15-251

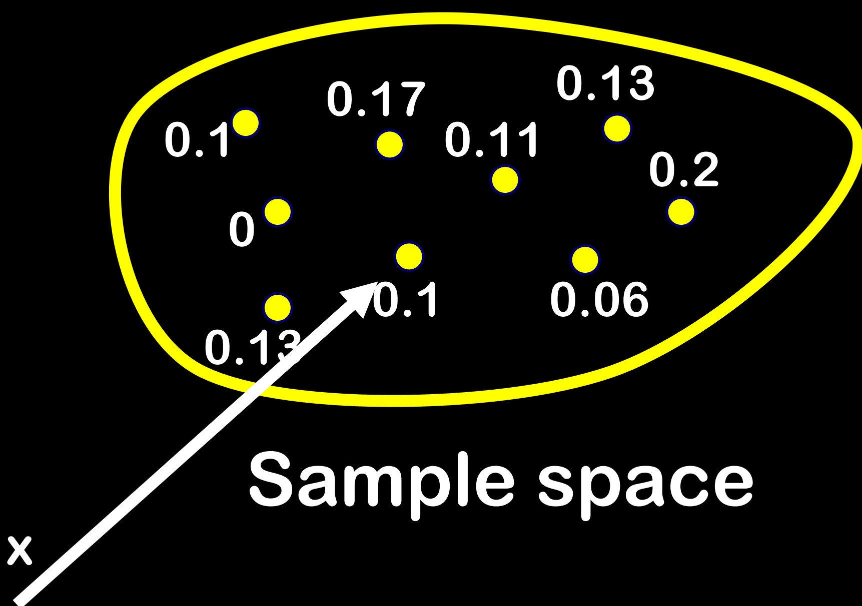
Great Theoretical Ideas in Computer Science

Finite Probability Distribution

A (finite) probability distribution **D** is a finite "sample space" **S** of elements or "samples", where each sample **x in S** has a non-negative real weight or probability **p(x)**

The weights must satisfy:

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

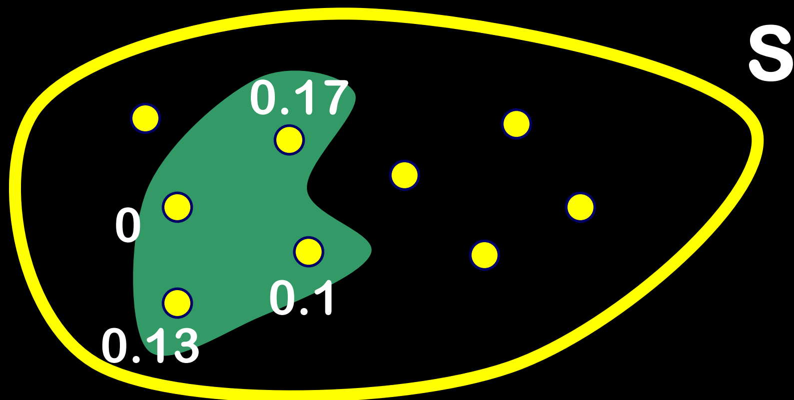


weight or probability of x
D(x) = p(x) = 0.1

Events

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

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

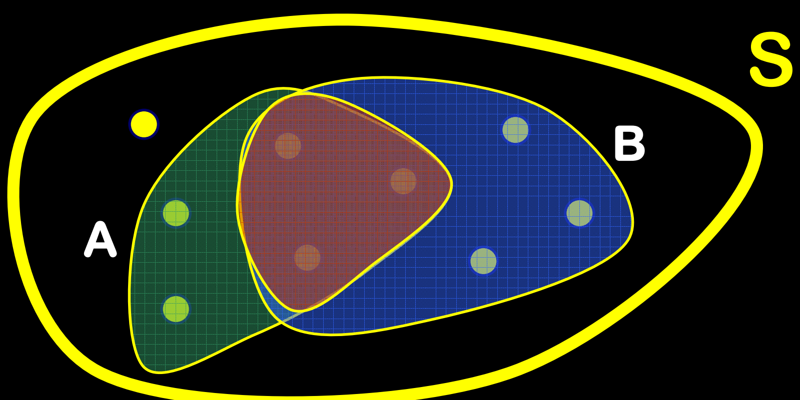


$$Pr_D[E] = 0.4$$

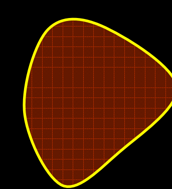
Conditional Probability

The probability of event A given event B is written $\Pr[\, A \mid B \,]$ and is defined to be =

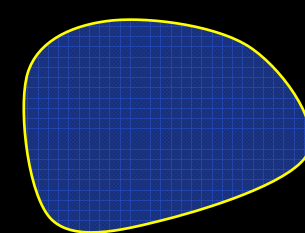
$$\frac{\Pr[\, A \cap B \,]}{\Pr[\, B \,]}$$



proportion of $A \cap B$



to B



Conditional Probability

The probability of event A given event B is written $\Pr[\, A \mid B \,]$ and is defined to be =

