bayesian networks

CSCI 373

$$P(x_{\female}, x_{\male}, x, y_{\female}, y_{\male}, y, z_{\female}, z_{\male}, z)$$

46656 rows

can we express this more compactly?

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1 \mid x_2, x_3, x_4) \overset{\boxed{y|x}}{=} \frac{P(x_1, x_2, x_3, x_4)}{P(x_2, x_3, x_4)}, \quad \text{so} \quad P(x_1, x_2, x_3, x_4) = P(x_1 \mid x_2, x_3, x_4)\, \underline{P(x_2, x_3, x_4)}$$

but what is this?

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1 \mid x_2, x_3, x_4) \overset{\boxed{y|x}}{=} \frac{P(x_1, x_2, x_3, x_4)}{P(x_2, x_3, x_4)}, \quad \text{so} \quad P(x_1, x_2, x_3, x_4) = P(x_1 \mid x_2, x_3, x_4)\, \underline{P(x_2, x_3, x_4)}$$

$$P(x_2 \mid x_3, x_4) \overset{\boxed{y|x}}{=} \frac{P(x_2, x_3, x_4)}{P(x_3, x_4)}, \quad \text{so} \quad \underline{P(x_2, x_3, x_4)} = P(x_2 \mid x_3, x_4)\, \underline{P(x_3, x_4)}$$

and what is this?

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1 | x_2, x_3, x_4) \overset{\boxed{y|x}}{=} \frac{P(x_1, x_2, x_3, x_4)}{P(x_2, x_3, x_4)}, \quad \text{so} \quad P(x_1, x_2, x_3, x_4) = P(x_1 | x_2, x_3, x_4)\, \underline{P(x_2, x_3, x_4)}$$

$$P(x_2 | x_3, x_4) \overset{\boxed{y|x}}{=} \frac{P(x_2, x_3, x_4)}{P(x_3, x_4)}, \quad \text{so} \quad \overline{P(x_2, x_3, x_4)} = P(x_2 | x_3, x_4)\, \underline{P(x_3, x_4)}$$

$$P(x_3 | x_4) \overset{\boxed{y|x}}{=} \frac{P(x_3, x_4)}{P(x_4)}, \quad \text{so} \quad \overline{P(x_3, x_4)} = P(x_3 | x_4)\, P(x_4)$$

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1, x_2, x_3, x_4) = P(x_1 | x_2, x_3, x_4) \, \underline{P(x_2, x_3, x_4)}$$

$$\overline{P(x_2, x_3, x_4)} = P(x_2 | x_3, x_4) \, \underline{P(x_3, x_4)}$$

$$\overline{P(x_3, x_4)} = P(x_3 | x_4) \, P(x_4)$$

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1, x_2, x_3, x_4) = P(x_1 \mid x_2, x_3, x_4)\, \underline{P(x_2, x_3, x_4)}$$

$$\overline{P(x_2, x_3, x_4)} = P(x_2 \mid x_3, x_4)\, \underline{P(x_3 \mid x_4)\, P(x_4)}$$

$$\overline{P(x_3, x_4)} = P(x_3 \mid x_4)\, P(x_4)$$

idea: re-express joint probability as a product of conditional probabilities

$$P(x_1, x_2, x_3, x_4) = P(x_1 | x_2, x_3, x_4)\, \underline{P(x_2 | x_3, x_4)\, P(x_3 | x_4)\, P(x_4)}$$

$$\overline{P(x_2, x_3, x_4)} = P(x_2 | x_3, x_4)\, \underline{P(x_3 | x_4)\, P(x_4)}$$

$$\overline{P(x_3, x_4)} = P(x_3 | x_4)\, P(x_4)$$

this is called the chain rule of probability

$$P(x_1, x_2, x_3, x_4) = P(x_1 | x_2, x_3, x_4)\, P(x_2 | x_3, x_4)\, P(x_3 | x_4)\, P(x_4)$$

$$\boxed{P(x_1, \dots, x_n) = \prod_{i=1}^{n} P(x_i | x_{i+1}, \dots, x_n)}$$

given variables $X_1, \dots, X_n$ with domains $D(X_i)$ and joint probability $P(x_1, \dots, x_n)$

$$P(x_1, \dots, x_m) = \sum_{x_{m+1} \in D(X_{m+1})} \cdots \sum_{x_n \in D(X_n)} P(x_1, \dots, x_n)$$

$$P(x_{m+1}, \dots, x_n \mid x_1, \dots, x_m) = \frac{P(x_1, \dots, x_n)}{P(x_1, \dots, x_m)}$$

$$P(x_{m+1}, \dots, x_n \mid x_1, \dots, x_m) = \frac{P(x_1, \dots, x_m \mid x_{m+1}, \dots, x_n)\, P(x_1, \dots, x_m)}{P(x_{m+1}, \dots, x_n)}$$

$$P(x_1, \dots, x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i+1}, \dots, x_n)$$

46656 rows

$$P(x_{♀}, x_{♂}, x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z)$$

can we express this more compactly?

46656 rows

$$P(x_{♀}, x_{♂}, x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z)$$

$$\begin{aligned} = \; & P(x_{♀} \mid x_{♂}, x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\ \cdot \; & P(x_{♂} \mid x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\ \cdot \; & P(x \mid y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\ \cdot \; & P(y_{♀} \mid y_{♂}, y, z_{♀}, z_{♂}, z) \\ \cdot \; & P(y_{♂} \mid y, z_{♀}, z_{♂}, z) \\ \cdot \; & P(y_{♂} \mid z_{♀}, z_{♂}, z) \\ \cdot \; & P(z_{♀} \mid z_{♂}, z) \\ \cdot \; & P(z_{♂} \mid z) \\ \cdot \; & P(z) \end{aligned}$$

did this save us space?

