RELATION

A relation is a math tool that describes associations between sets. This part focuses on binary relations.

Given sets A and B, their relation is a subset of their cross product;

$a \; f \; b$

domain
MIT Faculty

relation
in charge of

codomain
MIT courses whole

$a \sim_R b$

Equations have properties, those are symmetry, transitive, reflective and their opposites.
Relations that have the properties above are called equivalence relations and are particularly common.
A relation with a property, can't have its opposite, for example a symmetric relation can't be also asymmetric.
Equivalence equations falls in the partial order relations quite commonly

Reflexivity
likes itself

Transitivity
if x likes y and y likes z. then x likes z

Symmetry
$\forall x, y \in A$ xRy yRx
if x likes y then y likes x

Partitions

Given the equiv. relation $R: A \rightarrow A$, an equivalence class is a set related to a particular element x, by the relation R

A partition then, is a collection of disjoint, non empty subsets, whose union forms A. The subsets are called blocks of the partition

The equivalence classes of an equivalence relation on a set A form a partition of A

Partial orders.

Weak partial orders are transitive, asym., reflexive
strong ones are transitive, asymmetric and irreflexive. ~~Ref~~ Reflexivity makes the difference here.
These are denoted with $\preceq$ or $\sqsubseteq$ for weak
$\prec$ $\sqsubset$ are strong.
Partial orders are partial because there could be elements with no relation between them

Posets

Given a p order $\preceq$ on the set A, the pair $(A, \preceq)$ is a poset. In graph theory a poset is a directional graph with vertices A and edges $\preceq$

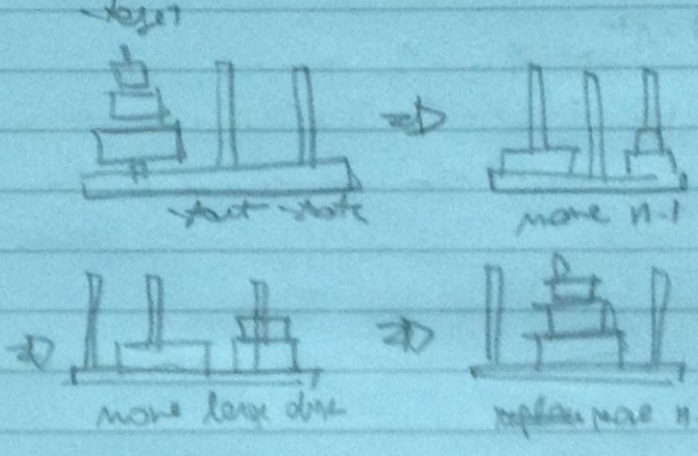
A poset is usually visualised with the Hasse diagram. Normal poset graphs ~~display~~ tend to contain several edges. Using Hasse we could ~~choose~~ choose to not to draw some of those.

Reading

TOWER OF HANOI

minimum #

Recursive solution T_n of [illegible] n



The goal is to move the others to another spot

part 1   #2   #3   upper bound

$$T_n = T_{n-1} + 1 + T_{n-1}$$

You can't move better than $2T_{n-1} + 1$ because you can't put a bigger disk on top of a smaller

T_n is recursively tied to its predecessors so what is the closed form

→ A simple way is to guess the solution and after, prove it

guess $2^n - 1$

CHECK

Prove by induction that $t_n = 2^{n-1}$

base $T_1 = 1$ ✓

ind step $t_n = t_{n-1} + 1$

$= (2^{n-1} - 1) + 1$

$= 2^n - 1$

$T_{64} = 2^{64} - 1 = 18$ bil bil steps

that was verify in class

PLUG N CHUG

The first step is to expand the recurrence equation by plugging (the recurrence) and chugging (simplify) until a pattern appears.

Too much can make the pattern hard to spot.

step 1 ~~[illegible]~~

$t_n = 2t_{n-1} + 1$

$= 2(2t_{n-2} + 1) + 1$ plug

$= 4t_{n-2} + 2 + 1$ chug

$= 4(2t_{n-3} + 1) + 2 + 1$ plug

~~$= 4(2t_{n-3} + 1) + 4 + 2 + 1$ chug~~

$= 8t_{n-3} + 4 + 2 + 1$ chug

→

$T_n = 2^k T_{n-k} + 2^{k-1} + 2^{k-2} + \dots + 2^2 + 2^1 + 2^0$

~~$T_n = 2^k T_{n-k} + 2^{k-1} + 2^{k-2} + \dots +$~~

$= 2^k T_{n-k} + 2^k - 1$

~~now we have a pattern, next is step~~

step 2

verify the general formula, one more round

$T_n = 2^k T_{n-k} + 2^k - 1$

$= 2^k(2T_{n-(k+1)} + 1) + 2^k - 1$

$= 2^{k+1} T_{n-(k+1)} + 2^{k+1} - 1$

the result is .. k+1, verified

step 3

figure out a closed form

$k = n - 1$ (such that T_{n-k} is T_1)

$t_n = 2^{n-1} T_1 + 2^{n-1} - 1$

$= 2^{n-1} \cdot 1 + 2^{n-1} - 1$

$= 2^n - 1$

We got the problem from opposite direction

MERGE SORT

Given ~~[illegible]~~ a list of n numbers.

