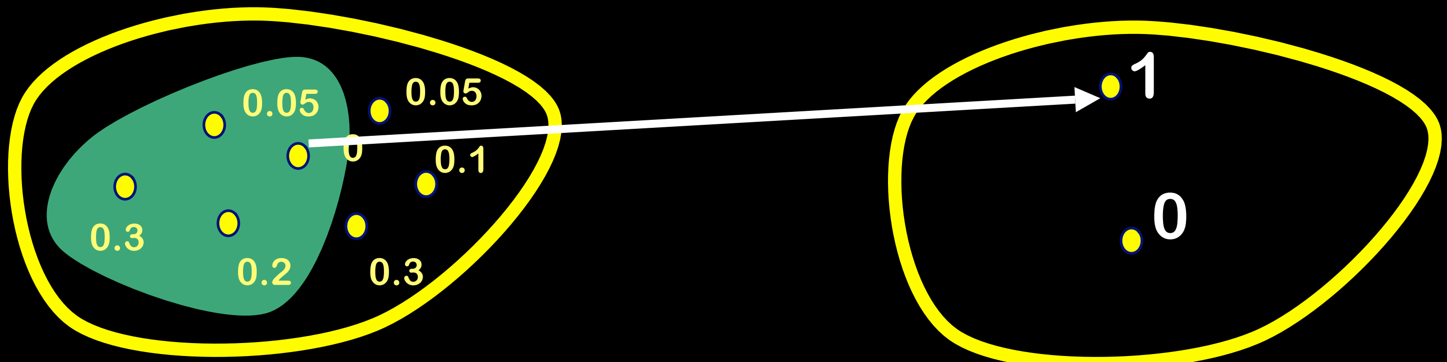
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event A, can define the indicator random variable for A:

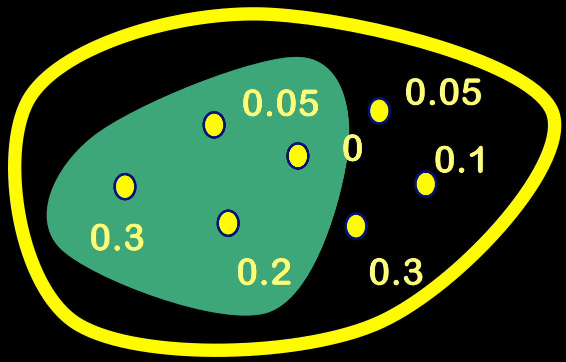
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

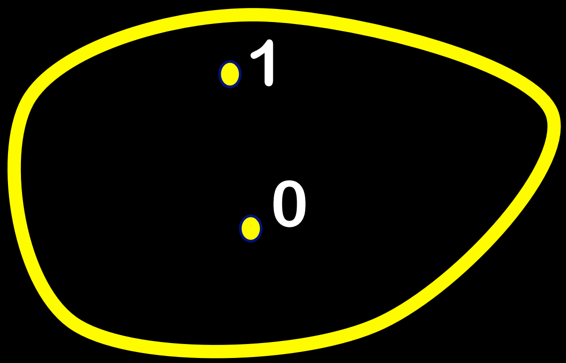


From Events to Random Variables

For any event **A**, can define the indicator random variable for **A**:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

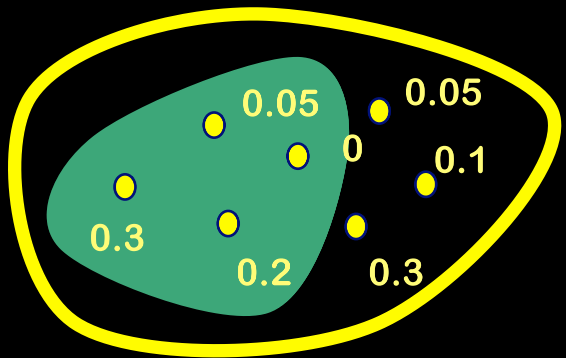


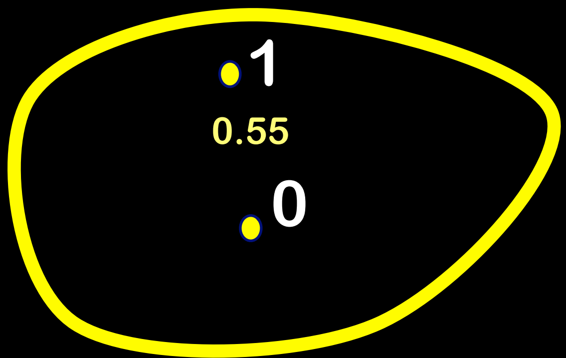


From Events to Random Variables

For any event A, can define the indicator random variable for A:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

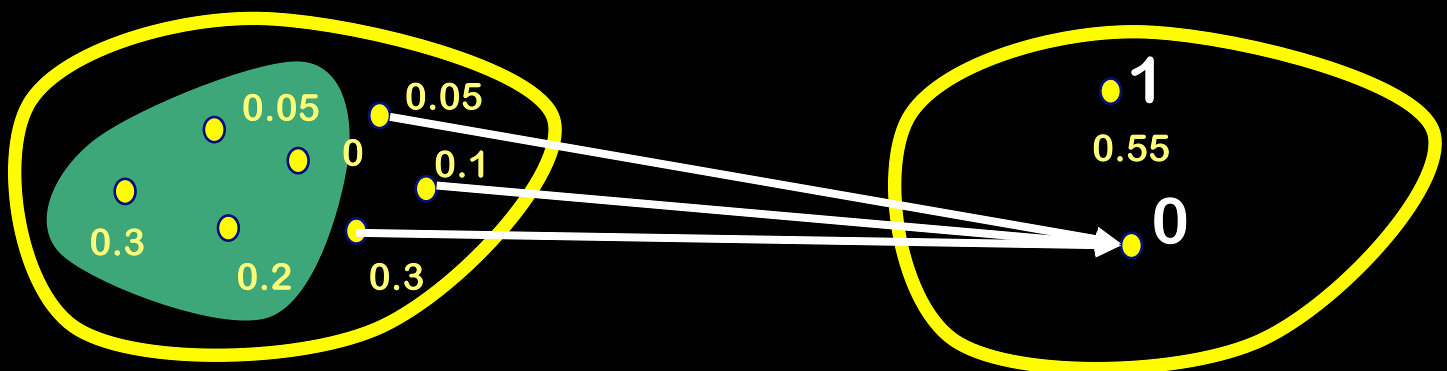




From Events to Random Variables

For any event A, can define the indicator random variable for A:

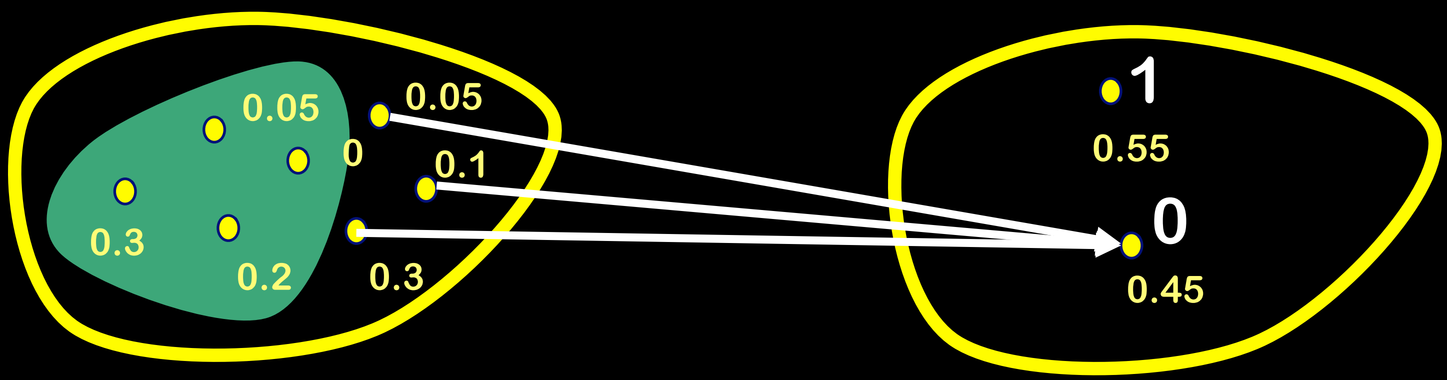
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



From Events to Random Variables

For any event **A**, can define the indicator random variable for **A**:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



Definition: Expectation

Definition: Expectation

The expectation, or expected value of a random variable X is written as E[X], and is

Definition: Expectation

The expectation, or expected value of a random variable X is written as E[X], and is

$$E[X] = \sum_{x \in S} \Pr(x)\, X(x) = \sum_{k} k \Pr[X = k]$$

Definition: Expectation

The expectation, or expected value of a random variable X is written as E[X], and is

$$E[X] = \sum_{x \in S} \Pr(x)\, X(x) = \sum_{k} k \Pr[X = k]$$

X is a function
on the sample space S

Definition: Expectation

The expectation, or expected value of a random variable X is written as E[X], and is

$$E[X] = \sum_{x \in S} \Pr(x)\, X(x) = \sum_{k} k \Pr[X = k]$$

X is a function on the sample space S

X has a distribution on its values

A Quick Calculation…

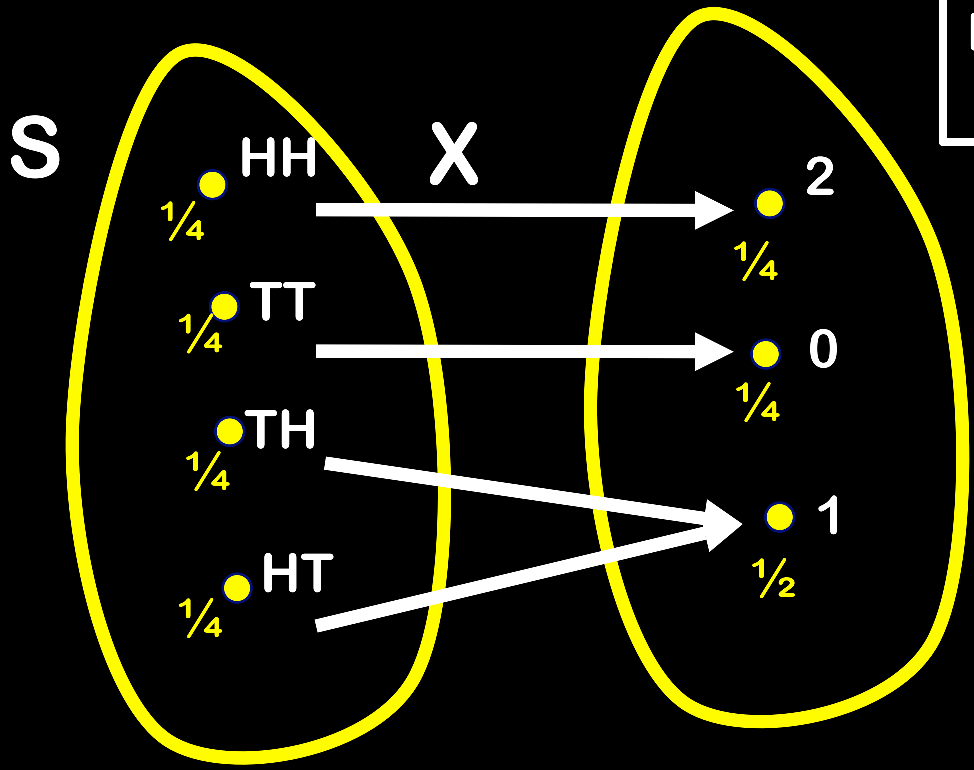
What if I flip a coin 2 times? What is the expected number of heads?

$$\boxed{E[X] = \sum_{x \in S} \Pr(x)\, X(x) = \sum_{k} k \Pr[X = k]}$$

A Quick Calculation…

What if I flip a coin 2 times? What is the expected number of heads?

$$E[X] = \sum_{x \in S} Pr(x)\, X(x) = \sum_{k} k\, Pr[X = k]$$



A Quick Calculation…

What if I flip a coin 3 times? What is the expected number of heads?

A Quick Calculation…

What if I flip a coin 3 times? What is the expected number of heads?

$$E[X] = (1/8)\times 0 + (3/8)\times 1 + (3/8)\times 2 + (1/8)\times 3 = 1.5$$

A Quick Calculation…

What if I flip a coin 3 times? What is the expected number of heads?

$E[X] = (1/8)\times 0 + (3/8)\times 1 + (3/8)\times 2 + (1/8)\times 3 = 1.5$

But $\Pr[\, X = 1.5 \,] = 0$

A Quick Calculation…

What if I flip a coin 3 times? What is the expected number of heads?

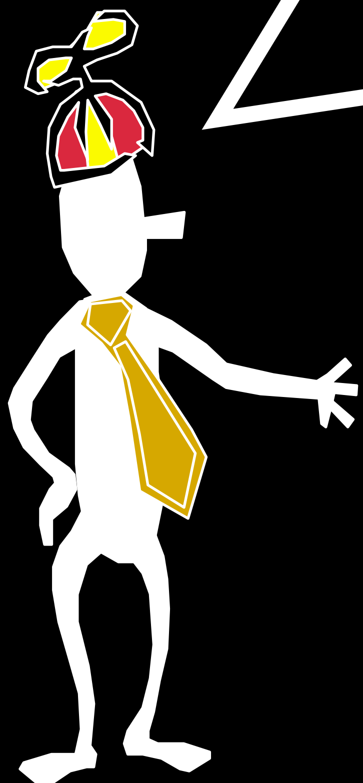
$E[X] = (1/8)\times 0 + (3/8)\times 1 + (3/8)\times 2 + (1/8)\times 3 = 1.5$

But $\Pr[\, X = 1.5 \,] = 0$

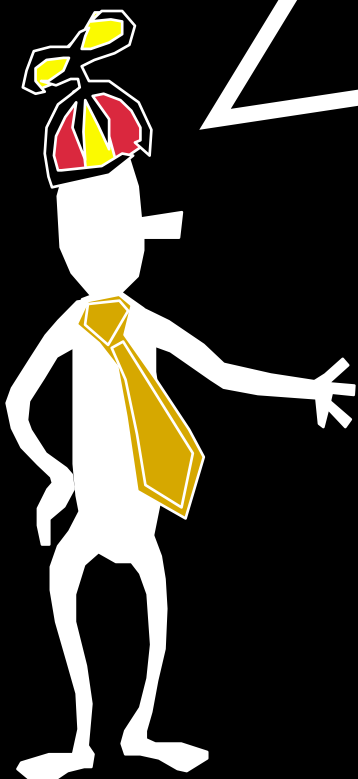
Moral: don't always expect the expected.
$\Pr[\, X = E[X] \,]$ may be 0 !

Type Checking

You can average the values of a random variable.



Type Checking



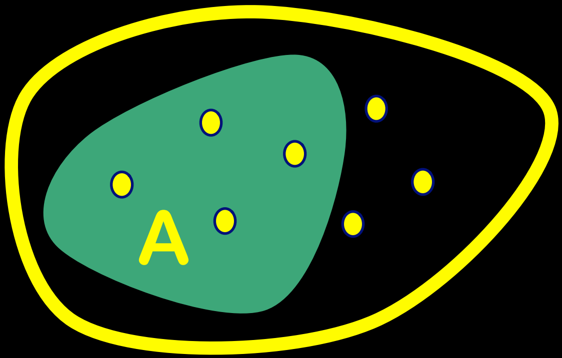
You can average the values of a random variable.

