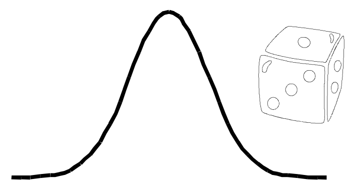
15-251

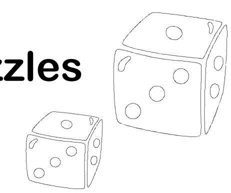
Great Theoretical Ideas in Computer Science

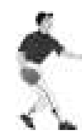
Probability Theory I

Lecture 11 (September 29, 2009)



Some Puzzles





Teams A and B are equally good

In any one game, each is equally likely to win

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

Flip coins until either 4 heads or 4 tails

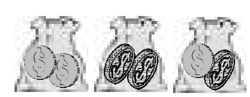
Is this more likely to take 6 or 7 flips?

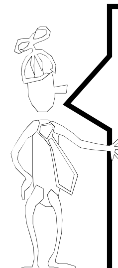
6 and 7 Are Equally Likely

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

Silver and Gold

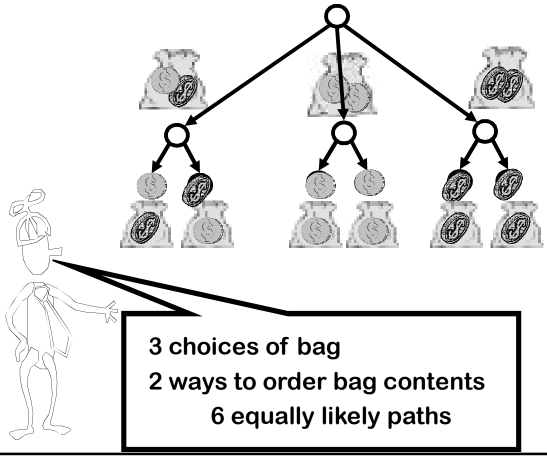


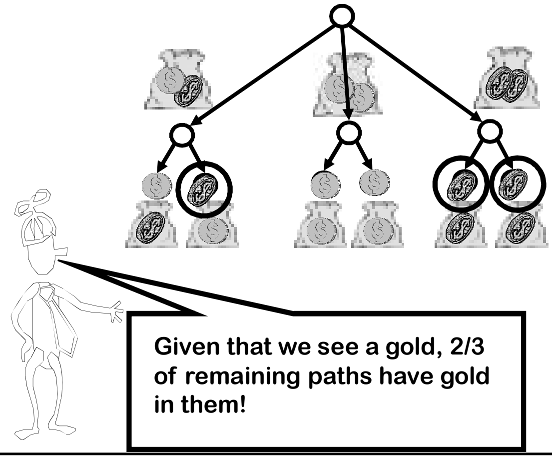


A bag has two silver coins, another has two gold coins, and the third has one of each

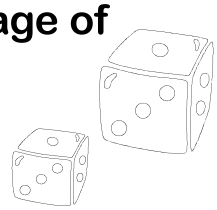
One bag is selected at random. One coin from it is selected at random. It turns out to be gold

What is the probability that the other coin is gold?

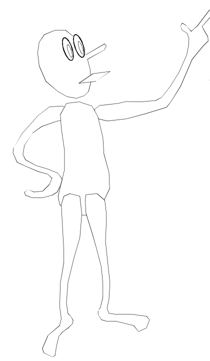




The Language of Probability



Language of Probability



The formal language of probability is a very important tool in describing and analyzing probability distribution

Finite Probability Distribution

A (finite) probability distribution D is a finite set S of elements, where each element t in S has a non-negative real weight, proportion, or probability p(t)

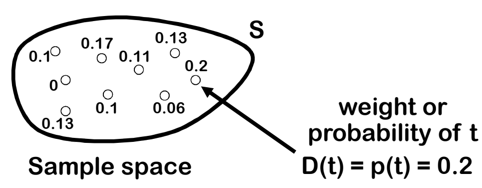
The weights must satisfy:

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

For convenience we will define D(t) = p(t)

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

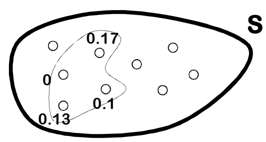
Sample Space



Events

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

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

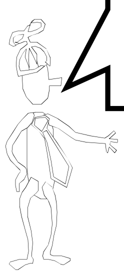


$$Pr_D[E] = 0.4$$

Uniform Distribution

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

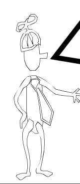
$$Pr_D[E] = \sum_{t \in E} p(t) = \frac{|E|}{|S|}$$



A fair coin is tossed 100 times in a row

What is the probability that we get exactly half heads?