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

A and B are **independent** events if

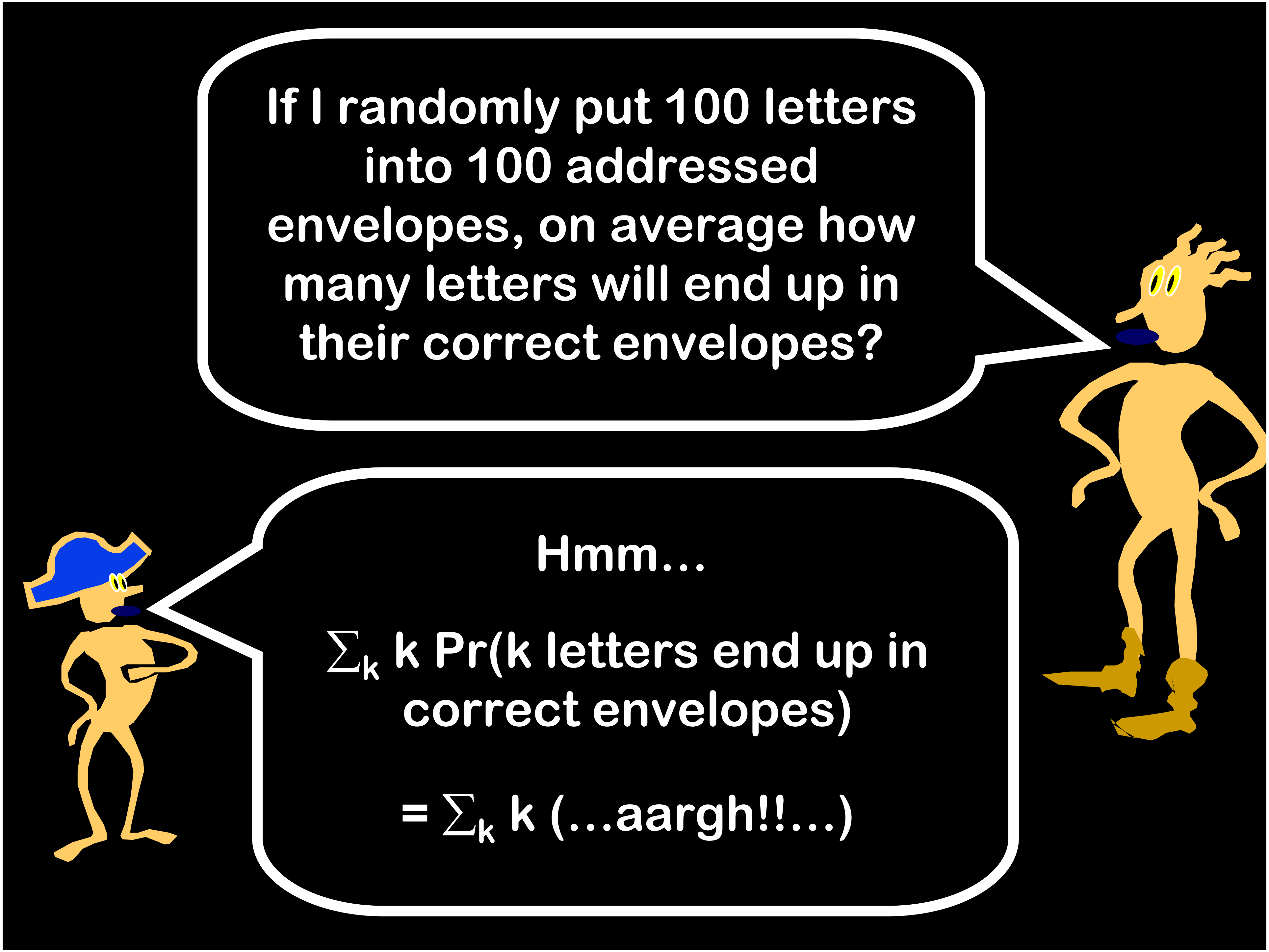
$$\Pr[\, A \mid B \,] = \Pr[\, A \,]$$

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

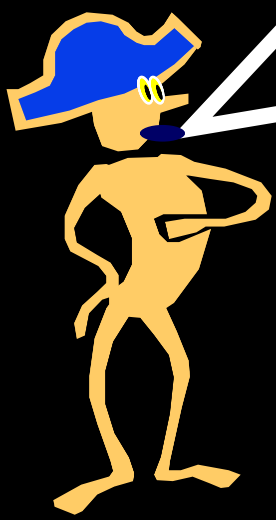
$$\Leftrightarrow \Pr[\, B \mid A \,] = \Pr[\, B \,]$$

Lecture 11 (October 2, 2007)

Today, we will 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?
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?



$\sum_k$ k Pr(exactly k collisions)

= $\sum_k$ k (…aargh!!!!…)

The new tool is called "Linearity of Expectation"

Random Variable

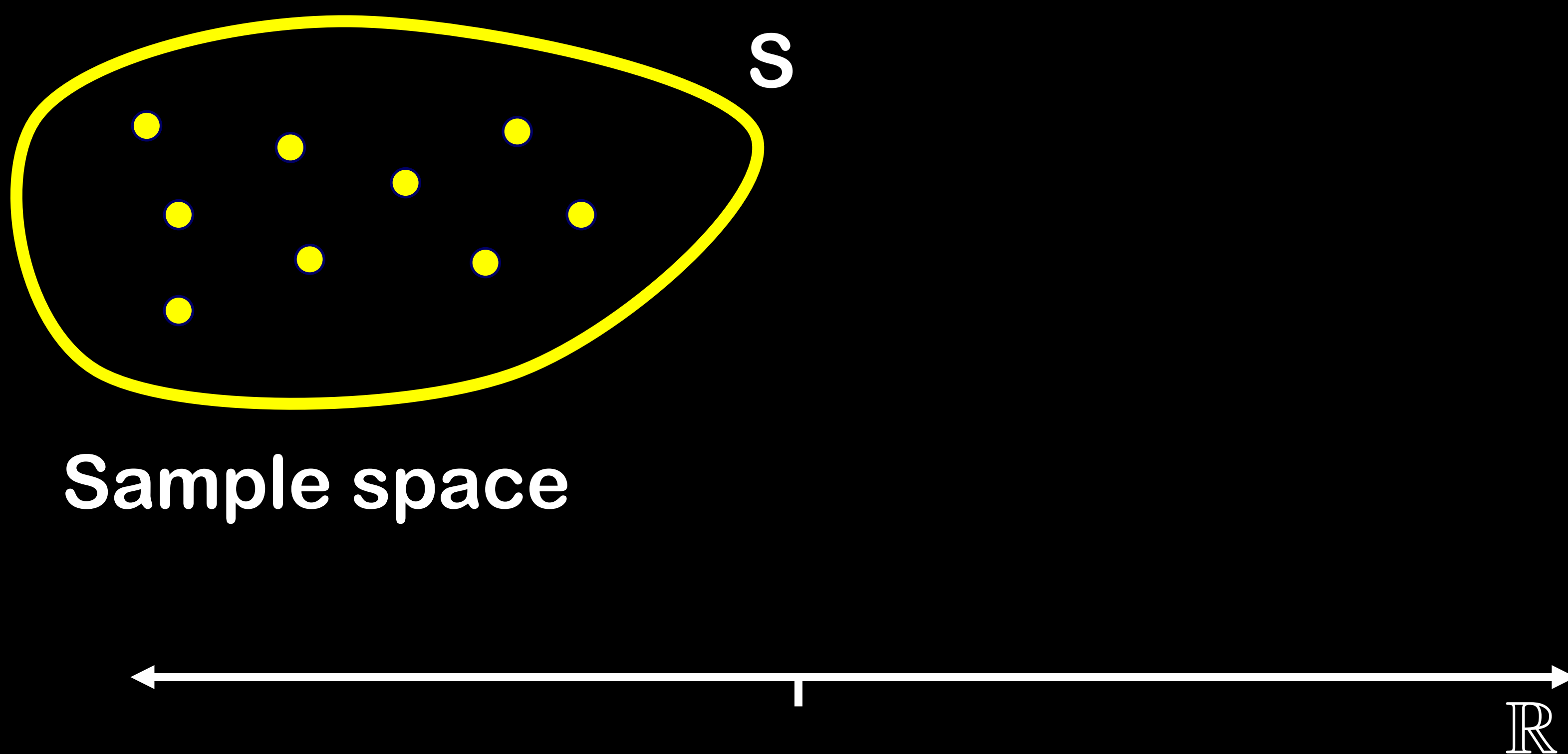
To use this new tool, we will also need to understand the concepts of
Random Variable
and
Expectations

Today's lecture: not too much material, but need to understand it well

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:

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

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

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

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

Tossing a Fair Coin n Times

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

D = uniform distribution on S

$\Rightarrow D(x) = (½)^n$ 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

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

Think of a R.V. as

Input to the function is random

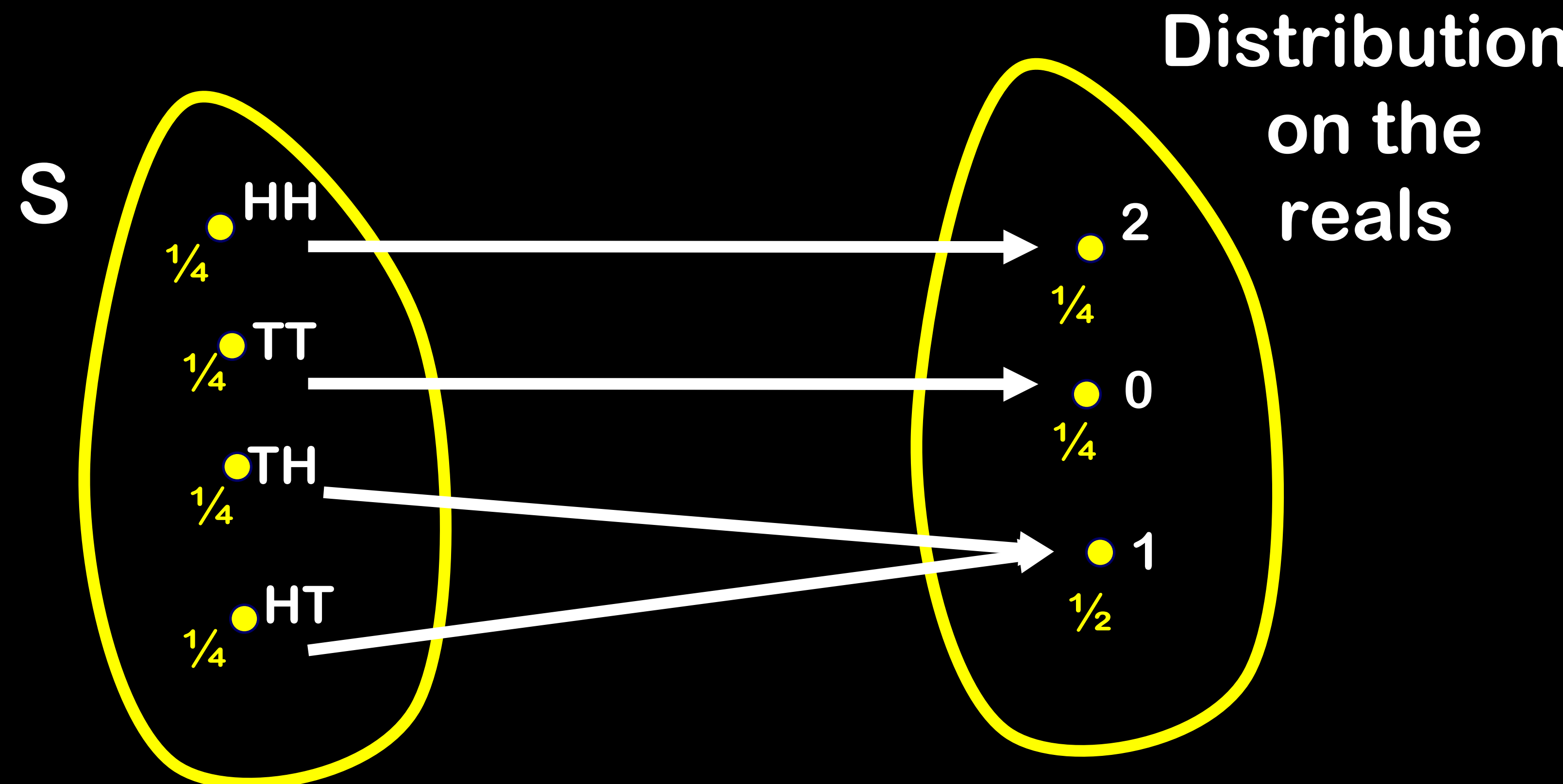
A function from S to the reals $\mathbb{R}$

