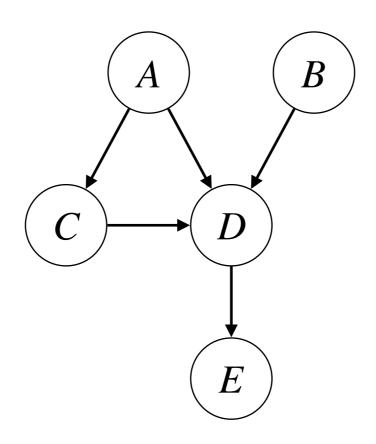
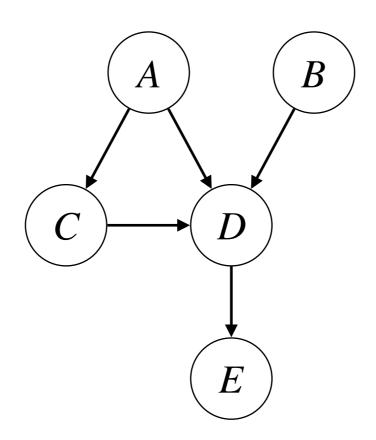


Directed acyclic graph (DAG)

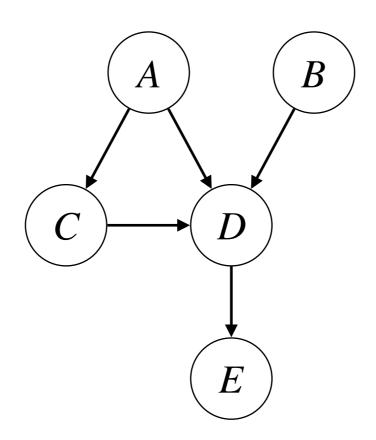


Directed acyclic graph (DAG)



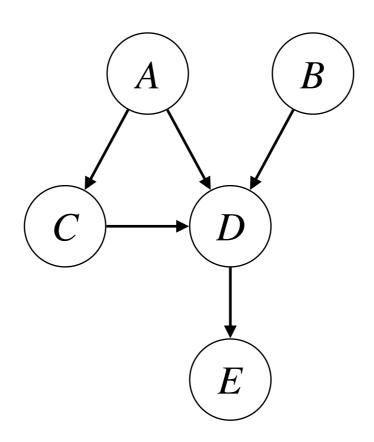
Each node is a random variable.

Directed acyclic graph (DAG)



Each node is a random variable.

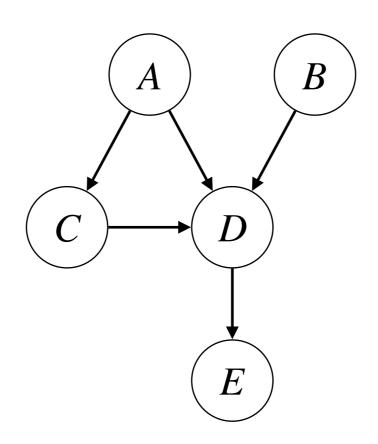
Directed acyclic graph (DAG)



Each node is a random variable.

The graph defines a joint distribution over the variables.

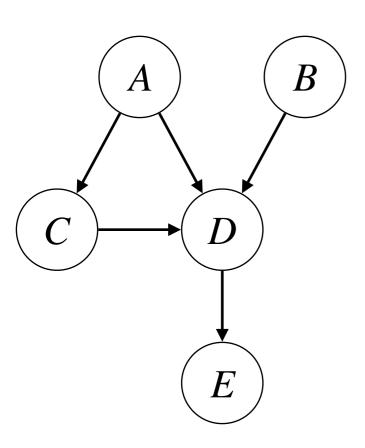
Directed acyclic graph (DAG)



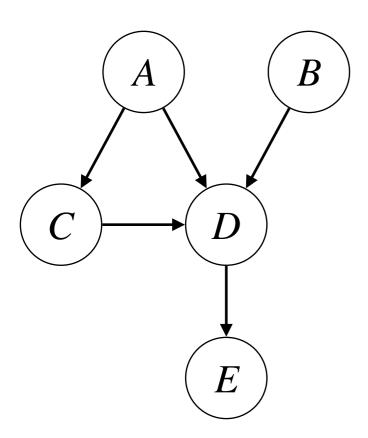
Each node is a random variable.