46656 rows

$$P(x_{♀}, x_{♂}, x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z)$$

$$\begin{aligned}
= \; & P(x_{♀} \mid x_{♂}, x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\
\cdot \; & P(x_{♂} \mid x, y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\
\cdot \; & P(x \mid y_{♀}, y_{♂}, y, z_{♀}, z_{♂}, z) \\
\cdot \; & P(y_{♀} \mid y_{♂}, y, z_{♀}, z_{♂}, z) \\
\cdot \; & P(y_{♂} \mid y, z_{♀}, z_{♂}, z) \\
\cdot \; & P(y_{♂} \mid z_{♀}, z_{♂}, z) \\
\cdot \; & P(z_{♀} \mid z_{♂}, z) \\
\cdot \; & P(z_{♂} \mid z) \\
\cdot \; & P(z)
\end{aligned}$$

no, not yet,

but...

$$P(y, z) = P(z \mid y) \cdot P(y)$$

y	z	$P(y, z)$
A	A	
A	B	
A	AB	
A	O	
B	A	
B	B	
B	AB	
B	O	
AB	A	
AB	B	
AB	AB	
AB	O	
O	A	
O	B	
O	AB	
O	O	

y	z	$P(z \mid y)$
A	A	
A	B	
A	AB	
A	O	
B	A	
B	B	
B	AB	
B	O	
AB	A	
AB	B	
AB	AB	
AB	O	
O	A	
O	B	
O	AB	
O	O	

y	$P(y)$
A	
B	
AB	
O	

do we have to specify all these values?

$$P(y, z) = P(z \mid y) \cdot P(y)$$

y	z	$P(y,z)$
A	A	
A	B	
A	AB	
A	O	
B	A	
B	B	
B	AB	
B	O	
AB	A	
AB	B	
AB	AB	
AB	O	
O	A	
O	B	
O	AB	
O	O	×

15 values

y	z	$P(z \mid y)$
A	A	
A	B	
A	AB	
A	O	×
B	A	
B	B	
B	AB	
B	O	×
AB	A	
AB	B	
AB	AB	
AB	O	×
O	A	
O	B	
O	AB	
O	O	×

12 values

y	$P(y)$
A	
B	
AB	
O	×

3 values

46655 rows to specify

$$P(x_{\female}, x_{\male}, x, y_{\female}, y_{\male}, y, z_{\female}, z_{\male}, z)$$

$$
\begin{aligned}
= \; & P(x_{\female} \mid x_{\male}, x, y_{\female}, y_{\male}, y, z_{\female}, z_{\male}, z) \\
\cdot \; & P(x_{\male} \mid x, y_{\female}, y_{\male}, y, z_{\female}, z_{\male}, z) \\
\cdot \; & P(x \mid y_{\female}, y_{\male}, y, z_{\female}, z_{\male}, z) \\
\cdot \; & P(y_{\female} \mid y_{\male}, y, z_{\female}, z_{\male}, z) \\
\cdot \; & P(y_{\male} \mid y, z_{\female}, z_{\male}, z) \\
\cdot \; & P(y_{\male} \mid z_{\female}, z_{\male}, z) \\
\cdot \; & P(z_{\female} \mid z_{\male}, z) \\
\cdot \; & P(z_{\male} \mid z) \\
\cdot \; & P(z)
\end{aligned}
$$

no, not yet,

but...

$$P(z, z_{\female}, z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$

$$= P(z \mid z_{\female}, z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$
$$\cdot P(z_{\female} \mid z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$
$$\cdot P(z_{\male} \mid y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$
$$\cdot P(y \mid y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$
$$\cdot P(y_{\female} \mid y_{\male}, x, x_{\female}, x_{\male})$$
$$\cdot P(y_{\male} \mid x, x_{\female}, x_{\male})$$
$$\cdot P(x \mid x_{\female}, x_{\male})$$
$$\cdot P(x_{\female} \mid x_{\male})$$
$$\cdot P(x_{\male})$$

first, observe that we can put the variables in any order before applying the chain rule

$$P(z, z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$

$$= P(z \mid z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$
$$\cdot\, P(z_{♀} \mid z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$
$$\cdot\, P(z_{♂} \mid y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$
$$\cdot\, P(y \mid y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$
$$\cdot\, P(y_{♀} \mid y_{♂}, x, x_{♀}, x_{♂})$$
$$\cdot\, P(y_{♂} \mid x, x_{♀}, x_{♂})$$
$$\cdot\, P(x \mid x_{♀}, x_{♂})$$
$$\cdot\, P(x_{♀} \mid x_{♂})$$
$$\cdot\, P(x_{♂})$$

next, observe:

$Z_{♀} = \text{A}$ $Z_{♂} = \text{O}$

$Z = \text{A}$

$$Z \perp\!\!\!\perp \{X, X_{♀}, X_{♂}, Y, Y_{♀}, Y_{♂}\} \mid \{Z_{♀}, Z_{♂}\}$$

i.e. $P(z \mid z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂}) = P(z \mid z_{♀}, z_{♂})$

$$P(z, z_{\female}, z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$

$$\begin{aligned}
= \; & P(z \mid z_{\female}, z_{\male}) \\
\cdot \; & P(z_{\female} \mid z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(z_{\male} \mid y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y \mid y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y_{\female} \mid y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y_{\male} \mid x, x_{\female}, x_{\male}) \\
\cdot \; & P(x \mid x_{\female}, x_{\male}) \\
\cdot \; & P(x_{\female} \mid x_{\male}) \\
\cdot \; & P(x_{\male})
\end{aligned}$$

what other conditional independence relationships are there?

$$P(z, z_{\female}, z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male})$$

$$\begin{aligned}
= \; & P(z \mid z_{\female}, z_{\male}) \\
\cdot \; & P(z_{\female} \mid z_{\male}, y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(z_{\male} \mid y, y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y \mid y_{\female}, y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y_{\female} \mid y_{\male}, x, x_{\female}, x_{\male}) \\
\cdot \; & P(y_{\male} \mid x, x_{\female}, x_{\male}) \\
\cdot \; & P(x \mid x_{\female}, x_{\male}) \\
\cdot \; & P(x_{\female} \mid x_{\male}) \\
\cdot \; & P(x_{\male})
\end{aligned}$$

what other conditional independence relationships are there?