Or think of the induced distribution on $\mathbb{R}$

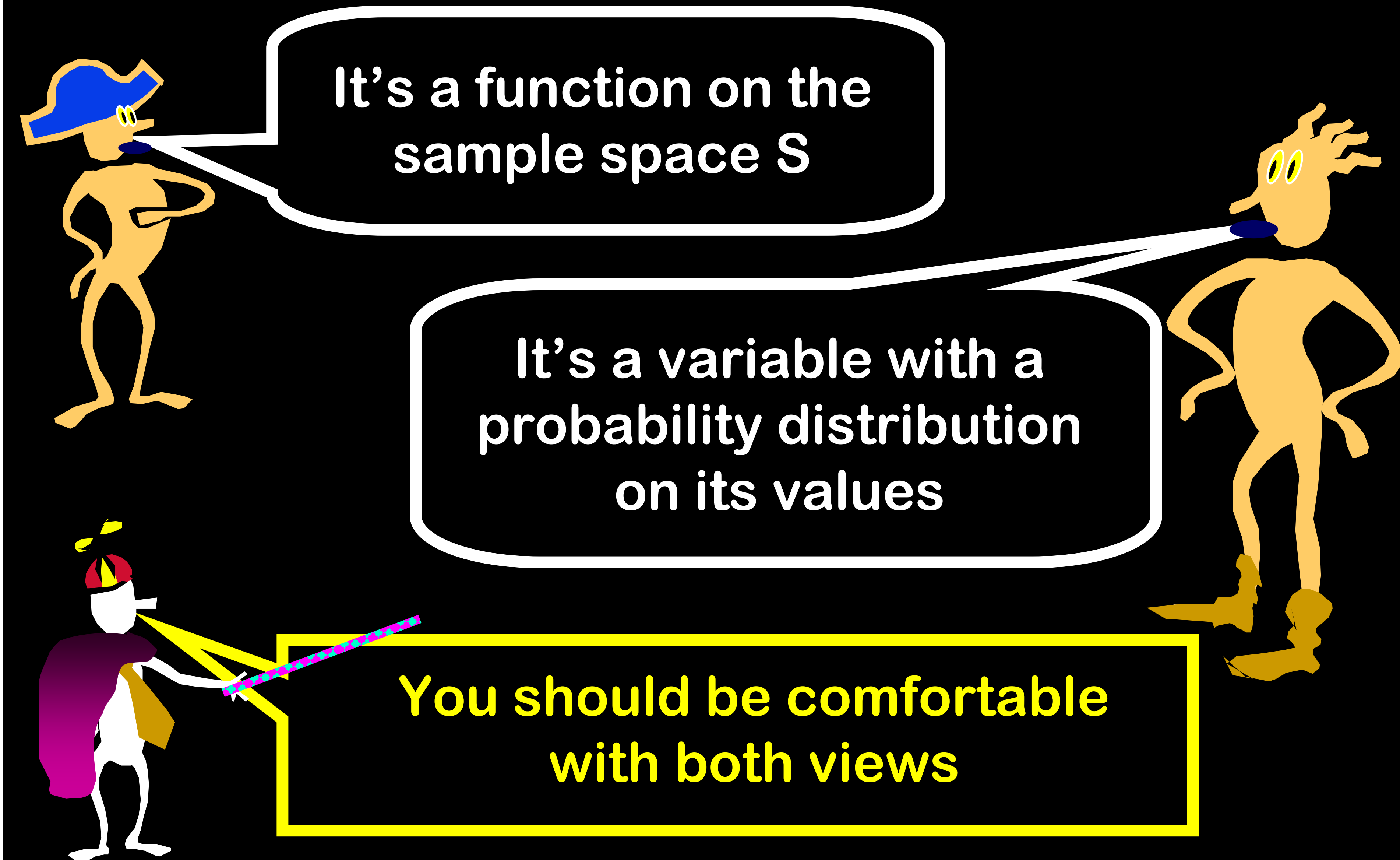
Randomness is "pushed" to the values of the function