The graph defines a joint distribution over the variables.

P(A, B, C, D, E)

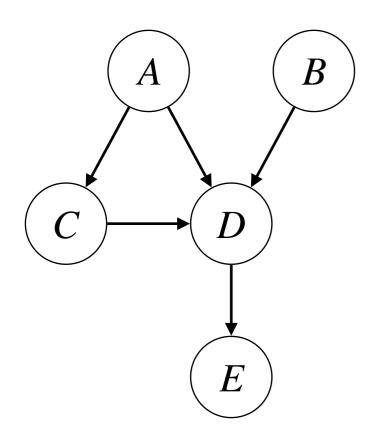


$$P(A, B, C, D, E) =$$



joint distribution = product of local distributions

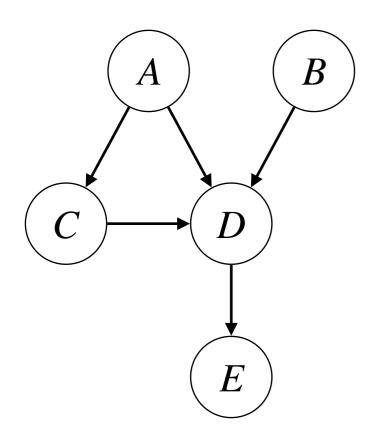
$$P(A, B, C, D, E) =$$



product of 5 local distributions

joint distribution = product of local distributions

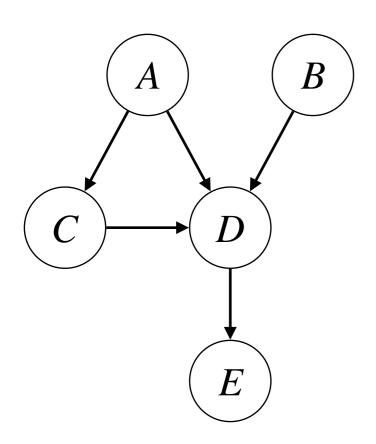
$$P(A, B, C, D, E) =$$



product of 5 local distributions

joint distribution = product of local distributions

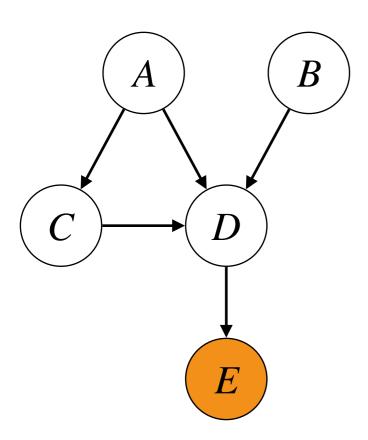
$$P(A, B, C, D, E) =$$



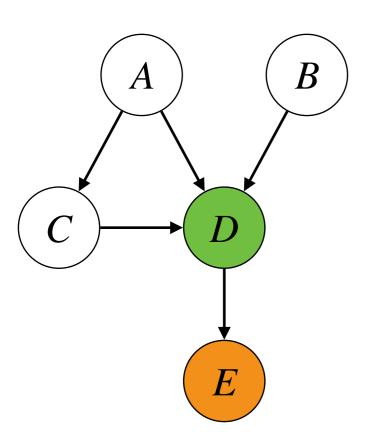
product of 5 local distributions

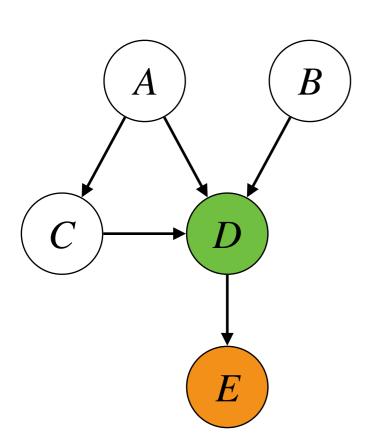
(one for each node in the graph)

