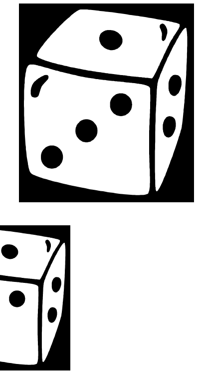
15-251

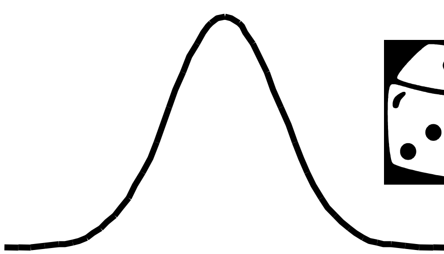
Great Theoretical Ideas in Computer Science

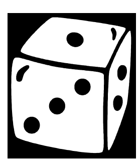
Some Puzzles



Probability Theory: Counting in Terms of Proportions

Lecture 10, September 25, 2008





Teams A and B are equally good

In any one game, each is equally likely to win

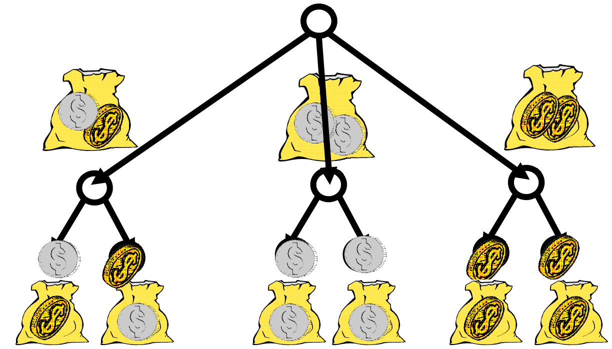
What is the most likely length of a “best of 7” series?

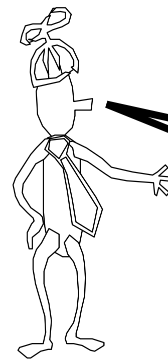
Flip coins until either 4 heads or 4 tails
Is this more likely to take 6 or 7 flips?

6 and 7 Are Equally Likely

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

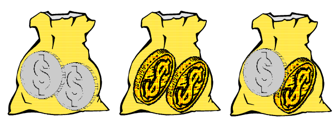
½ chance it ends 4 to 2; ½ chance it doesn't

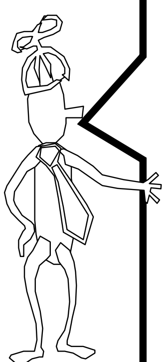




3 choices of bag
2 ways to order bag contents
6 equally likely paths

Silver and Gold

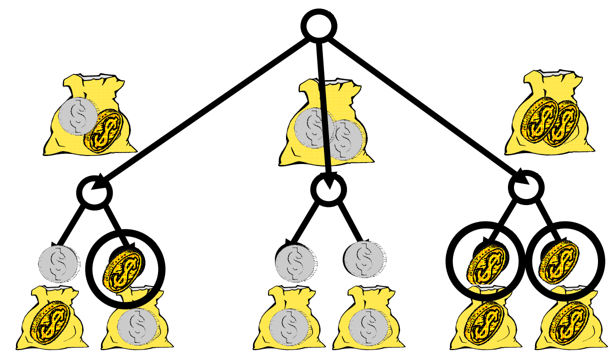


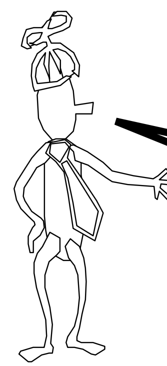


One bag has two silver coins, another has two gold coins, and the third has one of each

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?





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

? ?

So, sometimes, probabilities can be counter-intuitive

Finite Probability Distribution

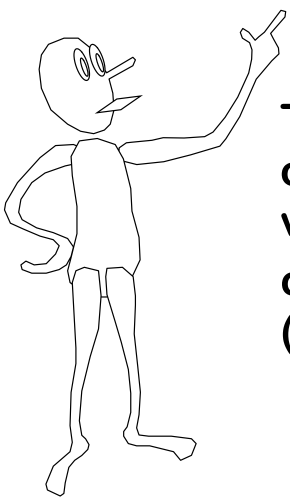
A (finite) probability distribution p is a finite set S of elements, together with a non-negative real weight, or probability p(x) for each element x in S

The weights must satisfy:

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

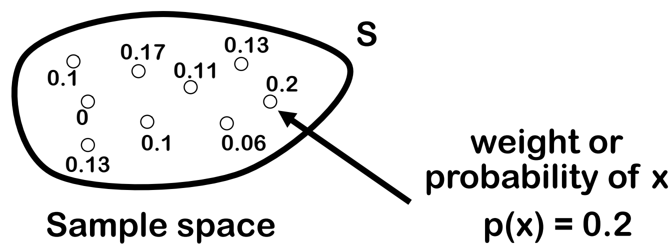
S is often called the sample space and elements x in S are called samples

Language of Probability



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

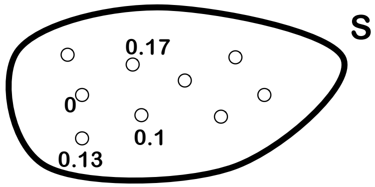
Sample Space



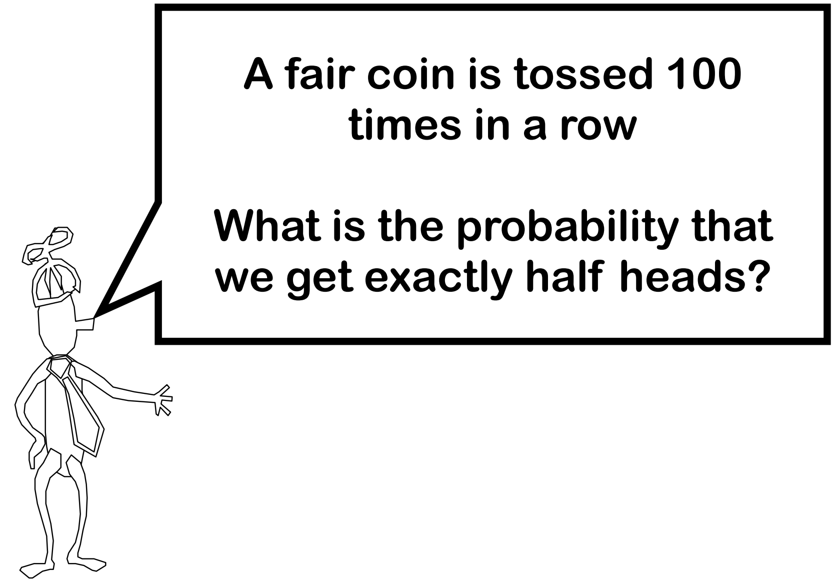
Events

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

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



$$Pr_D[E] = 0.4$$

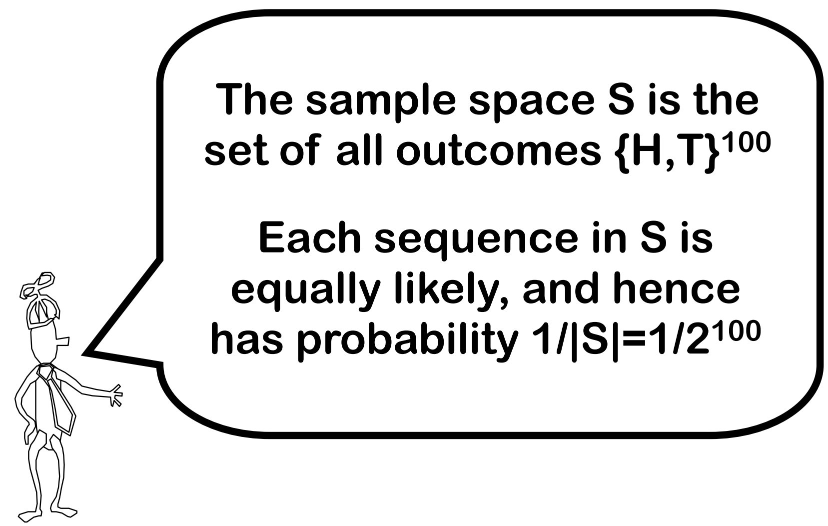


Uniform Distribution

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

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

Using the Language

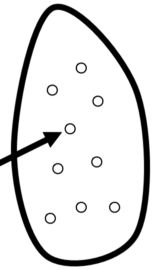


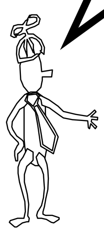
Visually

S = all sequences of 100 tosses

x = HHTTT……TH

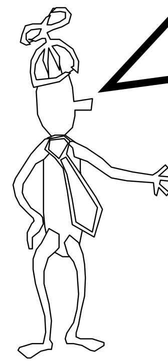
p(x) = 1/|S|





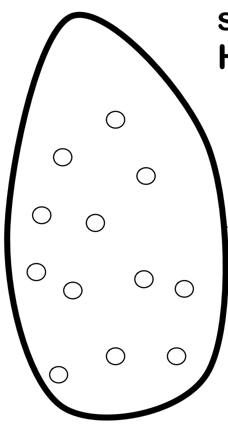
Suppose we roll a white die and a black die

What is the probability that sum is 7 or 11?



