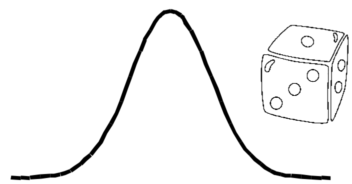
15-251

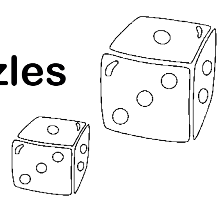
Great Theoretical Ideas in Computer Science

Probability Theory: Counting in Terms of Proportions

Lecture 10 (September 27, 2007)



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?

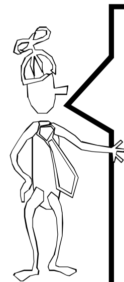
Flip coins until either 4 heads or 4 tails
Is this more likely to take 6 or 7 flips?

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

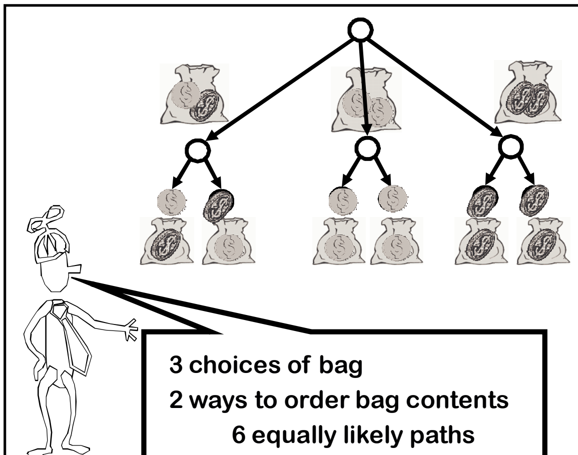
Silver and Gold

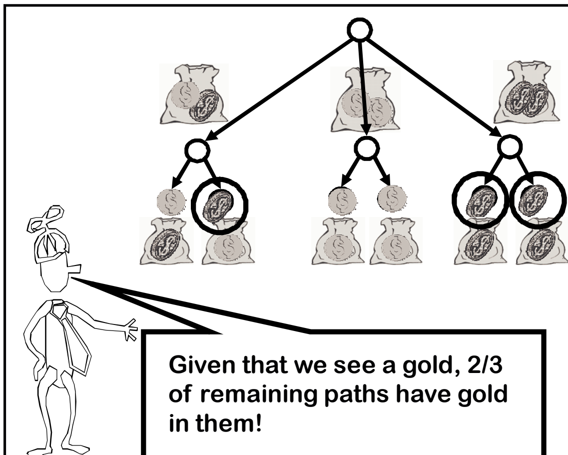


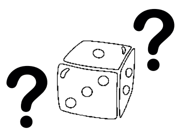
One bag has two silver coins, another has two gold coins, and the third has one of each

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?

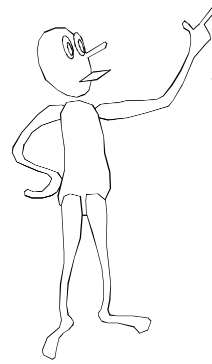






So, sometimes, probabilities can be counter-intuitive

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 x in S has a positive real weight, proportion, or probability p(x)

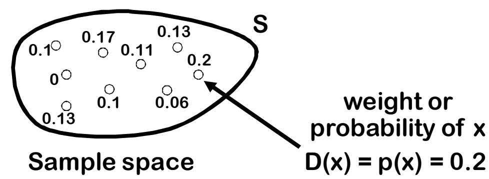
The weights must satisfy:

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

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

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

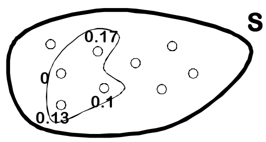
Sample Space



Events

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

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

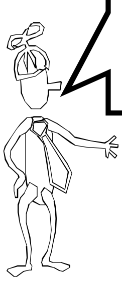


$Pr_D[E] = 0.4$

Uniform Distribution

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

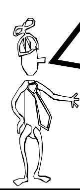
$$Pr_D[E] = \sum_{x \in E} p(x) = \frac{|E|}{|S|}$$



A fair coin is tossed 100 times in a row

What is the probability that we get exactly half heads?

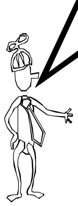
Using the Language



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

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

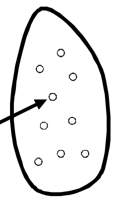
Visually



S = all sequences of 100 tosses

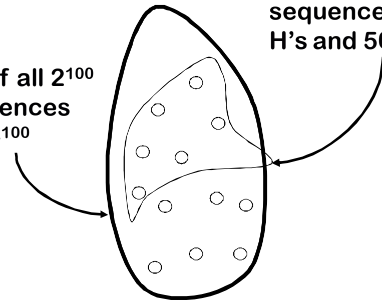
x = HHTTT……TH

p(x) = 1/|S|



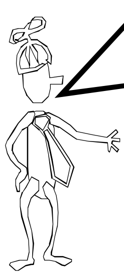
Event E = Set of sequences with 50 H's and 50 T's

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



Probability of event E = proportion of E in S

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



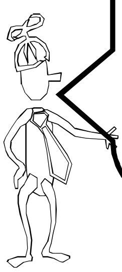
Suppose we roll a white die and a black die

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



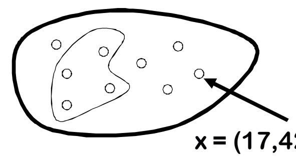
23 people are in a room

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



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

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

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

$|W| = 366^{23}$

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

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

The Descendants of Adam

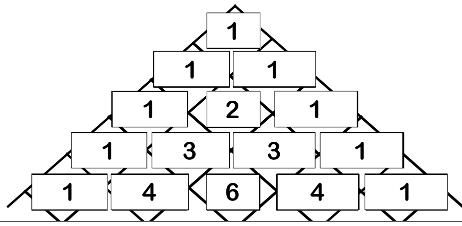
Adam was X inches tall

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: $\binom{n}{i} / 2^n$

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

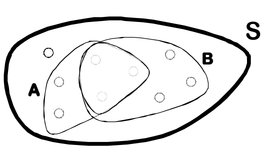
$$\frac{\binom{n}{i}}{2^n}$$

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

More Language Of Probability

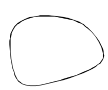
The probability of event A given event B is written Pr[A | B] and is defined to be =

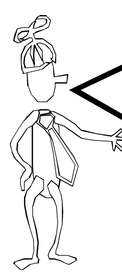
$$\frac{\Pr[A \cap B]}{\Pr[B]}$$



proportion of $A \cap B$

to B





Suppose we roll a white die and black die

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

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

Independence!

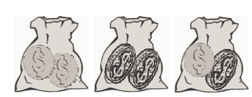
$A_1, A_2, \ldots, A_k$ are independent events if knowing if some of them occurred does not change the probability of any of the others occurring

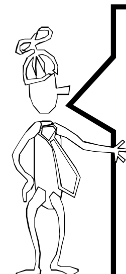
E.g., $\{A_1, A_2, A_3\}$ are independent events if:

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

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

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

$Pr[G_1] = 1/2$

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?

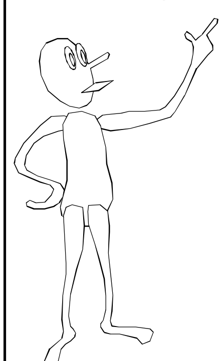
Monty Hall Problem

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 chose the correct door	we win if we chose 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!

Study Bee

Binomial Distribution

Definition

Language of Probability

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