$$P(A, B, C, D, E) =$$



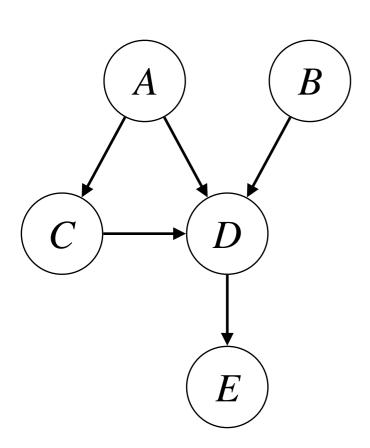
$$P(A, B, C, D, E) =$$





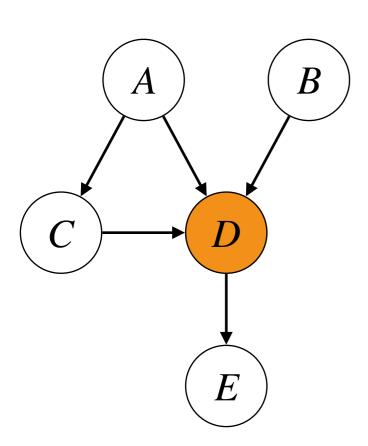
$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$



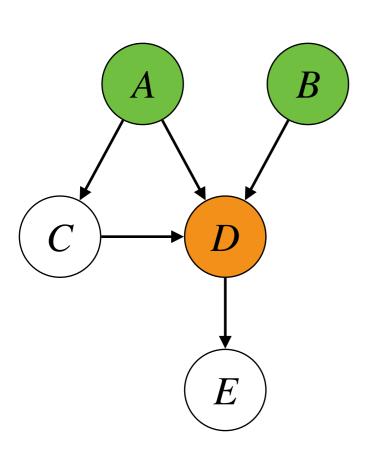
$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$



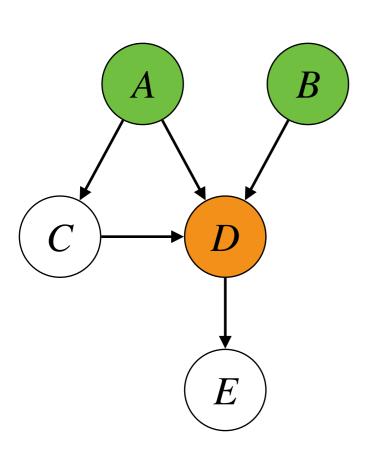
$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$



$$P(A, B, C, D, E) =$$

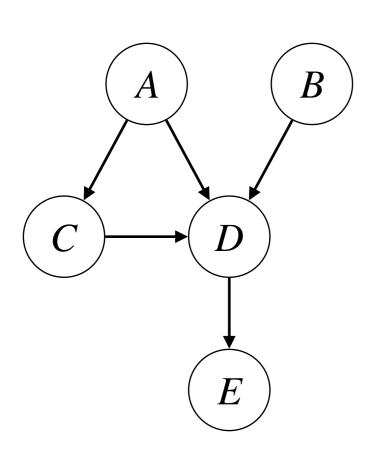
$$P(E \mid D)$$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

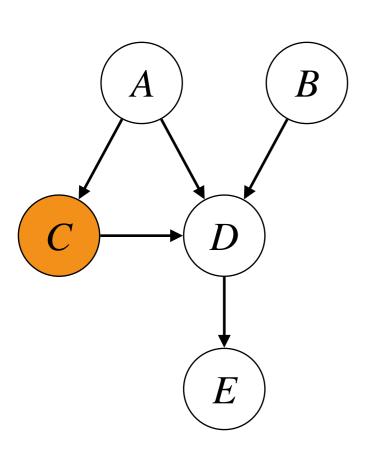
$$P(D \mid A, B)$$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

