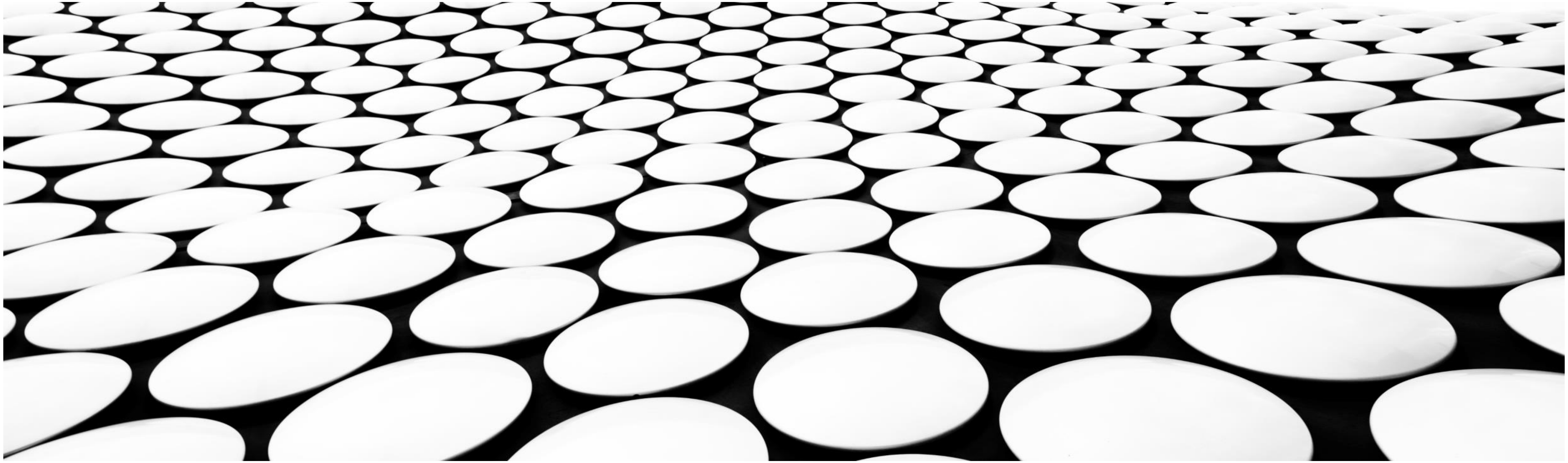
5-BAYESIAN NETWORK

ABDELBASSET B



Probability : reminder

Definition : Probability is a number that measures how likely an event is to happen.

It ranges from 0 to 1:

- 0 → the event is impossible
- 1 → the event is certain

- Values in between (like 0.2, 0.5, 0.9) show increasing likelihood

Formal definition: If all outcomes are equally likely, the probability of an event A is:

$$\mathrm{P(A)} = \frac{\text{number of favorable outcomes}}{\text{number of all possible outcomes}}$$

Probability : reminder

Example:
If you roll a fair six-sided die, the probability of getting a 4 is: $P(4) = ?$

A family is trying to guess whether it will rain tomorrow.

Med says:
"There's a 50% chance it will rain tomorrow."

Sara says:
"I read an article that says in our city, 60% of winter days have rain."

The weather station says:
"Based on radar and satellite data, the chance of rain tomorrow is 85%."

Which probability is correct?

Probability : reminder

Probability measures how uncertain we are about something.
It is not a property of the real world itself — it is a measure of what we know (or don't know).

✅ What probability measures **Our degree of belief**

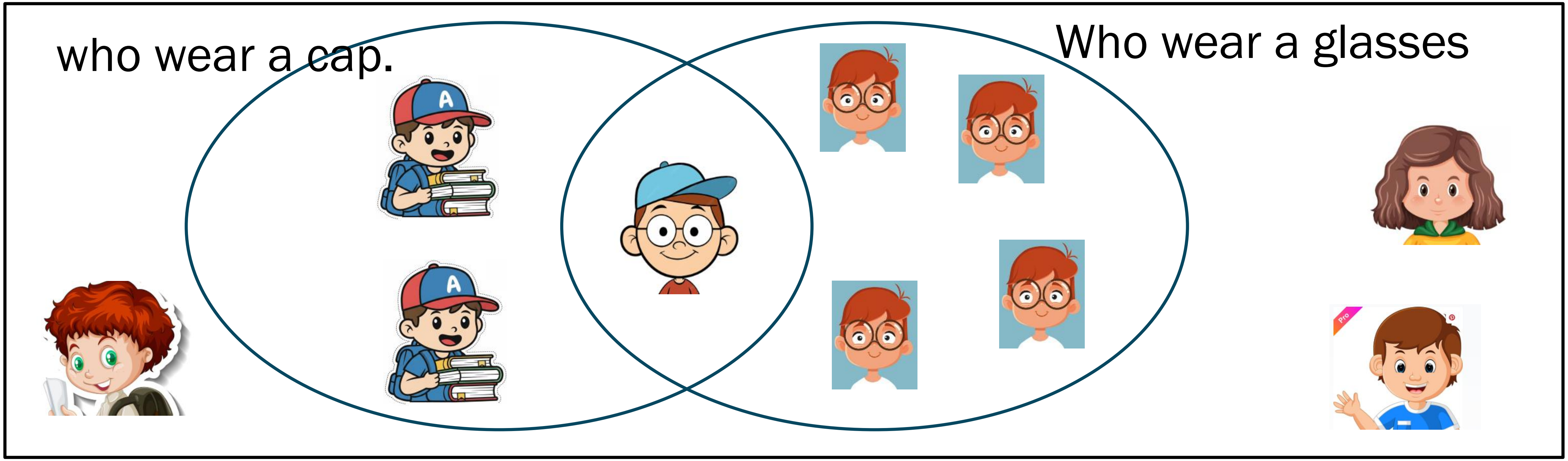
- If we know very little → the probability is more uncertain (like Med's 50% guess).
- If we know more → the probability becomes more accurate (like Sara's population data).
- If we have strong evidence → the probability becomes very high (like the 90% weatherman).

Probability : reminder

Conditional probability

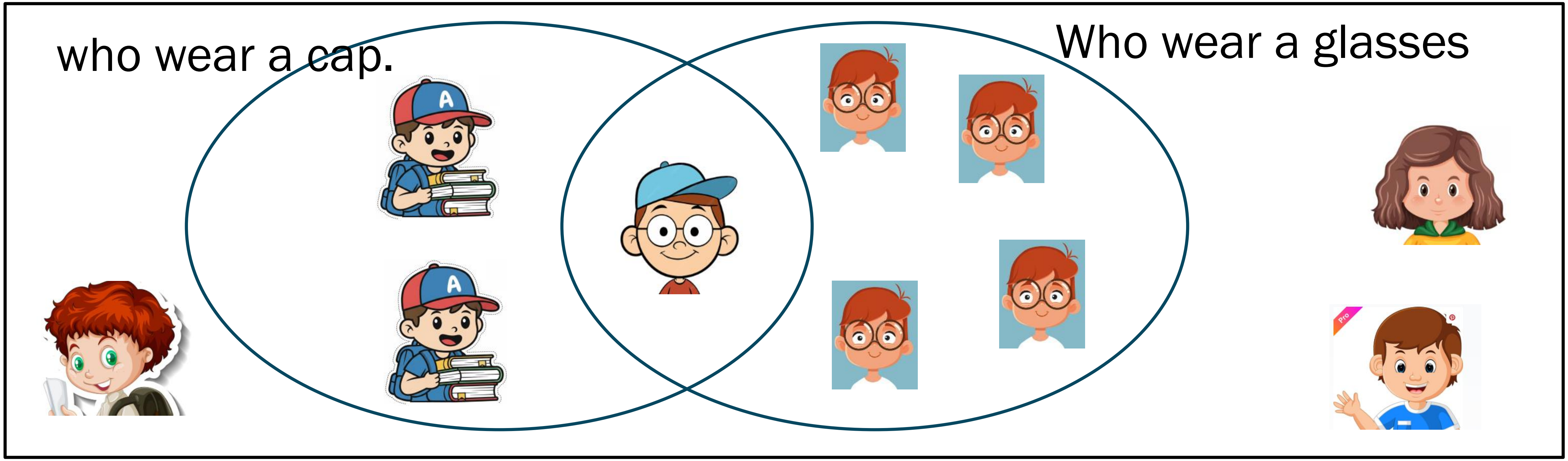
Let’s imagine that we have two sets:

- The set of students who wear a cap.
- The set of students who wear a glasses



Probability : reminder

Conditional probability

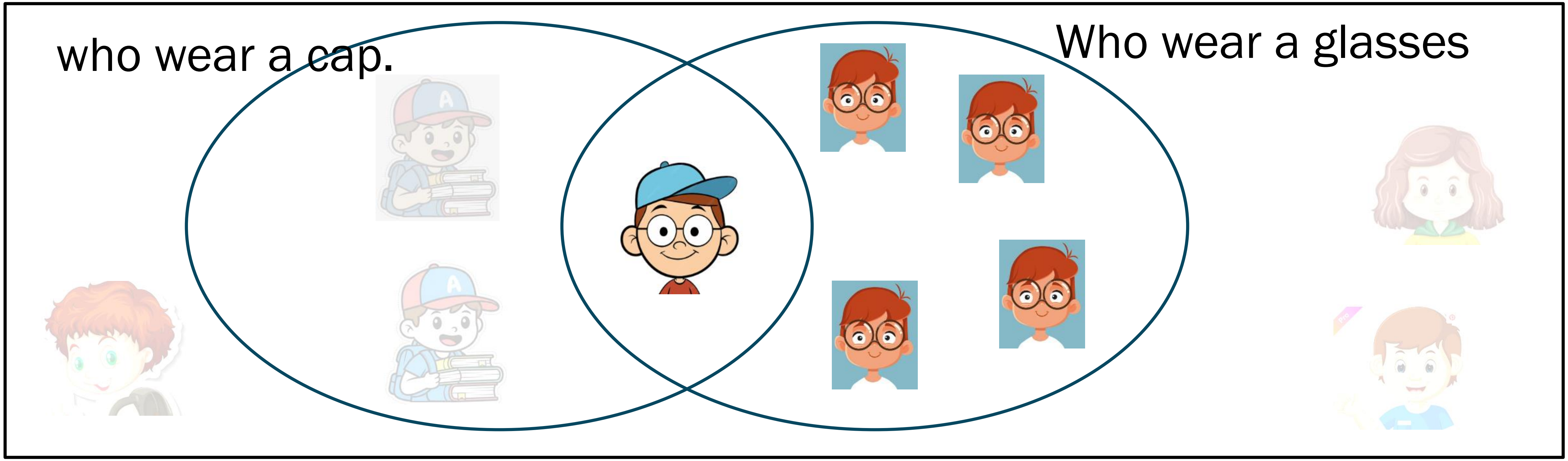


What is the probability of P(cap) = 3/10

What is the probability of P(cap/glasses) =

Probability : reminder

Conditional probability

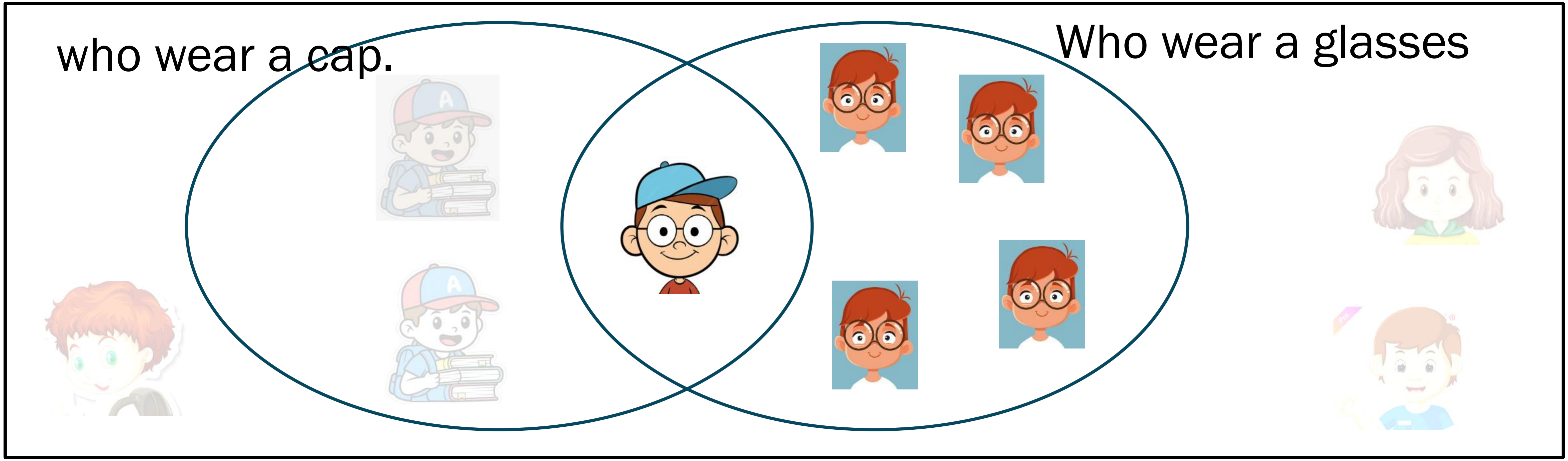


What is the probability of P(cap) = 3/10

What is the probability of P(cap | glasses) = 1/5

Probability : reminder

Conditional probability



$$P(cap \cap glasses) = \frac{|cap \cap glasses|}{|glassea|} = \frac{P(cap \cap glasses)}{\mathrm{P}(glassea)}$$

