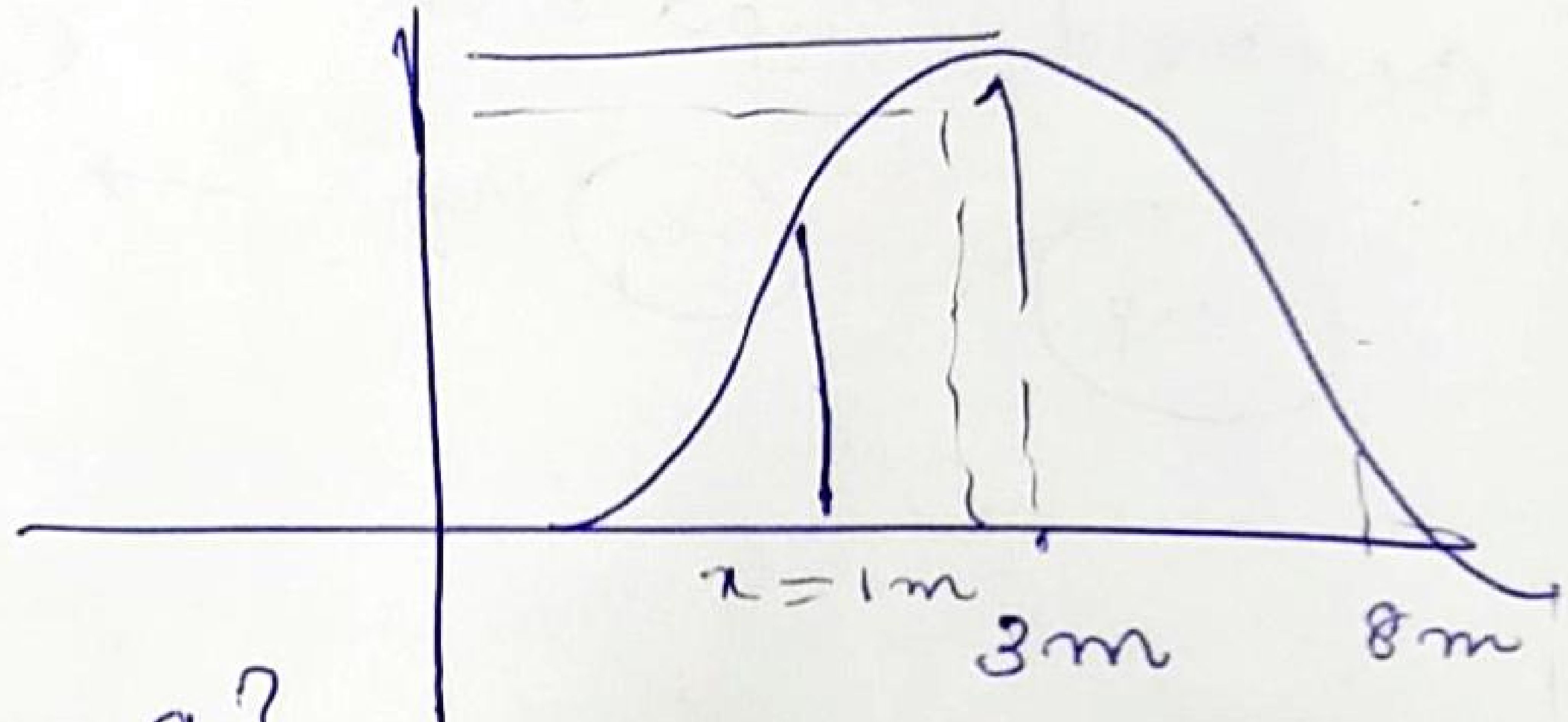
Lec-25

Bayes-rules: Recap



$p\overset{x=}{(3)} \Rightarrow 0.5$

$p(x = \cancel{2}.9) \Rightarrow 0.4$

$p(x = 3 \mid x = 2.9) = 0.9$?