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

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



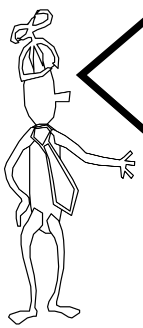
Probability of event E = proportion of E in S

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

Same Methodology!

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) }

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



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?

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

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

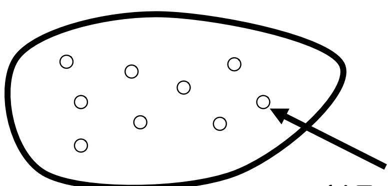
$|W| = 366^{23}$

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

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

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 $|E|$? Count $|\overline{E}|$ instead!

Sons of Adam

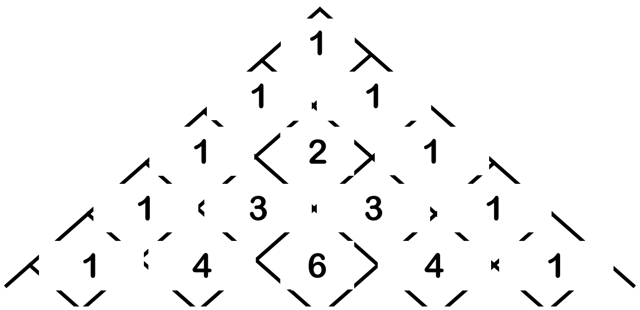
Adam was X inches tall

He had two sons:

One was X+1 inches tall

One was X-1 inches tall

Each of his sons had two sons …



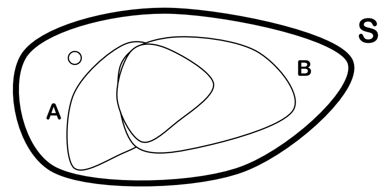
In the n^{th} generation there will be 2^n males, each with one of n+1 different heights:
$h_0, h_1, \ldots, h_n$

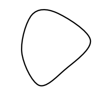
$h_i = (X-n+2i)$ occurs with proportion: $\binom{n}{i} / 2^n$

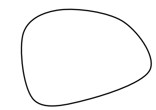
More Language Of Probability

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 $A \cap B$ 

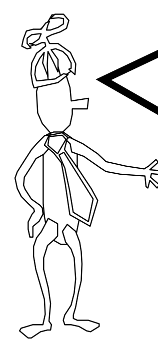
to B 

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



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) }

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

event A = {white die = 1}   event B = {total = 7}

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:

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

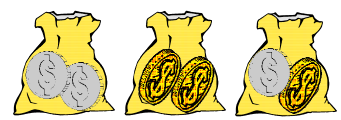
$\Pr[A_1 \mid A_2] = \Pr[A_1]$ $\Pr[A_1 \mid A_3] = \Pr[A_1]$
$\Pr[A_2 \mid A_1] = \Pr[A_2]$ $\Pr[A_2 \mid A_3] = \Pr[A_2]$
$\Pr[A_3 \mid A_1] = \Pr[A_3]$ $\Pr[A_3 \mid A_2] = \Pr[A_3]$

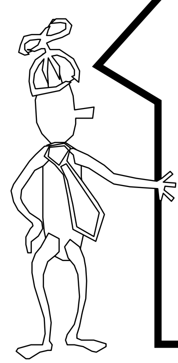
Independence!

A and B are independent events if

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

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

$Pr[G_1] = 1/2$

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

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

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random

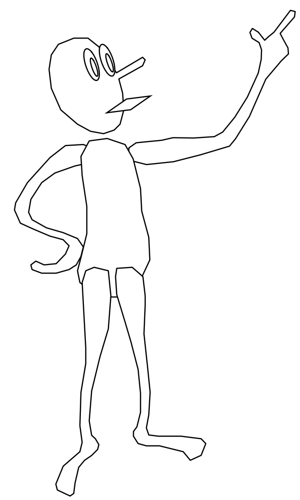
You select some door

Announcer opens one of others with no prize

You can decide to keep or switch

What to do?

Why Was This Tricky?

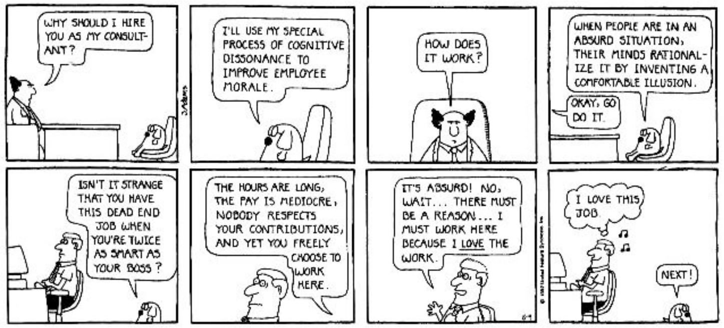


We are inclined to think:

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

But his action is not independent of yours!