Probability : reminder

Bayes theorem simple version

$$P(H|E) = {}^{P(H\cap E)}\!/_{P(E)}$$

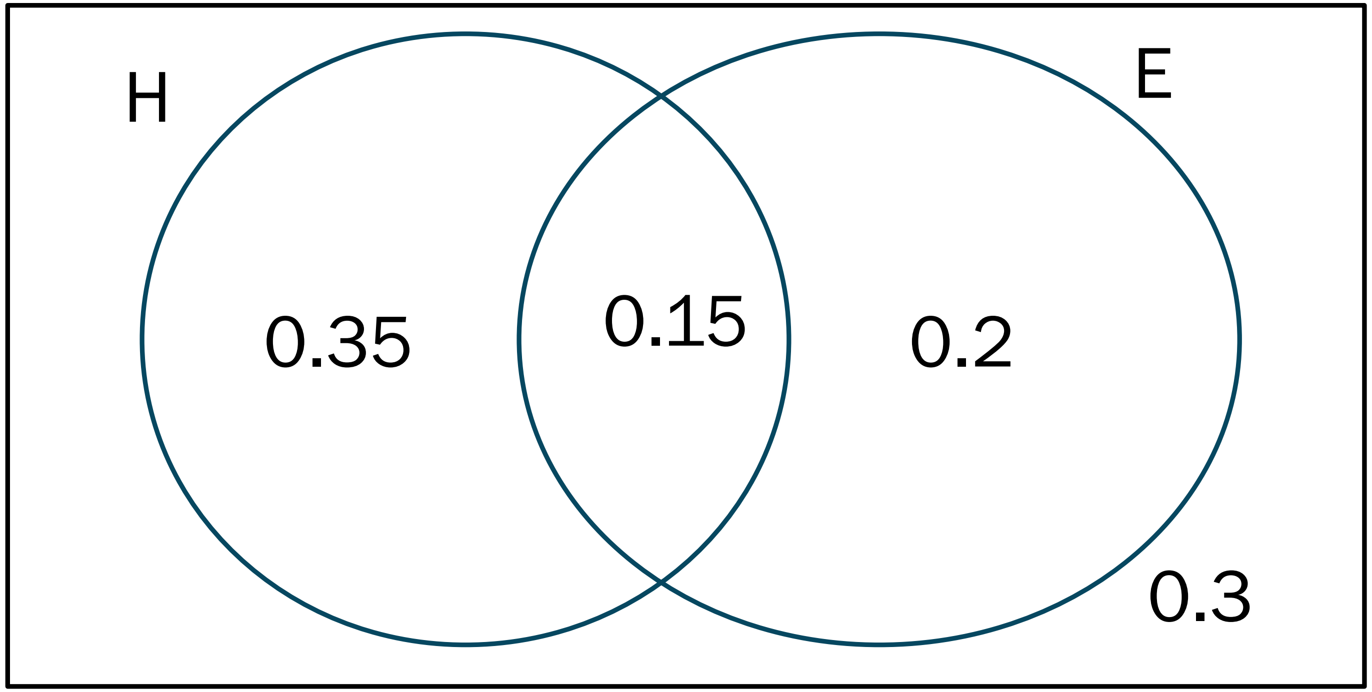
H : hypotheses to be found

E : The evidence

$$P(H|E) = ?$$

$$P(H|E) = \frac{0.15}{0.2+0.15} = \frac{0.15}{0.35} = \frac{3}{7}$$

Venn diagram



Probability : reminder

Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

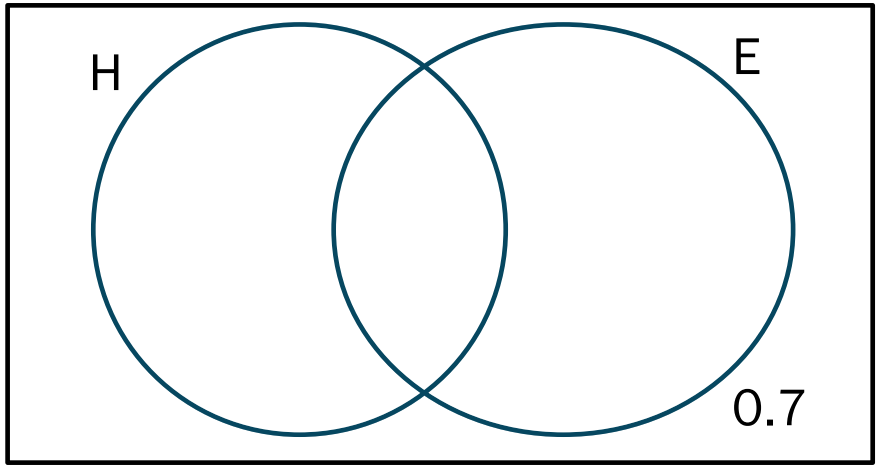
What is E

Khaoula is late $P(E) = 0.25$

What is H

Amira is late $P(H) = 0.2$

Where to put the probability 0.7$= P(\neg H \cap \neg E)$



Probability : reminder

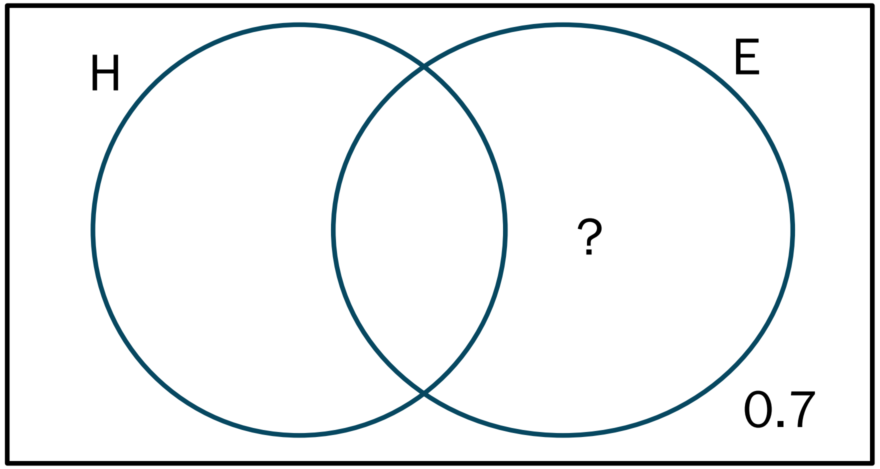
Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

What is E

Khaoula is late $P(E) = 0.25$ $? = 1 - (0.7 + 0.2)$



What is H

Amira is late $P(H) = 0.2$

Probability : reminder

Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

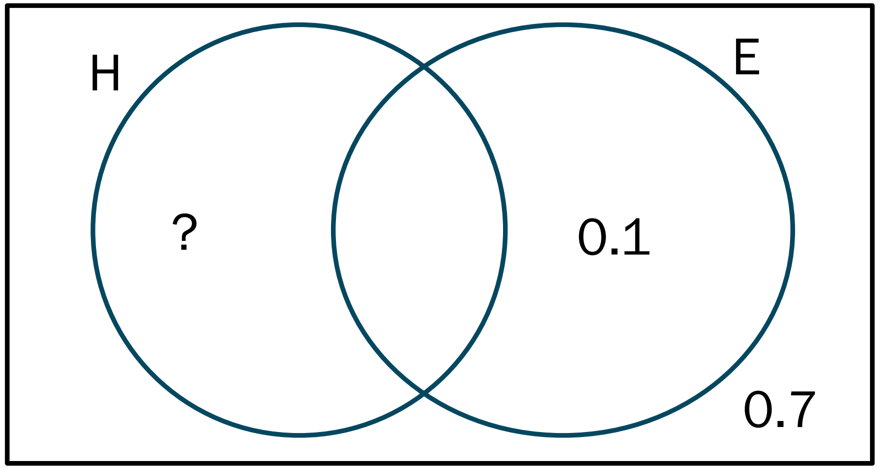
Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

What is E

Khaoula is late $P(E) = 0.25 \quad ? = 1 - (0.7 + 0.2)$

What is H

Amira is late $P(H) = 0.2 \quad ? = 1 - (0.7 + 0.25)$



Probability : reminder

Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

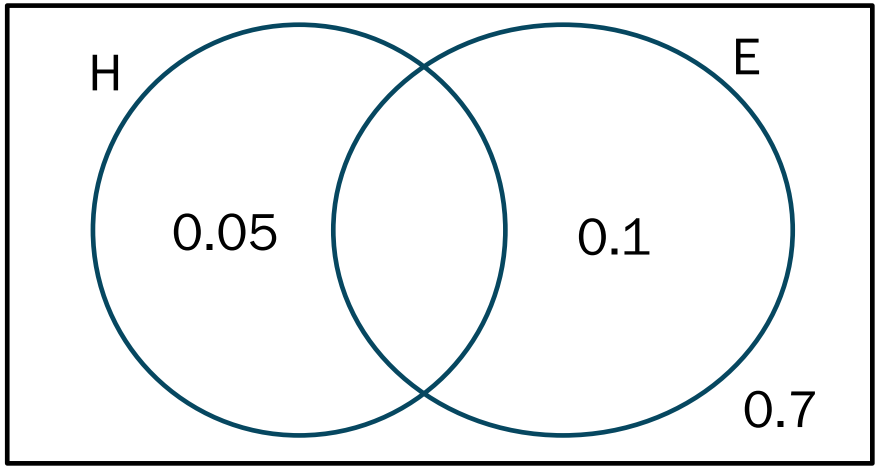
Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

What is E

Khaoula is late $P(E) = 0.25 \quad ? = 1 - (0.7 + 0.2)$

What is H

Amira is late $P(H) = 0.2 \quad ? = 1 - (0.7 + 0.25)$



Probability : reminder

Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

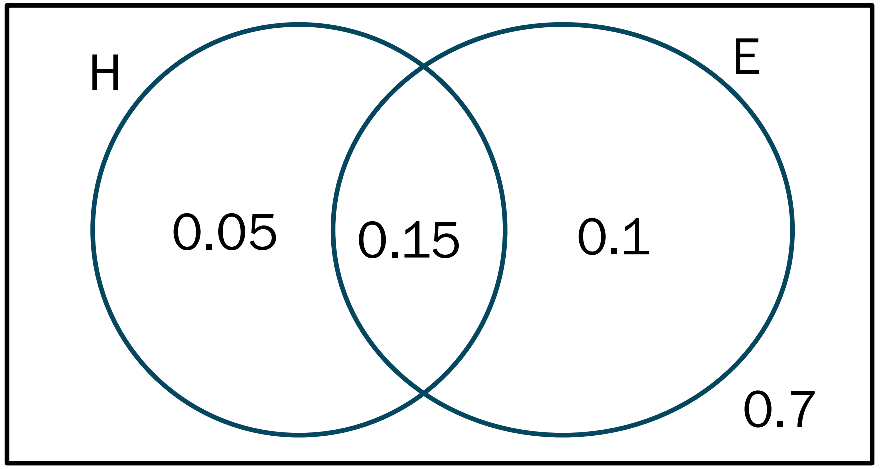
What is E

Khaoula is late $P(E) = 0.25 \quad ? = 1 - (0.7 + 0.2)$

What is H

Amira is late $P(H) = 0.2 \quad ? = 1 - (0.7 + 0.25)$

$P(H \cap E)$= 0.25-0.1 or 0.2-0.05 or 1-(0.7+0.1+0.05)