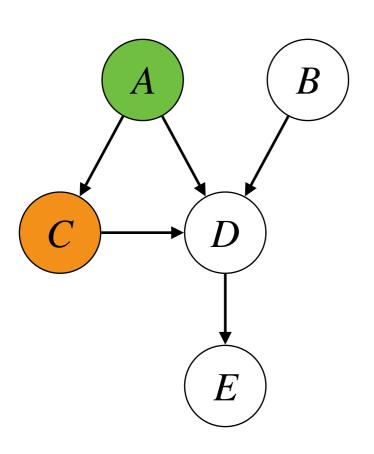
$$P(D \mid A, B)$$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

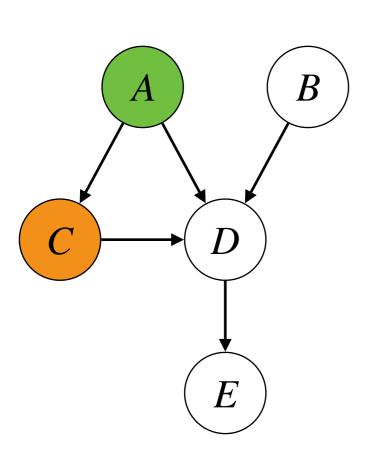
$$P(D \mid A, B)$$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

$$P(D \mid A, B)$$

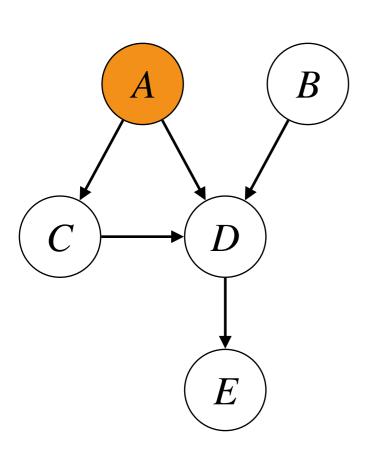


$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

$$P(D \mid A, B)$$

$$P(C \mid A)$$

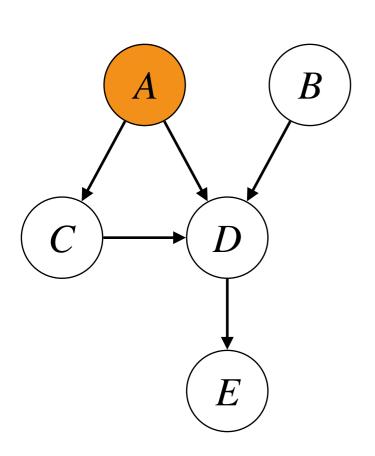


$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

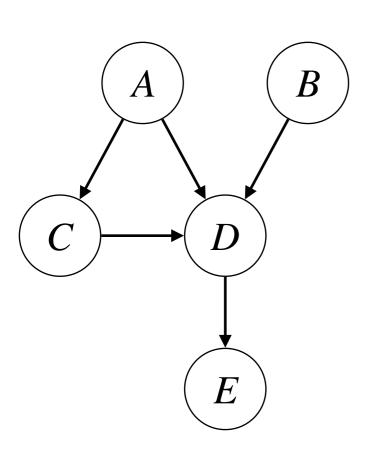
$$P(D \mid A, B)$$

$$P(C \mid A)$$



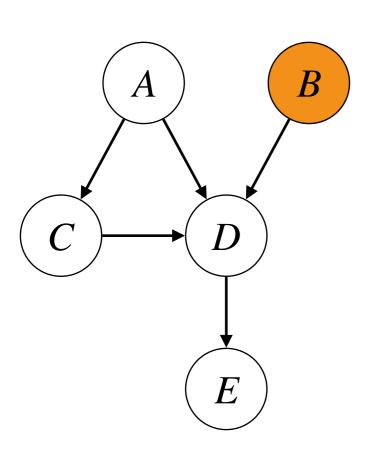
$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$
 $P(D \mid A, B)$
 $P(C \mid A)$
 $P(A)$