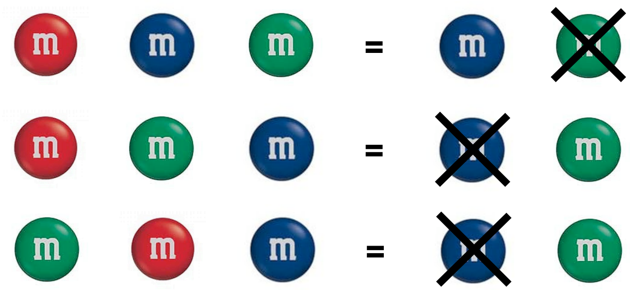
Cognitive Dissonance



Monty Meets Monkeys

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.

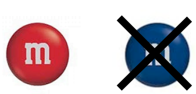


Monty Meets Monkeys

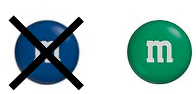
(from article by John Tierney)

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

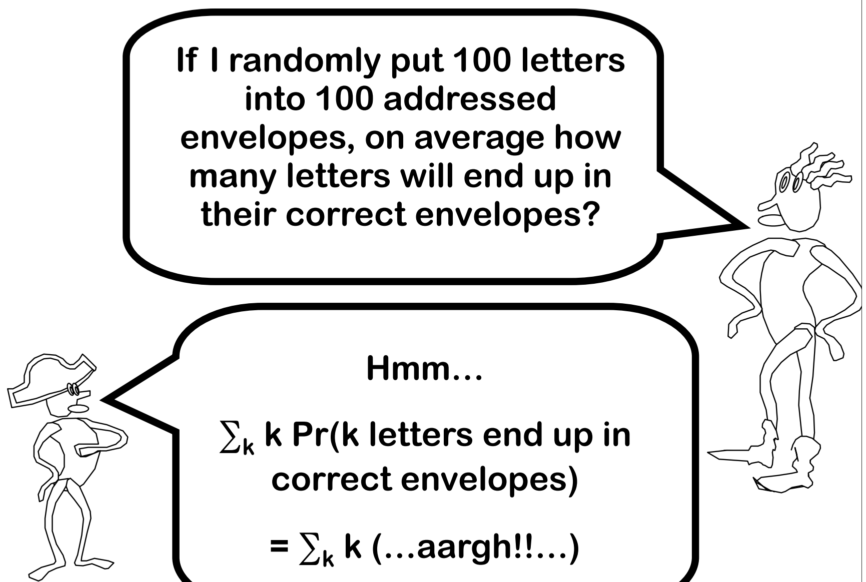
The monkey is given a choice of red or blue candy. It chooses red.



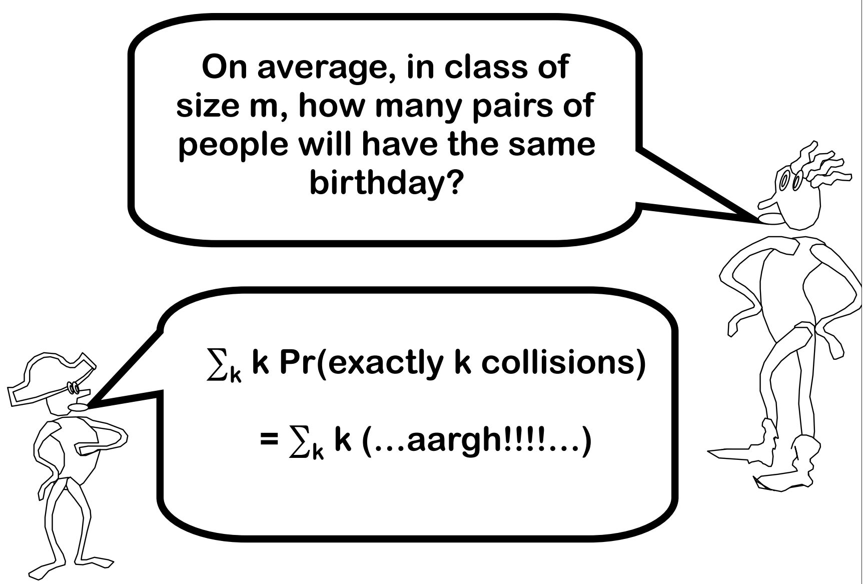
If the monkey is then given a choice of blue or green, it is more likely to choose green.