Does it mean

So

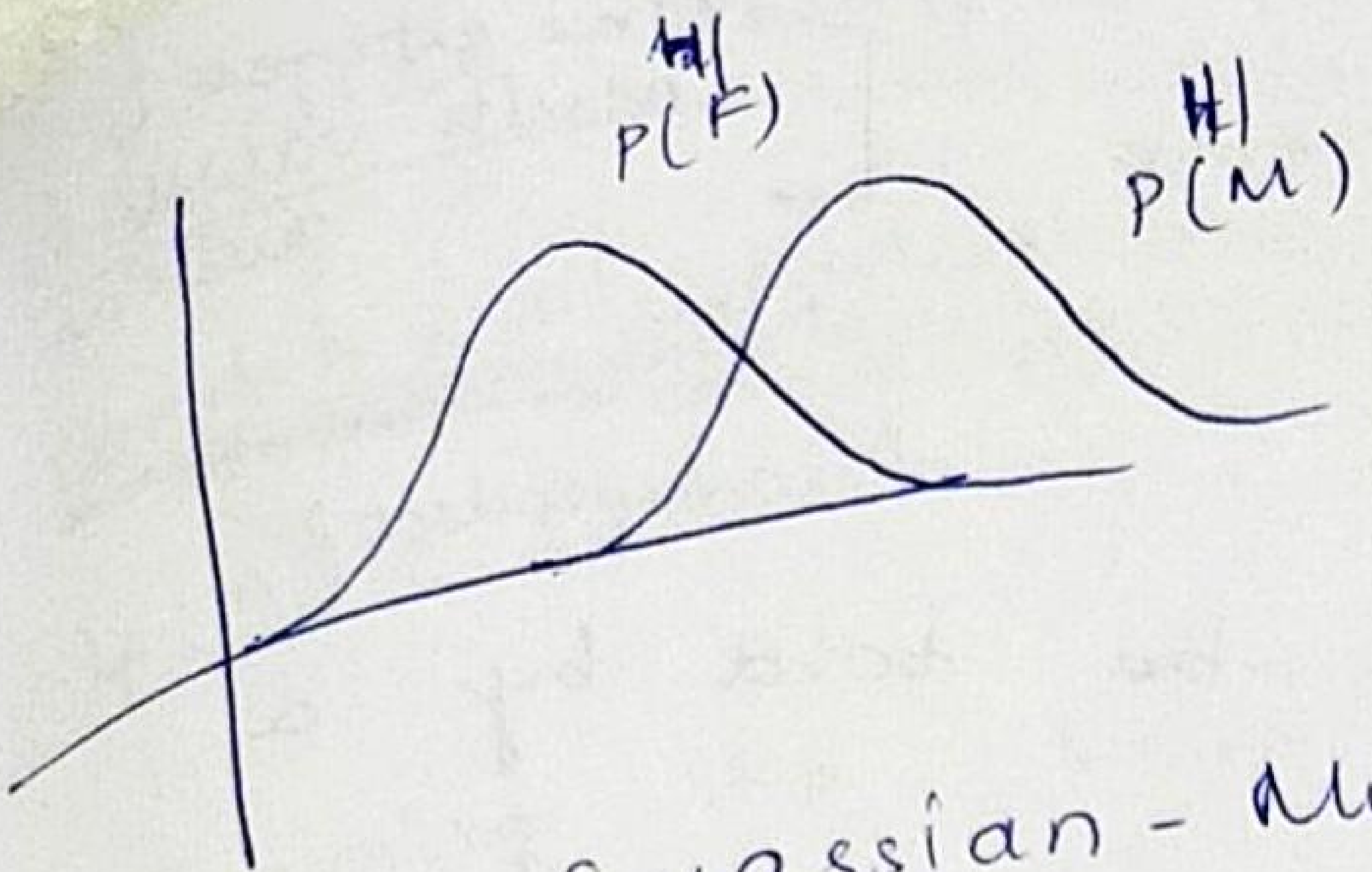
$\int f(x)\,dx$ is the real probability

Probability density:

$$P(A \mid \bar{x}) = \frac{p(\bar{x} \mid A) \times P(A)}{p(\bar{x})}$$

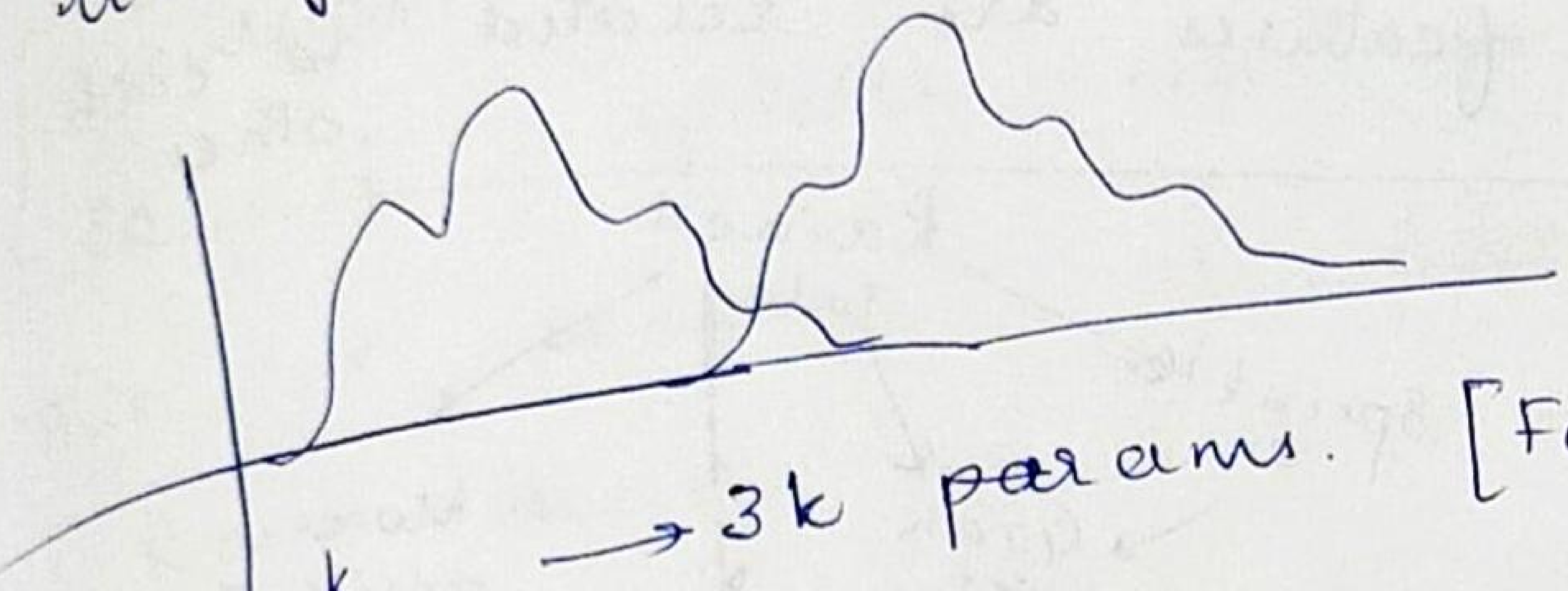
---

female, male → height distro



Height → | Prob. density

Using Guassian - Mixture - model is better,



k Guassians → 3k params. [For univariate]

for multivariate (5 dim)

k Guassians → $k\left[\underset{\mu}{d} + \underset{\Sigma}{d^2} + \underset{\pi}{1}\right]$

kd^2 params.

And rule of thumb, if b params to be estimated ⇒ 10b should be there samples

⇒ $10kd^2$ samples

Bayesian decision making, also P(d|h) always exists

P(d|h): d → data, h → hypothesis

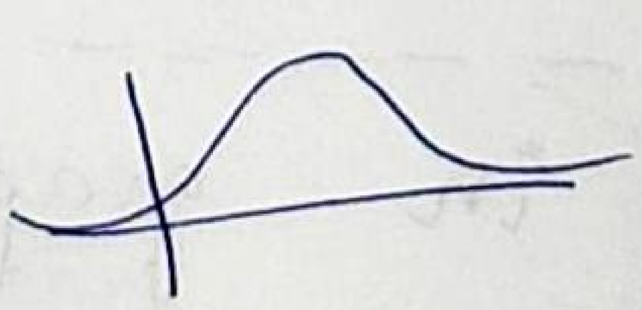
$P(x|A) = P(x_1|A) \times P(x_2|A) \cdots$