The algorithm splits the list in two halves, then sorts them, ~~after combine them compare~~ ~~[illegible] the lowest values each~~

~~After that the list are merged while during the lowest value of each list~~

Both halves are sorted recursively. then they are merged creating a sorted list by ~~adding the lowest element or~~ ~~taking between the halves at each step~~ emitting the smaller from the two leading terms

Max # comparisons

$T_n =$ ~~maximum # of comparisons~~

$T_n = T_{n/2} + T_{n/2} + n - 1$ ← merge

order 1st half

order 2nd ~~[illegible]~~ half

GUESS AND VERIFY

$T_1 = 0$

$T_2 = 2T_1 + 2 - 1 = 1$

$T_4 = 2T_2 + 4 - 1 = 5$

$T_8 = 2T_4 + 8 - 1 = 17$

$T_{16} = 2T_8 + 16 - 1 = 49$

No clear pattern

PLUG AND CHUG

$T_n = 2T_{n/2} + n - 1$

$= 2(2T_{n/4} + \frac{n}{2} - 1) + (n-1)$

$= 4T_{n/4} + (n-2) + (n-1)$

$= 4(2T_{n/8} + n/4 - 1) + (n-2) + (n-1)$

$= 8T_{n/8} + (n-4) + (n-2) + n - 1)$

pattern is

$2^k T_{n/2^k} + kn + 1 - 2^k$

$= 2^k(2^k T_{n/2^k} + kn + 1 - 2^k) + kn - 2^k + 1$

$= 2^{k+1} T_{n/2^{k+1}} + (k+1)n - 2^{k+1} + 1$

$k+1$ ✓

write using known terms

if $k = \log n$ then $T_{n/2^k} = T_1$

$T_n = n\log n - n + 1$

~~LINEAR RECURRENCES~~

~~"cookbook" methods, follow the recipe and get the answer ~~formula the recurrence~~~~

~~We use these to solve recurrences~~

~~the total number of ways to climb~~
~~a stair n~~

~~$$f(n) = f(n-1) + f(n-2) \quad \text{for } n>2$$~~

~~This is a linear recurrence, called fibo recurrence.~~

~~A linear recurrence usually has ~~test is~~ the form~~

~~$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \dots a_d f(n-d)$$~~

DIVIDE AND CONQUER

Merge sort is an example of a divide and conquer algorithm.

Divide and conquer recurrences ~~for~~ have the form

$$T(n) = \sum_{i=1}^{k} a_i T(b_i n) + g(n)$$

where a is positive, b is ~~both~~ between 0 and 1, and g(n) is ~~non-negative~~ nonnegative

the solution is given by the Akra-Bazzi formula, which looks really ugly

$$T_n = O\left(n^2\left(1 + \int_1^n \frac{u^2}{u^3}\,du\right)\right)$$

$$= O(n^2(1 + \log n))$$

$$= O(n^2 \log n)$$

Graduate student job problem

total # jobs = m (fixed over time)

- Each professor generates 1 graduate (prof) per year

- Expect 1st year profs produce 0

- No retirements

When do the jobs get filled?

Boundary cond: 1st professor hired in year 1st

~~Ex~~ Let $f(n)$ = # profs during year n.

$f(0)=0$ $f(2)=1$ $f(4)=3$

$f(1)=1$ $f(3)=2$ $f(5)=5$

for $n \geq 2$ $f(n) = \underset{\text{prior}}{f(n-1)} + \underset{\text{new}}{f(n-2)}$

USED IN LOTS OF FIELDS

FIBO recurrence.

Has been studied to model the population growth of rabbits

A recurrence is linear if it's of the form

$$f(n) = a_1 f(n-1) + a_2 f(n-2) \ldots a_d f(n-d)$$

$$= \sum_{i=1}^{d} a_i f(n-i),$$ for fixed a_i, d

d is the order
(Fibo is order two)

solution
$f(n) = \alpha^n$ for constant α

$$f(n) = f(n-1) + f(n-2)$$

$$\alpha^n = \alpha^{n-1} + \alpha^{n-2}$$

$$\alpha^2 = \alpha + 1$$

$$\alpha^2 - \alpha - 1 = 0$$

$$\alpha = \frac{1 \pm \sqrt{1+4}}{2}$$

$$f(n) = \alpha_1^n \text{ or } \alpha_2^n \qquad \alpha_1 = \frac{1+\sqrt{5}}{2} \quad \alpha_2 = \frac{1-\sqrt{5}}{2}$$

golden ratio

If $f(n) = \alpha_1^n$ & $f(n) = \alpha_2^n$ are solutions to a linear recurrence (w/o satisfying start boundary cond), then $f(n) = c_1\alpha_1^n + c_2\alpha_2^n$ is also a solution for any constant c_1 & c_2

$\Rightarrow f(n) = c_1\left(\frac{1+\sqrt{5}}{2}\right) + c_2\left(\frac{1-\sqrt{5}}{2}\right)$ is a solution to fibo recurrence