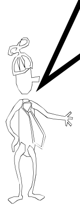
Using the Language



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

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

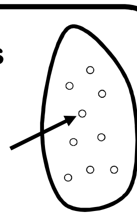
Visually

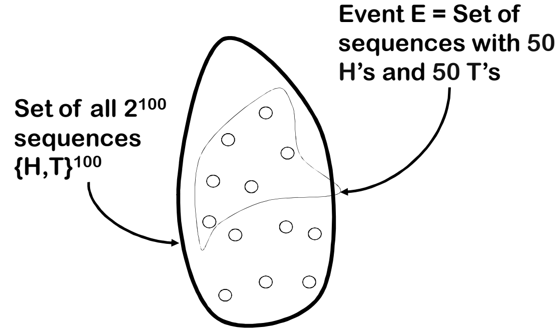


S = all sequences of 100 tosses

t = HHTTT……TH

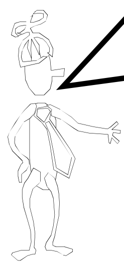
p(t) = 1/|S|





Probability of event E = proportion of E in S

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



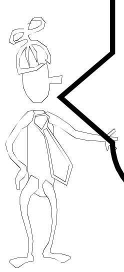
Suppose we roll a white die and a black die

What is the probability that sum is 7 or 11?

Same Methodology!

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



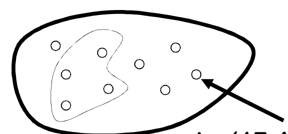
23 people are in a room

Suppose that all possible birthdays are equally likely

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

And The Same Methods Again!

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



t = (17,42,363,1,…, 224,177)

23 numbers

Event E = { t $\in$ W | two numbers in t are same }

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

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

$|\overline{E}|$ = (366)(365)…(344)

$|W| = 366^{23}$

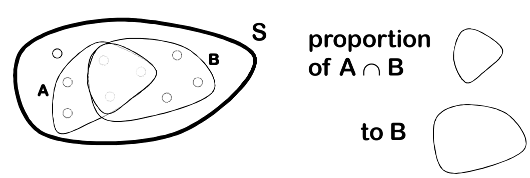
$$\frac{|\overline{E}|}{|W|} = 0.494\ldots$$

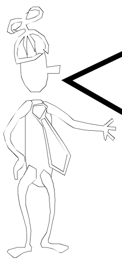
$$\frac{|E|}{|W|} = 0.506\ldots$$

More Language Of Probability

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

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





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}

Independence!

A and B are independent events if

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

$\Leftrightarrow$

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

$\Leftrightarrow$

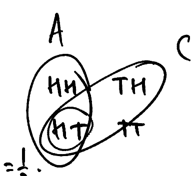
$\Pr[B \mid A] = \Pr[B]$

Two fair coins are flipped

A = {first coin is heads}

C = {two coins have different outcomes}

Are A and C independent? Yes!



Pr[A] = $\frac{1}{2}$

Pr[C] = $\frac{1}{2}$

Pr[A | C] = $\frac{\Pr[A \cap C]}{P[C]} = \frac{1/4}{1/2} = \frac{1}{2} = \Pr[A]$

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:

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

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

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

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

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

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

Two coins are flipped

A = {first coin is heads}

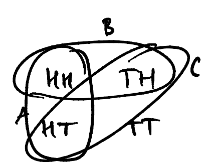
B = {second coin is heads}

C = {two coins have different outcomes}

A&B independent?