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



The new tool is called "Linearity of Expectation"



Random Variable

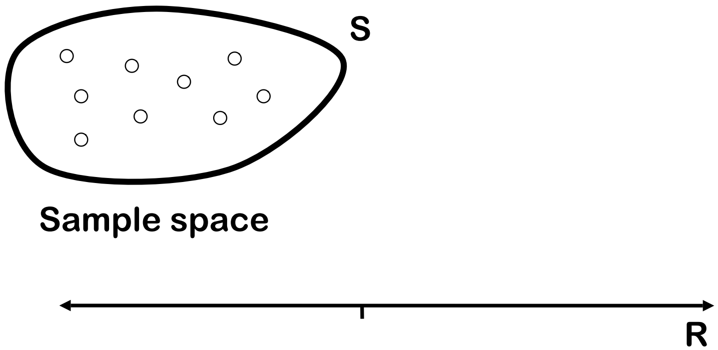
To use this new tool, we will also need to understand the concepts of Random Variable and Expectations

Basic, 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



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

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

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

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

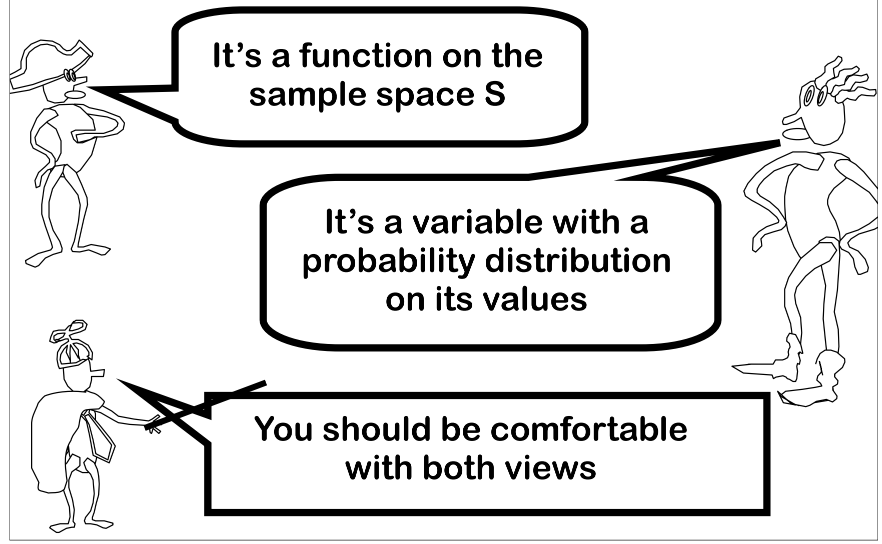
Input to the function is random

A function from S to the reals R

Or think of the induced distribution on R

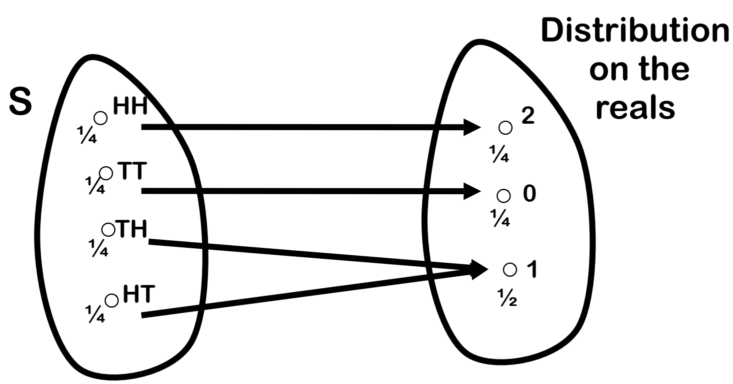
Randomness is “pushed” to the values of the function