Determine the constant

$$f(0) = 0 = c_1(\ )^0 + c_2(\ )^0$$

$$c_2 = -c_1$$

$$f(1) = 1 = c_1\left(\frac{1+\sqrt{5}}{2}\right) + c_2\left(\frac{1-\sqrt{5}}{2}\right)^1$$

$$= c_1\left(\frac{1+\sqrt{5}}{2}\right) - c_1\left(\frac{1-\sqrt{5}}{2}\right)$$

$$= c_1\left(\frac{2\sqrt{5}}{2}\right)$$

$$\Rightarrow c_1 = \frac{1}{\sqrt{5}}, \quad c_2 = \frac{-1}{\sqrt{5}}$$

$$f(n) = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

formula for the nth Fibo number

$0 = = =$

"Graduate problem"

All m tables filled when $f(n) \geq m$

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n + \delta(n) \geq m$$

$$\Rightarrow \left(\frac{1+\sqrt{5}}{2}\right) \geq \sqrt{5}\,(m - \delta(n))$$

$$n \geq \frac{\log(\sqrt{5}\,(m-\delta(n)))}{\log\left(\frac{1+\sqrt{5}}{2}\right)} = O(\log m)$$

Also if a can not
α equal two

Solving general linear recurrence

bdry cond

$f(n) = \sum_{i=1}^{d} a_i f(n-i)$ $f(0) = b_0$

$f(1) = b_1, \dots b_{d-1}$

$f(n) = \alpha^n$

$\alpha^n = a_1\alpha^{n-1} + a_2\alpha^{n-2} \dots a_d\alpha^{n-d}$

$\alpha^d = a_1\alpha^{d-1} + a_2\alpha^{d-2} \dots a_d$

$\alpha^d - a_1\alpha^{d-1} - a_2\alpha^{d-2} - a_d$ ← characteristic eqn

Simple roots are different

$\alpha^1 \dots \alpha^d$

Soln $f(n) = c_1\alpha_1^n + c_2\alpha_2^n \dots c_d\alpha_d^n$

← before applying bdry condition

ASYMPTOTIC NOT

Given f(n), we measure the of that function as n go to infinity (limit)

Small o ← strict

nonnegative

$f(x) = o(g(x)) \Rightarrow \lim_{x\to\infty} f(x)/g(x) = 0$

w/ limits

Big oh

Big oh, is the most freq used

$f = O(g)$ ← f(x) is less than or equal g(x)

$\lim_{x\to\infty} f(x)/g(x) < \infty$

We allow constant factors to and ~~willing to~~ ~~ignore~~ small x_0 (such that $x > x_0$)

Omega

The running time is at least

$f = \Omega(g)$

for all constant and for x_0 (such that $x > x_0$) the symbol means, f(x) is greater or equal to g(x). Not taking count of constants or small x_0

COUNTING

same set — not a set

S = {a, b, c} = {b, a, c} ~~{a, [illegible]}~~

cardinality → |S| = 3

when the collection is ordered is a sequence, the collection does not contain duplicate elements

(a, b, c) or (a, b, a)

a permutation of a set S is a sequence that contains every element in S exactly once

{a, b, c}

(a, b, c)
(b, c, a)
(c, a, b)

# of permutations is n!

Def f: x → y

surjective if every element of y is mapped to x at least once

if every elem of x is not mapped to a y is not a function

injective

.. at most once

meaning every y in the range at most one x.

bijective

exactly once

bijective iff injective and surjective

Mapping rule

$f: x \to y$ surjective $\Rightarrow |x| \geq |y|$

$f: x \to y$ injective $\Rightarrow |x| \leq |y|$

$f: x \to y$ bijective $\Rightarrow |x| = |y|$ ~~Bijection rule~~ Bijection rule

Ex x = all ways to select twelve donuts from five varieties

0 0 | | 000 000 | 0 0 | 0 0

001 1000 0001 001   we map to 0,1 sequence. bijective mapping

y = set of all 16 bit sequence with four ones

Ex
bijection from subsets of X {1.. n} to n bit sequences   by bijection rule |x| = |y|

$S \xrightarrow{f} (b_1 \dots b_n)$ via $b_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases}$

2^n = # n bit sequences = # subsets of an n-element set

Gen. pigeon hole

If $|x| > k|y|$ $\forall_{f: x \to y}\ \exists_{k+1}$ different elements of x that are mapped to the same elements in y

Ex (k=1, pigeon hole principle)

$\underbrace{> n \text{ pigeons}}_{x}$, fly into n holes $\Rightarrow$ at least two will fly into the same hole

Ex

Boston has about 500 000 $\underbrace{\text{non bald people}}_{x}$
there exist three people
such that they have the same # of hair

$\underbrace{\# \text{Hair}}_{y} \leq 200000$ $|x| > 2|y|$

$\exists_{k+1}$ people with # same
(2+1)

Ex
Pick 10 arbitrary double digit number
05, 10, 52, 54, 92, 80, 25, 34, 22, 60
There will be two subsets with the same sum

X = collection of subsets of #'s

$X = 2^{10} = 1024$
digits

$|x| > |y|$

$y = \{0, 1 \dots 10 \cdot 99\}$
all possible sums
= 990

Def A "k-to-1" function $f: X \to Y$ maps exactly k elements of X to every element of Y (for example ear to person is 2 to 1)