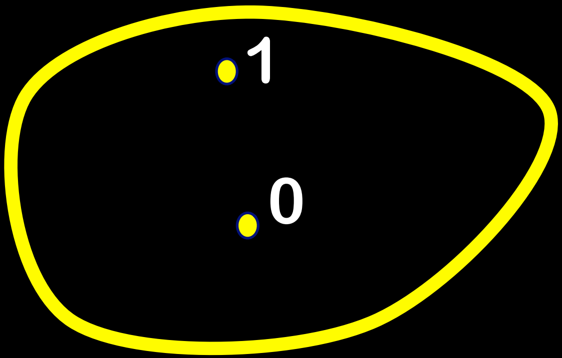
If you are computing an expectation, the thing whose expectation you are computing is a random variable

Indicator R.V.s: $E[X_A] = \Pr(A)$

For any event A, can define the indicator random variable for A:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

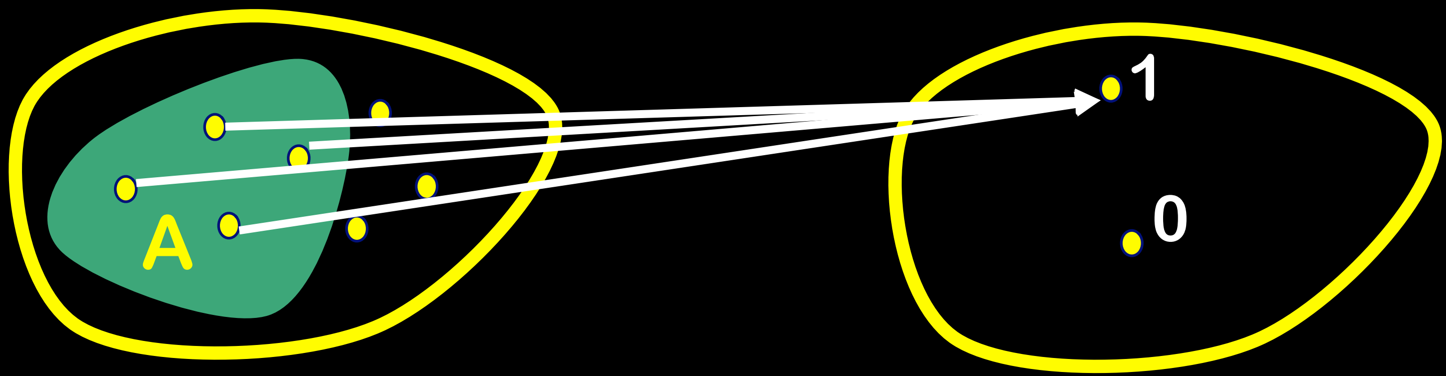




Indicator R.V.s: $E[X_A] = Pr(A)$

For any event A, can define the indicator random variable for A:

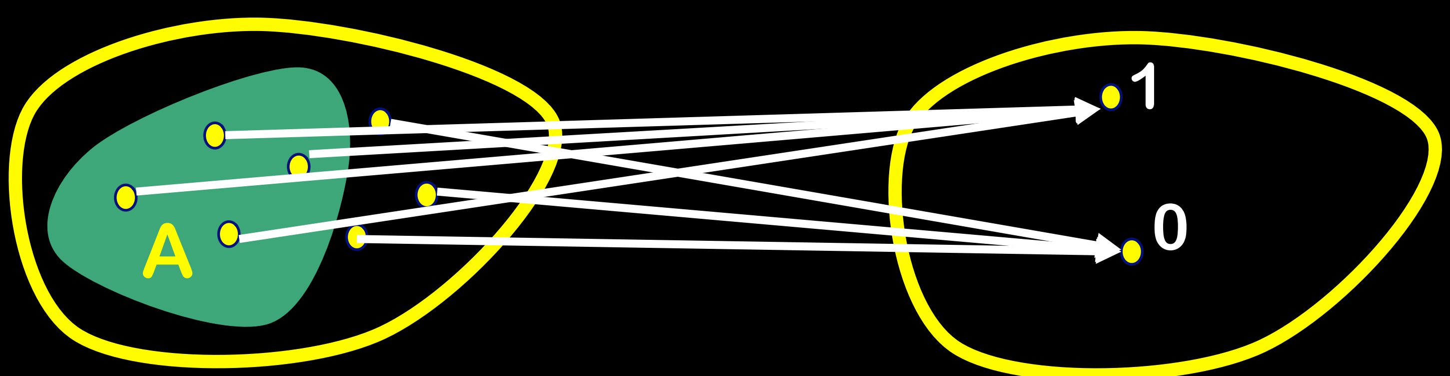
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



Indicator R.V.s: $E[X_A] = Pr(A)$

For any event A, can define the indicator random variable for A:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

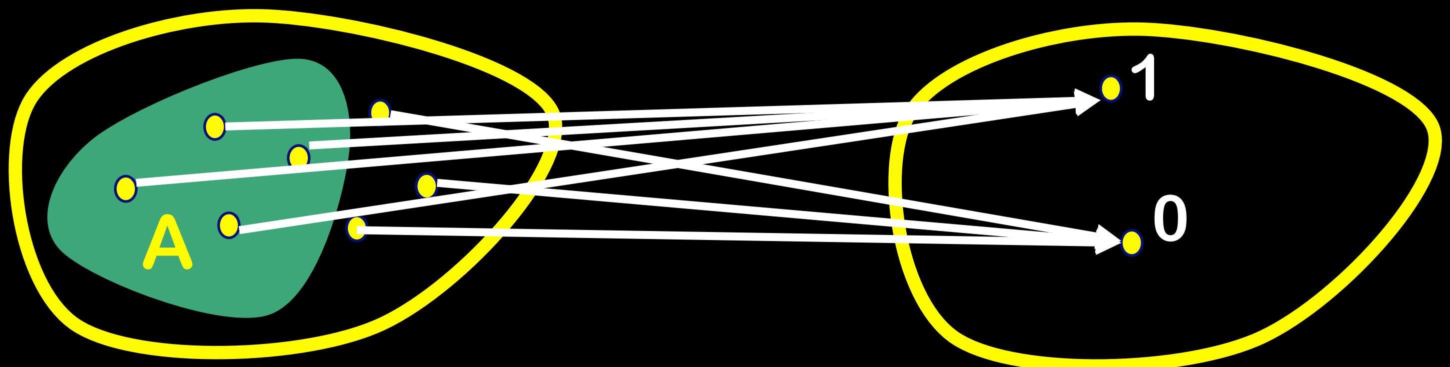


Indicator R.V.s: $E[X_A] = \Pr(A)$

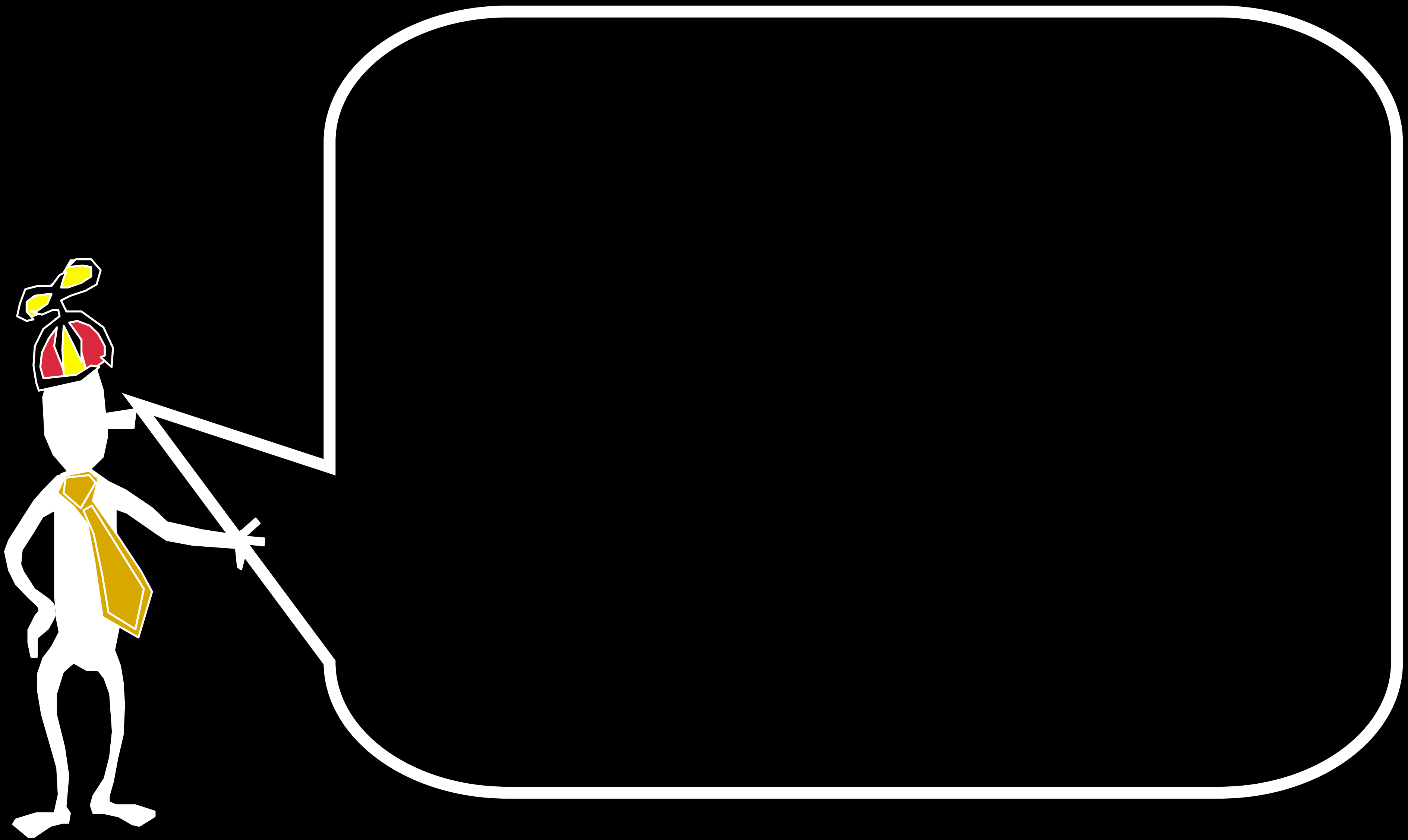
For any event A, can define the indicator random variable for A:

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

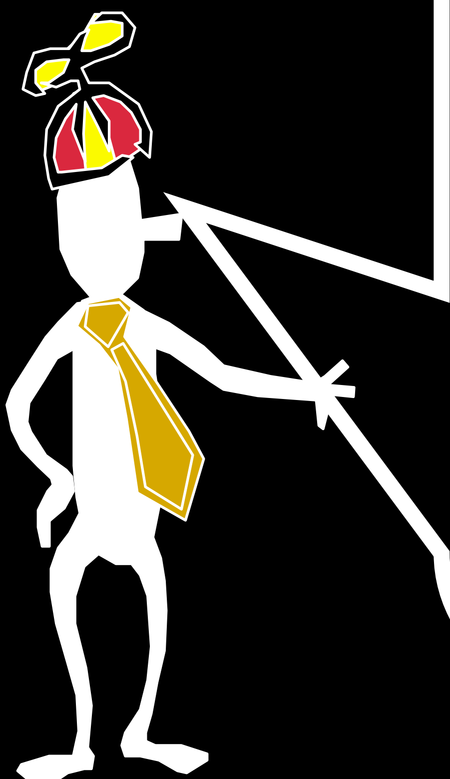
$$\boxed{E[X_A] = 1 \times \Pr(X_A = 1) = \Pr(A)}$$



Adding Random Variables



Adding Random Variables



If X and Y are random variables (on the same set S), then $Z = X + Y$ is also a random variable

Adding Random Variables

If X and Y are random variables (on the same set S), then $Z = X + Y$ is also a random variable

$$Z(x) = X(x) + Y(x)$$

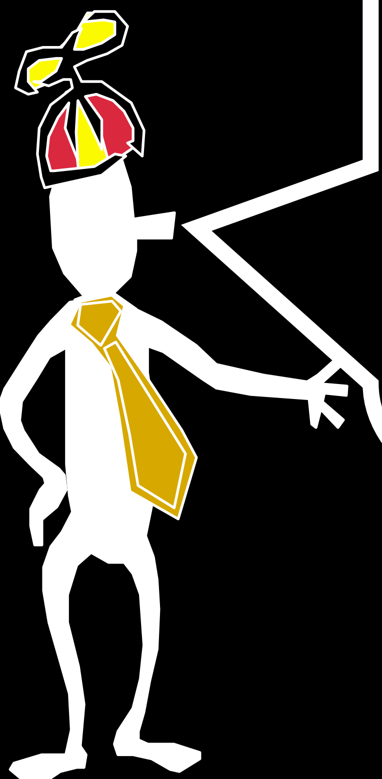
Adding Random Variables

If X and Y are random variables (on the same set S), then $Z = X + Y$ is also a random variable

$$Z(x) = X(x) + Y(x)$$

E.g., rolling two dice.
X = 1st die, Y = 2nd die,
Z = sum of two dice

Adding Random Variables



Example: Consider picking a random person in the world. Let X = length of the person's left arm in inches. Y = length of the person's right arm in inches. Let **Z = X+Y**. **Z** measures the combined arm lengths

Independence

Two random variables X and Y are independent if for every a,b, the events X=a and Y=b are independent