Probability : reminder

Bayes theorem

Example ***Amira*** and ***Khaoula*** are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

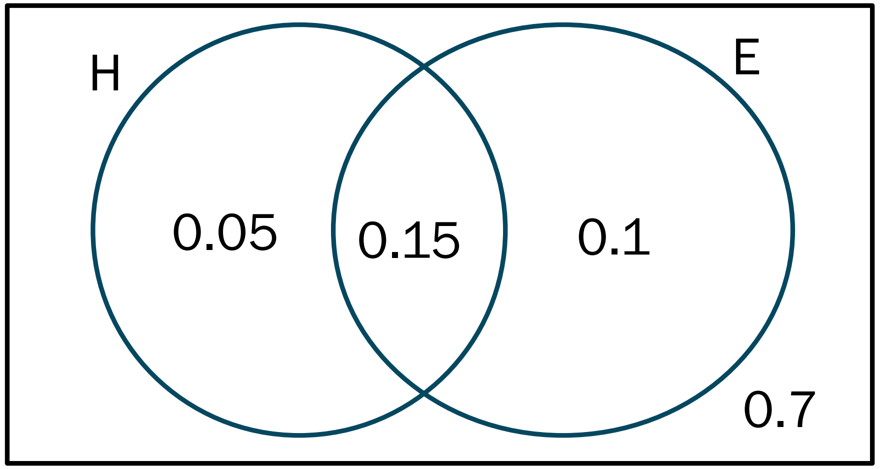
E = Khaoula is late $P(E) = 0.25$

H = ***Amira*** is late $P(H) = 0.2$

$P(\neg H \cap \neg E) = 0.7$

What we are searching for?

$$P(H|E)$$



Probability : reminder

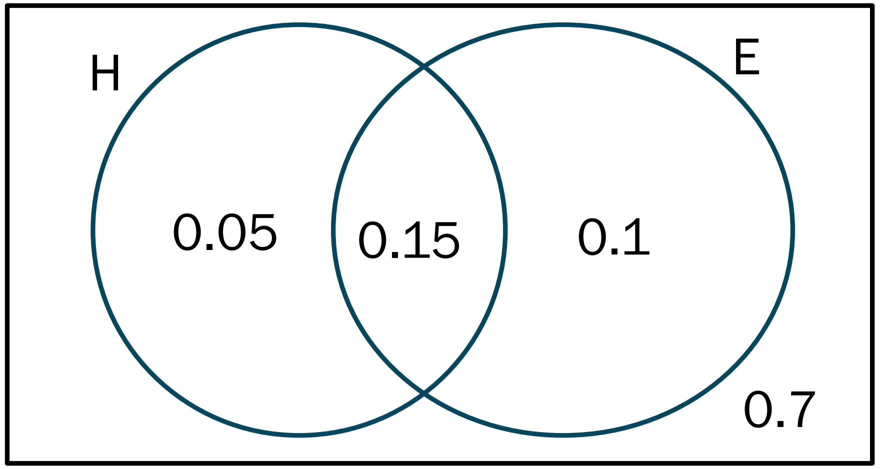
Bayes theorem

Example *Amira* and *Khaoula* are sometimes late for school, 70% of the time neither of them is late. ***Amira*** is late 20% of the time, while ***khaoula*** is late 25% of the time.

Last Monday ***Khaoula*** was late. Find the probability that ***Amira*** is late

$$P(H|E) = {}^{P(H\cap E)}/_{P(E)}$$

$$P(H|E) = {}^{0.15}/_{0.25} = {}^{3}/_{5} = 60\%$$



Probability : reminder

Random variables

Definition : A variable that can take different values, and each value has a probability.

Example 1: Flipping a coin

X = 1 if it shows a head.

X = 0 if it shows a tail

Example 2 : sunshine $S \in \{0,1\}$, rain $R \in \{0,1\}$

Probability : reminder

Random variables: sunshine $S \in \{0,1\}$, rain $R \in \{0,1\}$

Joint distribution: A joint distribution describes the probability of two or more random variables happening together.

A **joint distribution** answers questions like:

•*What is the probability that X = x* ***and*** *Y = y at the same time?*

It tells you how two (or more) variables behave **together**, not separately.

	s	r	P{S = s, R =r}
	0	0	0.20
P{S,R}	0	1	0.08
	1	0	0.70
	1	1	0.02

Probability : reminder

Random variables: sunshine $S \in \{0,1\}$, rain $R \in \{0,1\}$

Marginal distribution: shows the probability of **one variable by itself**, ignoring the other variables.

- You get it by summing (or integrating) the joint probabilities over the other variables.
- In other words, you "marginalize out" the other variable(s).

	s	r	P{S = s, R =r}
	0	0	0.20
P{S,R}	0	1	0.08
	1	0	0.70
	1	1	0.02

	s	P{S = s}
P{S}	0	0.28
	1	0.72

Probability : reminder

Random variables: sunshine $S \in \{0,1\}$, rain $R \in \{0,1\}$

Conditional distribution: describes the probability of one variable given that another variable has a specific value.

$$P(X=x|Y=y) = P(X=x,Y=y) / P(Y=y)$$

	s	r	P{S = s, R =r}
	0	0	0.20
P{S,R}	0	1	0.08
	1	0	0.70
	1	1	0.02

	s	P{S}R=1}
P{S}R=1}	0	0.8
	1	0.2

Probability : reminder

Random variables: sunshine $S \in \{0,1\}$, rain $R \in \{0,1\}$

Marginal Conditional distribution: is the **conditional distribution of one variable**, after we have *marginalized* (summed out) the others..

Let's expand our example

Random variables: X = (S, R, T, A)

- **Observe evidence (traffic in autumn): T = 1; A = 1**
- **Interested in query (rain?): R**

$$\mathbb{P}(\underbrace{R}_{\text{query}} \mid \underbrace{T=1, A=1}_{\text{condition}})$$

(S is **marginalized out**)

Bayes theorem

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

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

$$P(B \cap A) = P(B|A) * P(A)$$

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)}$$

Bayesian network :

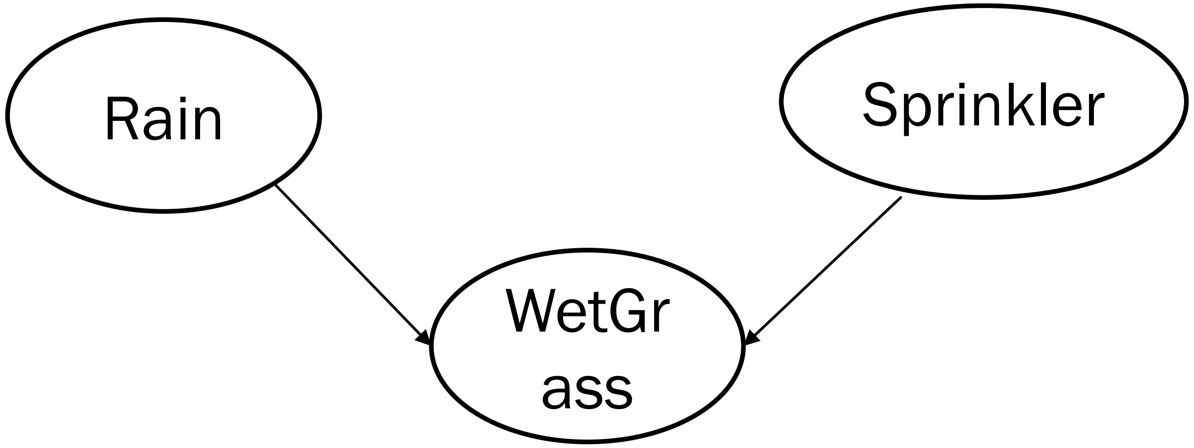
Random variables: We define three random variables:

Rain (R) ∈ {true, false}

Sprinkler (S) ∈ {true, false}

WetGrass (W) ∈ {true, false}

- **Rain → WetGrass**
- **Sprinkler → WetGrass**
- **Rain and Sprinkler are independent** (no arrow between them)



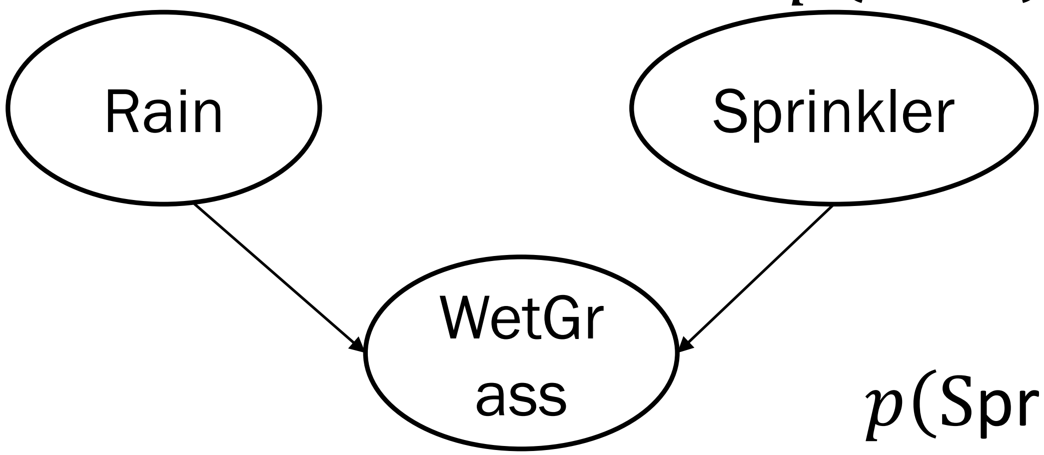
This is called the **Naïve Bayes structure.**

Bayesian network :

Let’s suppose that the probability Rain = the probability of sprinkler =ε

Rain	P(R)
1	ε
0	1-ε

$$p(Rain) = \varepsilon * [R = 1] + (1 - \varepsilon) * [R = 0]$$



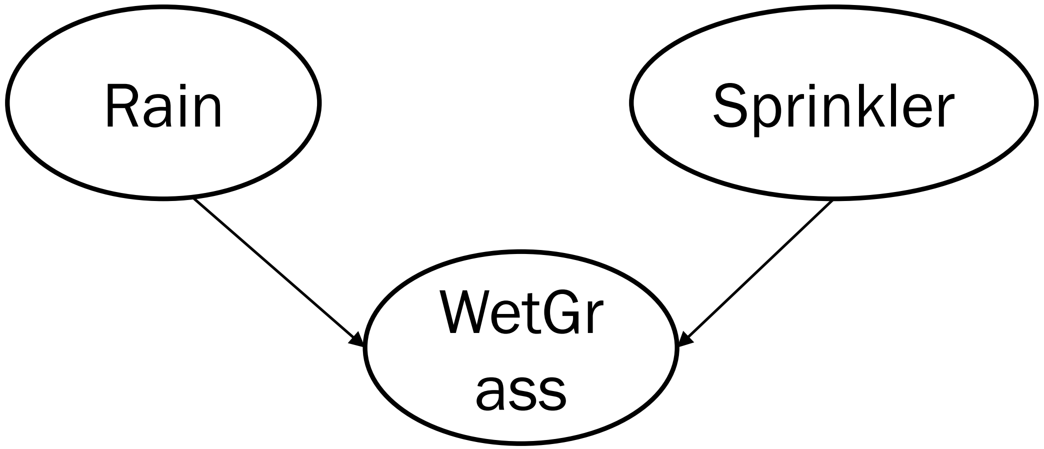
sprinkler	P(R)
1	ε
0	1-ε

$$p(\text{Sprinkler}) = \varepsilon * [S = 1] + (1 - \varepsilon) * [S = 0]$$

Bayesian network :

P(W | R,S)=[W=(R, S)]

R	S	W	P(W | R,S)
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1



Conditional Probability Table.

Bayesian network :

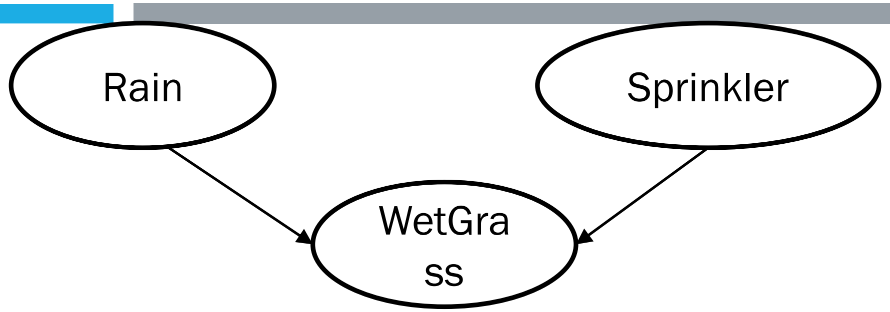
P(W | R,S)=[W=(R, S)]

R	S	W	P(W | R,S)	P(W=w,R=r,S =s)
0	0	0	1	(1-ε)(1-ε)
0	0	1	0	0
0	1	0	0	0
0	1	1	1	(1-ε) ε
1	0	0	0	0
1	0	1	1	ε(1-ε)
1	1	0	0	0
1	1	1	1	$ε^2$

Rain	**P(R)**
1	ε
0	1-ε

sprinkler	**P(R)**
1	ε
0	1-ε

Bayesian network :



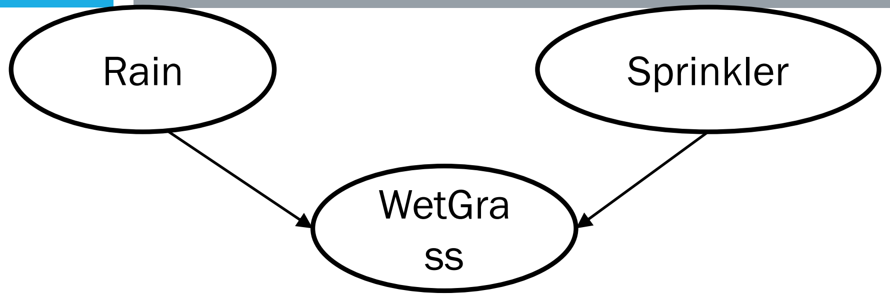
$P(W \mid R,S)=[W=(R, S)]$

R	S	W	P(W=w,R=r,S =s)
0	0	0	$(1-\varepsilon)(1-\varepsilon)$
0	0	1	0
0	1	0	0
0	1	1	$(1-\varepsilon)\ \varepsilon$
1	0	0	0
1	0	1	$\varepsilon(1-\varepsilon)$
1	1	0	0
1	1	1	ε^2

P(Rain) = it is a marginal distribution
Sum of all the probability where Rain =1

$$0+ \varepsilon(1-\varepsilon)+0+\varepsilon^2 =\varepsilon-\varepsilon^2+\varepsilon^2= \varepsilon$$

Bayesian network :



$P(W \mid R,S)=[W=(R, S)]$

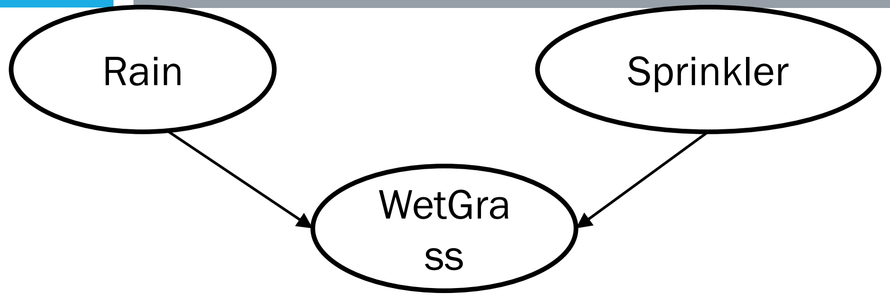
R	S	W	P(W=w,R=r,S =s)
0	0	0	$(1-\varepsilon)(1-\varepsilon)$
0	0	1	0
0	1	0	0
0	1	1	$(1-\varepsilon)\,\varepsilon$
1	0	0	0
1	0	1	$\varepsilon(1-\varepsilon)$
1	1	0	0
1	1	1	ε^2