Division rule (generalize bijection rule)

if f is kTo1 then $|X| = k|Y|$ $\frac{|X|}{k} = |Y|$

Ex how many ways there are to place two rooks in a chess in such a way that no row or column ~~is~~ there is shared

X = sequences (r_1, c_1, r_2, c_2) all possible sequences such that $r_1 \neq r_2$, $c_1 \neq c_2$

$$\underset{\substack{\uparrow \\ \text{maps the sequence} \\ \text{for the position} \\ \text{of the rooks}}}{f} : X \to Y \nwarrow \text{valid configuration}$$

$$|Y| = \frac{|X|}{2} = \frac{(8 \cdot 7)^2}{2}$$

$$\text{eg } (\underset{\substack{\uparrow \\ 8}}{r_1}, \cancel{r_2}\, \underset{\substack{\uparrow \\ 8}}{c_1}, \underset{\substack{\uparrow \\ 7}}{r_2}, \underset{\substack{\uparrow \\ 7}}{c_2})$$

choices 8 · 8 · 7 7

$\{(8, 11)(11, 8)\}$ is the same configuration

Generalized product rule

Let S be the set of length ~~k~~ k sequences

If there are

n_1 possible 1st entries

2nd n_2 for each previous entries

3rd n_3 for each prev and previous entries

n_kth

then $|S| = n_1 \cdots n_k$

Ex committee $(\underset{\text{leader}}{x}, \underset{\substack{\uparrow \\ \text{secretary}}}{y}, \underset{\text{auditor}}{z})$ selected from n members

n ways to choose x

n-1 ways to choose y (except x)

n-2 ways to choose z (except x, y)

$n(n-1)(n-2)$

Product rule

Def $A_1 \times A_2 \ldots \times A_n = \{(a_1, \ldots a_n) : a_1 \in A_1, a_2 \in A_2, \ldots a_n \in A_n$

prod rule

$$|A_1 \times A_2 \ldots \times A_n| = |A_1| \cdot |A_2| \ldots |A_n|$$

Sum rule: If $A_1 \ldots A_n$ are disjointed A_n

$$|A_1 \cup A_2 \cup \ldots A_n| = |A_1| + |A_2| + |A_3| \ldots |A_n|$$

Ex Passwords 6-8 symbols

1st symbol is a letter

other symbols are letters or digits

$F = \{a, b, c, \ldots A, B, C \ldots\}$, $|F| = 52$

first symbol

$S = \{\ldots \ldots 0, 1 \ldots 9\}$ $|S| = 62$

$P = (F \times S \times S \times S \times S \times S) \cup (F \times S^6) \cup (F \times S^7)$

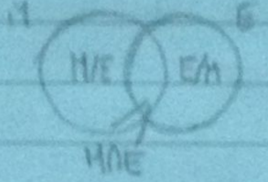
possible passwords

$|P| = |F \times S^5| + |F \times S^6| + |F \times S^7| =$

$|F| \cdot |S^5| + |F| \cdot |S^6| + |F| \cdot |S^7|$

$\approx 1.8 \cdot 10^{14}$

Inclusion-Exclusion



$|M| = |M/E| + |M \cap E|$

$|E| = |E/M| + |M \cap E|$

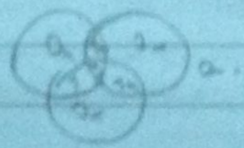
$|M \cup E| = |M/E| + |M/E| + |E/M|$

$|M| + |E|$ counts $|M \cap E|$ twice   include

$|M \cup E| = |M| + |E| - |M \cap E|$ exclude

$|M| + |E| + |S|$ b.

$|M \cap E \cap S|$ a



include

$|M \cup E \cup S| = |M| + |E| + |S|$

$- |M \cap E| - |M \cap S| - |E \cap S|$

exclude

$+ |M \cap E \cap S|$

include

Bookkeeper rule

Ex 16 bit seq such that four bits are 1

$\frac{10!}{4! \cdot 6!}$ bookkeeper rule

subset rule

k element subsets of an element set $\binom{n}{k}$

k elements from a n-element set

Binomial theorem

$$\forall n \; (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Ex $n=2$ $\underbrace{a^2}_{1a} + \underbrace{ab + ba}_{1ba\ 1ab} + \underbrace{b^2}_{1b}$ $a^3 + a^2b + aba + ba^2 + ab^2 + bab + b^2a + b^3$ $n=3$

The monty problem

should you switch? YEP

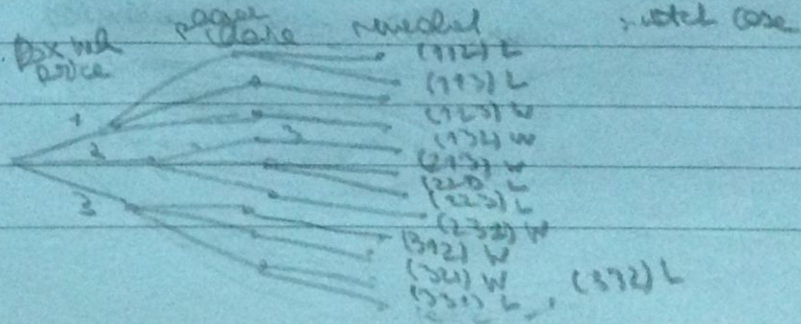
Def The sample space for an experiment is the set of all possible outcomes that could happen