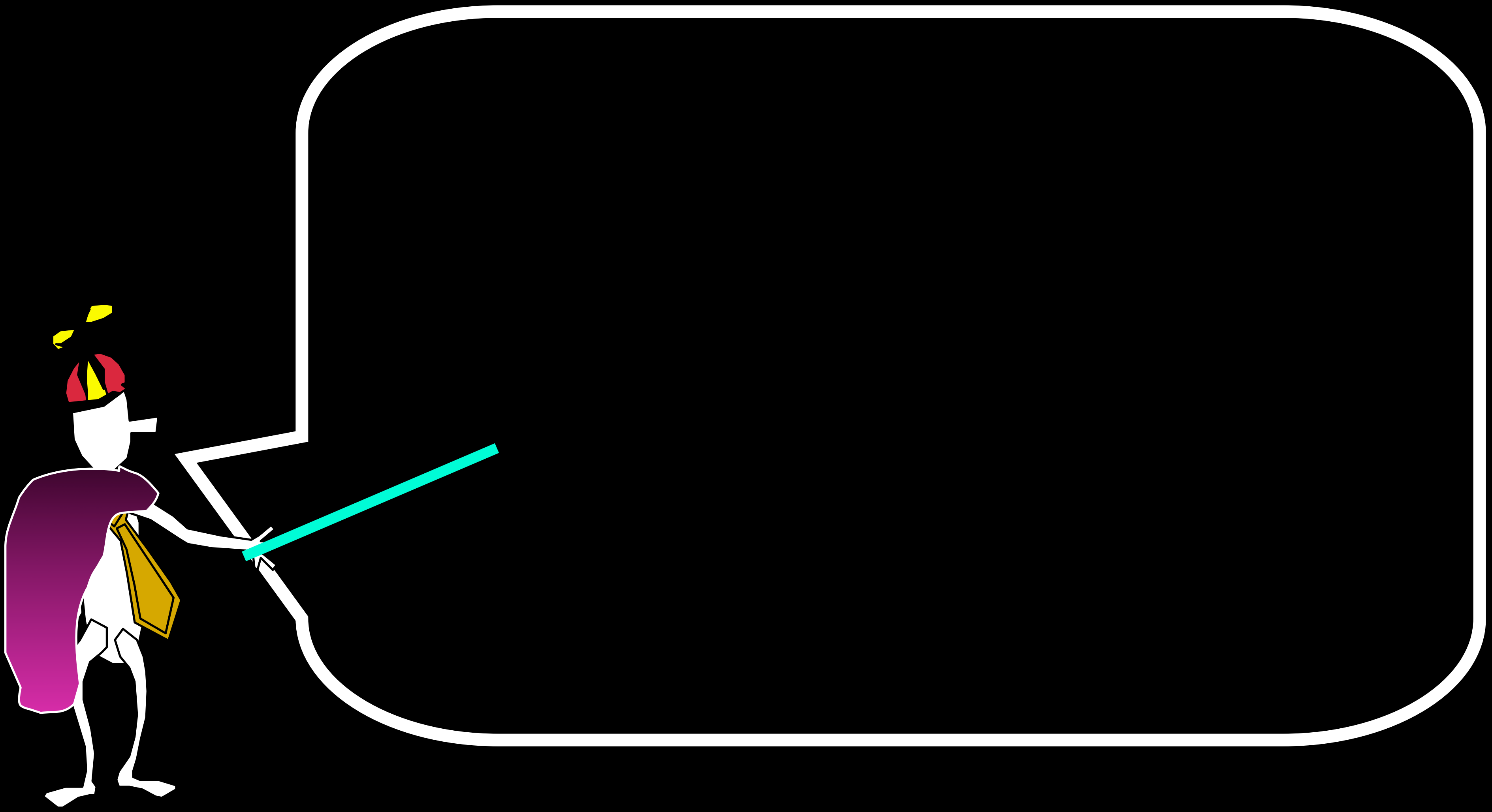
Independence

Two random variables X and Y are independent if for every a,b, the events X=a and Y=b are independent

How about the case of
X=1st die, Y=2nd die?
X = left arm, Y=right arm?

Linearity of Expectation

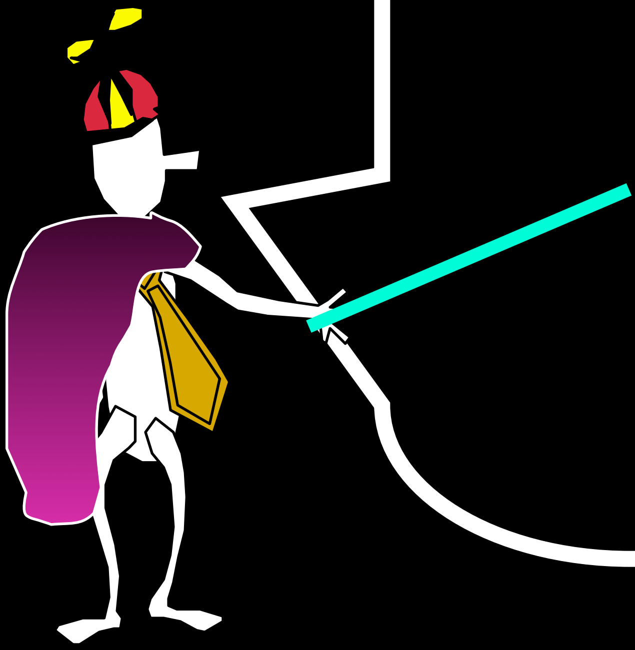
Expectatus linearis



Linearity of Expectation

Expectatus linearis

If $Z = X+Y$, then

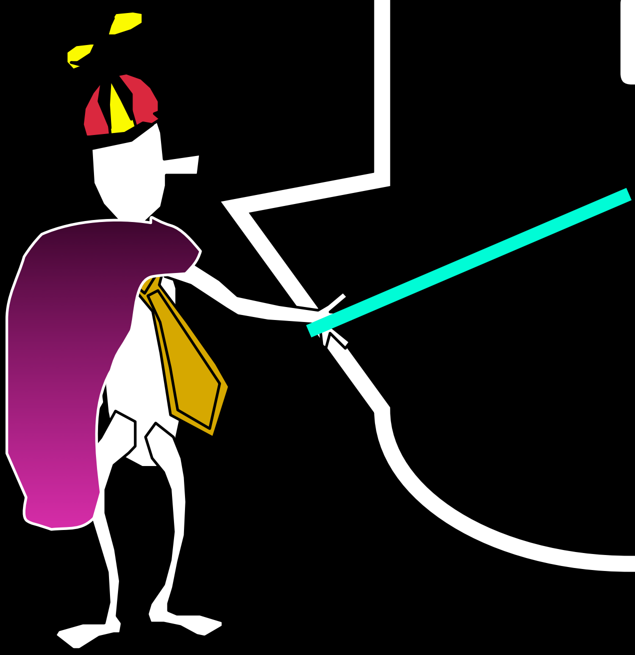


Linearity of Expectation

Expectatus linearis

If $Z = X+Y$, then

$$E[Z] = E[X] + E[Y]$$



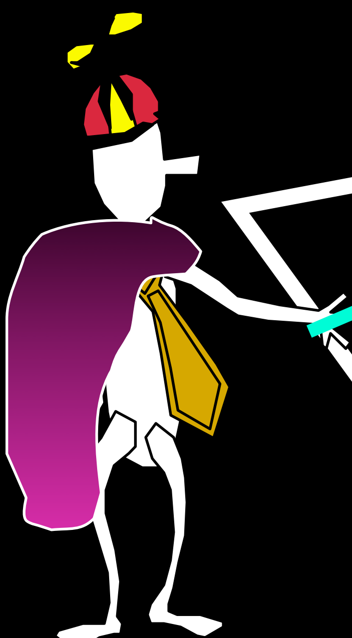
Linearity of Expectation

Expectatus linearis

If $Z = X+Y$, then

$$E[Z] = E[X] + E[Y]$$

Even if X and Y are not independent!



$$E[Z] = \sum_{x \in S} Pr[x]\, Z(x)$$

$$E[Z] = \sum_{x \in S} \Pr[x]\, Z(x)$$

$$= \sum_{x \in S} \Pr[x]\, (X(x) + Y(x))$$

$$E[Z] = \sum_{x \in S} \Pr[x]\, Z(x)$$

$$= \sum_{x \in S} \Pr[x]\, (X(x) + Y(x))$$

$$= \sum_{x \in S} \Pr[x]\, X(x) + \sum_{x \in S} \Pr[x]\, Y(x))$$

$$
\begin{aligned}
E[Z] &= \sum_{x \in S} \Pr[x]\, Z(x) \\
&= \sum_{x \in S} \Pr[x]\, (X(x) + Y(x)) \\
&= \sum_{x \in S} \Pr[x]\, X(x) + \sum_{x \in S} \Pr[x]\, Y(x)) \\
&= E[X] + E[Y]
\end{aligned}
$$

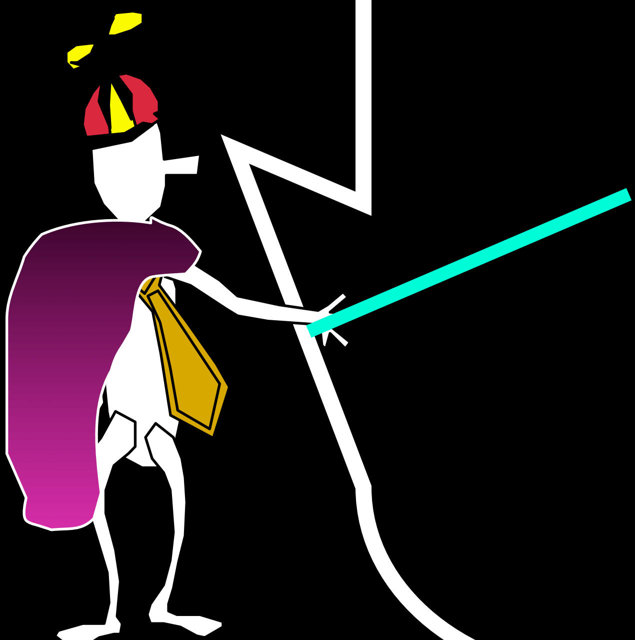
Linearity of Expectation

E.g., 2 fair flips:

X = 1st coin is heads,

Y = 2nd coin is heads.

$Z = X+Y$ = total # heads



Linearity of Expectation

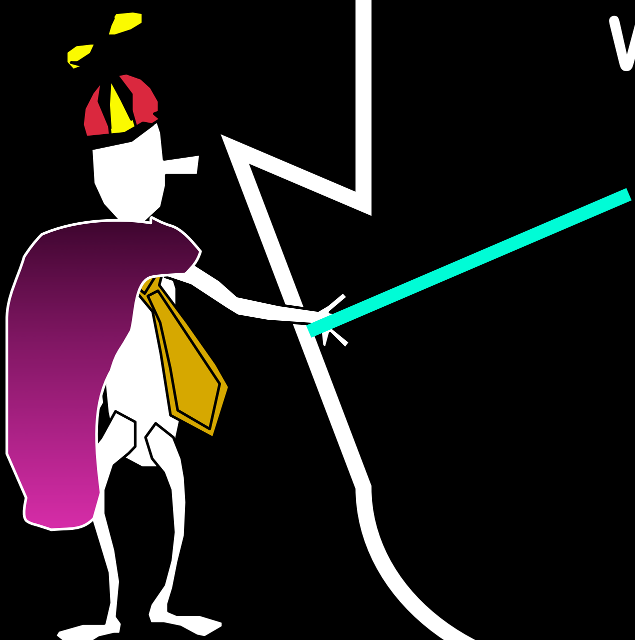
E.g., 2 fair flips:

X = 1st coin is heads,

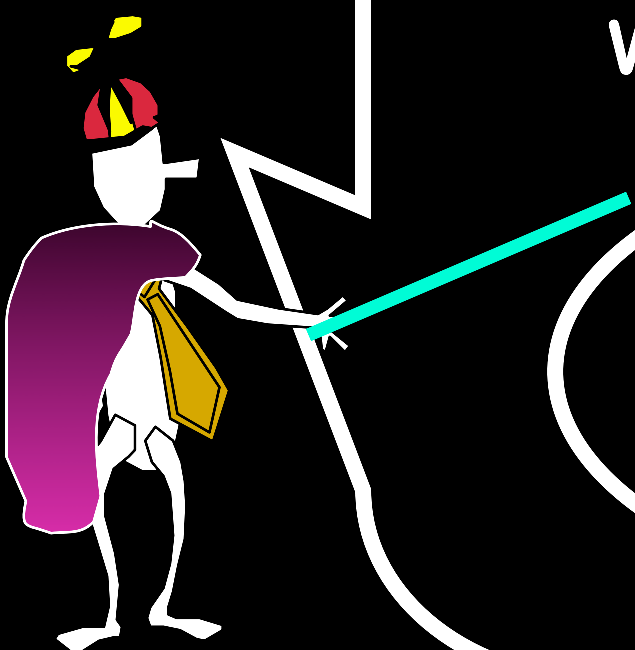
Y = 2nd coin is heads.

Z = X+Y = total # heads

What is E[X]? E[Y]? E[Z]?



Linearity of Expectation



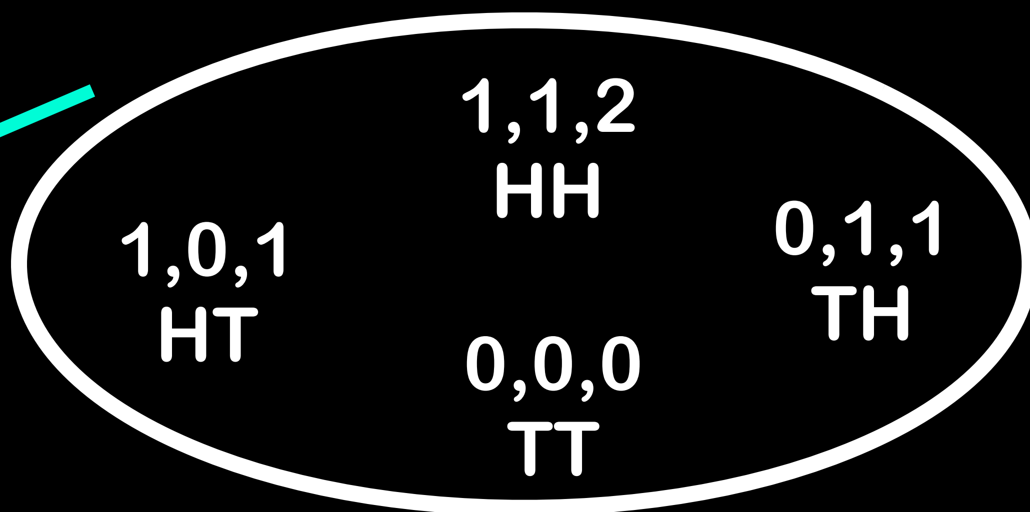
E.g., 2 fair flips:

X = 1st coin is heads,

Y = 2nd coin is heads.

Z = X+Y = total # heads

What is E[X]? E[Y]? E[Z]?

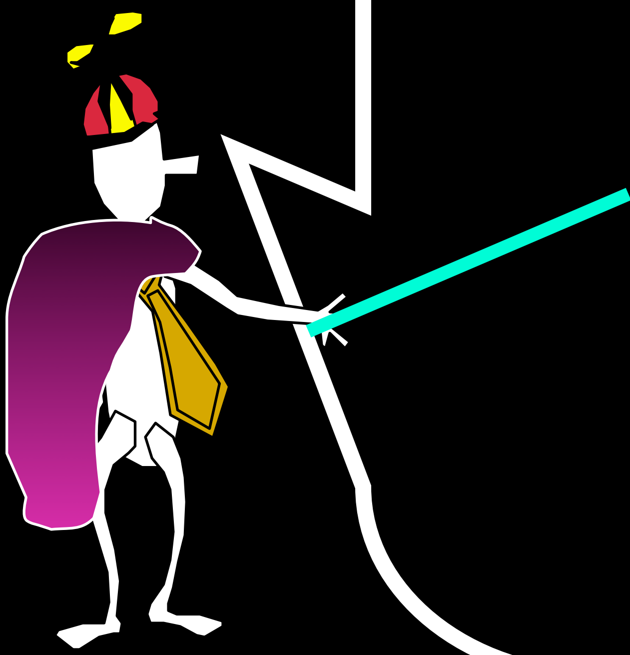


Linearity of Expectation

E.g., 2 fair flips:

X = at least one coin is heads

Y = both coins are heads, Z = X+Y



Linearity of Expectation

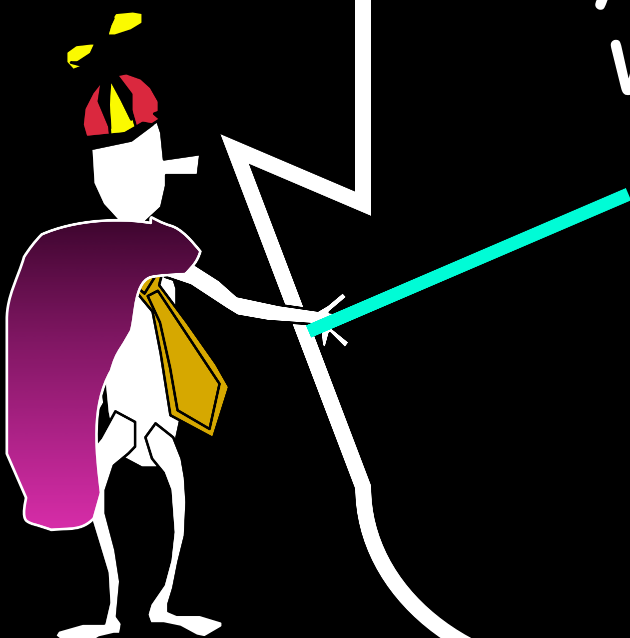
E.g., 2 fair flips:

X = at least one coin is heads

Y = both coins are heads, Z = X+Y

Are X and Y independent?