It’s a Floor Wax And a Dessert Topping



Two Coins Tossed

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



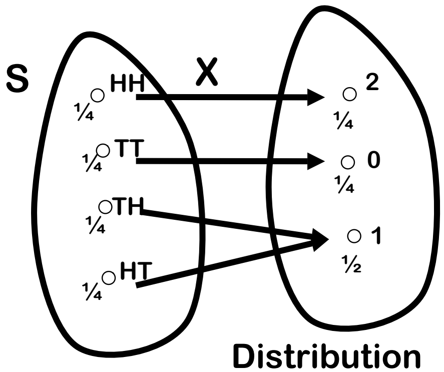
From Random Variables to Events

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

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

$\Pr(X = 1)$

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

$= \Pr(\{TH, HT\}) = ½$

Definition: Expectation

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

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

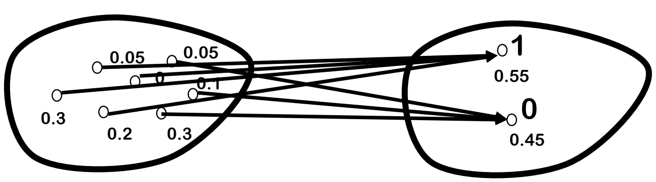
X is a function on the sample space S

X has a distribution on its values

From Events to Random Variables

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

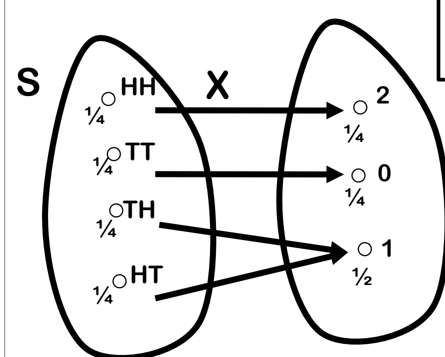
$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$



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

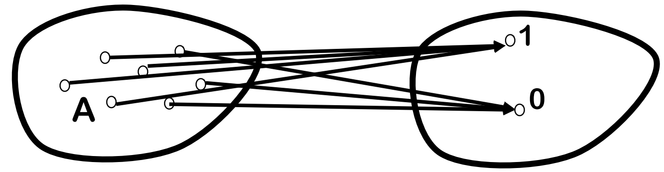
Moral: don't always expect the expected.
$\Pr[\, X = E[X] \,]$ may be 0 !

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

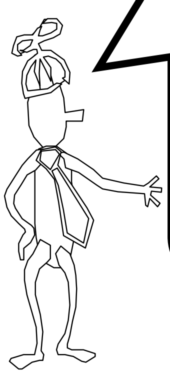
For any event A, can define the indicator random variable for A:

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

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



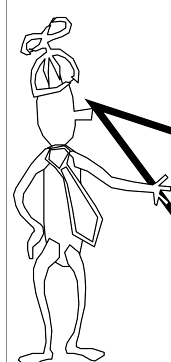
Type Checking



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

Adding Random Variables

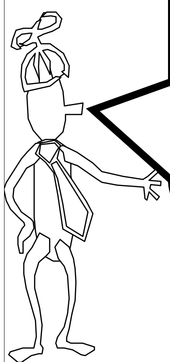


If X and Y are random variables (on the same set S), then
$Z = X + Y$ is also a random variable

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

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

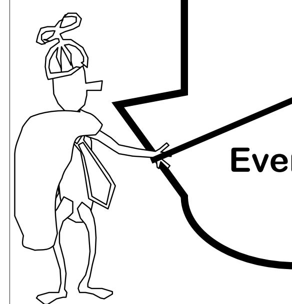
Adding Random Variables



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

Linearity of Expectation

Expectatus linearis

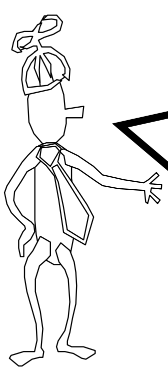


If Z = X+Y, then

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

Even if X and Y are not independent!

Independence

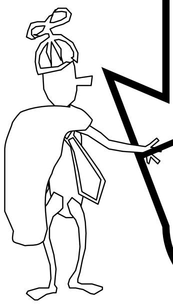


Two random variables X and Y are independent if for every a,b, the events X=a and Y=b are independent

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

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

Linearity of Expectation



E.g., 2 fair flips:
X = 1st coin is heads,
Y = 2nd coin is heads.
Z = X+Y = total # heads

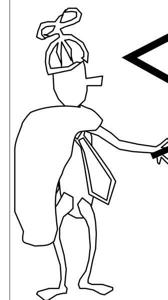
What is E[X]? E[Y]? E[Z]?