Def An outcome (sample point) consists of all the info about the experiment after it has been performed including the value of all random choices

Def An outcome of the Monty Hall game when the contestant switches consists of: box with prize
box chosen by contestant
box revealed

Ex sample pt (2,1,3) is where
- the prize is in box 2
- the player picks box 1
- carol reveals box 3

Ex construct sample space with all the outcomes



Def a probability space consists of a sample space and a probability function

"Pr" : $S \to \mathbb{R}$ s.t.

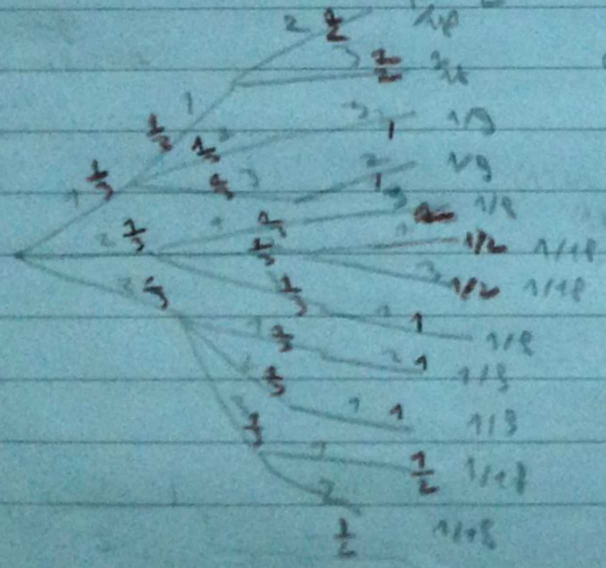
1) $\forall w \in S, \ 0 \le Pr(w) \le 1$

2) $\sum_{w \in S} Pr(w) = 1$

Interpretation: $\forall w \in S$, $Pr(w)$ = probability that w will be the outcome

Assumptions

1) The prize is in each box with prob 1/3
2) The player makes a random choice
3) No matter where the prize is, if carol has a choice she picks a box w/ probability 1/2



all possible outcomes

prev page

Prob 1

Def The probability of a sample point is the product of the probabilities on the path in the tree leading to the sample point

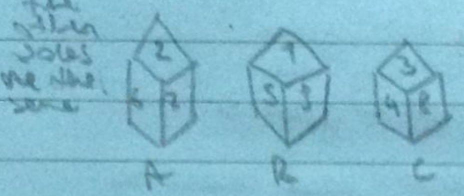
winning is $6\left(\frac{1}{9}\right) = \frac{2}{3}$

Def an event is a subset of sample space

Def the probability that an event E occurs is

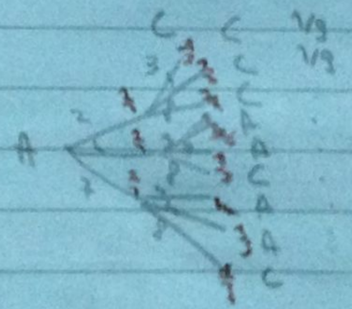
$$\sum_{w \in E} Pr(w)$$

Ex $Pr(E_L) = 6 \cdot \frac{1}{18} = \frac{1}{3}$



$Pr(A=2) = \frac{1}{3}$

$Pr(A=6) = \frac{1}{3}$



Def A sample space is uniform if every sample point has the same probability

$Pr(w) = \frac{1}{|S|}$

sample space

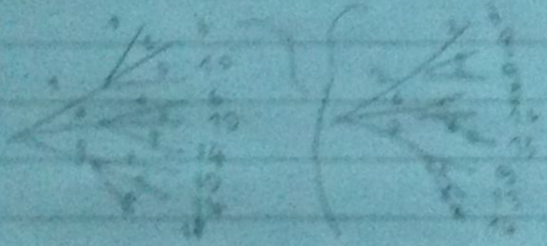
$A > B > C > A$

rolling twice, reverse.

(more rolling cancels out that many dice is likely to be better)

[illegible]

[illegible] table



11, 1+1

0+1+5+2+5+8
+5+8+9

$\frac{42}{81}$

A>B>C>D

Conditional Probability

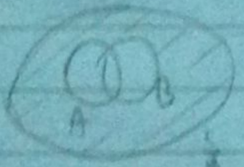
Pr(A|B) = Probability of A given B

Ex. B = event that Carol places the prize in box 1

A = the contestant chooses box 1

Assume $Pr(A|B) = \frac{1}{3}$

Def If $Pr(B) > 0$ then $Pr(A|B) = \frac{Pr(A \wedge B)}{Pr(B)}$



$Pr(B|B) = \frac{Pr(B \wedge B)}{Pr(B)} = 1$

sample space

[illegible] [illegible] [illegible]

product rule $Pr(A \wedge B) = Pr(B)\,Pr(A|B)$

$1/2 \cdot 1/3 = \frac{1}{3}$

General prod rule

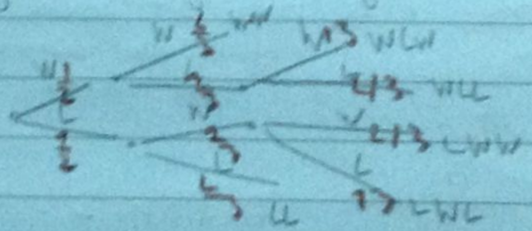
$Pr(A_1 \wedge A_2 \wedge \dots \wedge A_n) = Pr(A_1)\,Pr(A_2|A_1)\,Pr(A_3|A_2 \wedge A$