What is E[X]? E[Y]? E[Z]?



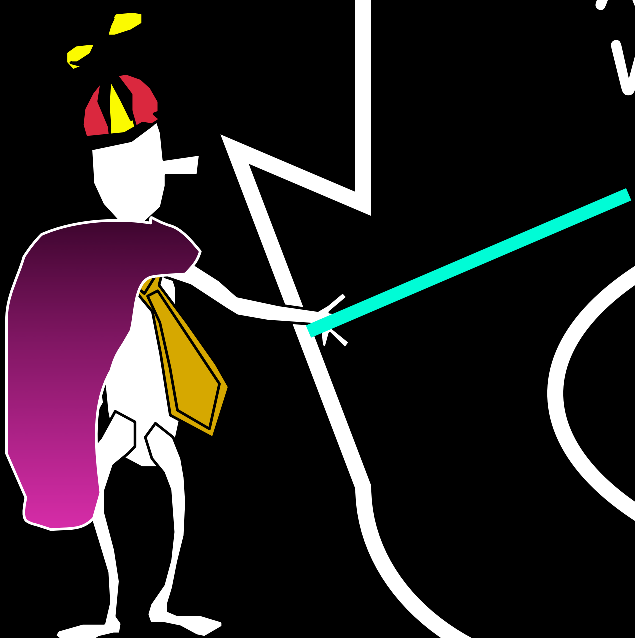
Linearity of Expectation

E.g., 2 fair flips:

X = at least one coin is heads

Y = both coins are heads, Z = X+Y

Are X and Y independent?
What is E[X]? E[Y]? E[Z]?



1,1,2
HH

1,0,1
HT

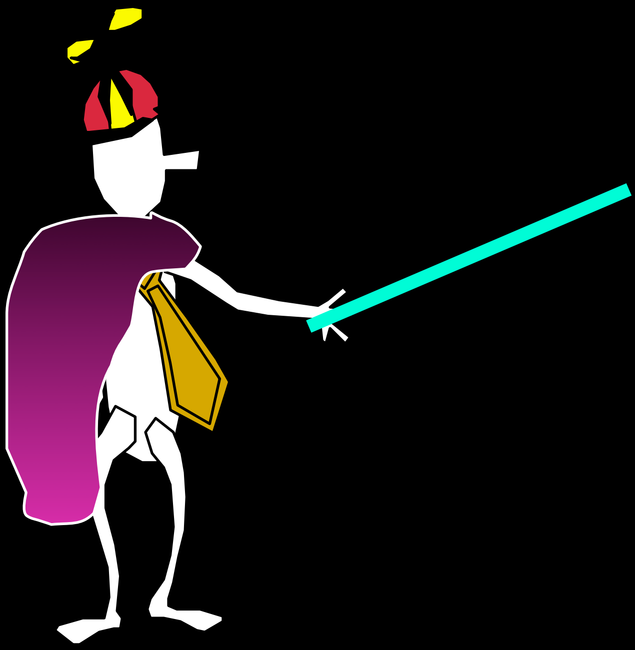
1,0,1
TH

0,0,0
TT

By Induction

$$E[X_1 + X_2 + \ldots + X_n] =$$

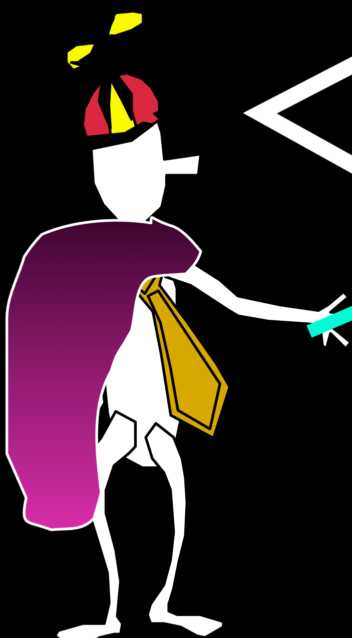
$$E[X_1] + E[X_2] + \ldots. + E[X_n]$$



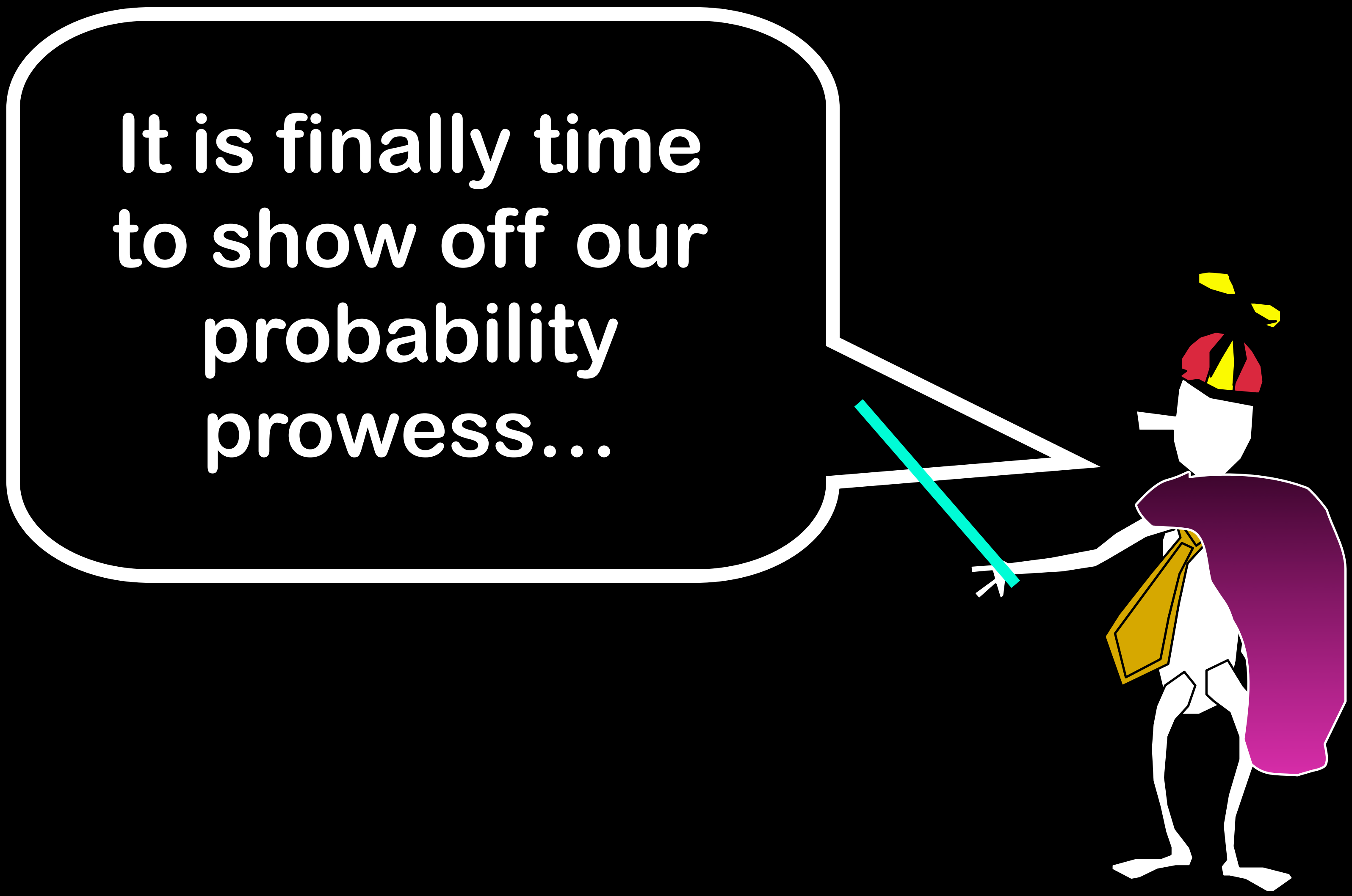
By Induction

$$E[X_1 + X_2 + \ldots + X_n] =$$

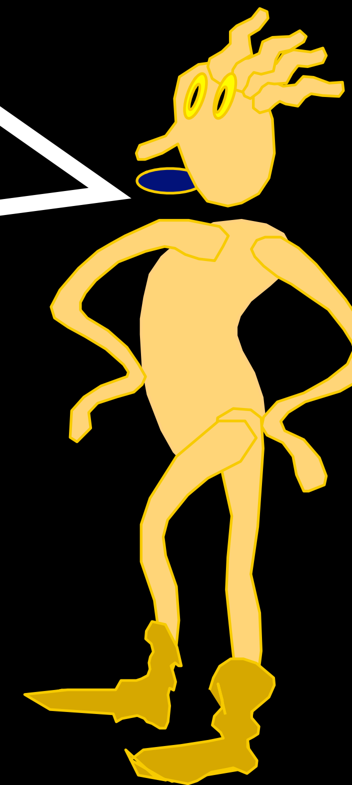
$$E[X_1] + E[X_2] + \ldots. + E[X_n]$$



The expectation
of the sum
=
The sum of the
expectations


It is finally time to show off our probability prowess...

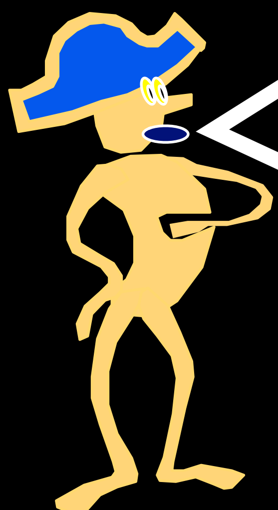
If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?



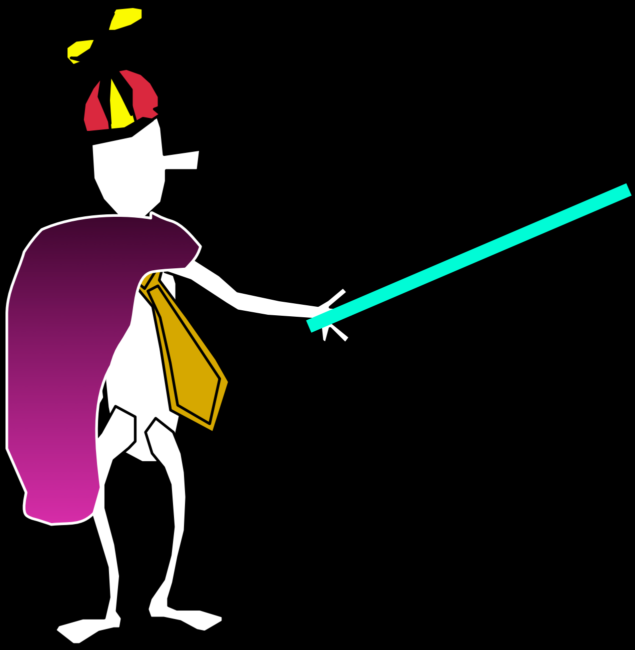
Hmm…

$\sum_k k \Pr(k$ letters end up in correct envelopes$)$

$= \sum_k k$ (…aargh!!…)

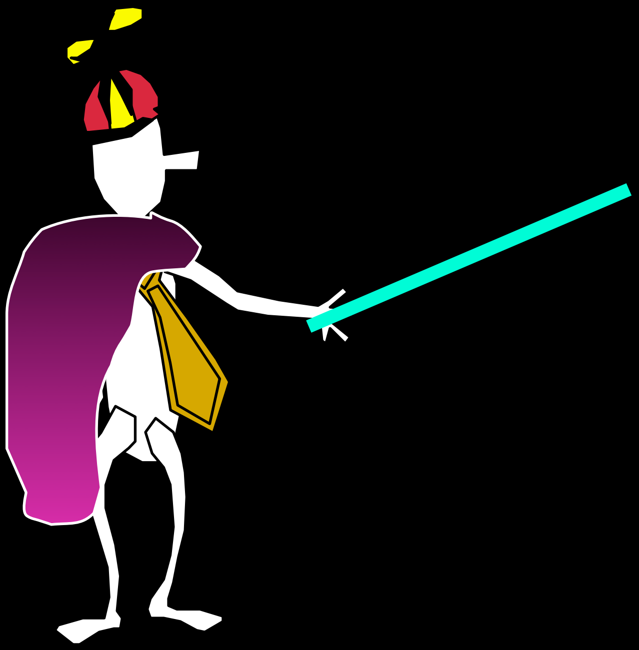


Use Linearity of Expectation



Use Linearity of Expectation

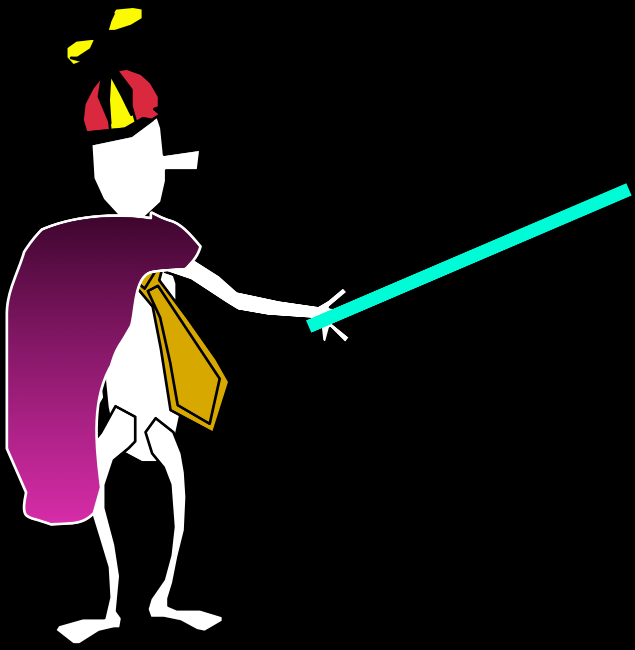
Let A_i be the event the i^{th} letter ends up in its correct envelope



Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

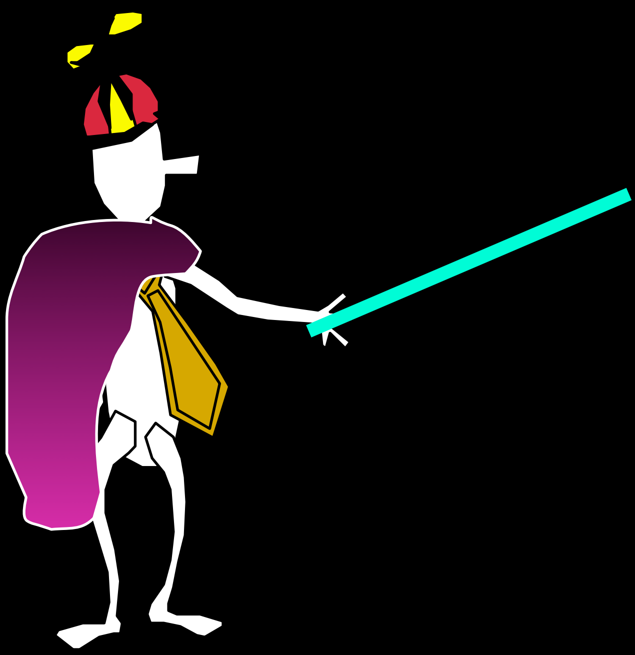


Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

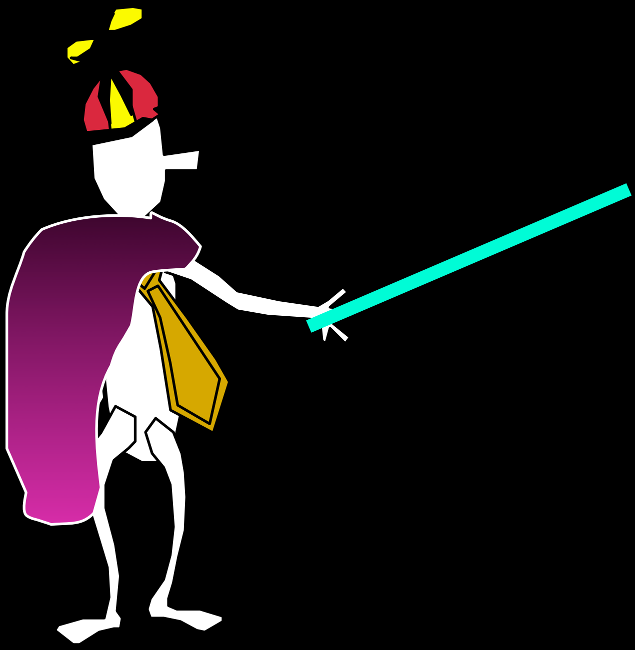


Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$



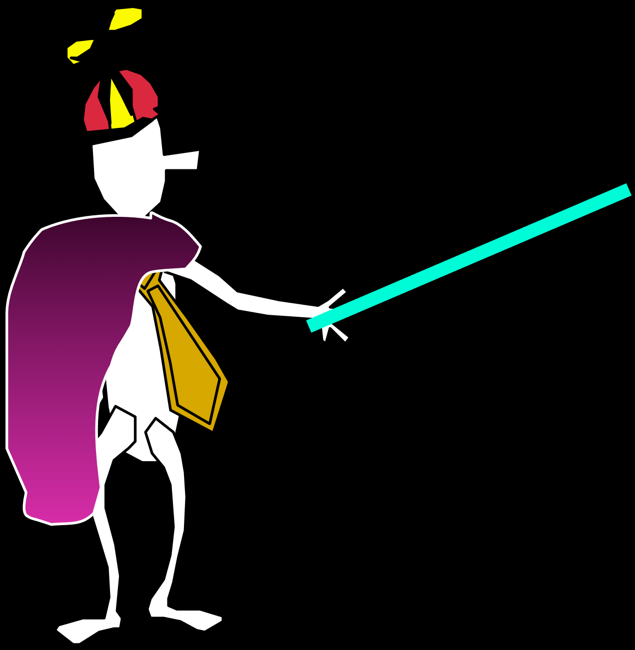
Let $Z = X_1 + \ldots + X_{100}$

Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$



Let $Z = X_1 + \ldots + X_{100}$

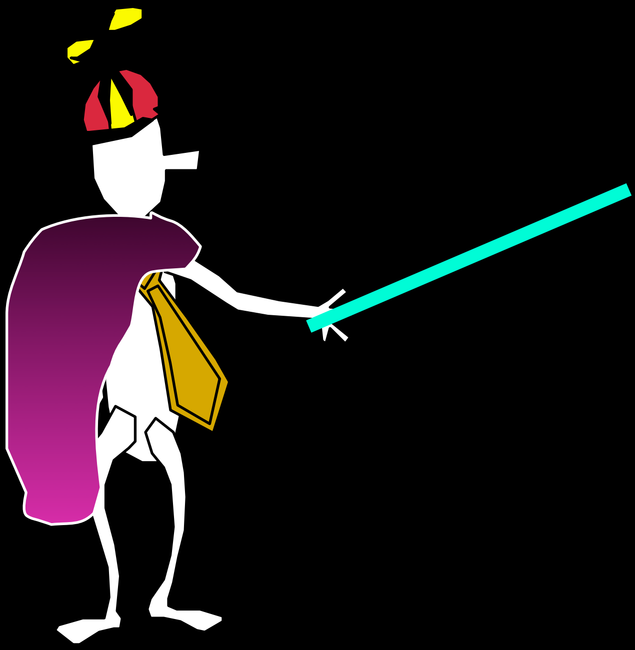
We are asking for **E[Z]**

Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$



Let $Z = X_1 + \ldots + X_{100}$

We are asking for $E[Z]$

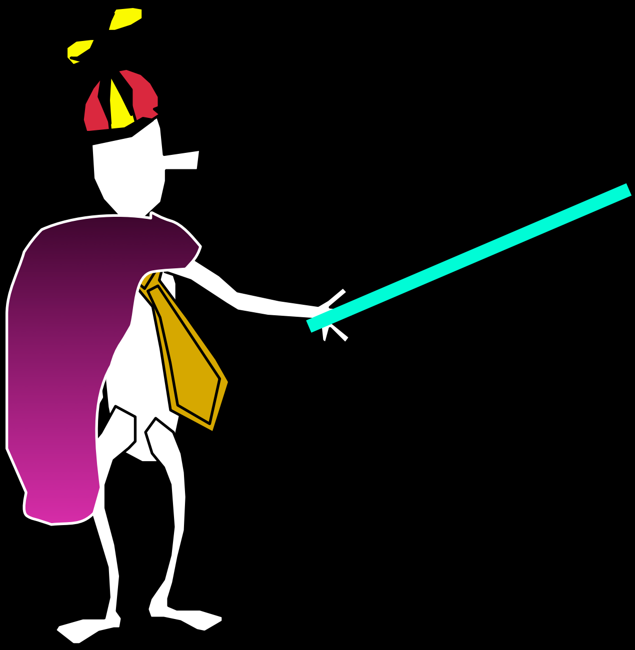
$E[X_i] = \Pr(A_i) = 1/100$

Use Linearity of Expectation

Let A_i be the event the i^{th} letter ends up in its correct envelope

Let X_i be the indicator R.V. for A_i

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$



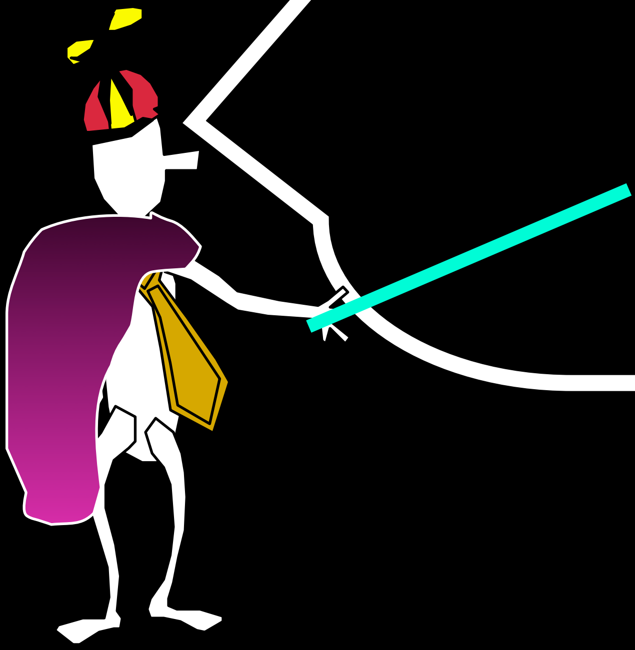
Let $Z = X_1 + \ldots + X_{100}$

We are asking for **E[Z]**

$E[X_i] = \Pr(A_i) = 1/100$

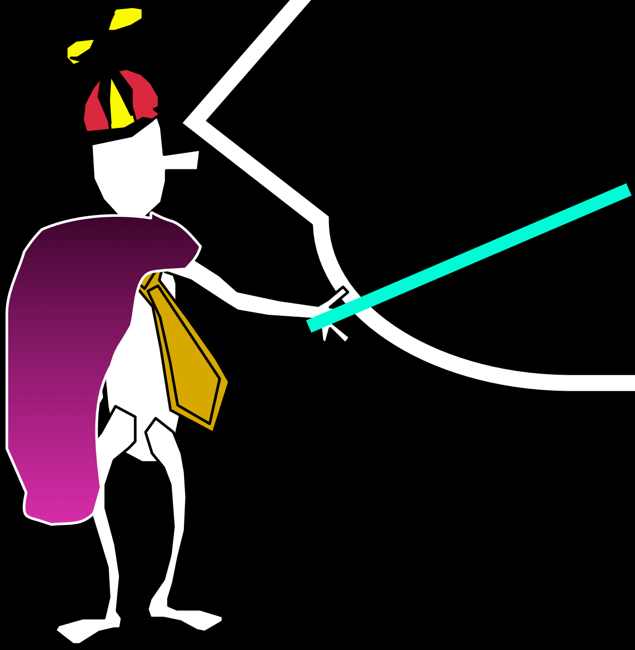
So $E[Z] = 1$

So, in expectation, 1 letter will be in the correct envelope



So, in expectation, 1 letter will be in the correct envelope

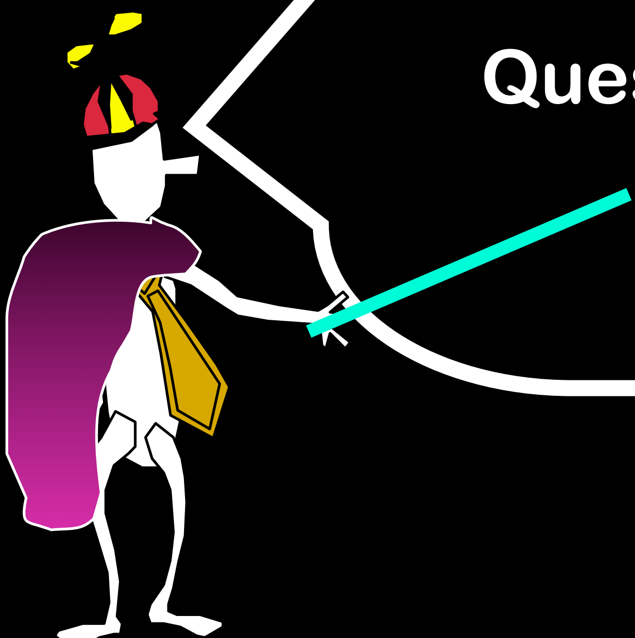
Pretty neat: it doesn't depend on how many letters!

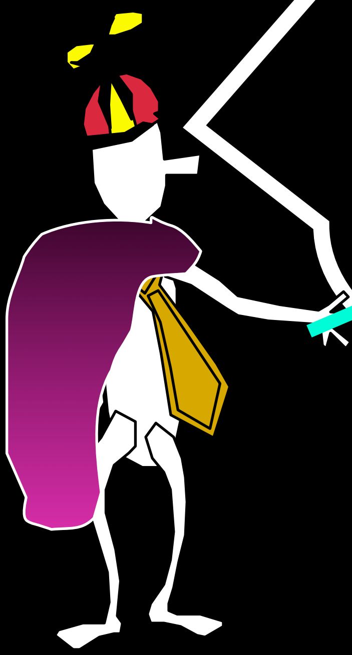


So, in expectation, 1 letter will be in the correct envelope

Pretty neat: it doesn't depend on how many letters!

Question: were the X_i independent?





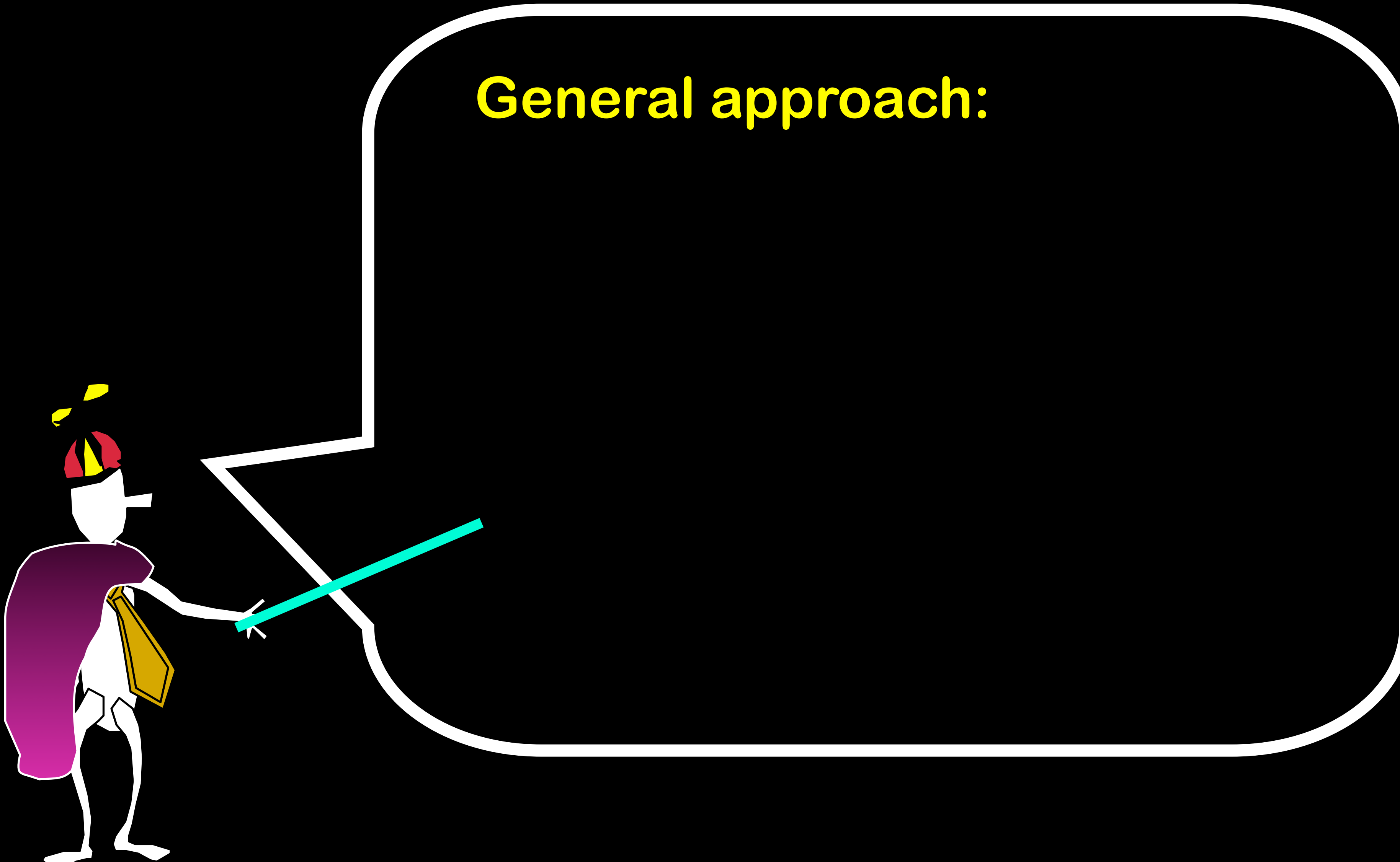
So, in expectation, 1 letter will be in the correct envelope

Pretty neat: it doesn't depend on how many letters!