Two Coins Tossed

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



It's a Floor Wax And a Dessert Topping



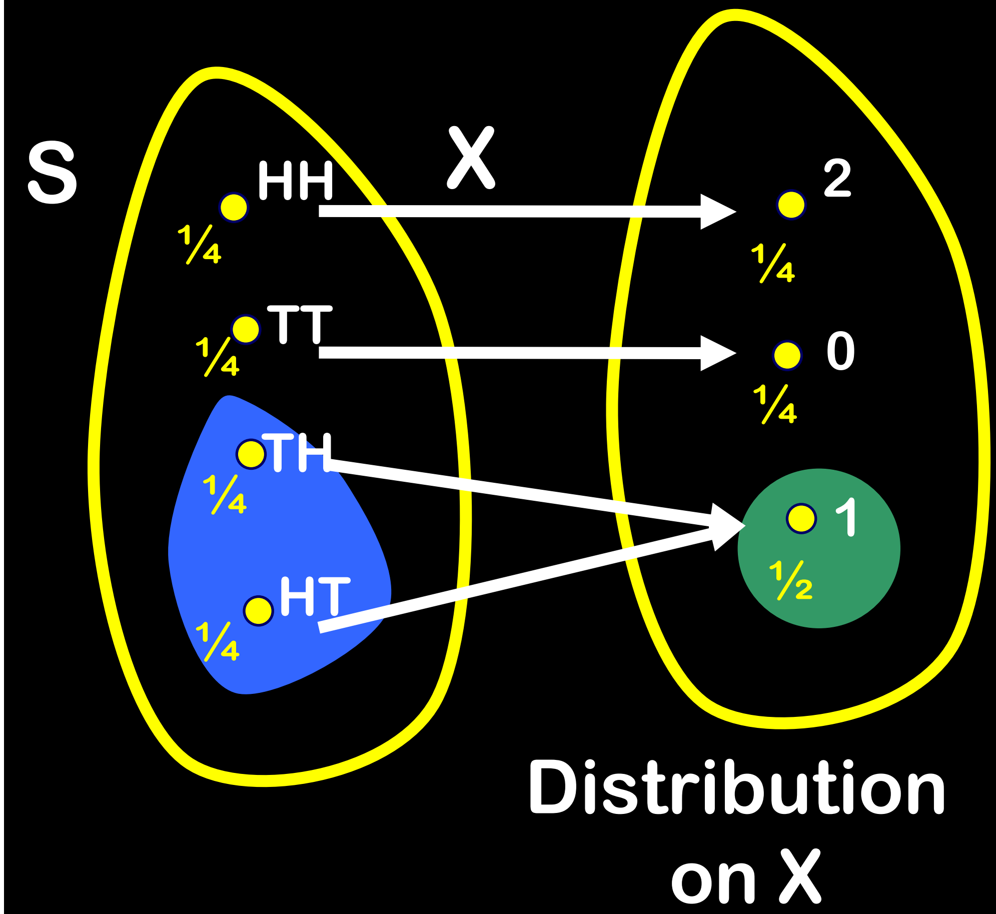
From Random Variables to Events

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

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

Two Coins Tossed

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



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

$\Pr(X = 1)$

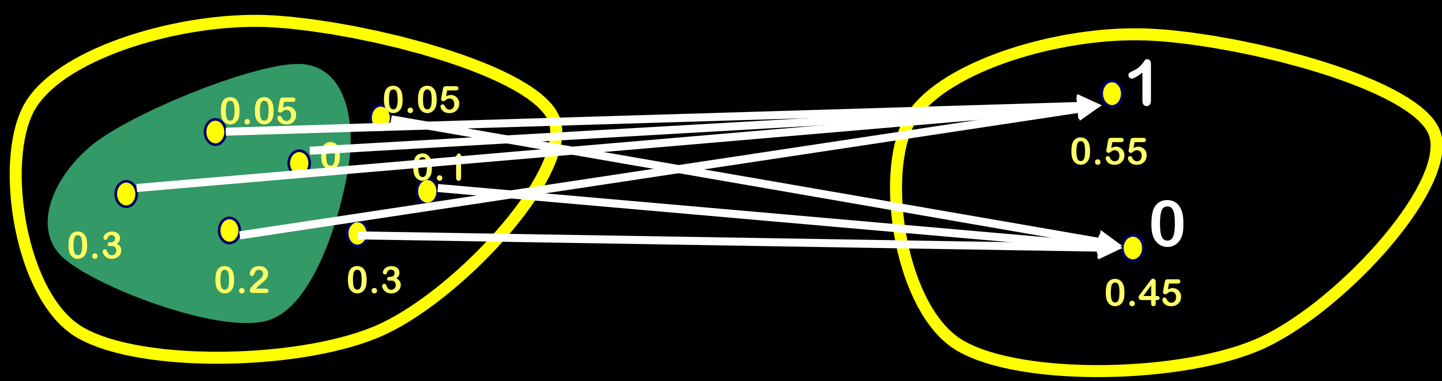
$= \Pr(\{x \in S \mid X(x) = 1\})$

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

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

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

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

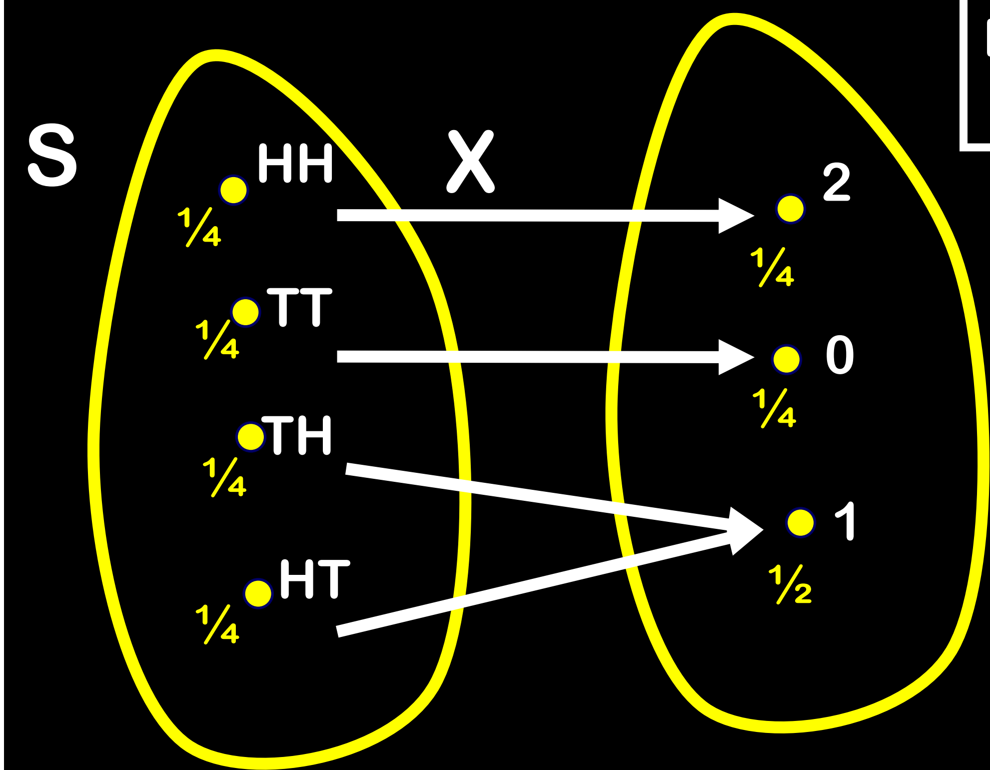
X is a function on the sample space S

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?

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



A Quick Calculation…

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

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

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

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

Type Checking

A Random Variable is the type of thing you might want to know an expected value of

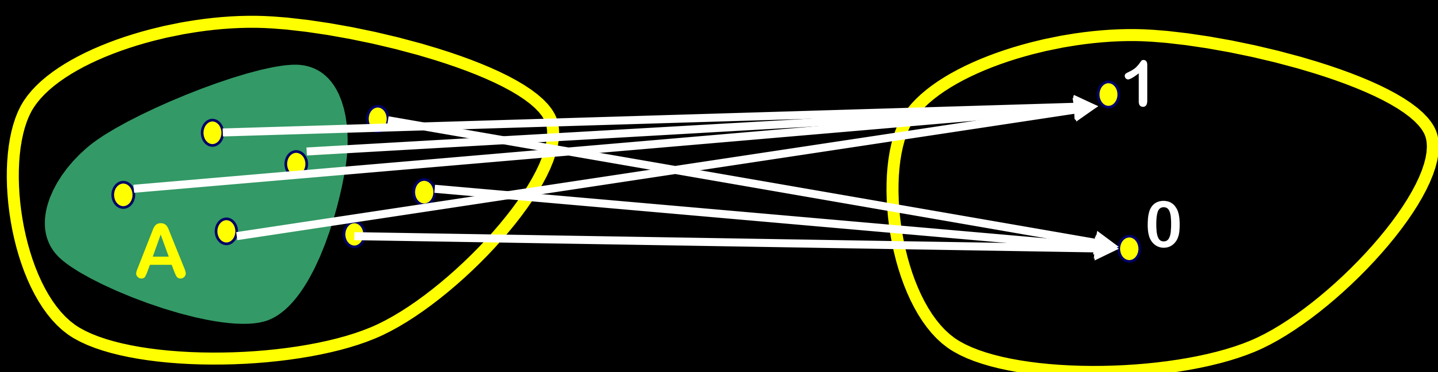
If you are computing an expectation, **the thing whose expectation you are computing is a random variable**