So independent 1D - density functions

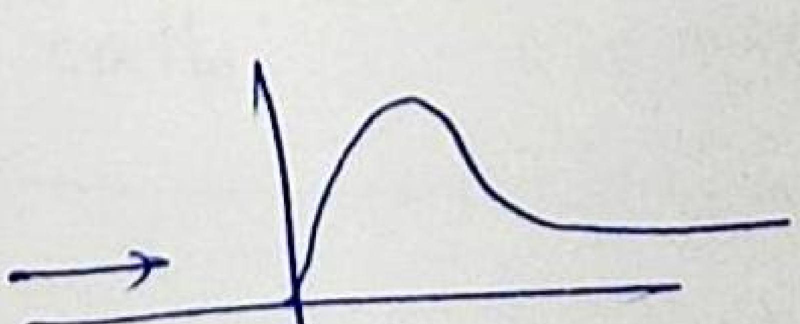
so $O(k \times d)$ params only.
(Some const)

and each $P(x_i|A)$ can be different probability density fns

eg $P(x_1|$ Height →



$P(x_2|$
↓
Length of hair →



Probabilistic Graphical models

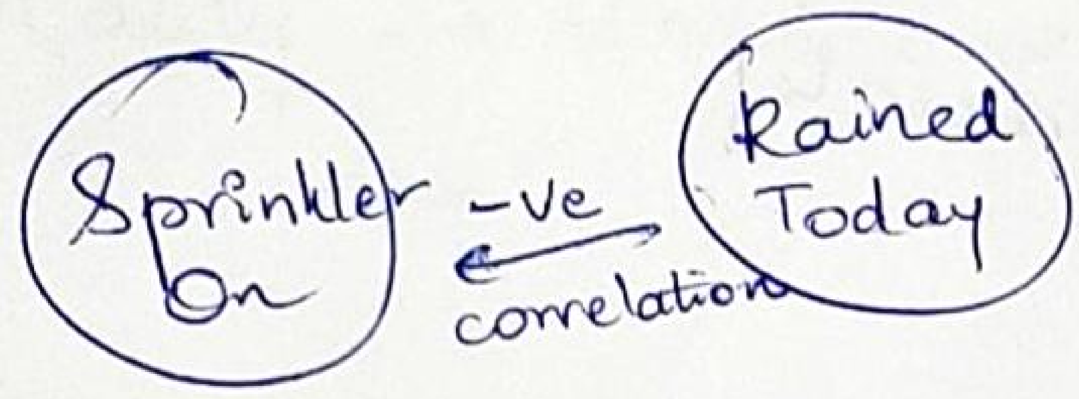
→ Both ends (kd^2, ~~kd~~) are extreme — One extreme is multi-variate guassian; Other is each variable is independent !!