1,1,2 HH
1,0,1 HT
0,1,1 TH
0,0,0 TT

By Induction

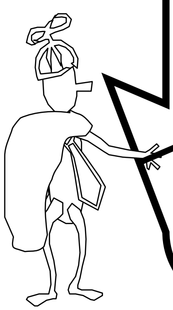
$$E[X_1 + X_2 + \ldots + X_n] =$$

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



The expectation of the sum
=
The sum of the expectations

Linearity of Expectation

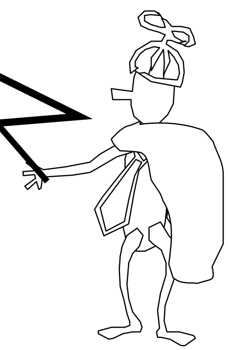


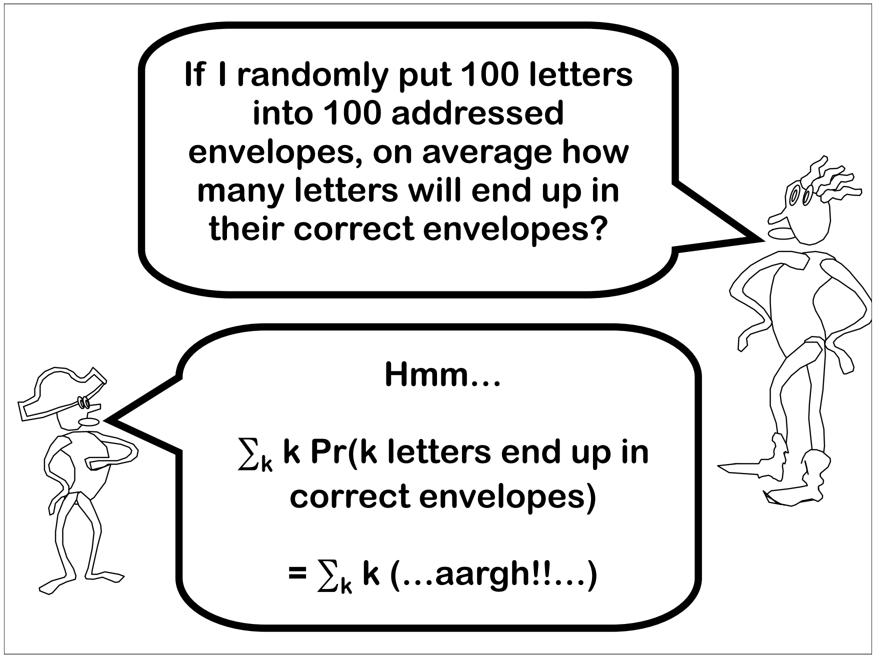
E.g., 2 fair flips:
X = at least one coin is heads
Y = both coins are heads, Z = X+Y

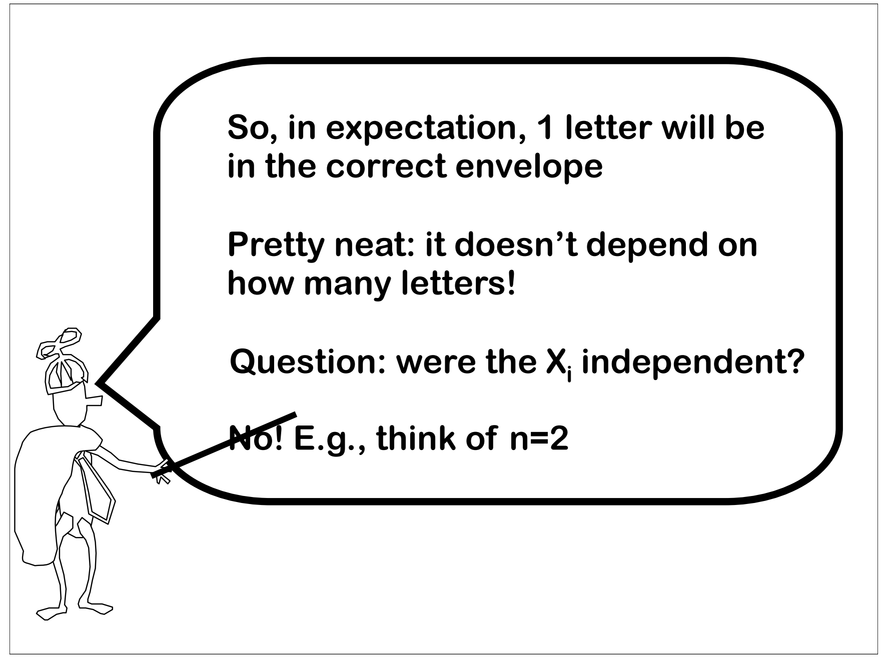
Are X and Y independent?
What is E[X]? E[Y]? E[Z]?