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

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

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

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



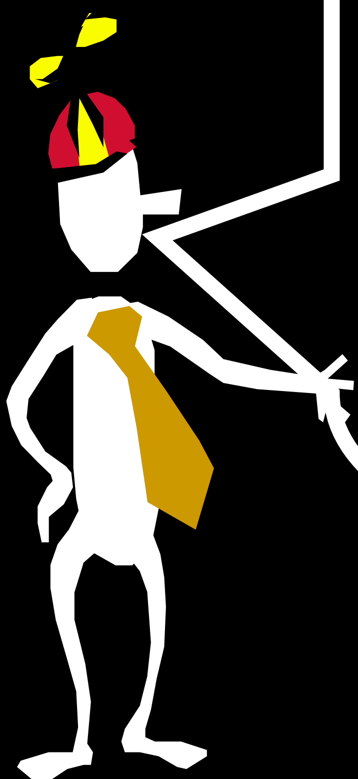
Adding Random Variables

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

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

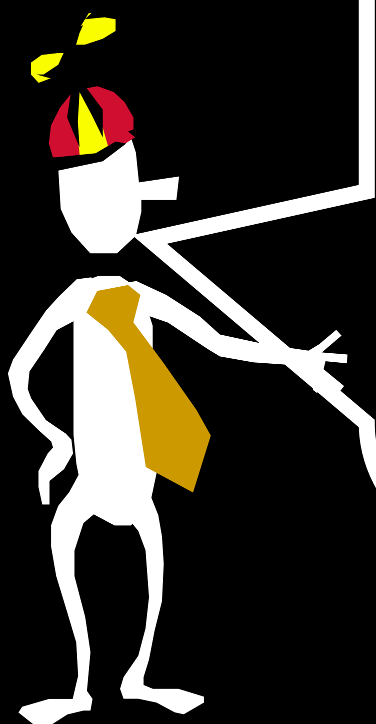
E.g., rolling two dice.
X = 1st die, Y = 2nd die,
Z = sum of two dice

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

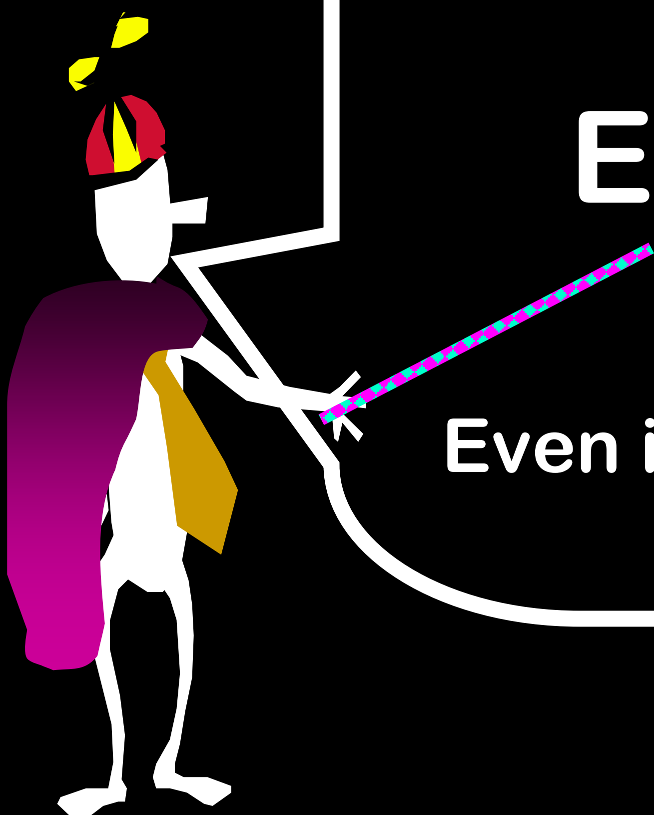
Independence



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

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

Linearity of Expectation



If $Z = X+Y$, then

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

Even if X and Y are not independent!

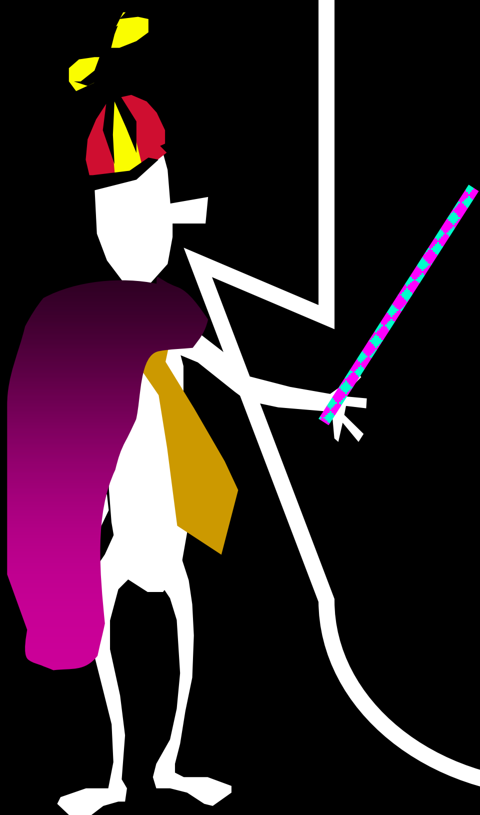
$$E[Z] = \sum_{x \in S} Pr[x]\, Z(x)$$

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

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

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

Linearity of Expectation



E.g., 2 fair flips:

X = 1st coin, Y = 2nd coin

Z = X+Y = total # heads

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

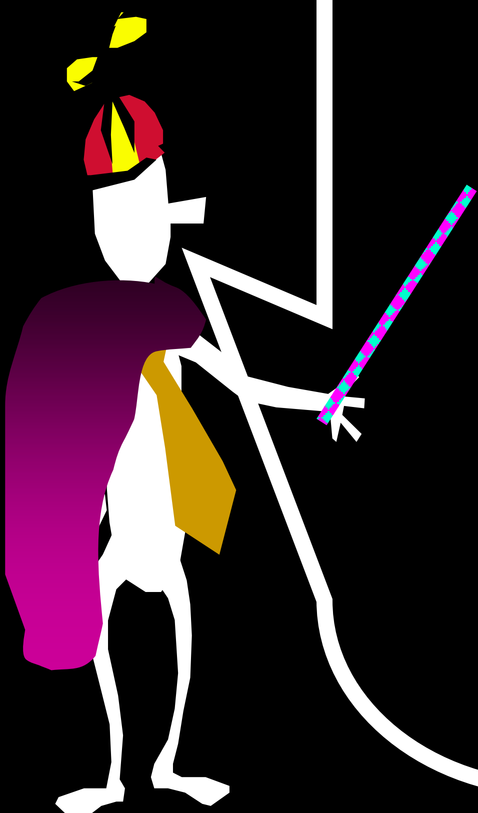
1,1,2
HH

1,0,1
HT

0,1,1
TH

0,0,0
TT

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

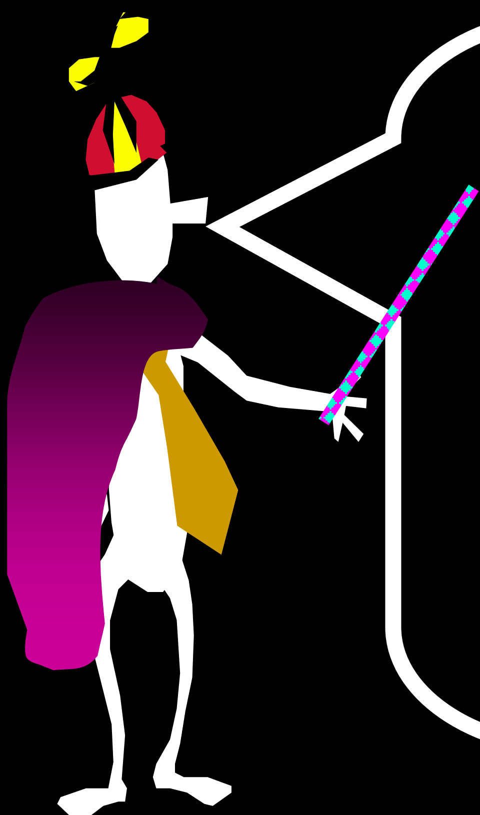
1,0,1
TH

0,0,0
TT

By Induction

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

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

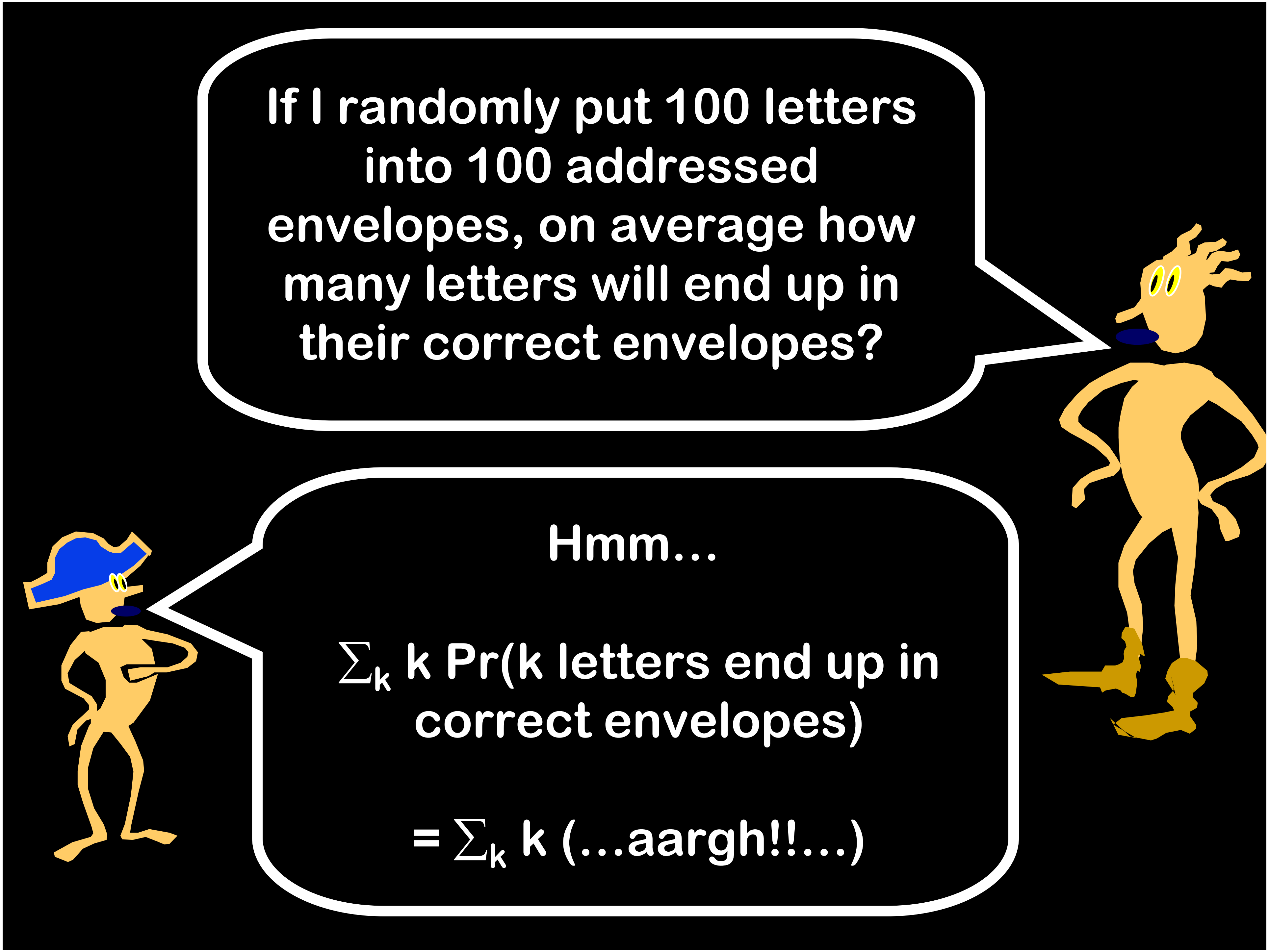


The expectation of the sum

=

The sum of the expectations

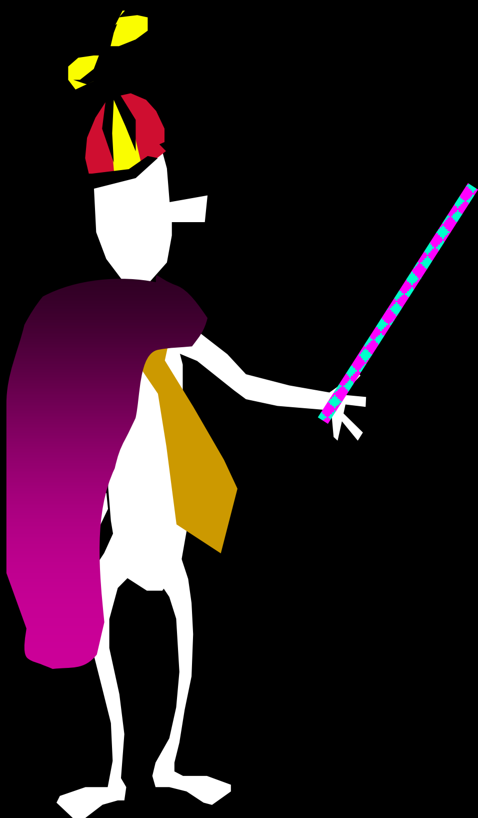

It is finally time to show off our probability prowess…


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

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



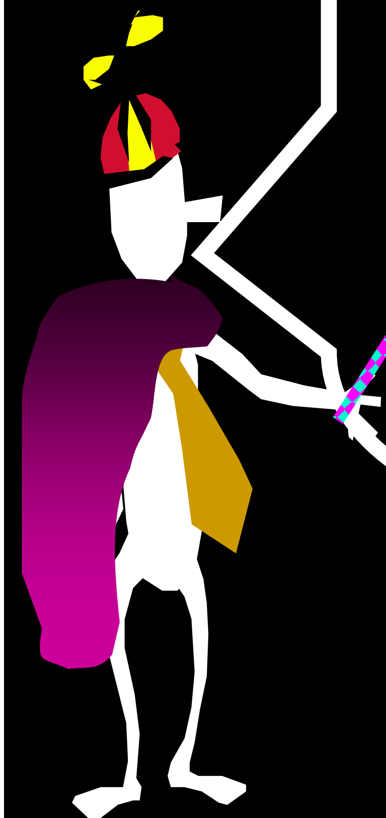
$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

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

We are asking for $E[Z]$

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

So $E[Z] = 1$



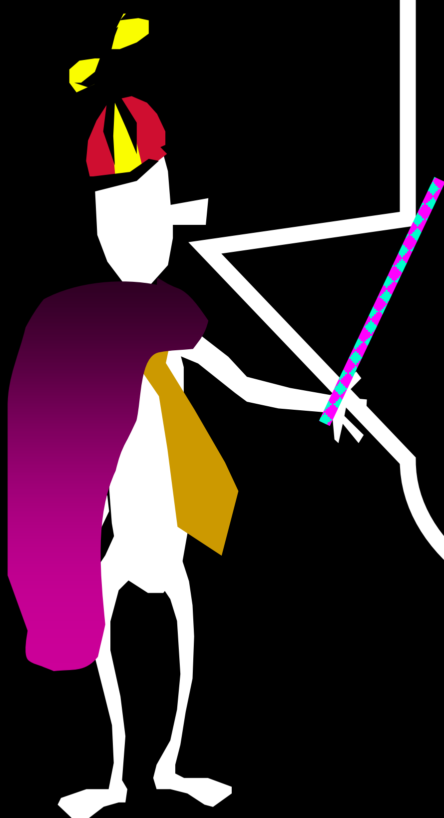
So, in expectation, 1 letter will be in the same correct envelope

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



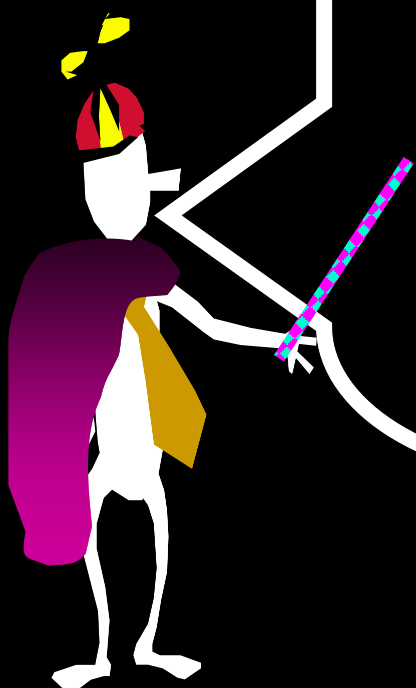
General approach:

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?