$Pr(A_4|A_3 \wedge A_2 \wedge A_1) \dots Pr(A_n$

Ex

In a best 2 out of 3 series, the prob of winning 1st game is 1/2, the prob of winning a game following a win is 2/3, winning after a loss is $\frac{1}{3}$

game 1 game 2 game 3



product rule

$Pr(WLW) = Pr(W\ 1st) \cdot \frac{1}{2}$

$Pr(L\ 2nd / W\ 1st)$ $\frac{1}{3}$

$Pr(W\ 3rd / W\ 1st \wedge L\ 2nd)$ $\frac{1}{3}$

$\frac{1}{18}$

$Pr(WW) = Pr(W\ 1st)\ Pr(W\ 2nd | W$

$\frac{1}{2} \cdot \frac{2}{3} =$

A = win series
B = event win 1st game

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{\frac{1}{3} + \frac{1}{8}}{\frac{1}{3} + \frac{1}{8} + \frac{1}{3}} = \frac{7}{9}$$

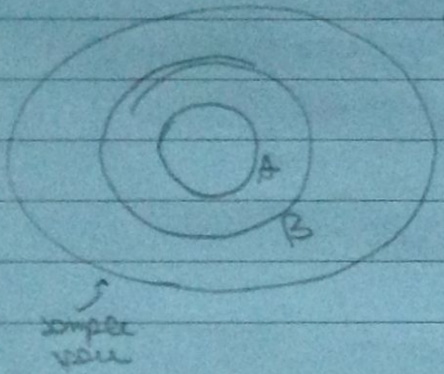
A∩B
WW
WLW

a posteriori probability

Pr(B|A) where B precedes A in time

$$\frac{Pr(B \cap A)}{Pr(A)} = \frac{\frac{1}{3} + \frac{1}{8}}{\frac{9}{8}} = \frac{7}{9}$$

probability of winning the 1st game, given that you won the series

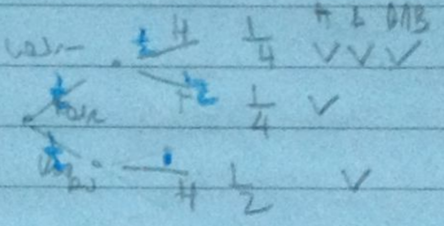


$Pr(A|B) < 1$
$Pr(B|A) = 1$

$Pr(A|B) = Pr(B|A)$ ⇒ they are equal when
$Pr(A) = Pr(B)$
or
$Pr(A \cap B) = 0$

$$\frac{Pr(A|B)}{Pr(B)} = \frac{Pr(B|A)}{Pr(A)}$$

Ex Suppose we have two coins
- Fair coin $Pr(H) = Pr(T) = \frac{1}{2}$
- Unfair coin $Pr(H) = 1$, $Pr(T) = 0$



A = choose fair coin
B = result H

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Medical testing

10% has the disease

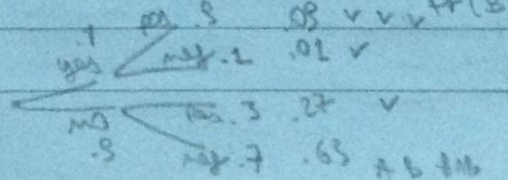
If you have the disease there is a 10% chance that the test is negative

If you don't " there's 30% that the test is positive

Take a random person, the test is positive, what's the chance you really have it

A = event person has disease
B = event that person test is positive

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{.09}{.09 + .27} = \frac{1}{4}$$



Pr(test is correct) =
.09 + .63 = .72

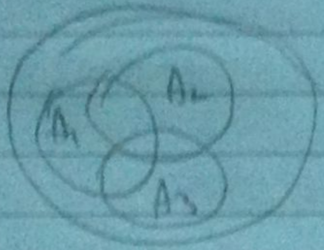
prob

Carnival game (wk reading)

Throw three dices, pick a number, what's the probability that it comes out?

$$Pr(A_1 \cup A_2 \cup A_3) = Pr(A_1) + Pr(A_2) + Pr(A_3) - Pr(A_1 \cap A_2) - Pr(A_1 \cap A_3) - Pr(A_2 \cap A_3) + Pr(A_1 \cap A_2 \cap A_3)$$

$$\approx .421$$



In the same way

$$Pr(A \cup B | C) = Pr(A|C) + Pr(B|C) - Pr(A \cap B | C)$$

$Pr(A|B \cup C) = Pr(A|B) + Pr(A|C)$

False

Events A applicant is admitted
F_{cs} " a female & cs
F_{EE} " " " EE
M_{cs} " a male & cs
M_{EE} " " " EE

