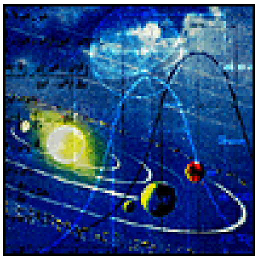
Great Theoretical Ideas In Computer Science
Victor Adamchik CS 15-251
Carnegie Mellon University

Probability - I




The Story

The theory of probability was originally developed by a French mathematician Blaise Pascal.
Pascal had a friend, fond of gambling, who asked him for strategies.

Probability was invented to analyze gambling.

This is not "Great Theoretical Ideas in Gambling", but in Computer Science.

Plan

Sample and Events
Conditional Probability
Bayes Law
Law of Total Probability
Use of Generating Functions

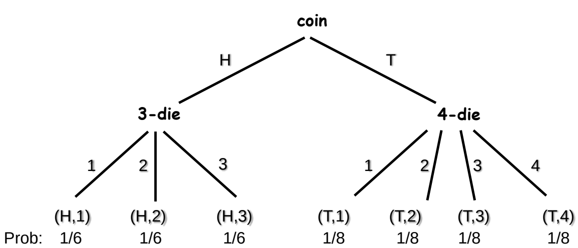
Example

Mary flips a fair coin. If it's a head, she rolls a 3-sided die. If it's a tail, she rolls a 4-sided die.

What is the probability die roll is 3 or higher?

Draw a probability tree.

A Probability Tree



Label the leaves with **"outcomes"**

Under each, write its probability: multiply along the path

Outcome:

A leaf in the probability tree.
I.e., a possible sequence of values of all calls to generators in an execution.

Sample Space:

The <u>set</u> of all outcomes.
E.g., { (H,1), (H,2), (H,3), (T,1), (T,2), (T,3), (T,4) }

Probability:

Each outcome has a nonnegative probability.
Sum of all outcomes' probabilities always 1.

Example

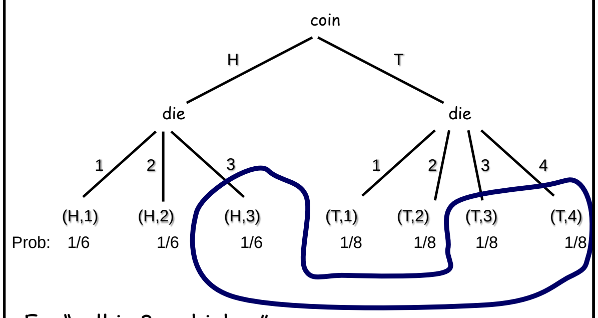
Mary flips a fair coin. If it's heads, she rolls a 3-sided die. If it's tails, she rolls a 4-sided die.

What is the probability die roll is 3 or higher?

Event:

A <u>subset</u> of outcomes.
In our example, E = { (H,3), (T,3), (T,4) }.

Pr[E] = sum of the probabilities of the outcomes in E.



E = "roll is 3 or higher"

Pr[E] = 1/6 + 1/8 + 1/8 = 5/12

We will consider experiments with a <u>finite</u> number of possible outcomes $w_1, w_2, \ldots, w_n$

The sample space Ω of the experiment is the set of all possible outcomes.

The roll of a die: $\Omega = \{1,2,3,4,5,6\}$

Each subset of a sample space is defined to be an event.

The event: $E = \{2,4,6\}$

Probability of a event

Let X be a random variable which denotes the value of the outcome of a certain experiment.

We will assign probabilities to the possible outcomes of an experiment.

We do this by assigning to each outcome w_j a nonnegative number $p(w_j)$ in such a way that
$p(w_1) + \ldots + p(w_n) = 1$

The function $p(w_j)$ is called the probability distribution function of the random variable X, $Pr[X = w_k]$

Probabilities

For any subset E of Ω, we define the probability of E to be the number Pr[E] given by

$$Pr[E] = \sum_{w_k \in E} p(w_k)$$

event

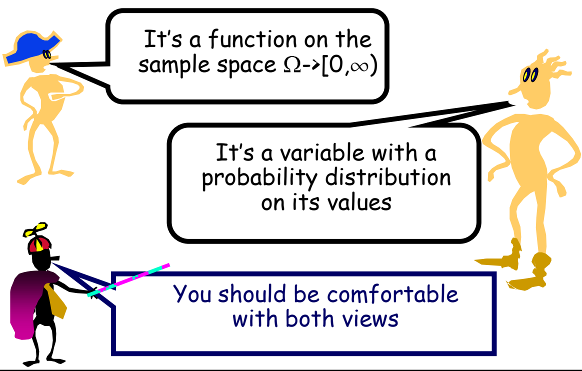
probabilities of outcomes

From Random Variables to Events

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

$Pr[E] = Pr[X=a] = Pr[\{t \in \Omega \mid X(t)=a\}]$

From Random Variables to Events



Random Variable

A coin is tossed ten times.