We could do this by summing

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

But now we know a better way!

Linearity of Expectation!

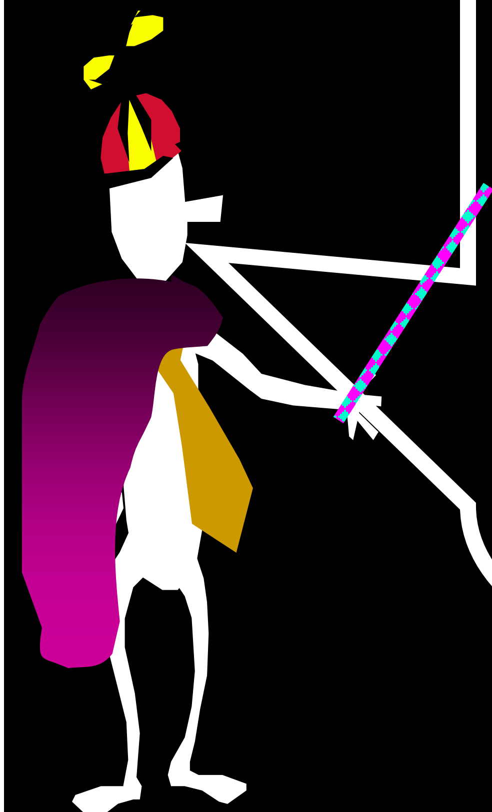
Let X = number of heads when n independent coins of bias p are flipped

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] × 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

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

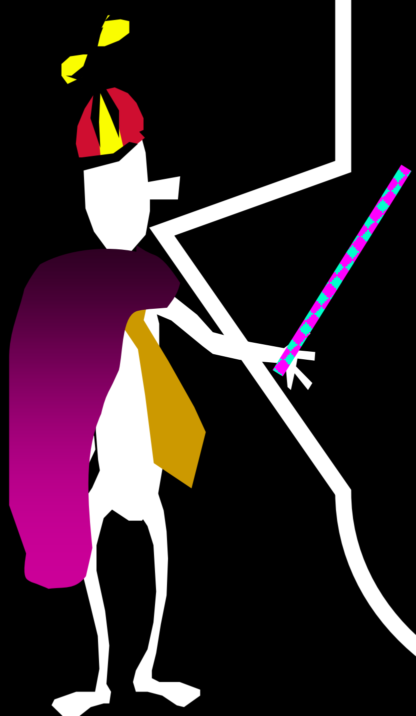
XY = indicator for "both are heads"

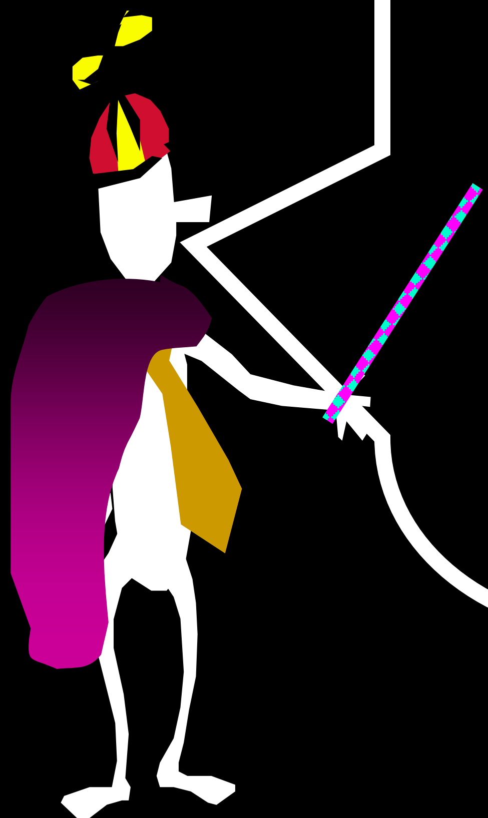
$E[X] = ½, E[Y] = ½, E[XY] = ¼$

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

No: $E[X^2] = ½, E[X]^2 = ¼$

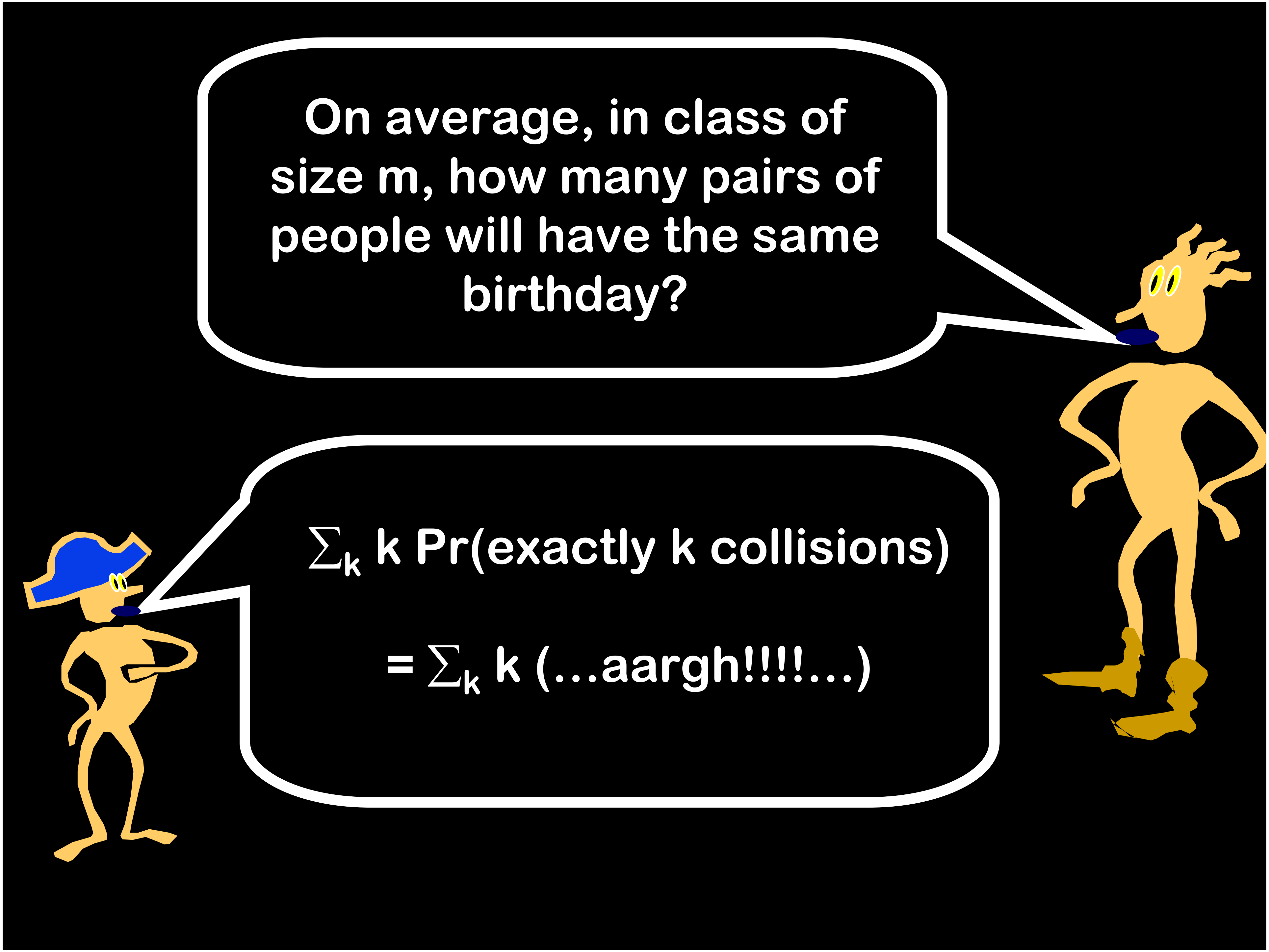
In fact, $E[X^2] – E[X]^2$ is called the variance of X





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

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(exactly k collisions)
= $\sum_k$ k (...aargh!!!!...)