$P(R=1 \mid W=1)=$

$$\frac{\varepsilon(1-\varepsilon)+\varepsilon^2}{\varepsilon(1-\varepsilon)+\varepsilon^2+(1-\varepsilon)\,\varepsilon}=\frac{1}{2-\varepsilon}$$

$$\approx 50\%$$

Bayesian network :



P(W | R,S)=[W=(R, S)]

R	S	W	P(W=w,R=r,S =s)
0	0	0	(1-ε)(1-ε)
0	0	1	0
0	1	0	0
0	1	1	(1-ε) ε
1	0	0	0
1	0	1	ε(1-ε)
1	1	0	0
1	1	1	ε^2

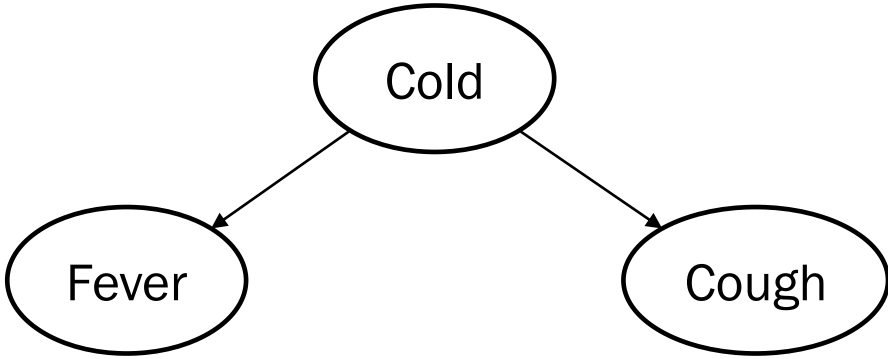
P(R=1|W=1, S=1)=

$$\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)\,\varepsilon} = \varepsilon$$

Bayesian network :

P(Fever | Cold)

cold	P(cold)
1	0.1
0	0.9

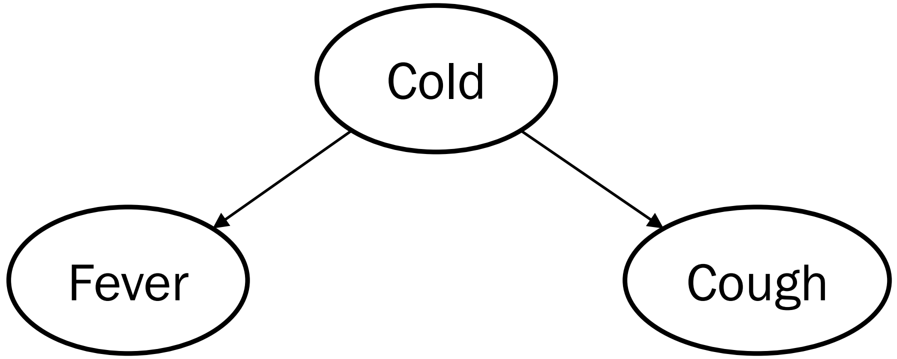


cold	Fever	P(Fever | Cold)
0	0	0.98
0	1	0.02
1	0	0.30
1	1	0.70

Bayesian network :

P(Cough | Cold)

cold	P(cold)
1	0.1
0	0.9



Cold	Cough	P(Cough | Cold)
0	0	0.95
0	1	0.05
1	0	0.10
1	1	0.90

Query 1: Question: What is the probability of having a fever?

$$P(F = 1) = ?$$

By the law of total probability and from the network:

$$P(F = 1) = P(F = 1\ and\ C = 0) + P(F = 1\ and\ C = 1)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

P(F=1 and C=0) = P(F=1|C=0)·P(C=0)

P(F=1 and C=0) = 0.02*0.9=0.018

P(F=1 and C=1) = P(F=1|C=1)·P(C=1)

P(F=1 and C=1) = 07*0.1=0.07

0.088=8.8%

cold	Fever	P(Fever | Cold)
0	0	0.98
0	1	0.02
1	0	0.30
1	1	0.70

Cold	Cough	P(Cough | Cold)
0	0	0.95
0	1	0.05
1	0	0.10
1	1	0.90

cold	**P(cold)**
1	0.1
0	0.9

Query 2:Question: Given someone has a fever, what's the probability they have a cold?

$$P(C = 1 \mid F = 1) = ?$$

In this example we don't have the information of $P(C = 1 \mid F = 1)$ however, we have $P(\mathrm{F} = 1 \mid \mathrm{C} = 1)$

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

$P(C=1|F=1) = P(F=1|C=1)\ P(C=1) / P(F=1)$

$$P(C = 1 \mid F = 1) = \frac{0.70 \times 0.1}{0.088} = \frac{0.07}{0.088} \approx 0.795$$

79%

cold	Fever	P(Fever | Cold)
0	0	0.98
0	1	0.02
1	0	0.30
1	1	0.70

Cold	Cough	P(Cough | Cold)
0	0	0.95
0	1	0.05
1	0	0.10
1	1	0.90

cold	**P(cold)**
1	0.1
0	0.9

Query 3:Question: Given someone has both fever and cough, what's the probability they have a cold?

$$P(C, = 1 \mid F = 1H = 1) = ?$$

$$P(C{=}1|F{=}1,H{=}1) = \frac{P(F{=}1,H{=}1|C{=}1) \cdot P(C{=}1)}{P(F{=}1,H{=}1)}$$

$$P(F{=}1,H{=}1|C{=}1) \cdot P(C{=}1)$$

Since F and H are conditionally independent given C:

$$P(F{=}1,H{=}1|C{=}1) = P(F{=}1|C{=}1) * P(H{=}1|C{=}1)$$

$$=0.70 \times 0.90 = 0.63 \quad * P(C{=}1)$$

$$=0.63*0.1=0.063$$

cold	Fever	P(Fever | Cold)
0	0	0.98
0	1	0.02
1	0	0.30
1	1	0.70

Cold	Cough	P(Cough | Cold)
0	0	0.95
0	1	0.05
1	0	0.10
1	1	0.90

cold	**P(cold)**
1	0.1
0	0.9

Query 3:Question: Given someone has both fever and cough, what's the probability they have a cold?

$$P(C' = 1 \mid F = 1 H = 1) = ?$$

$$P(C{=}1|F{=}1,H{=}1) = \frac{P(F{=}1,H{=}1|C{=}1)\cdot P(C{=}1)}{P(F{=}1,H{=}1)}$$

$$P(F{=}1,H{=}1|C{=}1)\cdot P(C{=}1) = 0.063$$

$$P(F{=}1,H{=}1)=0.0639$$

$$P(C{=}1|F{=}1,H{=}1)=0.063/0.0639$$

$$=0.986=98.6\%$$

cold	Fever	P(Fever | Cold)
0	0	0.98
0	1	0.02
1	0	0.30
1	1	0.70

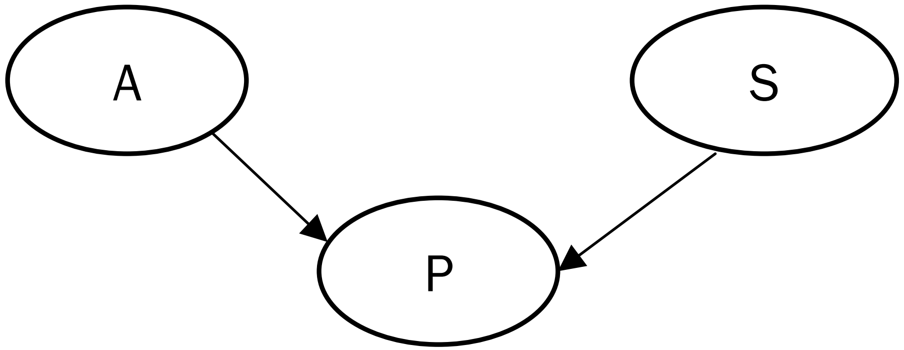
Cold	Cough	P(Cough | Cold)
0	0	0.95
0	1	0.05
1	0	0.10
1	1	0.90

cold	**P(cold)**
1	0.1
0	0.9

Imagine we are studying the academic performance of students in a course. We focus on three factors:

1. Attendance (A) 2. Studying (S) 3. Passing the Test (P)

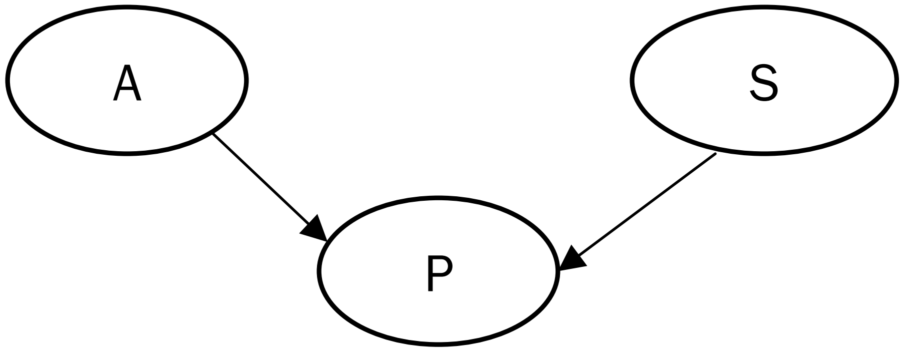
Obs	A (Attended)	S (Studied)	P (Passed)
1	1	1	1
2	1	1	1
3	1	1	0
4	1	0	0
5	1	1	1
6	1	0	0
7	0	0	0
8	0	0	1
9	0	1	0
10	0	0	0
11	0	1	1
12	1	1	1



Imagine we are studying the academic performance of students in a course. We focus on three factors:

1. Attendance (A) 2. Studying (S) 3. Passing the Test (P)

Obs	A (Attended)	S (Studied)	P (Passed)
1	1	1	1
2	1	1	1
3	1	1	0
4	1	0	0
5	1	1	1
6	1	0	0
7	0	0	0
8	0	0	1
9	0	1	0
10	0	0	0
11	0	1	1
12	1	1	1



Since A and S haven't parent variable, we can fill their CPT (Conditional Probability Table) just using P(A) or P(S) respectively

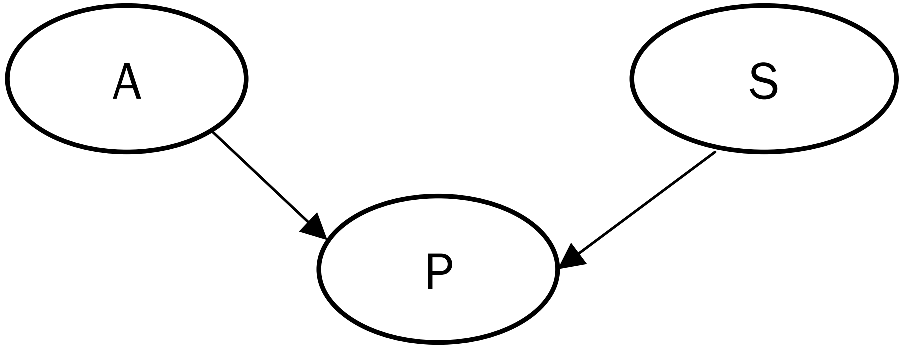
$$P(A=0) = \frac{5}{12} = 0.4167$$

$$P(A=1) = \frac{7}{12} = 0.5833$$

Imagine we are studying the academic performance of students in a course. We focus on three factors:

1. Attendance (A) 2. Studying (S) 3. Passing the Test (P)

Obs	A (Attended)	S (Studied)	P (Passed)
1	1	1	1
2	1	1	1
3	1	1	0
4	1	0	0
5	1	1	1
6	1	0	0
7	0	0	0
8	0	0	1
9	0	1	0
10	0	0	0
11	0	1	1
12	1	1	1



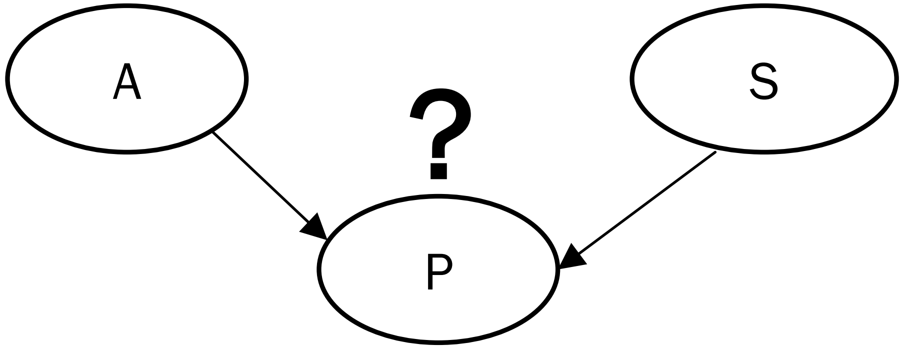
Since A and S haven't parent variable, we can fill their CPT (Conditional Probability Table) just using P(A) or P(S) respectively

$$P(S=0)=\frac{5}{12}=0.4167$$

$$P(S=1)=\frac{7}{12}=0.5833$$

Each node has its own CPT, till now we have the ones of A and S

A	P(A)
1	0.5833
0	0.4167



S	P(S)
1	0.5833
0	0.4167

Based on the structure of the given Bayesian network P is influenced by A and S, so we use the following computation to fill its CPT

$$P(A,S,P)=P(A)*P(S)*P(P|A,\ S)$$

Obs	A (Attended)	S (Studied)	P (Passed)
1	1	1	1
2	1	1	1
3	1	1	0
4	1	0	0
5	1	1	1
6	1	0	0
7	0	0	0
8	0	0	1
9	0	1	0
10	0	0	0
11	0	1	1
12	1	1	1

Marginal distribution of A ignore S and P

$P(A=1)= 7/12$ $P(A=0)= 5/12$

$P(S=1)=7/12$ $P(S=0)=5/12$

Conditional distribution

$P(P=1|A=1,\ S=1)=4/5$

$P(P=0|A=1,\ S=1)=1/5$

$P(P=1|A=1,\ S=0)=0$

$P(P=0|A=1,\ S=0)=2/2$

$P(P=1|A=0,\ S=1)=1/2$

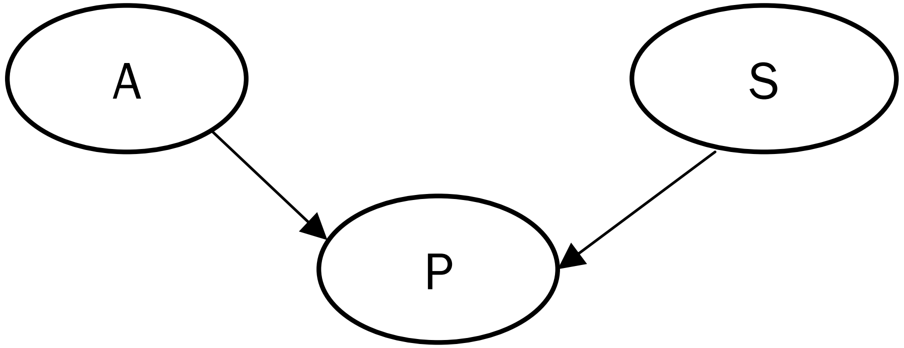
$P(P=0|A=0,\ S=1)=1/2$

$P(P=1|A=0,\ S=0)=1/3$

$P(P=0|A=0,\ S=0)=2/3$

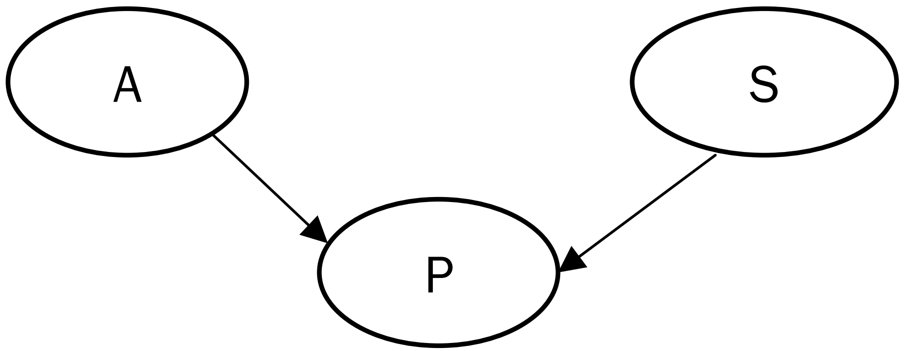
Each node has its own CPT, till now we have the ones of A and S