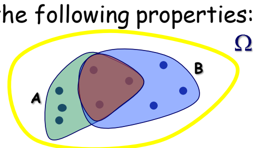
The random variable X is the number of tails that are noted.

X can only take the values 0, 1, ..., 10,

so X is a discrete random variable.

Theorem

The probabilities satisfy the following properties:



$Pr[\Omega] = 1$

$Pr[E] \geq 0$

$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$

$Pr[A^c] = 1 - Pr[A]$

$Pr[A] = Pr[A \cap B] + Pr[A \cap B^c]$

These should make sense if we think of events as sets.

Uniform Distribution

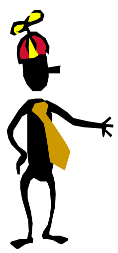
When a coin is tossed and the die is rolled, we assign an equal probability to each outcome.

The uniform distribution on a sample space containing n elements is the function p defined by

$$p(w) = 1/n$$

for every $w \in \Omega$

Example



Consider the experiment that consists of rolling a pair of standard dice.
What is the probability of getting a sum of 7 or a sum of 11?

We assume that each of 36 outcomes is equally likely.

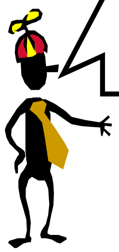
Example

E

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), F
(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) }

P(E)= 6 * 1/36   P(F)= 2 * 1/36

$P(E \cup F)$= 8/36



A fair coin is tossed 100 times in a row

What is the probability that we get exactly half heads?

Binomial Distribution

The sample space Ω is the set of all outcomes (sequences) $\{H,T\}^{100}$

Each sequence in Ω is equally likely, and hence has probability

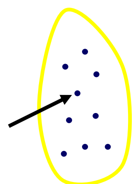
$1/|\Omega|=1/2^{100}$

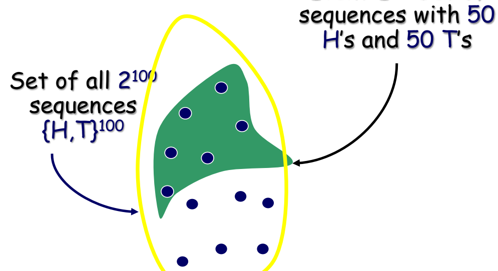
Visually

Ω = all sequences of 100 tosses

t = HHTTT……TH

P(t) = 1/|S|

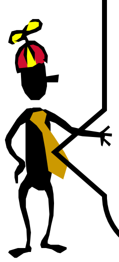




Probability of event E = proportion of E in S

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

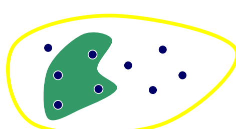
Birthday Paradox



How many people do we need to have in a room to make it a favorable bet (probability of success greater than 1/2) that two people in the room will have <u>the same birthday</u>?

Visually

Sample space $\Omega = 365^x$



We must find sequences that have no duplication of birthdays.

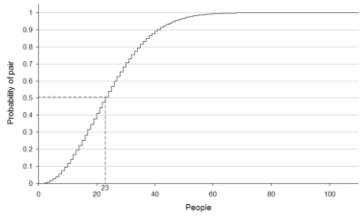
Event E = { $w \in \Omega$ | two numbers not the same }

$$\Pr[E]=1(1-\frac{1}{365})(1-\frac{2}{365})\ldots(1-\frac{x-1}{365})$$

Birthday Problem

Pr[in x students, some pair share a "bday"]=

$$1\text{-}(1-\frac{1}{365})(1-\frac{2}{365})\ldots(1-\frac{x-1}{365})$$



Birthday Problem

What if there are N possible "birthdays"?