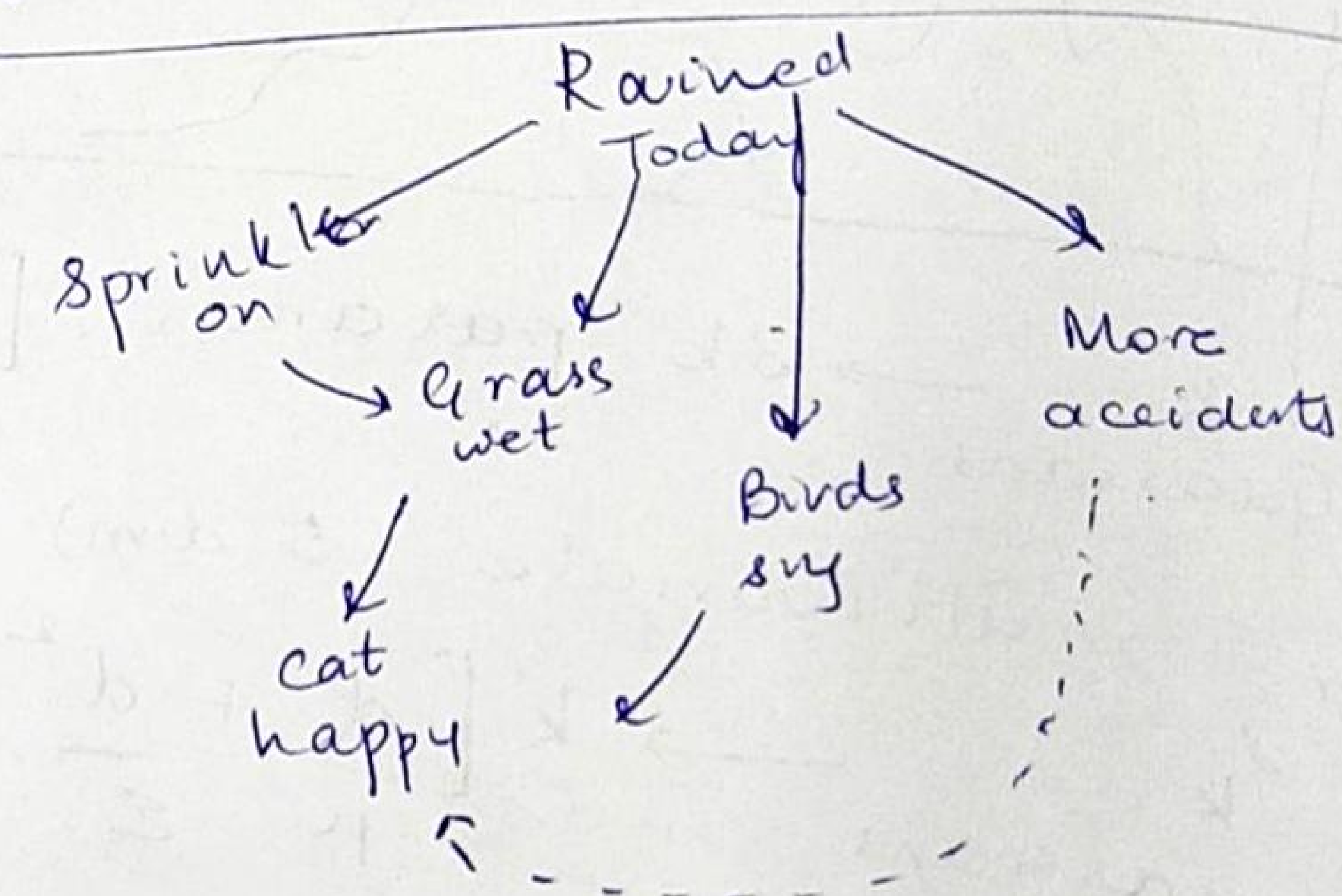
Bayesian

• So; PGM's came.

The dependency b/w features can be told by a human expert

• eg [In real world, some features are related to each other]





R	0
S	1
G	0
B	1
M	0
A	1

a 6 events → 2^6 values

To get the prob distribution curve needed

→ So, atleast 20 ~~x~~ 30 times (samples) is needed

$20 \times 2^6 = 20 \times 64$

$= 1280$

So, 1000's of days are required.

But graph can help,

Note No line ⇏ No correlation

eg: Sprinkler On ⇒ Rained is not there ⇒ No accidents

Correlation

• So, the lines in graph are causual, not correlations

If 2 features are independent

Cat	0	0	1	1
Accidn	0	1	0	1

? No need to calculate all possibilities

Just find independently

Bayesian - Belief network

Conditional independence:

$p(c)p(A) = P(C, A)$

(sprinkler)

- Knowing the value of one can tell something abt other (accidents).

So if sprinkler → off ⇒ accidents → more.

This is because of rain.