$$P(z, z_♀, z_♂, y, y_♀, y_♂, x, x_♀, x_♂)$$

$$
\begin{aligned}
= \; & P(z \mid z_♀, z_♂) \\
\cdot\; & P(z_♀ \mid x_♀, x_♂) \\
\cdot\; & P(z_♂ \mid y_♀, y_♂) \\
\cdot\; & P(y \mid y_♀, y_♂) \\
\cdot\; & P(y_♀) \\
\cdot\; & P(y_♂) \\
\cdot\; & P(x \mid x_♀, x_♂) \\
\cdot\; & P(x_♀) \\
\cdot\; & P(x_♂)
\end{aligned}
$$

did this save us space?

$$P(z, z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂}) \leftarrow \text{46656 rows}$$

$$
\left.
\begin{aligned}
= \; & P(z \mid z_{♀}, z_{♂}) && \leftarrow 4 \cdot 3 \cdot 3 = 36 \text{ rows} \\
\cdot \; & P(z_{♀} \mid x_{♀}, x_{♂}) && \leftarrow 3 \cdot 3 \cdot 3 = 27 \text{ rows} \\
\cdot \; & P(z_{♂} \mid y_{♀}, y_{♂}) && \leftarrow 3 \cdot 3 \cdot 3 = 27 \text{ rows} \\
\cdot \; & P(y \mid y_{♀}, y_{♂}) && \leftarrow 4 \cdot 3 \cdot 3 = 36 \text{ rows} \\
\cdot \; & P(y_{♀}) && \leftarrow 3 \text{ rows} \\
\cdot \; & P(y_{♂}) && \leftarrow 3 \text{ rows} \\
\cdot \; & P(x \mid x_{♀}, x_{♂}) && \leftarrow 4 \cdot 3 \cdot 3 = 36 \text{ rows} \\
\cdot \; & P(x_{♀}) && \leftarrow 3 \text{ rows} \\
\cdot \; & P(x_{♂}) && \leftarrow 3 \text{ rows}
\end{aligned}
\right\} \text{174 rows}
$$

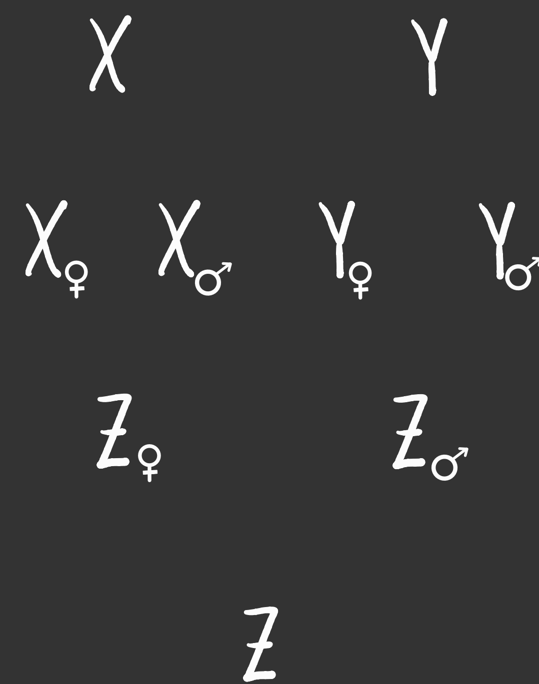
$$P(z, z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂})$$

$$= P(z \mid z_{♀}, z_{♂}) \cdot P(z_{♀} \mid x_{♀}, x_{♂}) \cdot P(z_{♂} \mid y_{♀}, y_{♂}) \cdot P(y \mid y_{♀}, y_{♂}) \cdot P(y_{♀}) \cdot P(y_{♂}) \cdot P(x \mid x_{♀}, x_{♂}) \cdot P(x_{♀}) \cdot P(x_{♂})$$

we can represent our independence assumptions with a graph

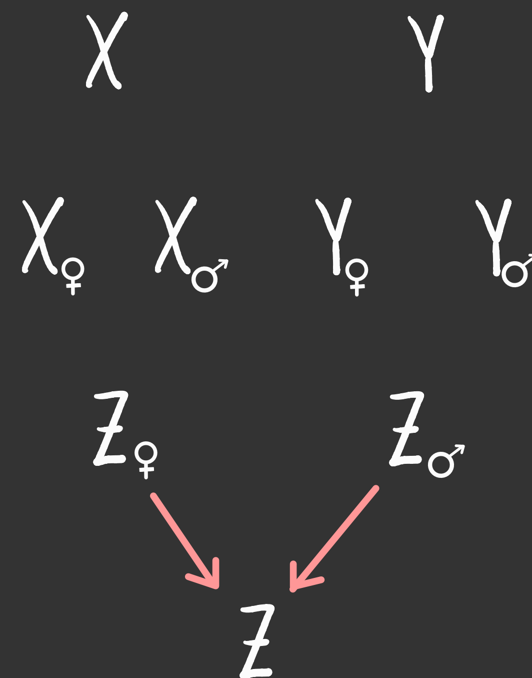
we create a vertex for each variable

$$
\begin{aligned}
P(z, z_{♀}, z_{♂}, y, y_{♀}, y_{♂}, x, x_{♀}, x_{♂}) = \; & P(z \mid z_{♀}, z_{♂}) \\
& \cdot P(z_{♀} \mid x_{♀}, x_{♂}) \\
& \cdot P(z_{♂} \mid y_{♀}, y_{♂}) \\
& \cdot P(y \mid y_{♀}, y_{♂}) \\
& \cdot P(y_{♀}) \\
& \cdot P(y_{♂}) \\
& \cdot P(x \mid x_{♀}, x_{♂}) \\
& \cdot P(x_{♀}) \\
& \cdot P(x_{♂})
\end{aligned}
$$