Pr[in m students, some pair share a "bday"]

$$=1\text{-}(1-\frac{1}{N})(1-\frac{2}{N})\ldots(1-\frac{m-1}{N})$$

For what value of m is this ≈ 1/2 ?

$$m \approx \sqrt{N}$$

Cryptographic Hash Functions

1991: Rivest publishes MD5. (k=128)

1993: NSA publishes SHA-0. (k=160)

1995: NSA publishes SHA-1. (k=160)

SHA-1 now used in SSL, PGP, …

2001: NSA also introduces SHA-2

Birthday Attack

Imagine trying to find a collision for SHA-1:
Take a huge number of strings, hash them all,
hope that two hash to the same 160 bits.

This is like the Birthday Problem with $N = 2^{160}$!
So # tries before good chance of collision:

$$\approx \sqrt{2^{160}} = 2^{80} = 1208925819614629174706176$$

Birthday Attack

A crypto hash function is considered
"broken" if you can beat the birthday attack.

Prof. Xiaoyun Wang (王小云)

2005: SHA-1 collisions in $<2^{69}$

Later (w/ coauthors): in $<2^{63}$

SHA-1 = broken

Infinite Sample Spaces

A coin is tossed until the first time that a head turns up.

$\Omega = \{1, 2, 3, 4, \ldots\}$

A distribution function: $m(n) = 2^{-n}$.

$$\Pr = \sum_{w} m(w) = \frac{1}{2} + \frac{1}{4} + \ldots = 1$$

Infinite Sample Spaces

Let E be the event that the first time a head turns up is after an even number of tosses.

E = {2, 4, 6, …}

$$\Pr[E] = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \ldots = \frac{1}{3}$$

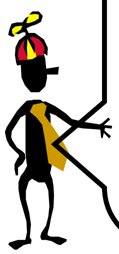
Infinite Sample Spaces

Suppose your experiment involves throwing a dart, which is equally likely to land anywhere in the interval [0,1]. What is the probability that the dart lands at exactly 0.5?

Probability of landing at any rational number is the same…

If that > 0, then the sum of all probabilities will be > 1. Thus, the probability that the dart lands at exactly 0.5 is zero.

Conditional Probability



Consider our voting example: three candidates A, B, and C are running for office. We decided that A and B have an equal chance of winning and C is only 1/2 as likely to win as A.

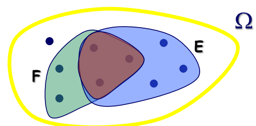
Suppose that before the election is held, A drops out of the race.
What are new probabilities to the events B and C

P(B|A)=2/3 P(C|A)=1/3

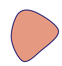
Conditional Probability

Definition. The probability of event F given event E is written P(F | E) and is defined by

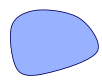
$$P(F \mid E) = \frac{P(F \cap E)}{P(E)}$$



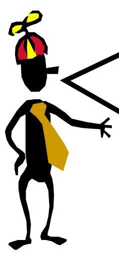
proportion of $F \cap E$



to E



We narrow our sample space to E.



Suppose we roll a white die and black die

What is the probability that the white die 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

$P(A \mid B) = P(A)$

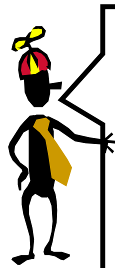
$\Leftrightarrow$

$P(A \cap B) = P(A)\,P(B)$

$\Leftrightarrow$

$P(B \mid A) = P(B)$

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

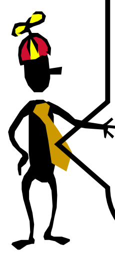
$Pr[G_2 \mid G_1] = Pr[G_1 \cap G_2] / Pr[G_1]$

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

$= 2/3$

Note: G_1 and G_2 are not independent

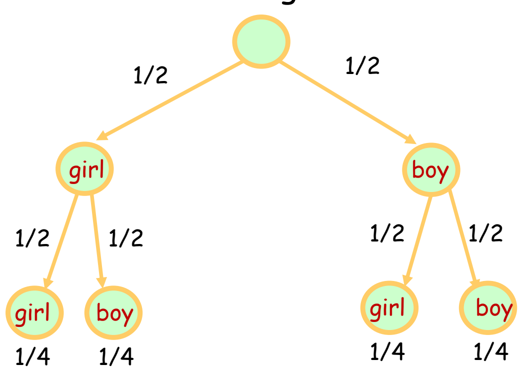
Another Paradox



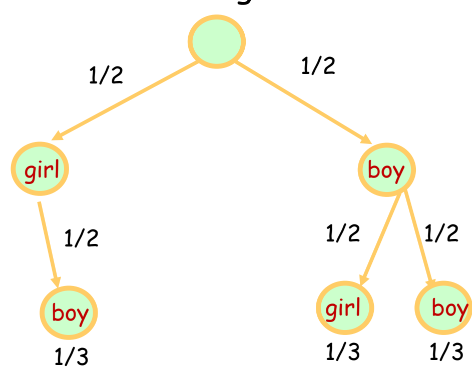
Consider a family with two children. Given that one of the children is a boy, what is the probability that both children are boys?