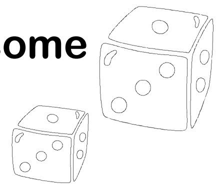
A&C independent?

B&C independent?

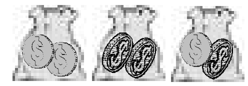


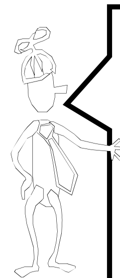
A&B&C independent? $\Pr[C \mid A \cap B] = 0 \neq \Pr[C]$

Let's (re)solve some problems….



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

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?

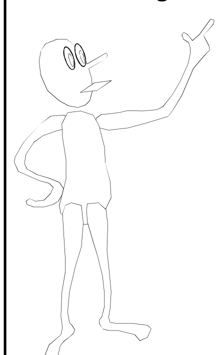
Monty Hall Problem

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 choose the correct door	we win if we choose the incorrect door
Pr[choosing correct door] = 1/3	Pr[choosing incorrect door] = 2/3

Why Was This Tricky?



We are inclined to think:

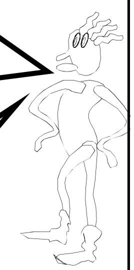
"After one door is opened, others are equally likely…"

But his action is not independent of yours!

Next, 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?

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



Pretty messy with direct counting…

The new tool is called "Linearity of Expectation"

But first, we need to pin down two concepts:
Random Variable and Expectation

Random Variable

Let S be sample space in a probability distribution
A [real valued] Random Variable is a real-valued function on S
Examples:

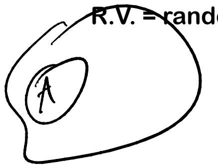
X = value of white die in a two-dice roll
X(3,4) = 3, X(1,6) = 1
Y = sum of values of the two dice
Y(3,4) = 7, Y(1,6) = 7

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

A function from S to the reals R

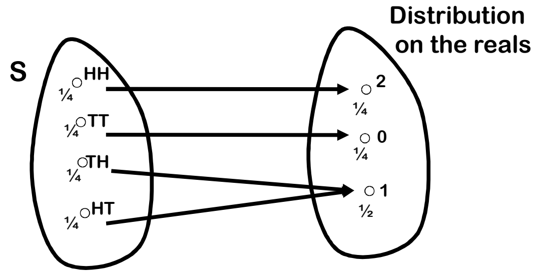
Input to the function is random

Or think of the induced distribution on R

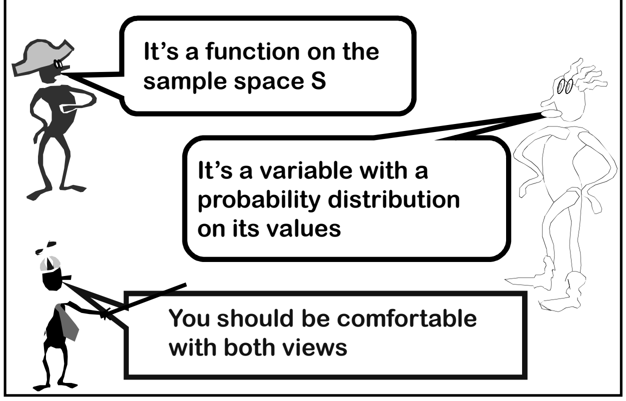
Randomness is "pushed" to the values of the function

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



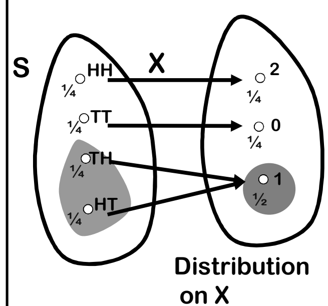
From Random Variables to Events

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

Pr(A) = Pr(X=a) = Pr({t ∈ S| X(t)=a})

Two Coins Tossed

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



Pr(X = a) = Pr({t ∈ S| X(t) = a})

Pr(X = 1)

= Pr({t ∈ S| X(t) = 1})

= Pr({TH, HT}) = ½

Definition: Expectation

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

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

X is a function on the sample space S

X has a prob. distribution on its values

A Quick Calculation…

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

E[X] = (1/8)×0 + (3/8)×1 + (3/8)×2 + (1/8)×3 = 1.5

But Pr[X = 1.5] = 0

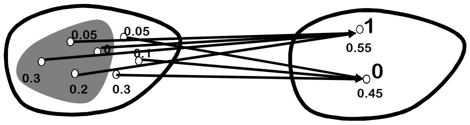
Moral: don't always expect the expected. Pr[X = E[X]] may be 0 !