the conditioning set of a variable are its parents in the graph

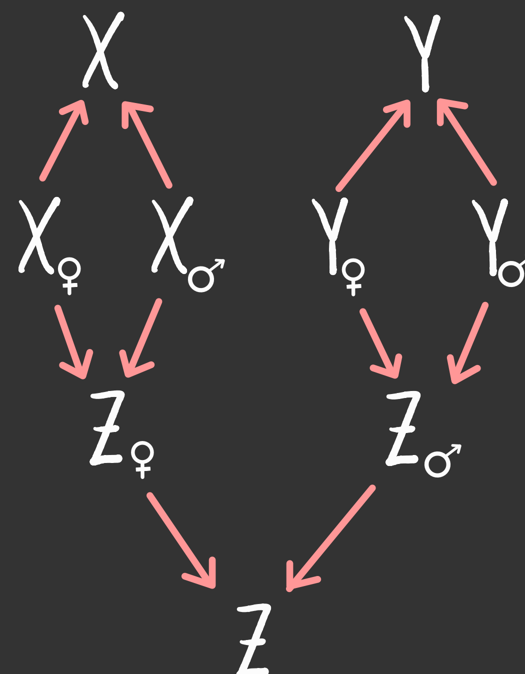
$$
\begin{aligned}
P(z, z_{\text{♀}}, z_{\text{♂}}, y, y_{\text{♀}}, y_{\text{♂}}, x, x_{\text{♀}}, x_{\text{♂}}) = \; & P(z \mid z_{\text{♀}}, z_{\text{♂}}) \\
& \cdot P(z_{\text{♀}} \mid x_{\text{♀}}, x_{\text{♂}}) \\
& \cdot P(z_{\text{♂}} \mid y_{\text{♀}}, y_{\text{♂}}) \\
& \cdot P(y \mid y_{\text{♀}}, y_{\text{♂}}) \\
& \cdot P(y_{\text{♀}}) \\
& \cdot P(y_{\text{♂}}) \\
& \cdot P(x \mid x_{\text{♀}}, x_{\text{♂}}) \\
& \cdot P(x_{\text{♀}}) \\
& \cdot P(x_{\text{♂}})
\end{aligned}
$$



this graph is called a

bayesian network

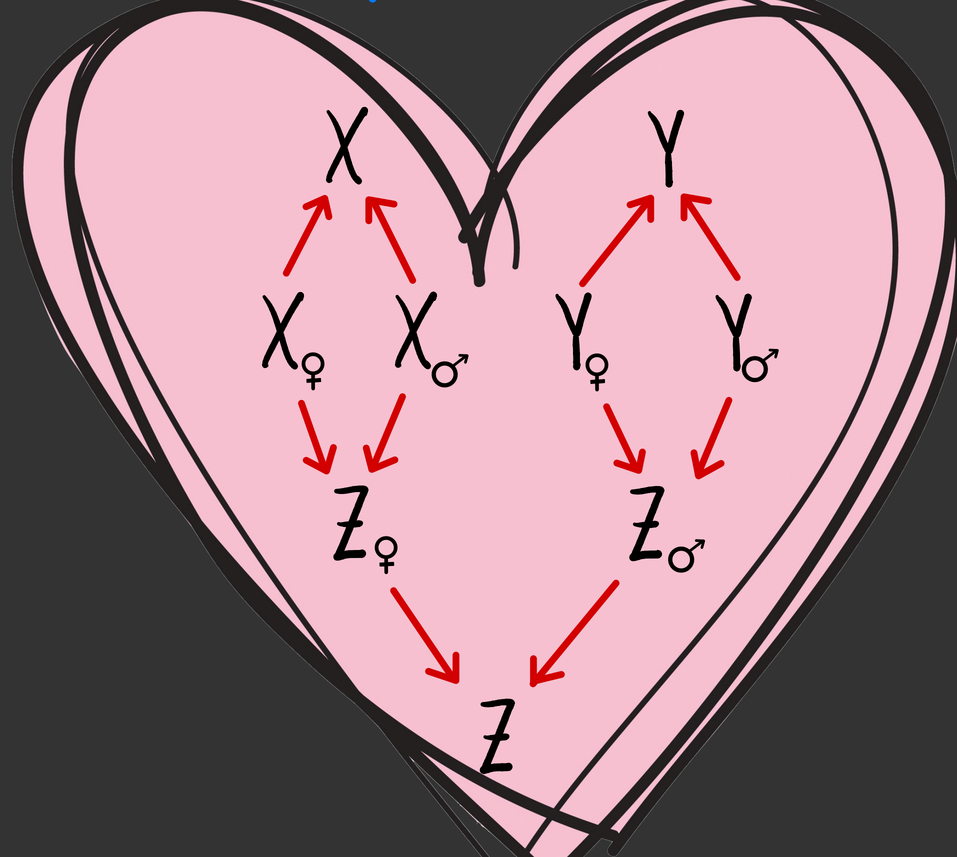
$$
\begin{aligned}
P(z, z_{\text{♀}}, z_{\text{♂}}, y, y_{\text{♀}}, y_{\text{♂}}, x, x_{\text{♀}}, x_{\text{♂}}) = \; & P(z \mid z_{\text{♀}}, z_{\text{♂}}) \\
& \cdot P(z_{\text{♀}} \mid x_{\text{♀}}, x_{\text{♂}}) \\
& \cdot P(z_{\text{♂}} \mid y_{\text{♀}}, y_{\text{♂}}) \\
& \cdot P(y \mid y_{\text{♀}}, y_{\text{♂}}) \\
& \cdot P(y_{\text{♀}}) \\
& \cdot P(y_{\text{♂}}) \\
& \cdot P(x \mid x_{\text{♀}}, x_{\text{♂}}) \\
& \cdot P(x_{\text{♀}}) \\
& \cdot P(x_{\text{♂}})
\end{aligned}
$$



this graph is called a

bayesian network

$$
\begin{aligned}
P(z, z_{\text{♀}}, z_{\text{♂}}, y, y_{\text{♀}}, y_{\text{♂}}, x, x_{\text{♀}}, x_{\text{♂}}) = \; & P(z \mid z_{\text{♀}}, z_{\text{♂}}) \\
\cdot \; & P(z_{\text{♀}} \mid x_{\text{♀}}, x_{\text{♂}}) \\
\cdot \; & P(z_{\text{♂}} \mid y_{\text{♀}}, y_{\text{♂}}) \\
\cdot \; & P(y \mid y_{\text{♀}}, y_{\text{♂}}) \\
\cdot \; & P(y_{\text{♀}}) \\
\cdot \; & P(y_{\text{♂}}) \\
\cdot \; & P(x \mid x_{\text{♀}}, x_{\text{♂}}) \\
\cdot \; & P(x_{\text{♀}}) \\
\cdot \; & P(x_{\text{♂}})
\end{aligned}
$$