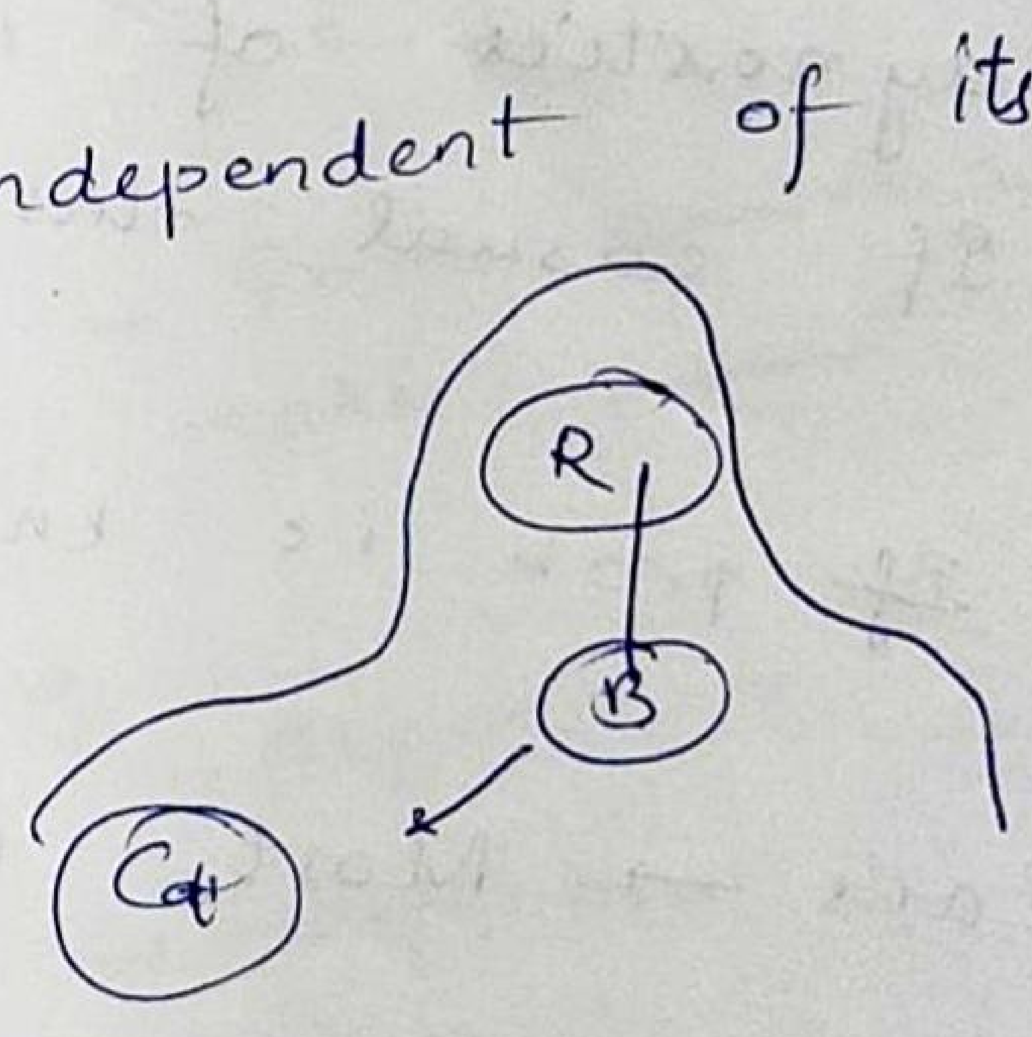
Rain decides accidents

Given the value of rain yes/no ⇒ knowing sprinkler on/off, doesnot affect accidents //

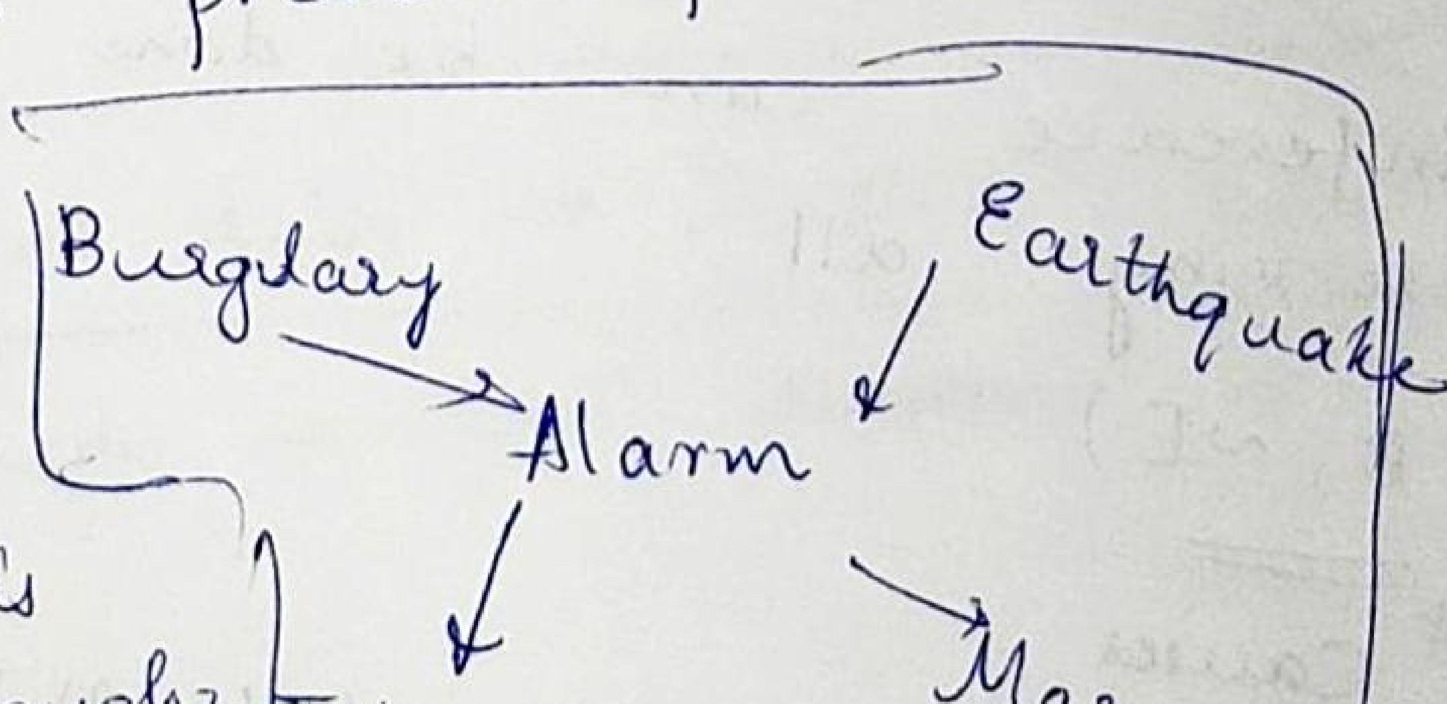
Similarly if rain is known
Grass wet, Birds sing,
⇒ Grass, Bird become independent