From Events to Random Variables

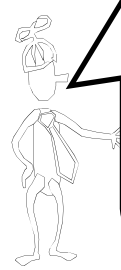
For any event A, can define the indicator random variable for A:

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

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



Type Checking

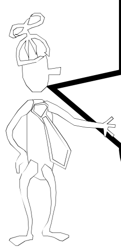


P(B) B must be an event

E(X) X must be a R.V.

cannot do P(R.V.) or E(event)

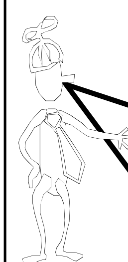
Independence



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?

Adding Random Variables

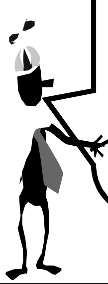


If X and Y are random variables (on the same set S), then Z = X + Y is also a random variable

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

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

Linearity of Expectation



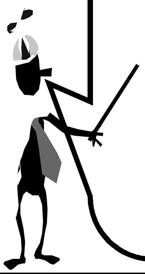
If Z = X+Y, then

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

Even if X and Y are not independent

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

Linearity of Expectation



E.g., 2 fair flips:

X = heads in 1st coin,

Y = heads in 2nd coin

Z = X+Y = total # heads

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

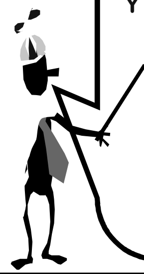
1,1,2 HH

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

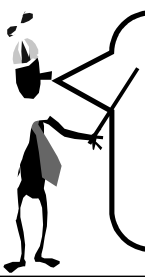
1,0,1 TH

0,0,0 TT

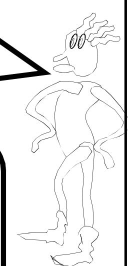
By Induction

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

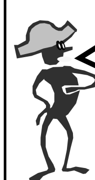
$E[X_1] + E[X_2] + \ldots. + E[X_n]$



The expectation of the sum = The sum of the expectations



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



Hmm…

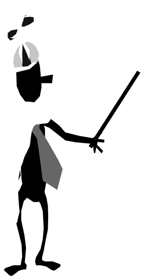
$\sum_k$ k Pr(exactly k letters end up in correct envelopes)

= $\sum_k$ 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



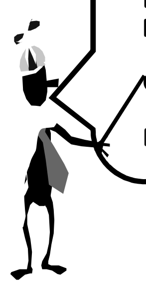
$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

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



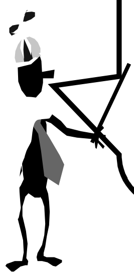
So, in expectation, 1 letter will be in the same correct envelope

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



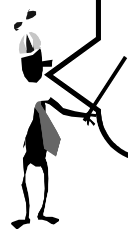
General approach:

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

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

$= n \cdot p$

But now we know a better way!

$\Pr[\text{get exactly } i \text{ heads}] = \binom{n}{i} p^i (1-p)^{n-i}$

Linearity of Expectation!

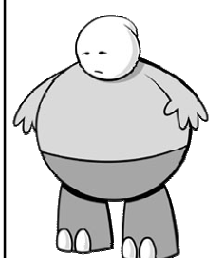
Let X = number of heads when n independent coins of bias p are flipped

Break X into n simpler RVs:

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

$E[X] = E[\sum_i X_i] = np$

$= \sum_i E[X_i] = \sum_i p.$



Here's What You Need to Know…

Language of Probability

- Events
- Pr [A | B]
- Independence

Random Variables

- Definition
- Indicator R.V.'s
- Two Views of R.V.s
- Events vs R.V.'s

Expectation

- Definition
- Linearity