$Pr(A|F_{cs}) < Pr(A|M_{cs})$ && $Pr(A|F_{EE}) < Pr(A|M_{EE})$

$Pr(A|F_{cs}+F_{EE}) > Pr(A|M_{cs}+M_{EE})$

They are both right, cause of the weight

$0/1 < 50/100$
$70/100 < 1/1$
$70/101 > 51/101$

Independence

Def An event A is independent of an event B if

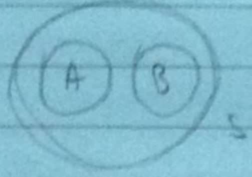
$$Pr(A|B) = Pr(A) \text{ or if } Pr(B) = 0$$

A is independent if knowing that B happened doesn't change the prob. that A is gonna happen

as if you know that B can't happen

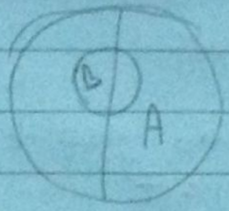
Ex flip two fair independent coins

A = flip 1st coin $Pr(A) = \frac{1}{2}$
B = flip 2nd coin $Pr(B) = \frac{1}{2}$
Then $Pr(A|B) = \frac{1}{2} = Pr(A)$



disjoint $\not\Rightarrow$ indep

$Pr(A|B) = 0 \neq Pr(A)$



$Pr(A|B) = Pr(A)$

indep

Thm (Product rule for indep events)

If A is indep of B, then

$\Pr(A\cap B) = \Pr(A)\Pr(B)$

Pf: - $\Pr(B)=0 \Rightarrow \Pr(A\cap B)=0=\Pr(A)\cdot\Pr(B)$

- $\Pr(B)>0 \rightarrow \Pr(A\cap B)=\Pr(B)\Pr(A\mid B)=\Pr(B)\Pr(A)$

Thm Symmetry of independence

If A is indep of B, then B is indep of A

Ex 2 fair indep coins

A = [illegible] coin matches

B = [illegible] coin is H

$\Pr(A\mid B) = \Pr(\text{2nd coin is H})$

$= \frac{1}{2}$

$\Pr(A) = \Pr(HH) + \Pr(TT)$

$= \frac{1}{4} + \frac{1}{4}$

$= \frac{1}{2}$

by theorem B is indep

Def Events $A_1, A_2 \ldots A_n$ are mutually if all of them does not influence the event you're looking at

$A_1, A_2 \ldots A_n$ are mutually indep iff $\forall J \subseteq [1,n]$, $\Pr(\bigcap_{i\in J} A_i) = \prod_{j\in J} \Pr(A_j)$

Ex A_1, A_2, A_3 are mutually indep if

n=3

Application of mutually indep of n=3

$\Pr(A_1\cap A_2) = \Pr(A_1)\Pr(A_2)$

$\Pr(A_1\cap A_3) = \Pr(A_1)\Pr(A_3)$

$\Pr(A_2\cap A_3) = \Pr(A_2)\Pr(A_3)$

$\Pr(A_1\cap A_2\cap A_3) = \Pr(A_1)\Pr(A_2)\Pr(A_3)$

Def Events $A_1, A_2 \ldots A_n$ are pairwise if $\forall_{i,j} (i\neq j)$

A_i & A_j are indep.

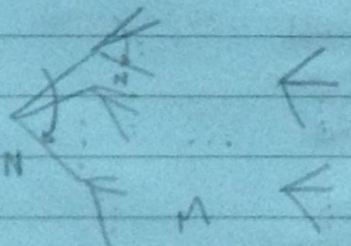
pairwise $\not\Rightarrow$ mutually indep

Birthday problem

N birthdays (ex N = 365)

M people (ex M = 100)

probability that two or more people have the same birthday



birthday sample space

$S = \{(b_1, b_2 \ldots b_M) \mid 1 \le b_i \le N\}$ birthdays

$|S| = N^M$

$\Pr((b_1, b_2 \ldots b_M)) = \left(\frac{1}{N}\right)^M$

It's easier to count diff birthdays:

$N(N-1)(N-2)\ldots(N-M+1) = \frac{N!}{(N-M)!}$

$\Pr(\text{diff birthdays}) = \frac{N!}{(N-M)!\,N^M}$

↑

Birthday principle

..is relevant in compsecurity

$h: L \to S$

ex. Hashtable

Def x collides with y if $h(x) = h(y)$ but $x \neq y$

LECT 22

Random variable R is a function

$R: S \to \mathbb{R}$

sample space

real #s

Ex toss 3 coins

$R = \#$ heads

$R(HTH) = 2$

$M = \begin{cases} 1 & \text{if all coins match} \\ 0 & \text{otherwise} \end{cases}$

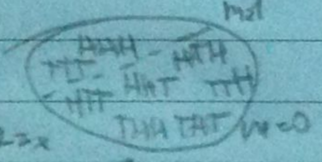
$M(HHT) = 0$

$M(TTT) = 1$

Def M is an indicator (/Bernoulli) random var. An indicator value is 0 or 1.

outcome

$\{w \mid R(w) = x\}$ when $R = x$



Def $\Pr(R = x) = \sum_{w:R(w)=x} \Pr(w)$

probability of that event happening

For $A \subseteq \mathbb{R}$ $\Pr(R \in A) = \sum_{a \in A} \Pr(R = a)$

Ex $A = \{1, 3\}$ $\Pr(R \in A) = \frac{1}{2}$

$\Pr(R = 2 \mid M = 1) = 0$ "nothing holds..."