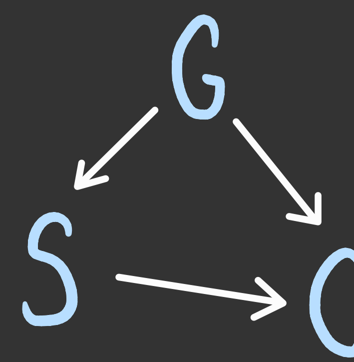


let S represent whether a person smokes

let C represent whether a person gets lung cancer

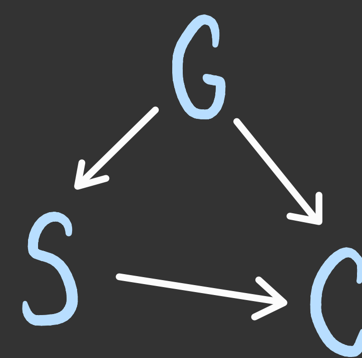
let G represent a particular gene

what claims does this bayesian network make?



$$P(g, s, c) = P(g)P(s \mid g)P(c \mid s, g)$$

what claims does this bayesian network make?

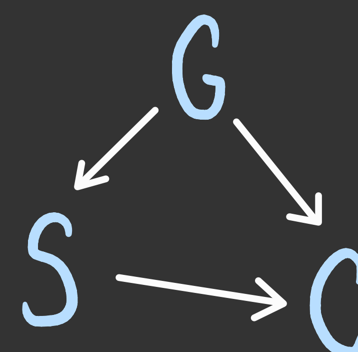


this is just the chain rule

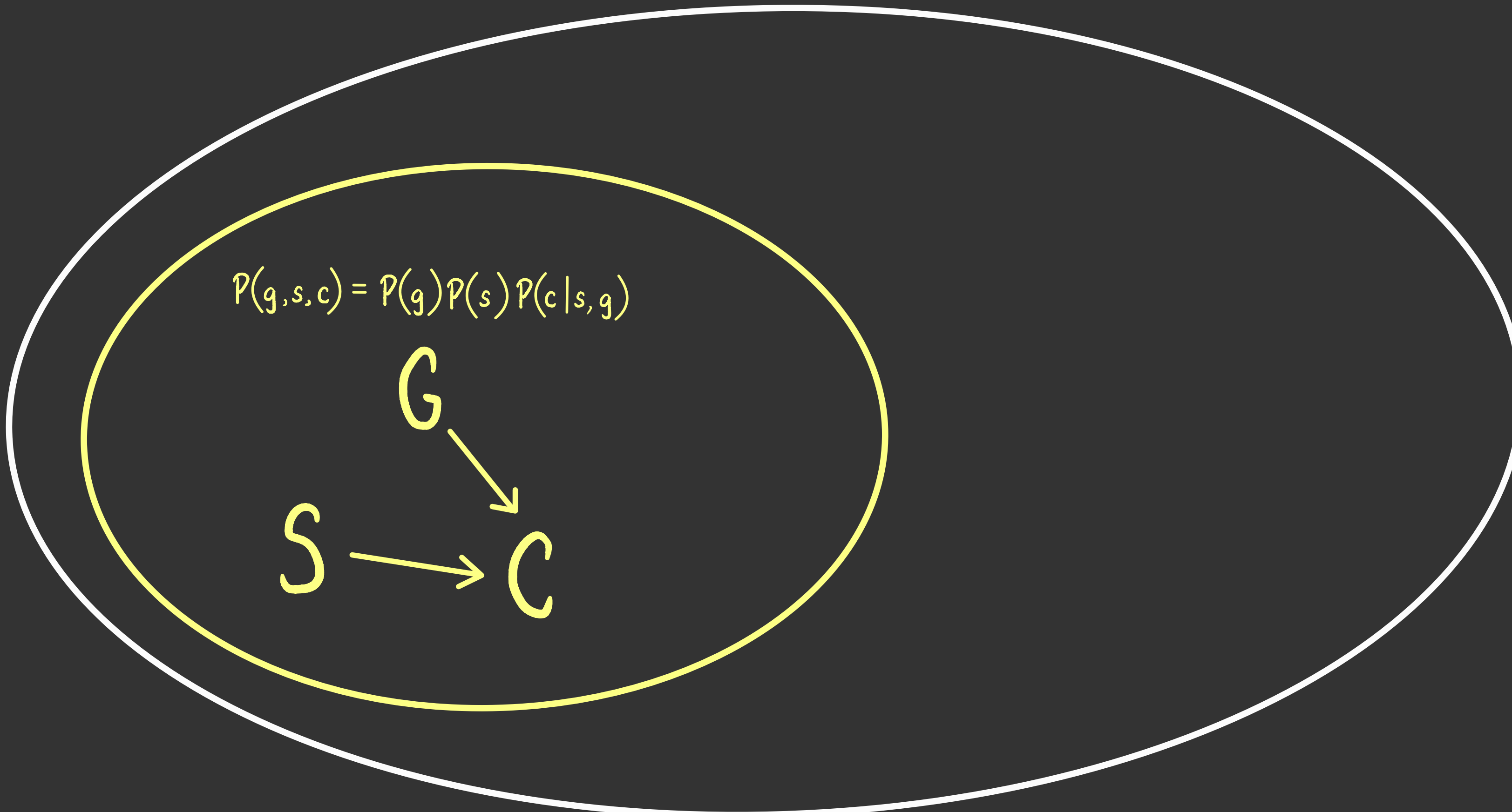
$$P(g, s, c) = P(g)\,P(s \mid g)\,P(c \mid s, g)$$