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

$E[X] = ?$

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

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

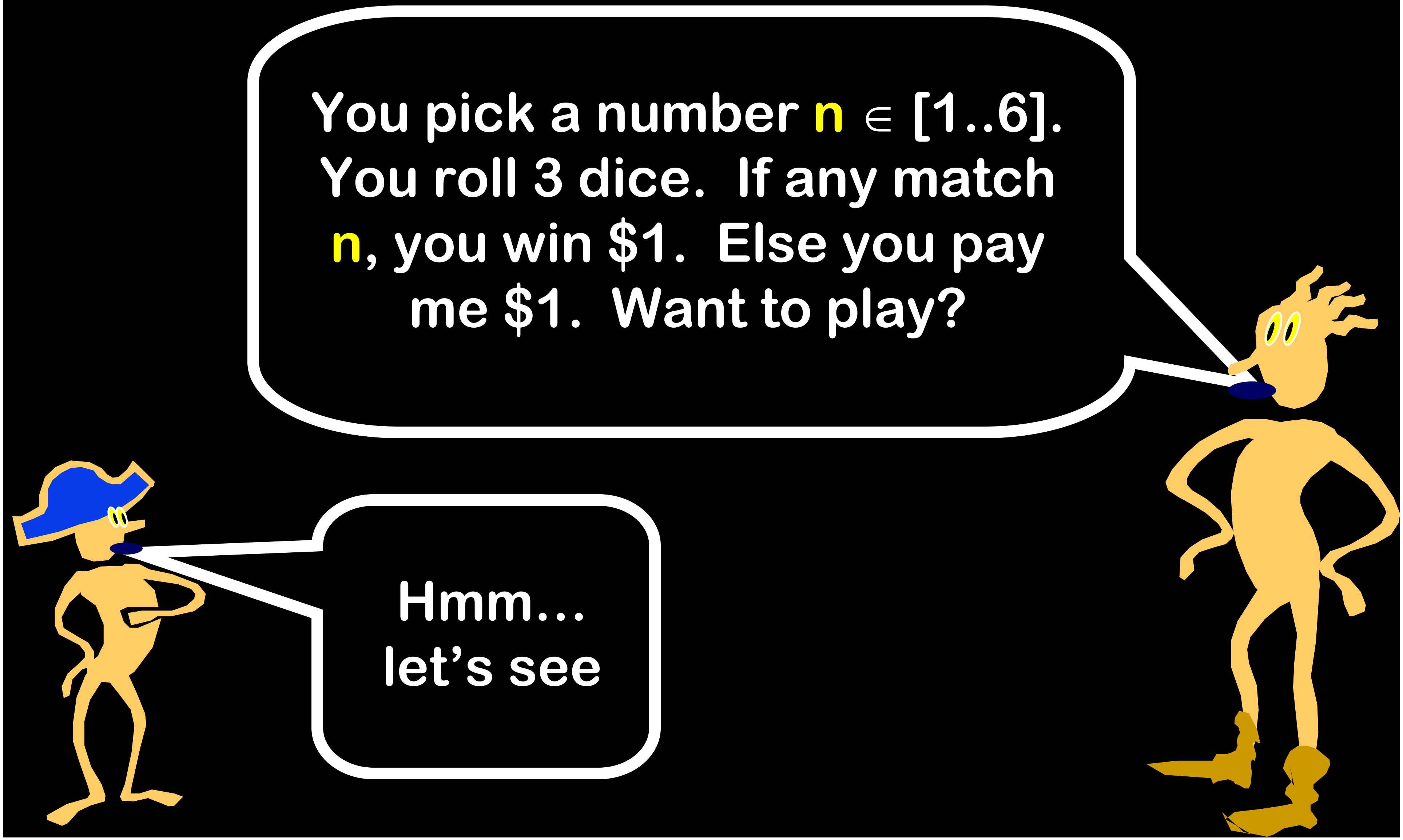
X = number of pairs of people with the same birthday

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

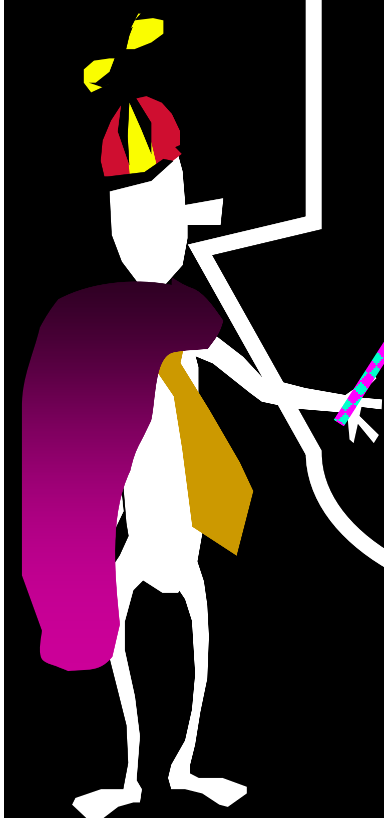
$$E[X_{jk}] = (1/366)\ 1 + (1 - 1/366)\ 0$$
$$= 1/366$$

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

Step Right Up…



Analysis



A_i = event that i-th die matches

X_i = indicator RV for A_i

Expected number of dice that match:

$E[X_1+X_2+X_3] = 1/6+1/6+1/6 = ½$

But this is not the same as

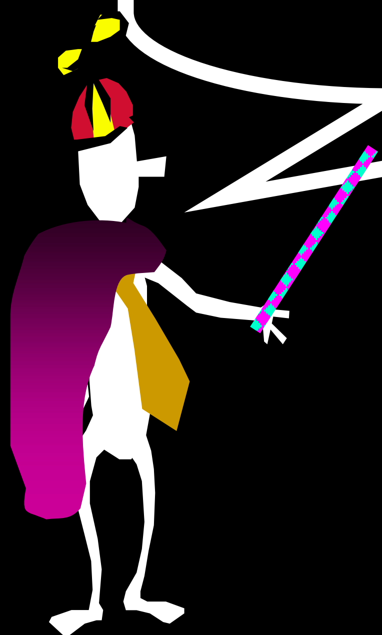
Pr(at least one die matches)

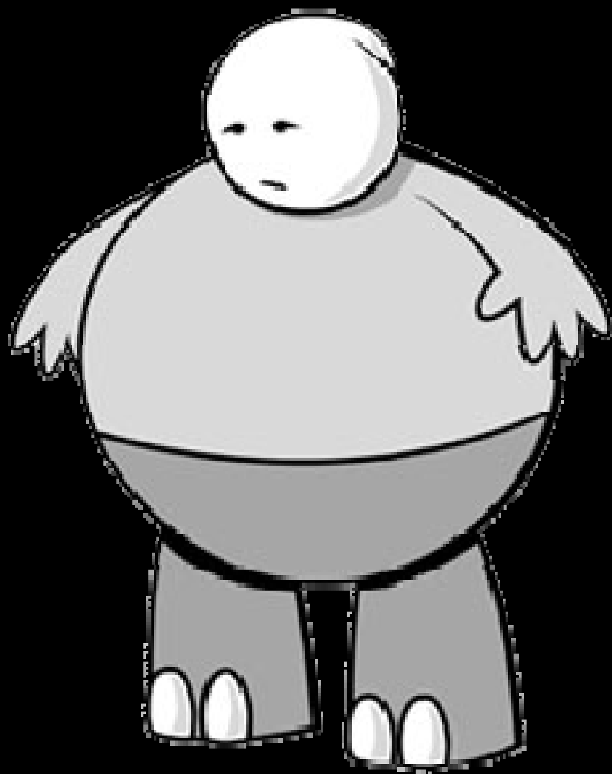
Analysis

Pr(at least one die matches)

$= 1 - \Pr(\text{none match})$

$= 1 - (5/6)^3 = 0.416$





Here's What You Need to Know…

Random Variables

Definition

Indicator r.v.s

Two Views of r.v.s

Expectation

Definition

Linearity

How to solve problems using r.v.s & expectations.