Def two rand vars R_1, R_2 are independent if

$\forall x_1, x_2 \in \mathbb{R}$ $\Pr(R_1 = x_1 \mid R_2 = x_2) = \Pr(R_1 = x_1)$

$\Pr(R_2 = x_2) = 0$

$\forall x_1, x_2 \in \mathbb{R}$ $\Pr(R_1 = x_1 \wedge R_2 = x_2) = \Pr(R_1 = x_1) \cdot \Pr(R_2 = x_2)$

Def $R_1, R_2 \ldots R_n$ are mutually indep

if $\forall x_1, x_2, \ldots x_n \in \mathbb{R}$,

$\Pr(R_1 = x_1 \wedge R_2 = x_2 \wedge \ldots R_n = x_n) = \Pr(R_1 = x_1) \ldots \Pr(R_n = x_n)$

generalization!

Def Given a r.v. R, the probability distribution function for R is

$f(x) = \Pr(R = x)$

Def the cumulative distribution function for R is

$F(x) = \Pr(R \leq x) = \sum_{y \leq x} \Pr(R = y)$

For a uniform RV on $[1, n]$

$f_n(k) = \frac{1}{n}$ for $1 \leq k \leq n$

$F_n(k) = \frac{k}{n}$

minute 36 on..

Envelope game

Winning strategy

0 Envelopes contain y & $z \in [0, n]$ while $y > z$

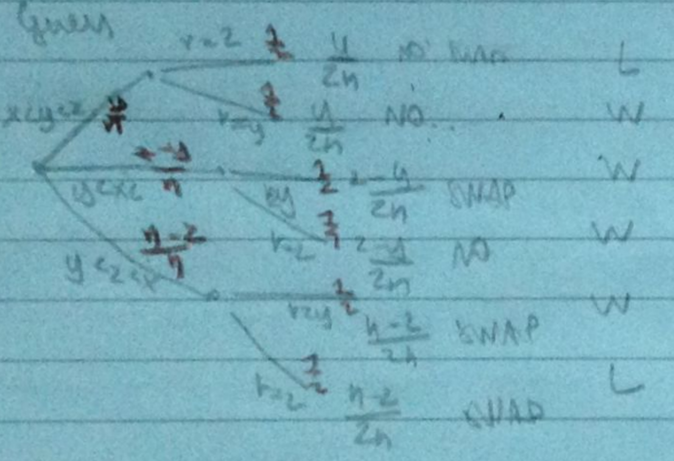
1 Player picks x uniformly in $\{0.5, 1.5, 2.5 \ldots n\tfrac{1}{2}\}$

2 Player is hoping that $x<y<z$

3 Player opens a random envelope with $R \in \{y, z\}$

4 Player swaps if $R<x$

Depends on guessing



UNIFORMAL DISTRIB.

$\Pr(W) = \ldots$

$= \frac{1}{2} + \frac{1}{2n}$

BINOMIAL DISTR

Unbiased binomial dist

$f_n(k) = \binom{n}{k} 2^{-n}$ $\quad n \geq 1 \quad 0 \leq k \leq n$

$(p = \frac{1}{2})$

General binom dist

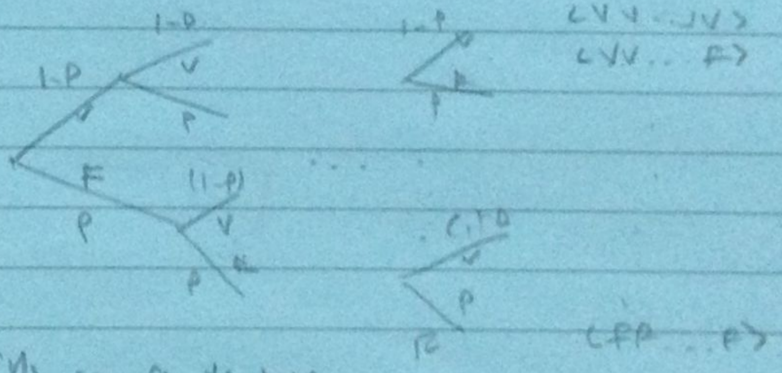
$f_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\quad 0<p<1$

Ex

n components, each fails indep with prob p $(0<p<1)$

R = # failures

Then $\Pr(R=k) = f_{n,p}(k)$ for $0 \leq k \leq n$



$\binom{n}{k}$ sample pts have
k failed components
each has prob $p^k(1-p)^{n-k}$
$\Pr(R=k) = \binom{n}{k} p^k (1-p)^{n-k} = f_{n,p}(k)$