1/3

Tree Diagram



Tree Diagram



Monty Hall Problem

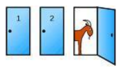
The host show hides a car behind one of 3 doors at random. Behind the other two doors are goats.

You select a door.

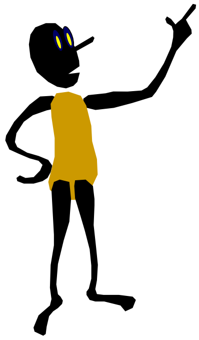
Announcer opens one of others with <u>no</u> prize.

You can decide to keep or switch.

What to do? To switch or not to switch?



This is Tricky!



We are inclined to think:

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

Monty Hall Problem

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

Each has probability 1/3

Staying

we win if we choose the correct door

Pr[choosing correct door] = 1/3

Switching

we win if we choose the incorrect door

Pr[choosing incorrect door] = 2/3

Monty Hall Problem

Let the doors be called X, Y and Z.

Let Cx, Cy, Cz be the events that the car is behind door X and so on.

Let Hx, Hy, Hz be the events that the host opens door X and so on.

Supposing that you choose door X, the possibility that you win a car if you switch is

$$\Pr[Hy \cap Cz] + P[Hz \cap Cy] =$$

$$\Pr[Hy|Cz]\, P[Cz] + \Pr[Hz|Cy]\, \Pr[Cy] =$$

$$1 \times 1/3 + 1 \times 1/3 = 2/3$$

Bayes' formula

Sometimes we need to know Pr[F|E], but all we know is the reverse direction Pr[E|F].

Theorem.

$$P[F\,|\,E] = \frac{\Pr[E\,|\,F]\Pr[F]}{\Pr[E]}$$

Proof.

$$\Pr[F\,|\,E] = \frac{\Pr[F \cap E]}{\Pr[E]} = \frac{\Pr[E \cap F]}{\Pr[E]} = \frac{\Pr[E\,|\,F]\Pr[F]}{\Pr[E]}$$

Law of Total Probability

Theorem. Let $F_1, F_2, \ldots, F_n$ partition of a sample space. Then

$$\Pr[E] = \sum_{k=1}^{n} \Pr[E \cap F_k] = \sum_{k=1}^{n} \Pr[E\,|\,F_k]\Pr[F_k]$$

Proof. Note $E \cap F_k$ are mutually exclusive

$$E = \bigcup_{k=1}^{n} E \cap F_k$$

Problem

A fair die is tossed and its outcome is denoted by X. After that, X independent fair coins are tossed and the number of heads obtained is denoted by Y.

Compute: Pr[Y=4]

Exercise

By the law of total probability

$$\Pr[Y=4]=\sum_{k=1}^{6}\Pr[Y\,|\,X_k]*\Pr[X_k]$$

$\Pr[X_k]$ is 1/6. Compute $\Pr[Y=4|X_k]$

$$\Pr[Y=4\,|\,X_k]=\binom{k}{4}/2^k$$

Use of Generating Functions

Problem

Compute the probability of rolling a sum of 18 on 4 standard dice.

Compute the probability of rolling a sum of 18 on 4 standard dice.

Create a generating polynomial for the problem

$$\left(\frac{x+x^2+\ldots+x^6}{6}\right)^4$$

Take the coefficient by x^{18}.

Note,

$$x+x^2+\ldots+x^6=x\frac{x^6-1}{x-1}$$

$$\frac{1}{(x-1)^4}=\sum_{k=0}^{\infty}\binom{3+k}{k}x^k$$

$$x^4(x^6-1)^4=x^6-4x^{10}+6x^{16}+\ldots$$

Thus, the coefficient by x^{18} is

$$\frac{1}{6^4}\left(\binom{3+14}{14}-4\binom{3+8}{8}+6\binom{3+2}{2}\right)=\frac{5}{81}$$

Here's What You Need to Know…

- Sample and Events
- Conditional Probability
- Bayes Law
- Law of Total Probability