which holds for **every** joint distribution

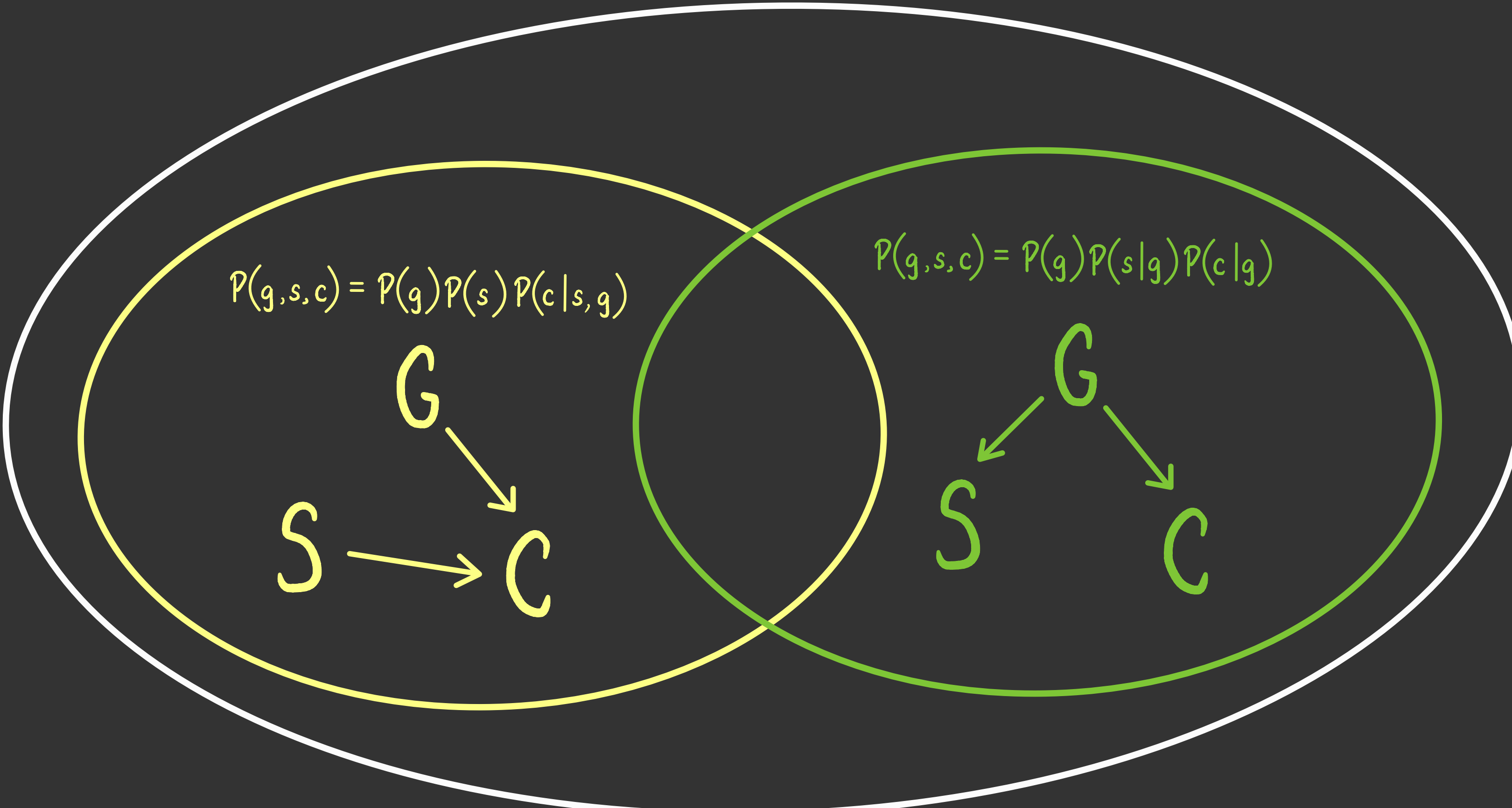
so this bayesian network makes <u>no</u> claims