Question: were the X_i independent?

No! E.g., think of n=2

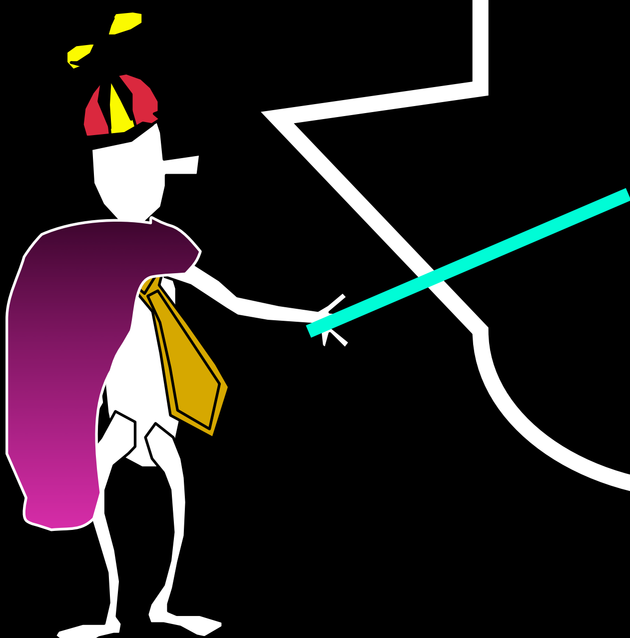
Use Linearity of Expectation



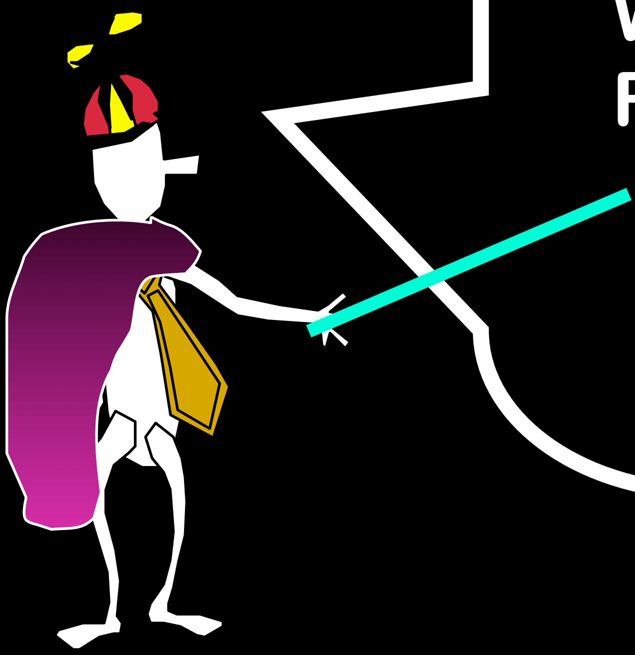
Use Linearity of Expectation

General approach:

View thing you care about as expected value of some RV



Use Linearity of Expectation

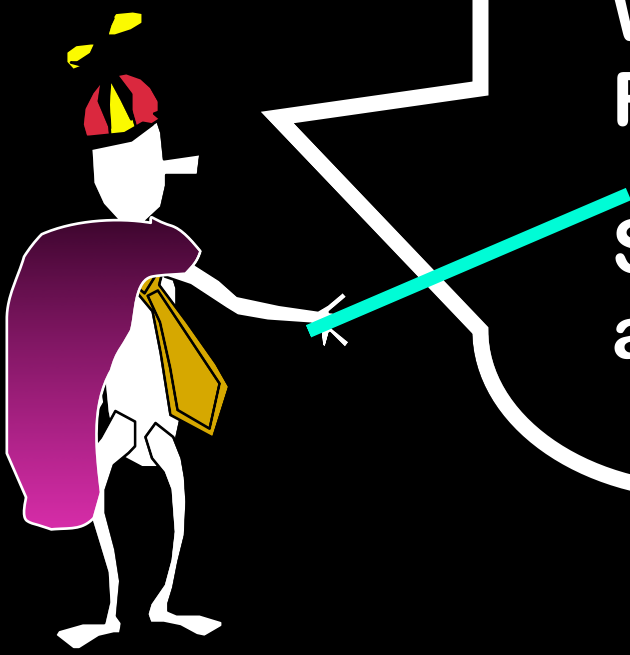


General approach:

View thing you care about as expected value of some RV

Write this RV as sum of simpler RVs (typically indicator RVs)

Use Linearity of Expectation



General approach:

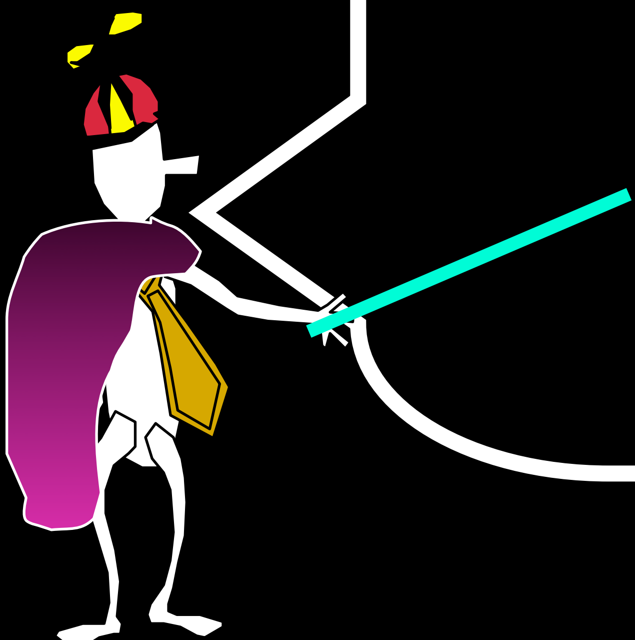
View thing you care about as expected value of some RV

Write this RV as sum of simpler RVs (typically indicator RVs)

Solve for their expectations and add them up!

Example

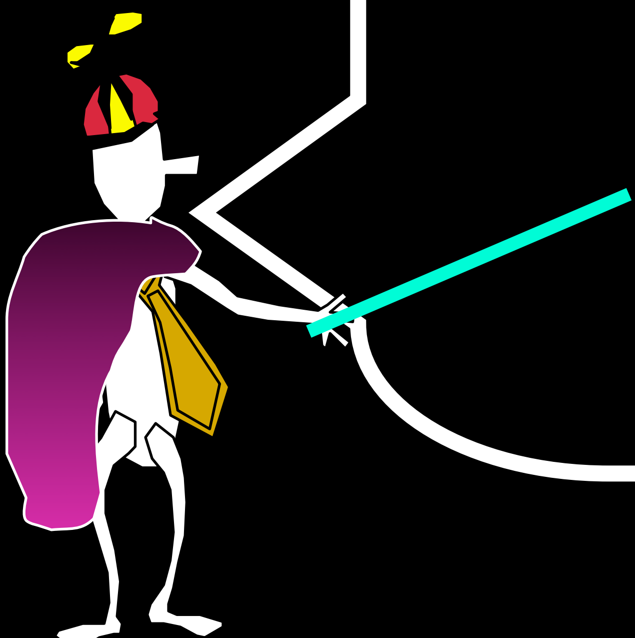
We flip n coins of bias p. What is the expected number of heads?



Example

We flip n coins of bias p. What is the expected number of heads?

We could do this by summing

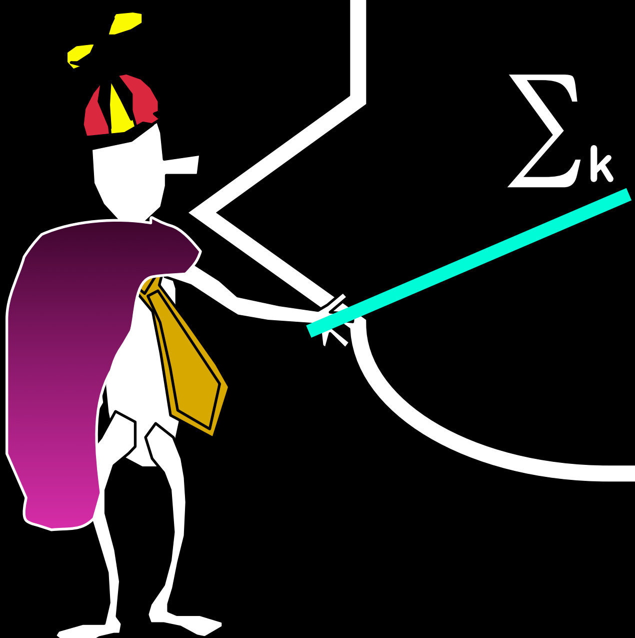


Example

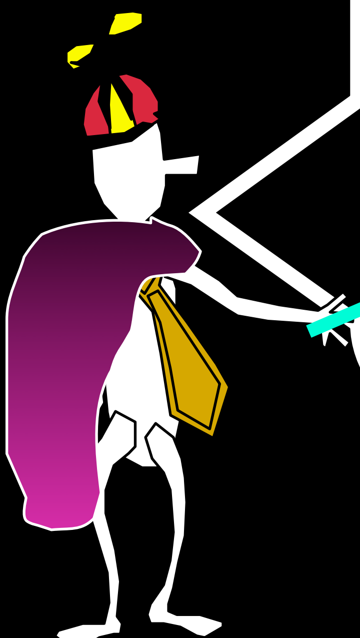
We flip n coins of bias p. What is the expected number of heads?

We could do this by summing

$$\sum_k k \Pr(X = k) = \sum_k k \begin{bmatrix} n \\ k \end{bmatrix} p^k(1-p)^{n-k}$$



Example



We flip n coins of bias p. What is the expected number of heads?

We could do this by summing

$$\sum_k k \Pr(X = k) = \sum_k k \binom{n}{k} p^k (1-p)^{n-k}$$

But now we know a better way!

Linearity of Expectation!

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

Break X into n simpler RVs:

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

Break X into n simpler RVs:

$$X_i = \begin{cases} 1 & \text{if the } j^{th} \text{ coin is tails} \\ 0 & \text{if the } j^{th} \text{ coin is heads} \end{cases}$$

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

Break X into n simpler RVs:

$$X_i = \begin{cases} 1 & \text{if the } j^{th} \text{ coin is tails} \\ 0 & \text{if the } j^{th} \text{ coin is heads} \end{cases}$$

$E[\,X\,] = E[\,\Sigma_i X_i\,] =$

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

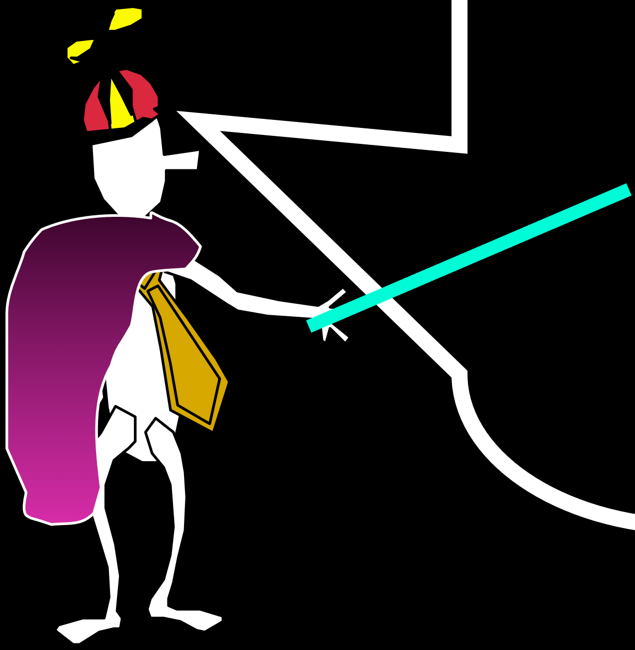
Break X into n simpler RVs:

$$X_i = \begin{cases} 1 & \text{if the } j^{th} \text{ coin is tails} \\ 0 & \text{if the } j^{th} \text{ coin is heads} \end{cases}$$

$E[\,X\,] = E[\,\Sigma_i X_i\,] = np$

What About Products?

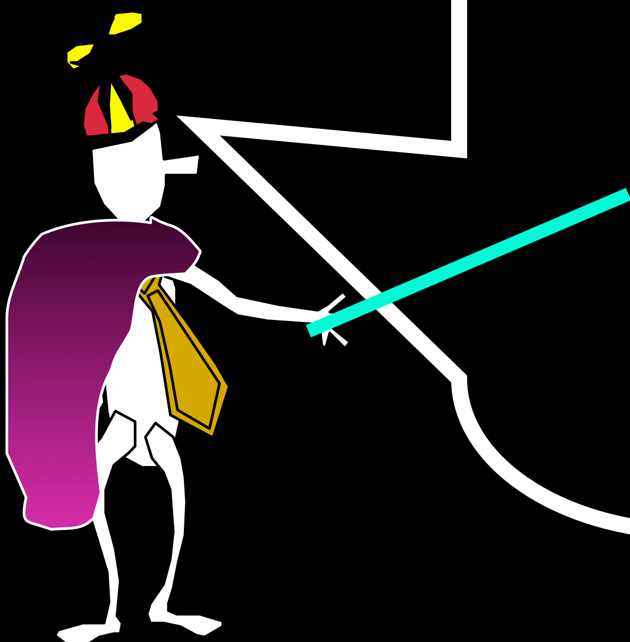
If $Z = XY$, then
$E[Z] = E[X] \times E[Y]$?



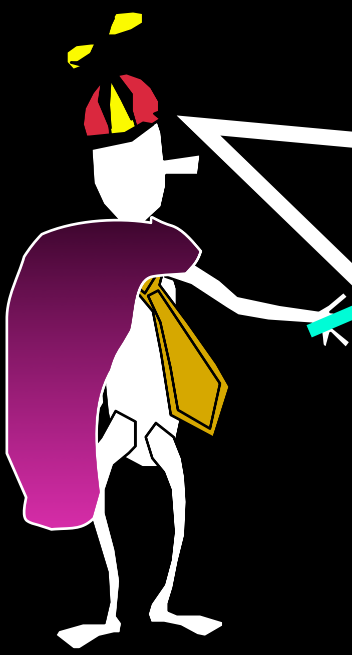
What About Products?