1,1,2 HH
1,0,1 HT
1,0,1 TH
0,0,0 TT

It is finally time to show off our probability prowess…





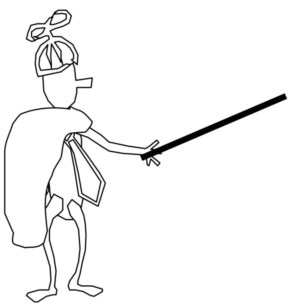


Use Linearity of Expectation

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

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



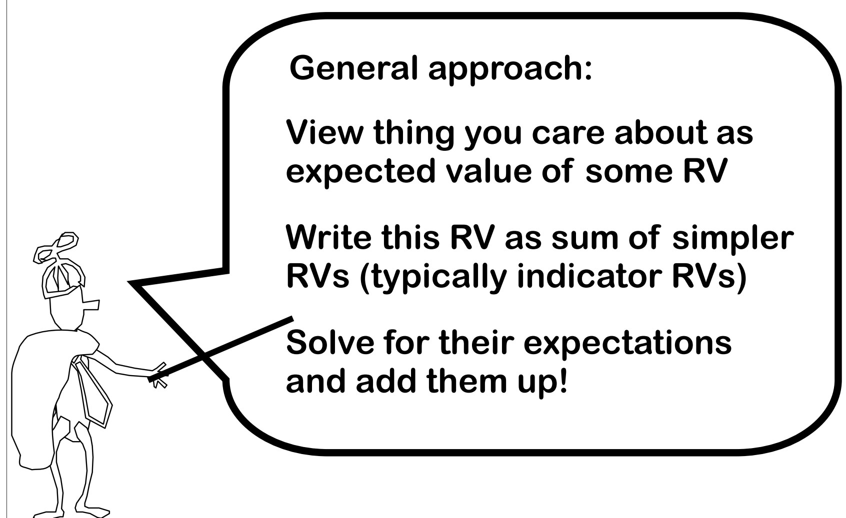
Let $Z = X_1 + \ldots + X_{100}$

We are asking for E[Z]

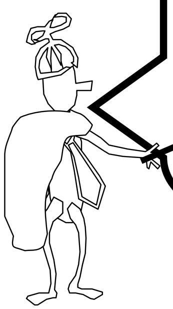
$E[X_i] = Pr(A_i) = 1/100$

So E[Z] = 1

Use Linearity of Expectation



Example



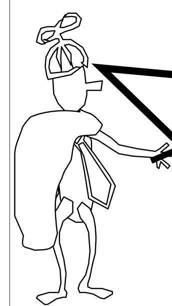
We flip n coins of bias p. What is the expected number of heads?

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!

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

Linearity of Expectation!

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

Break X into n simpler RVs:

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

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

But It's True If RVs Are Independent

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

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

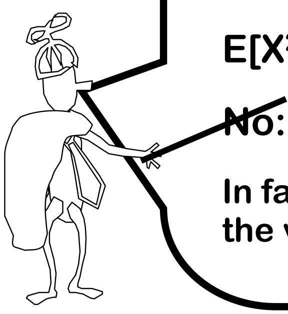
$E[XY] = \sum_c c \times \Pr(XY = c)$

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

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

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

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



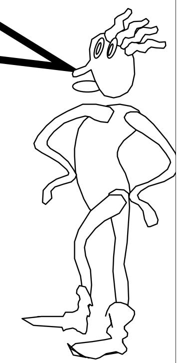
Example: 2 fair flips
X = indicator for 1st coin heads
Y = indicator for 2nd coin heads
XY = indicator for "both are heads"

$E[X] = ½, E[Y] = ½, E[XY] = ¼$

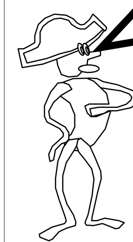
$E[X^2] = E[X]^2$?

No: $E[X^2] = ½, E[X]^2 = ¼$

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

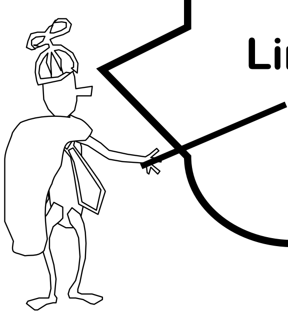


On average, in class of size m, how many pairs of people will have the same birthday?



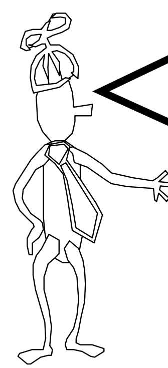
$\sum_k$ k Pr(exactly k collisions)

= $\sum_k$ k (…aargh!!!!…)



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

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

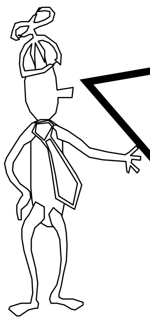


Use linearity of expectation

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] = ?



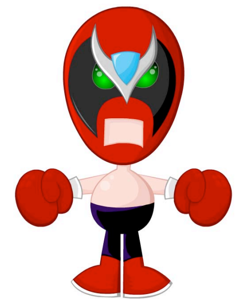
X = number of pairs of people with the same birthday

$E[X] = ?$

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

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



Here's What You Need to Know…

Language of Probability

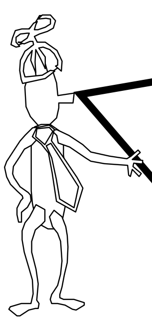
Sample Space
Events
Uniform Distribution
Pr [A | B]
Independence

Binomial Distribution

Definition

Random Variables

Two views
Expectation
Linearity of expectation



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

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

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