Consider
Rained → Bird → Cat
~~Given Rained~~

A variable is conditionally independent of its [non-desendants] given the value of its parents



BBN
Each node have probability distro functions



- Burgulary
  ↳ A few months is need to get enough data [Sees a lotta houses]

Earthquake
↳ A few years

$B \cap E$ → Very rare
↳ Takes a lot time to run

- Mary calls

A	P(M)
t	0.7
f	0.01

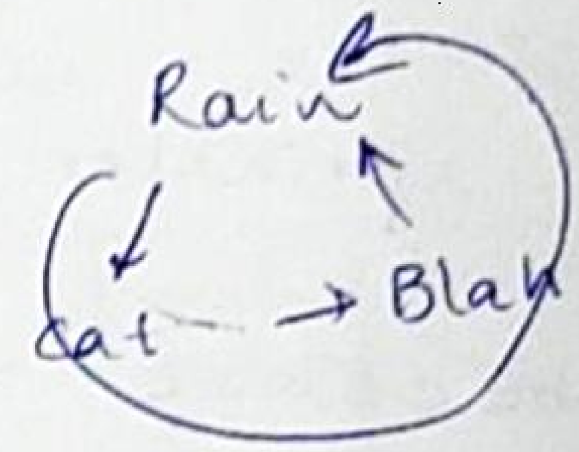
$\to P(M|A=t)$ — 0.3
$\to P(M|A=f)$ — 0.99

P(M does not happen | A)

These add to 1, So they are not explicitly written

Need not add to 1

Graph is a DAG: Cycles ⇒ Time travel



$P(E, B, A, J, M) = ?$
$\quad\;\; t \;\; f \;\; f \;\; f \;\; t$

E is indep of B, A, J, M

~~for~~ $E \to B$ · indep
$(B, E) \to A \to 0.95$

Diagnostics of machines:
If casual rlns are known → The reason for problem can be found

If prob is know → Most likely cause can be found

- Car's → Most complicated

- Casual inference can be done
[Not observing all

$P(\underline{J} | \underline{B, \sim E})$
Effect Causes

$= P(J|A) \times P(A|B, \sim E) + P(J|\sim A)\, P(\sim A | B, \sim E)$

$P(A|\underline{B})$
Alarm needs earthquake also

$P(B|A)$
observe an effect, what is the prob of cause?

$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} = \frac{P(A|B) \times P(B)}{P(A|B) \times P(B) + P(A|B') \times P(B')}$ → The "Bayes" part of BBN

- $P(B|A) = 0.374$, why?

① $P(E) > P(B)$

② If

B	E	P(A)
f	f	0.001

is more, $P(B|A)$ is less, because randomly alarm is ringing

Inter casual inference?

Can Burgulary and Earthquake affect?

Knowing decendants value (will affect earthquake)

and John calls → Alarm → Marycalls

~~col~~ Usually john calls when alarm ↑ and marycalls ↑ if alarm ↑.

If we know alarm ↓ and mary calls.

Caus Causs Eff → Not knowing → Causes combined makes
knowing →

Cause Eff Eff → Not knowing cause → dependent
knowing → independent

IBM Watson → Medical data → Doctors constructed BBN

~~If~~ If large enough → then ~~doctors~~ it can find causes also.