If $Z = XY$, then
$E[Z] = E[X] \times E[Y]$?

No!



What About Products?

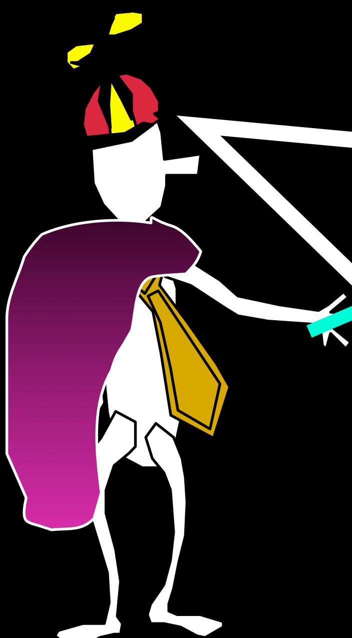


If $Z = XY$, then
$E[Z] = E[X] \times E[Y]$?

No!

X=indicator for "1st flip is heads"
Y=indicator for "1st flip is tails"

What About Products?



If $Z = XY$, then
$E[Z] = E[X] \times E[Y]$?

No!

X=indicator for "1st flip is heads"
Y=indicator for "1st flip is tails"

$E[XY]=0$

But It's True If RVs Are Independent

But It’s True If RVs Are Independent

Proof:

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times Pr(X=a)$

$E[Y] = \sum_b b \times Pr(Y=b)$

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times Pr(X=a)$

$E[Y] = \sum_b b \times Pr(Y=b)$

$E[XY] = \sum_c c \times Pr(XY = c)$

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times \Pr(X=a)$

$E[Y] = \sum_b b \times \Pr(Y=b)$

$$E[XY] = \sum_c c \times \Pr(XY = c)$$

$$= \sum_c \sum_{a,b:ab=c} c \times \Pr(X=a \cap Y=b)$$

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times Pr(X=a)$

$E[Y] = \sum_b b \times Pr(Y=b)$

$E[XY] = \sum_c c \times Pr(XY = c)$

$= \sum_c \sum_{a,b:ab=c} c \times Pr(X=a \cap Y=b)$

$= \sum_{a,b} ab \times Pr(X=a \cap Y=b)$

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times Pr(X=a)$

$E[Y] = \sum_b b \times Pr(Y=b)$

$$E[XY] = \sum_c c \times Pr(XY = c)$$

$$= \sum_c \sum_{a,b:ab=c} c \times Pr(X=a \cap Y=b)$$

$$= \sum_{a,b} ab \times Pr(X=a \cap Y=b)$$

$$= \sum_{a,b} ab \times Pr(X=a)\, Pr(Y=b)$$

But It's True If RVs Are Independent

Proof: $E[X] = \sum_a a \times Pr(X=a)$
$E[Y] = \sum_b b \times Pr(Y=b)$

$E[XY] = \sum_c c \times Pr(XY = c)$

$= \sum_c \sum_{a,b:ab=c} c \times Pr(X=a \cap Y=b)$

$= \sum_{a,b} ab \times Pr(X=a \cap Y=b)$

$= \sum_{a,b} ab \times Pr(X=a)\, Pr(Y=b)$

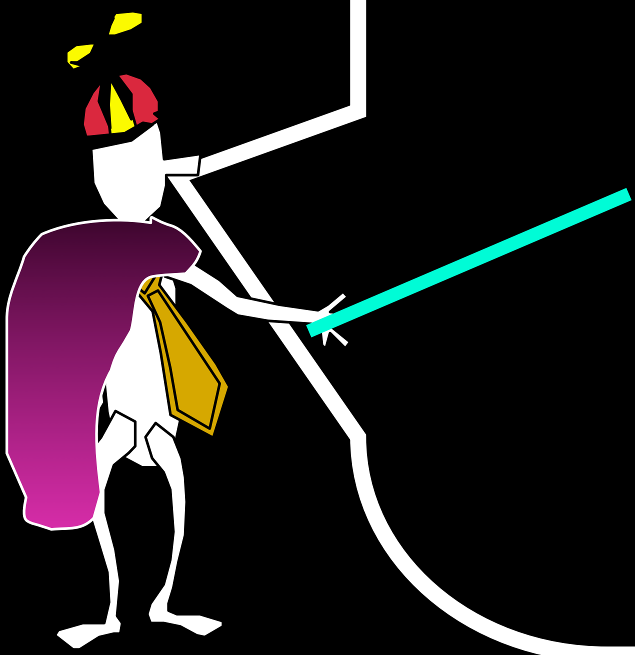
$= E[X]\, E[Y]$

Example: 2 fair flips

X = indicator for 1st coin heads

Y = indicator for 2nd coin heads

XY = indicator for “both are heads”



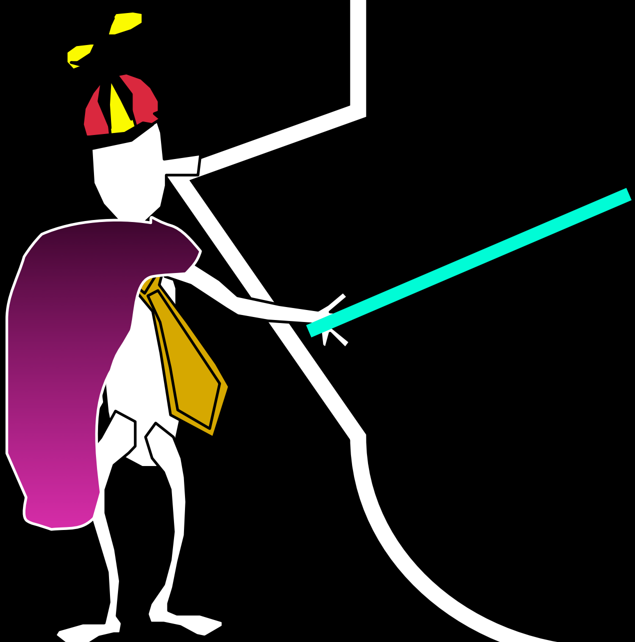
Example: 2 fair flips

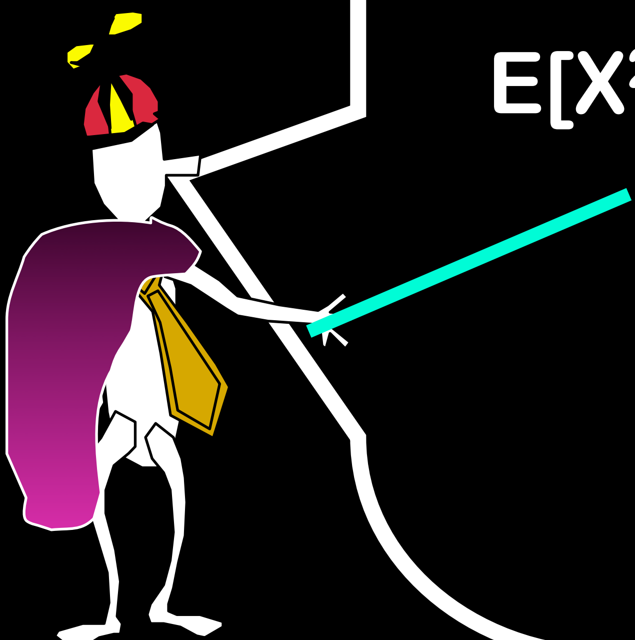
X = indicator for 1st coin heads

Y = indicator for 2nd coin heads

XY = indicator for "both are heads"

$E[X] = \frac{1}{2}$, $E[Y] = \frac{1}{2}$, $E[XY] = \frac{1}{4}$





Example: 2 fair flips

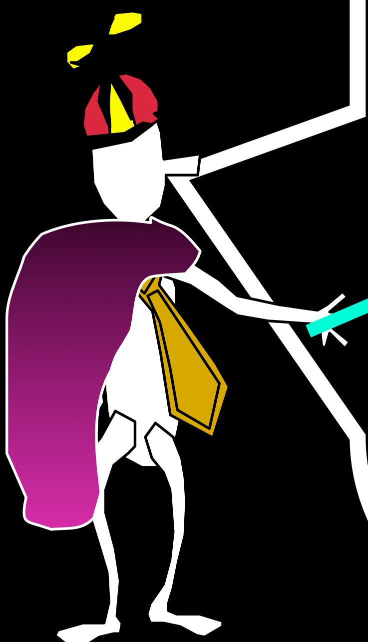
X = indicator for 1st coin heads

Y = indicator for 2nd coin heads

XY = indicator for "both are heads"

$E[X] = \frac{1}{2}$, $E[Y] = \frac{1}{2}$, $E[XY] = \frac{1}{4}$

$E[X^2] = E[X]^2$?

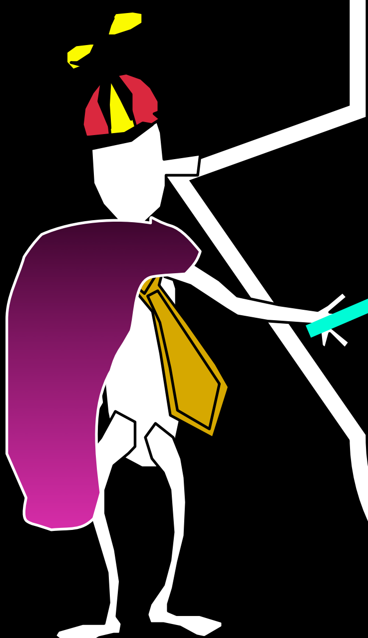


Example: 2 fair flips
X = indicator for 1st coin heads
Y = indicator for 2nd coin heads
XY = indicator for “both are heads”

$E[X] = \frac{1}{2}$, $E[Y] = \frac{1}{2}$, $E[XY] = \frac{1}{4}$

$E[X^2] = E[X]^2$?

No: $E[X^2] = \frac{1}{2}$, $E[X]^2 = \frac{1}{4}$



Example: 2 fair flips
X = indicator for 1st coin heads
Y = indicator for 2nd coin heads
XY = indicator for “both are heads”

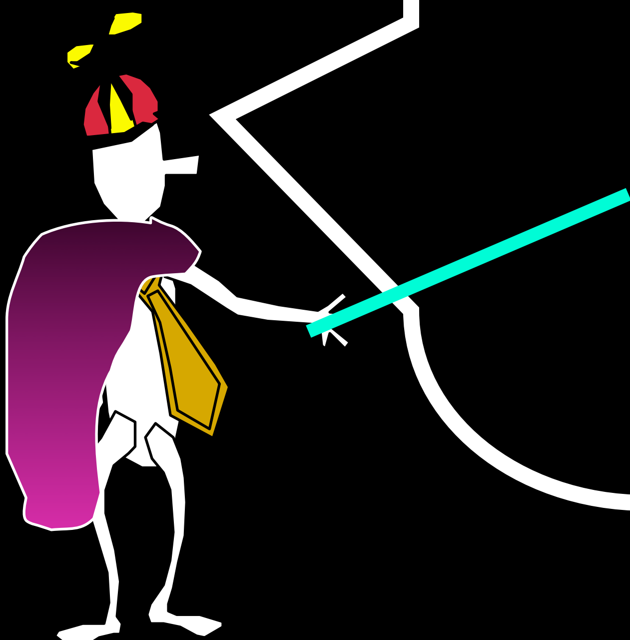
$E[X] = ½, E[Y] = ½, E[XY] = ¼$

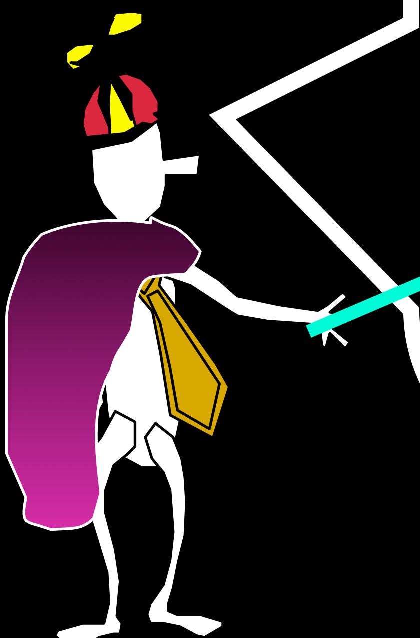
$E[X^2] = E[X]^2$?

No: $E[X^2] = ½, E[X]^2 = ¼$

In fact, $E[X^2] – E[X]^2$ is called the variance of X

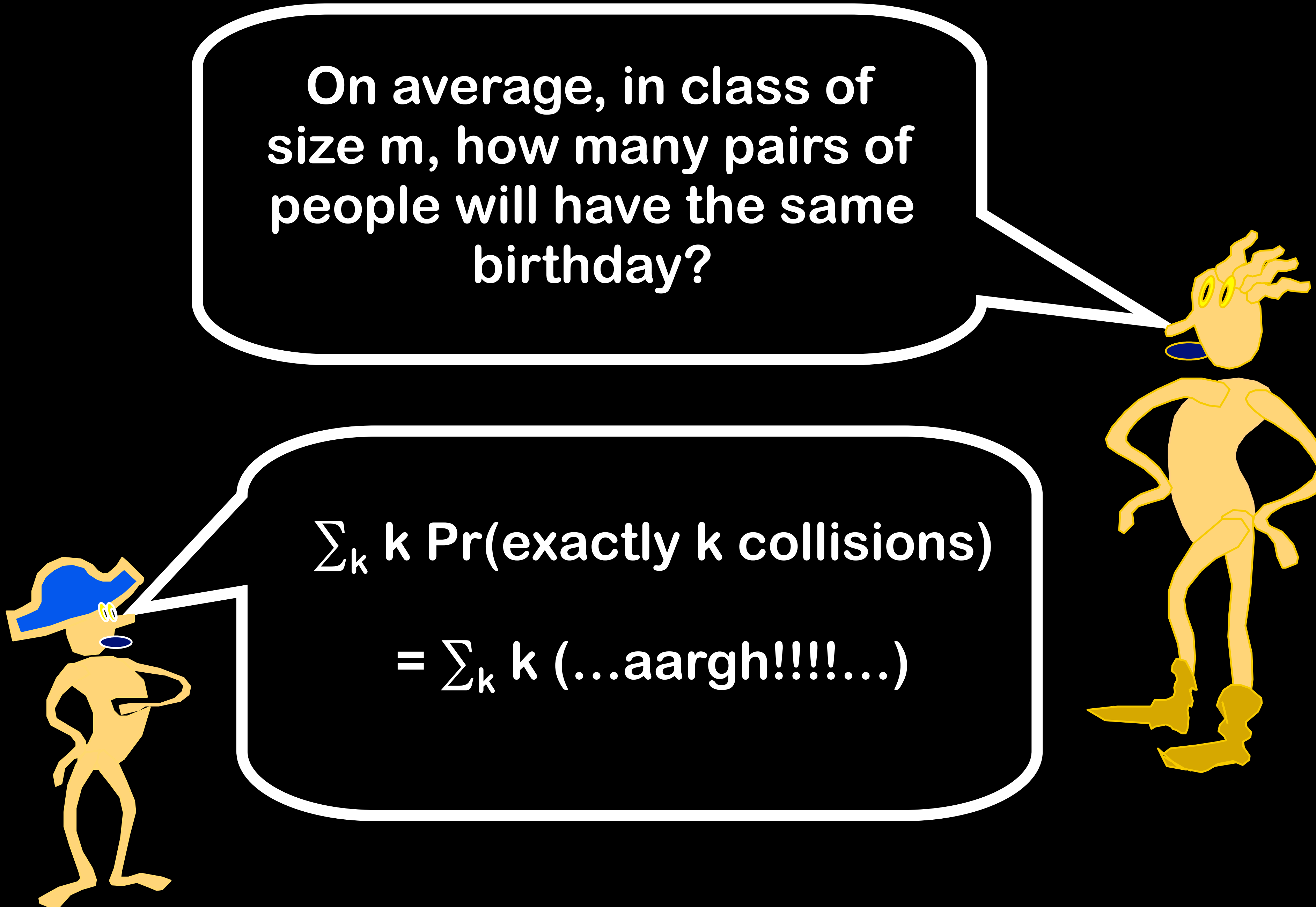
Most of the time, though, power will come from using sums