the space of all joint distributions $P(g, s, c)$



the space of all joint distributions $P(g, s, c)$



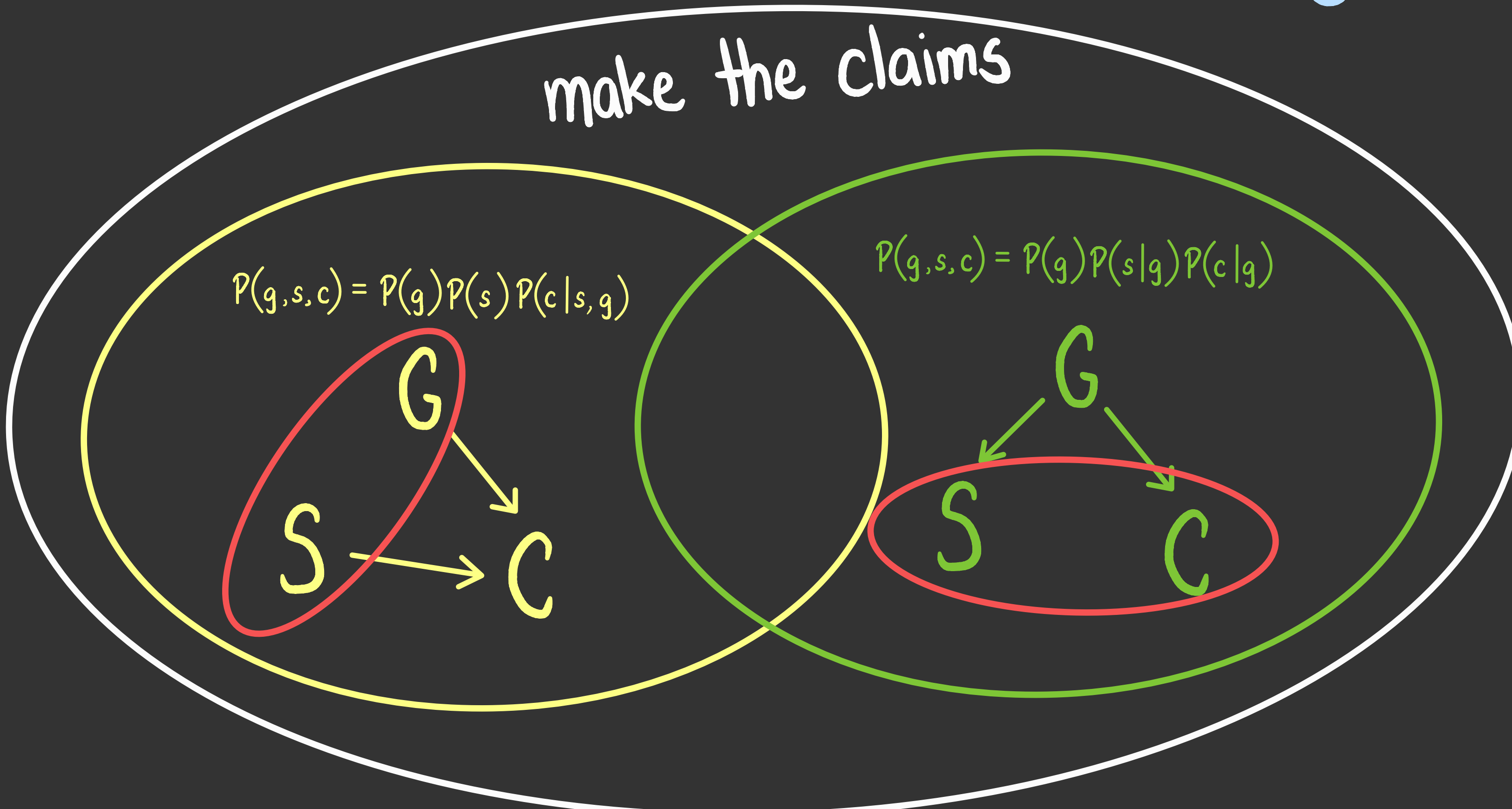
the space of all joint distributions $P(g,s,c)$

$$P(g,s,c) = P(g)P(s|g)P(c|s,g)$$

$$\begin{tikzcd}
 & G \arrow[dl] \arrow[dr] & \\
S \arrow[rr] & & C
\end{tikzcd}$$

excludes no distributions

in a bayesian network, the **absent edges**



$\boxed{X \rightarrow Y}$ $\boxed{X \leftarrow Y}$

$\boxed{X \quad Y}$

Consider the three

2-variable bayesian

networks

Consider the three 2-variable bayesian networks

$\boxed{X \rightarrow Y}$

claims that:

$$P(x,y) = P(x)P(y|x)$$

$\boxed{X \leftarrow Y}$

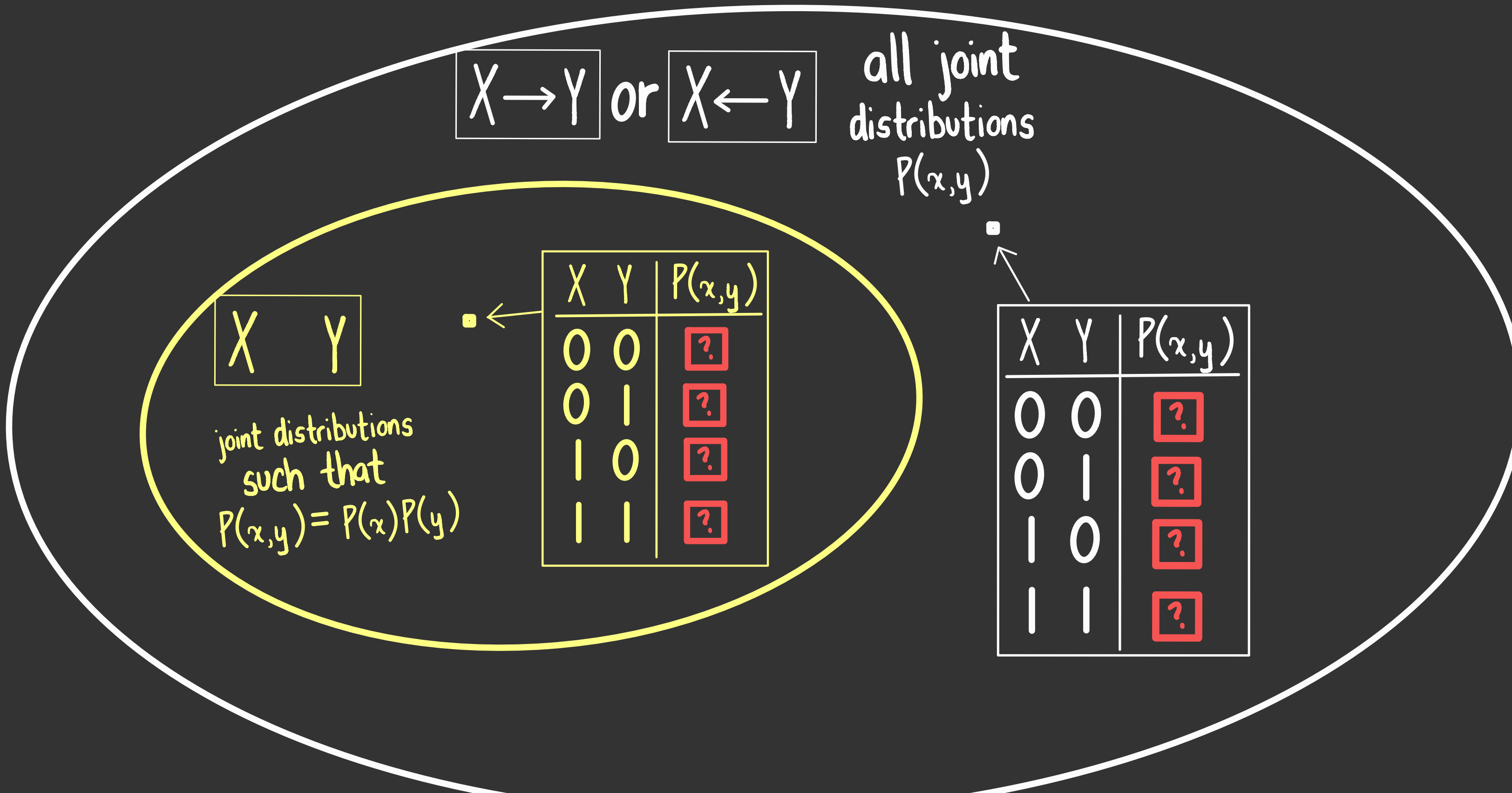
claims that:

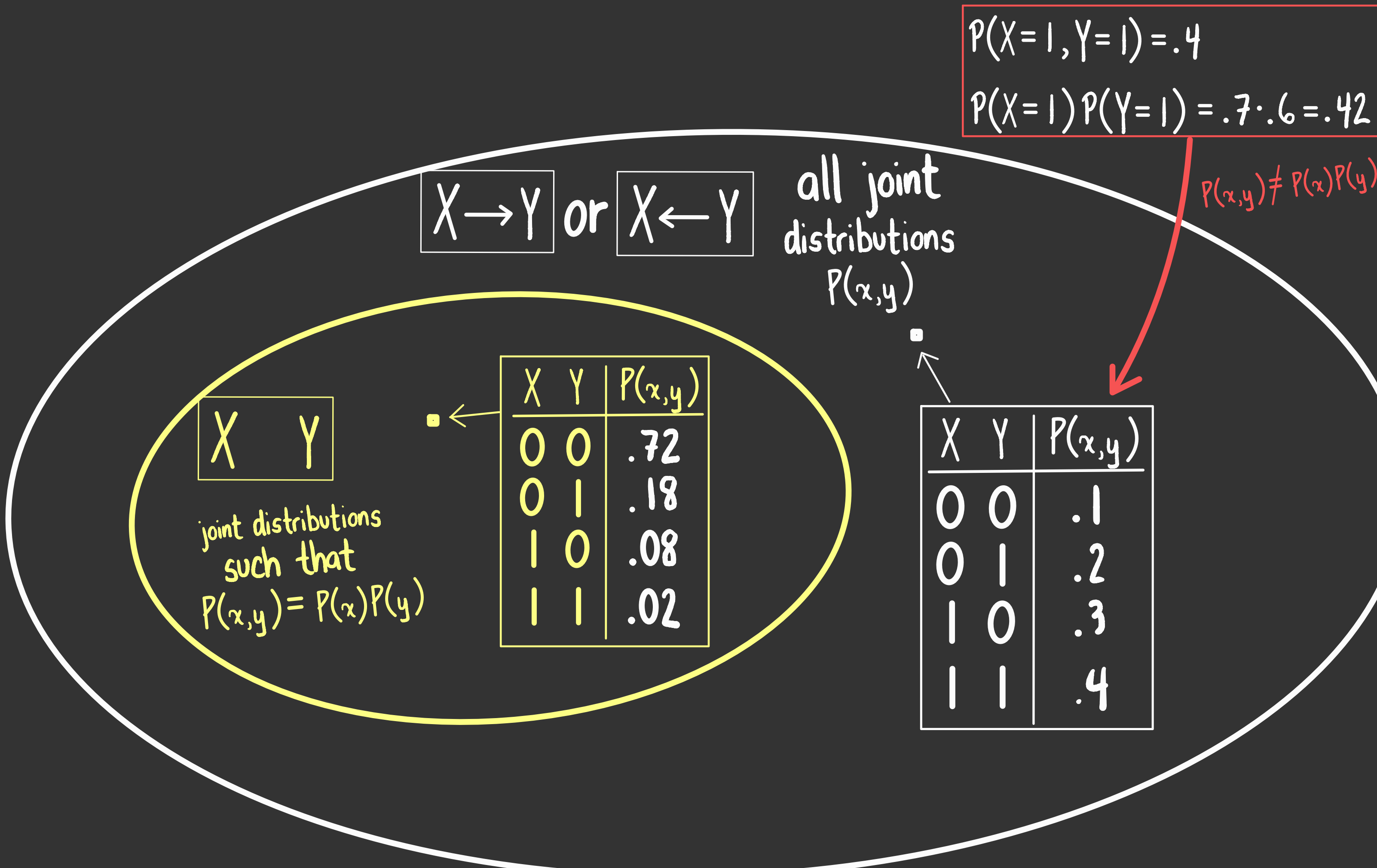
$$P(x,y) = P(y)P(x|y)$$

$\boxed{X \quad Y}$

claims that:

$$P(x,y) = P(x)P(y)$$


X → Y
or
X ← Y
all joint distributions P(x,y)
X Y
joint distributions such that P(x,y) = P(x)P(y)
X Y P(x,y)
0 0 ?
0 1 ?
1 0 ?
1 1 ?
X Y P(x,y)
0 0 ?
0 1 ?
1 0 ?
1 1 ?



$P(X=1, Y=1) = .4$
$P(X=1)P(Y=1) = .7 \cdot .6 = .42$
$P(x,y) \neq P(x)P(y)$
$X \rightarrow Y$ or $X \leftarrow Y$
all joint distributions $P(x,y)$
X Y
joint distributions such that $P(x,y) = P(x)P(y)$
X Y P(x,y)
0 0 .72
0 1 .18
1 0 .08
1 1 .02
X Y P(x,y)
0 0 .1
0 1 .2
1 0 .3
1 1 .4