A	P(A)
1	0.5833
0	0.4167



S	P(S)
1	0.5833
0	0.4167

A (Attended)	S (Studied)	P(P | S, A)
0	0	0.3333
1	0	0
0	1	0.5
1	1	0.8



Based on the structure of the given Bayesian network P is influenced by A and S, so we use the following eauqtion to compute the joint distribution:

$$P(A, S, P) = P(A) * P(S) * P(P|A,\ S)$$

$$P(A,S,P)=P(A)*P(S)*P(P|A,\ S)$$

$$P(A=0,S=0,P=0)=P(A=0)\ \ P(S=0)\ \ P(P=0|A=0,\ S=0)$$

$$=\frac{5}{12}*\frac{5}{12}*\frac{2}{3}\quad\approx 0.116$$

$$P(A=1,S=0,P=0)=P(A=1)\ \ P(S=0)\ \ P(P=0|A=1,\ S=0)$$

$$=\frac{7}{12}*\frac{5}{12}*1\ \approx 0.243$$

$$P(A=0,S=1,P=0)=P(A=0)\ \ P(S=1)\ \ P(P=0|A=0,\ S=1)$$

$$=\frac{5}{12}*\frac{7}{12}*\frac{1}{2}\quad\approx 0.1215$$

$$P(A=1,S=1,P=0)=P(A=1)\ \ P(S=1)\ \ P(P=0|A=1,\ S=1)$$

$$=\frac{7}{12}*\frac{7}{12}*\frac{1}{5}\quad\approx 0.0680$$

$$P(A,S,P) = P(A) * P(S) * P(P|A,\ S)$$

$$P(A=0,S=0,P=1) = P(A=0)\ P(S=0)\ \ P(P=1|A=0,\ S=0)$$

$$= \frac{5}{12} * \frac{5}{12} * \frac{1}{3} \approx 0.058$$

$$P(A=1,S=0,P=1) = P(A=1)\ P(S=0)\ \ P(P=1|A=1,\ S=0)$$

$$= \frac{7}{12} * \frac{5}{12} * 0 = 0$$

$$P(A=0,S=1,P=1) = P(A=0)\ P(S=1)\ \ P(P=1|A=0,\ S=1)$$

$$= \frac{5}{12} * \frac{7}{12} * \frac{1}{2} \approx 0.121$$

$$P(A=1,S=1,P=1) = P(A=1)\ P(S=1)\ \ P(P=1|A=1,\ S=1)$$

$$= \frac{7}{12} * \frac{7}{12} * \frac{4}{5} \approx 0.272$$

$$P(A,S,P)=P(A)*P(S)*P(P|A,\ S)$$

$P(A=0,S=0,P=0)=$ ≈0.116

$P(A=1,S=0,P=0)=$ ≈0.243

$P(A=0,S=1,P=0)=$ ≈0.1215

$P(A=1,S=1,P=0)=$ ≈0.0680

$P(A=0,S=0,P=1)=$ ≈0.058

$P(A=1,S=0,P=1)=$ =0

$P(A=0,S=1,P=1)=$ ≈0.121

$P(A=1,S=1,P=1)=$ ≈0.272

A (Attended)	S (Studied)	P (Passed)	P(A, S, P)
0	0	0	0.116
1	0	0	0.243
0	1	0	0.1215
1	1	0	0.0680
0	0	1	0.058
1	0	1	0
0	1	1	0.121
1	1	1	0.272

Inference based on BN

Calculate the probability that some one passes P(P=1)

$$P(P=1) = P(A=1, S=1, P=1) + P(A=0, S=1, P=1) + P(A=1, S=0, P=1) + P(A=0, S=0, P=1)$$

$$P(P=1) = 0.272 + 0.121 + 0 + 0.058$$

$$= 0.451$$

Diagnostic Inference (Cause from Effect)

- **Query: P(S=1 | P=1)** (Probability of a student **studied**, given he **passed**).

This tells you the effectiveness of studying as a cause of passing, filtering out those who failed. This calculation confirms whether the act of studying is a good predictor of a successful outcome.

$$\mathbf{P(S=1|P=1)} = \frac{P(S=1, P=1)}{P(P=1)}$$

Calculate the Numerator: P(S=1, P=1)

$$P(S=1, P=1) = P(A=1, S=1, P=1) + P(A=0, S=1, P=1)$$

=0.272+0.121 = 0.393 = 0.393 / 0.451 = 0.87

Probability of student get **passed**, given he **attended**

$$P(P=1 \mid A=1) = P(P=1,\ A=1) / P(A=1)$$

$$P(P=1, S = 1\ \ A=1)+ P(P=1, S = 0, A=1) / P(A=1)$$

$$= (0+0.272) / 0.5833$$
$$= 0.466$$

$$P(A=1,S=1 \mid P=1) = P(P=1, S=1, A=1) \,/\, P(P=1)$$

$$= 0.272 \,/\, 0.451$$

$$= 0.60$$

$P(A=1 \mid P=1) = P(P=1,\ A=1) / P(P=1)$

$P(P=1, S = 1\ \ A=1) + P(P=1, S = 0, A=1) / P(A=1)$

$= 0.272 / 0.451$

$= 0.60$

1. Predictive queries

- $P(P = 1 \mid A = 1)$ — Probability of passing if the student attended.
- $P(P = 1 \mid S = 0)$ — Probability of passing if the student did not study.
- $P(P' = 1 \mid A = 1,\ S = 0)$ — Probability of passing given attended but did not study.

2. Diagnostic queries

- $P(S = 1 \mid P = 0)$ — Probability they studied given that they failed.
- $P(A = 1 \mid P = 0)$ — Probability they attended given that they failed.

3. "Explaining away" queries

- $P(S' = 1 \mid P = 1,\ A = 0)$ — If they passed without attending, how likely is it that they studied?
- Compare $P(S = 1 \mid P = 1)$ vs $P(S' = 1 \mid P = 1, A = 1)$ — seeing $A = 1$ reduces the likelihood that $S = 1$ caused $P = 1$.

Definition of a Bayesian Network

A **Bayesian Network (BN)** is a **probabilistic graphical model** that represents a joint probability distribution over a set of random variables using:

1. A Directed Acyclic Graph (DAG)

- Each **node** is a random variable.
- Each **directed edge** represents a **direct probabilistic dependency**.
- If there is an edge $X \rightarrow Y$, then **X is a parent of Y**, and Y's distribution depends on X.

2. A set of Conditional Probability Distributions (CPDs)

For each variable X, the BN contains a CPD:

$$P\big(X \mid Parents(X)\big)$$

This table defines the probability of X given the values of its parents in the graph

Factorization Property (Key Idea)

A Bayesian Network compactly represents the **full joint distribution**:

$$P(X_1, X_2, \ldots, X_n)$$

by factorizing it into local conditional probabilities:

$$\boxed{P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i \mid Parents(X_i)\right)}$$

Steps to Build a Bayesian Network

1-Identify the variables

Decide what random variables your BN will include.

Example (medical diagnosis):

Disease (D), Test Result (T), Symptom (S)

Example (weather):

Weather (W), Sprinkler (S), Grass Wet (G)

2-Determine causal relationships (edges)

Draw arrows **from causes to effects.**

Example:

Weather → Sprinkler → Grass Wet

Rules:

- Arrows = direct influence
- **No cycles** (DAG)
- Each node will have **parents** = the nodes that point to it

Steps to Build a Bayesian Network

3-Define parents for each node

- Parents = the nodes that directly affect this node
- CPTs will depend only on parents