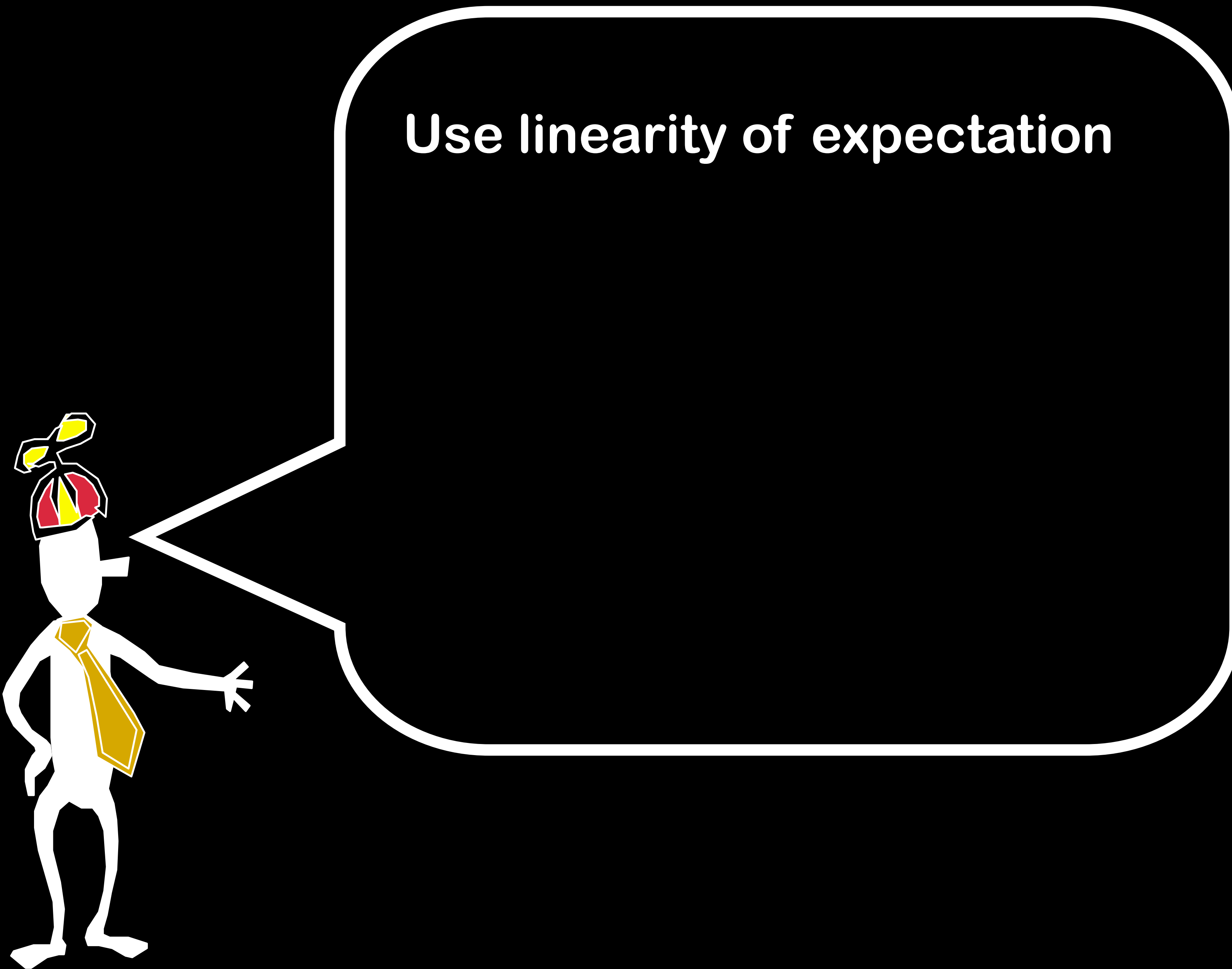


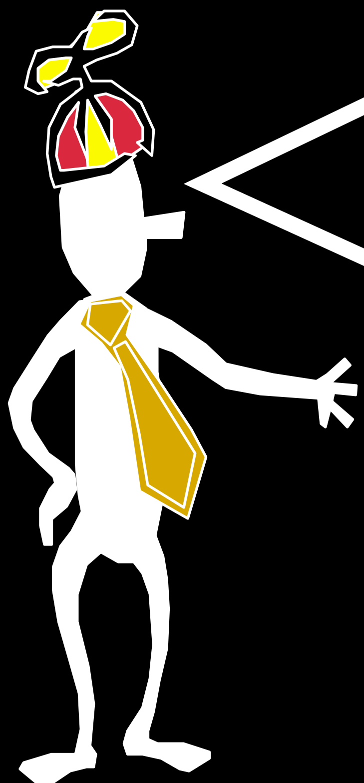


Most of the time, though, power will come from using sums

Mostly because **Linearity of Expectations** holds even if RVs are not independent

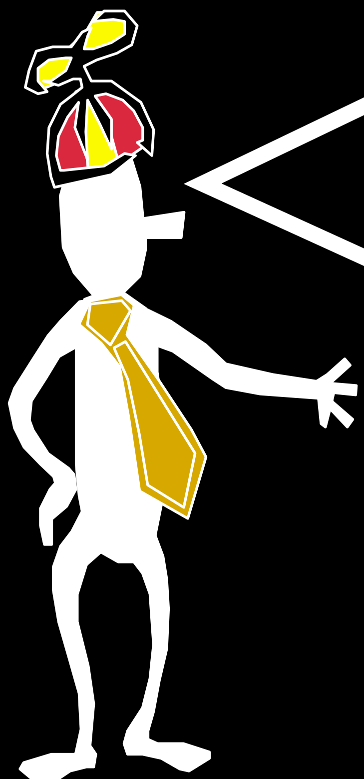

On average, in class of size m, how many pairs of people will have the same birthday?
$\sum_k k \Pr(\text{exactly } k \text{ collisions})$
$= \sum_k k$ (...aargh!!!!...)


Use linearity of expectation



Use linearity of expectation

Suppose we have m people each with a uniformly chosen birthday from 1 to 366



Use linearity of expectation

Suppose we have m people each with a uniformly chosen birthday from 1 to 366

X = number of pairs of people with the same birthday

Use linearity of expectation

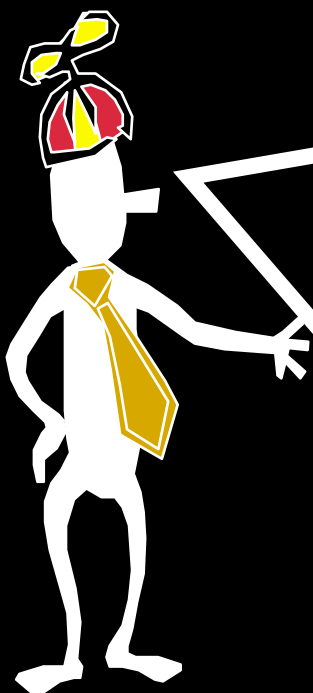
Suppose we have m people each with a uniformly chosen birthday from 1 to 366

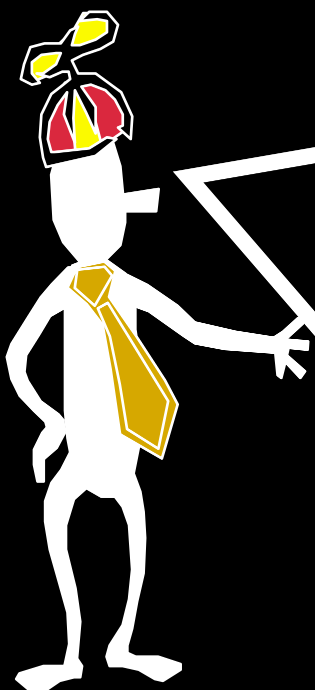
X = number of pairs of people with the same birthday

E[X] = ?

X = number of pairs of people with the same birthday

E[X] = ?

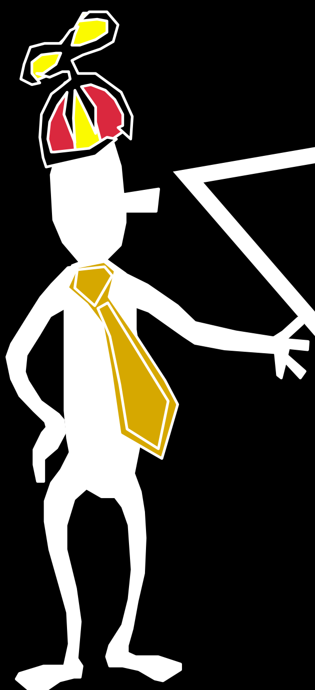




X = number of pairs of people with the same birthday

$E[X]$ = ?

Use m(m-1)/2 indicator variables, one for each pair of people



X = number of pairs of people with the same birthday

$E[X] = ?$

Use $m(m-1)/2$ indicator variables, one for each pair of people

$X_{jk} = 1$ if person j and person k have the same birthday; else 0

X = number of pairs of people with the same birthday

$E[X] = ?$

Use $m(m-1)/2$ indicator variables, one for each pair of people

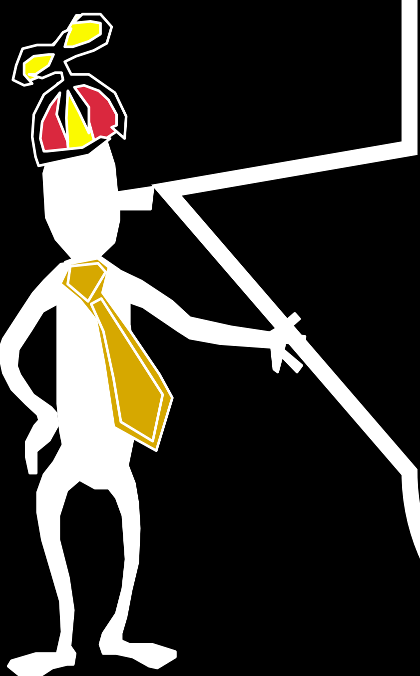
$X_{jk} = 1$ if person j and person k have the same birthday; else 0

$$E[X_{jk}] = (1/366)\ 1 + (1 - 1/366)\ 0 = 1/366$$

X = number of pairs of people with the same birthday

X_{jk} = 1 if person j and person k have the same birthday; else 0

$$E[X_{jk}] = (1/366)\,1 + (1 - 1/366)\,0 = 1/366$$

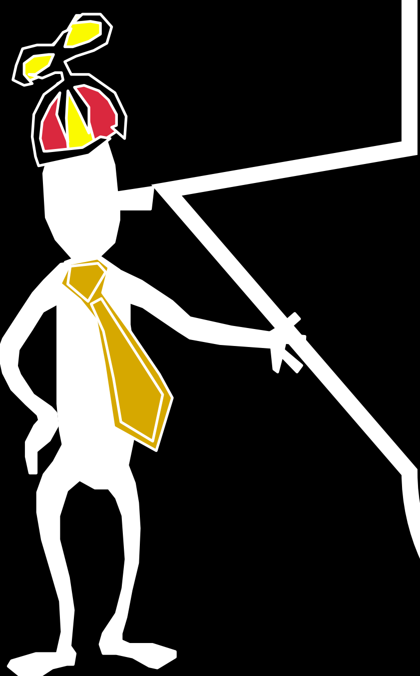


X = number of pairs of people with the same birthday

X_{jk} = 1 if person j and person k have the same birthday; else 0

$$E[X_{jk}] = (1/366)\,1 + (1 - 1/366)\,0$$
$$= 1/366$$

$$E[X] = E[\,\Sigma_{j \le k \le m} X_{jk}\,]$$



X = number of pairs of people with the same birthday

X_{jk} = 1 if person j and person k have the same birthday; else 0

$$E[X_{jk}] = (1/366)\,1 + (1 - 1/366)\,0$$
$$= 1/366$$

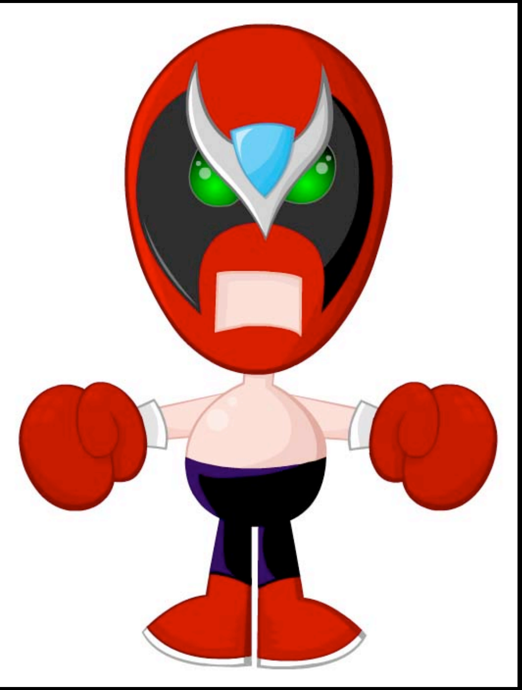
$$E[X] = E[\textstyle\sum_{j \le k \le m} X_{jk}]$$
$$= \textstyle\sum_{j \le k \le m} E[X_{jk}]$$

X = number of pairs of people with the same birthday

X_{jk} = 1 if person j and person k have the same birthday; else 0

$$E[X_{jk}] = (1/366)\ 1 + (1 - 1/366)\ 0$$
$$= 1/366$$

$$E[X] = E[\ \Sigma_{j \le k \le m} X_{jk}\ ]$$
$$= \Sigma_{j \le k \le m} E[\ X_{jk}\ ]$$
$$= m(m-1)/2 \times 1/366$$



Here's What You Need to Know…

Language of Probability

Sample Space

Events

Uniform Distribution

Pr [A | B]

Independence

Binomial Distribution

Definition

Random Variables

Two views

Expectation

Linearity of expectation