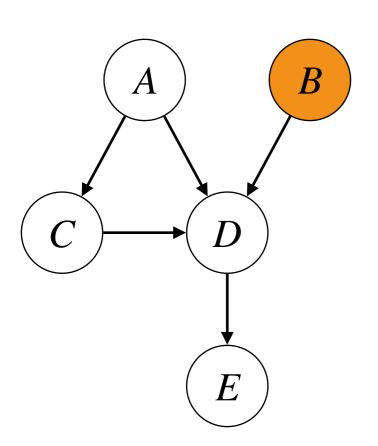
$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$
 $P(D \mid A, B)$
 $P(C \mid A)$
 $P(A)$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$
 $P(D \mid A, B)$
 $P(C \mid A)$
 $P(A)$

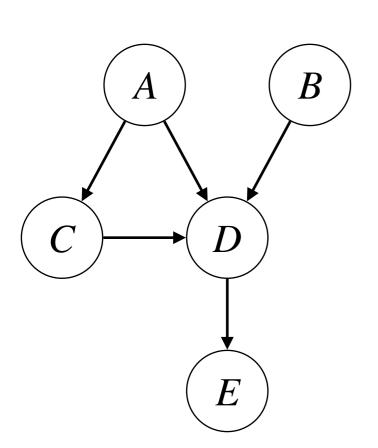


$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$

$$P(D \mid A, B)$$

$$P(C \mid A)$$



$$P(A, B, C, D, E) =$$

$$P(E \mid D)$$
 $P(D \mid A, B)$
 $P(C \mid A)$
 $P(A)$
 $P(B)$

Learning the parameters of a Bayesian network

data = {(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)} model = $p_G(g)$ $p_R(r \mid g)$ data = {(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)} model = $p_G(g)$ $p_R(r \mid g)$

data =
$$\{(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)\}$$

$$model = p_G(g) p_R(r \mid g)$$

g	$p_G(g)$
d	
c	

data =
$$\{(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)\}$$

$$model = p_G(g) p_R(r \mid g)$$

g	$p_G(g)$
d	
c	

data =
$$\{(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)\}$$

$$model = p_G(g) p_R(r \mid g)$$

g	$p_G(g)$
d	3/5
c	

data =
$$\{(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)\}$$

$$model = p_G(g) p_R(r \mid g)$$

g	$p_G(g)$
d	3 / 5
c	2/5

data = {(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)} model = $p_G(g)$ $p_R(r \mid g)$ data = {(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)} model = $p_G(g)$ $p_R(r \mid g)$

<u>g</u>	r	$p_R(r \mid g)$
d	1	
d	2	
d	3	
d	4	
d	5	
c	1	
C	2	
C	3	
c	4	
c	5	

<u>g</u>	r	$p_R(r \mid g)$
d	1	
d	2	
d	3	
d	4	
d	5	
c	1	
c	2	
c	3	
С	4	
C	5	

g	r	$p_R(r \mid g)$
d	1	
d	2	
d	3	
d	4	
d	5	
c	1	
C	2	
c	3	
c	4	
c	5	

<u>g</u>	r	$p_R(r \mid g)$
d	1	
d	2	
d	3	
d	4	2/3
d	5	1/3
c	1	
c	2	
c	3	
c	4	
C	5	

g	r	$p_R(r \mid g)$
d	1	0
d	2	0
d	3	0
d	4	2/3
d	5	1/3
С	1	
c	2	
С	3	
С	4	
c	5	

<u>g</u>	r	$p_R(r \mid g)$
d	1	0
d	2	0
d	3	0
d	4	2/3
d	5	1/3
С	1	
C	2	
C	3	
c	4	
c	5	

g	r	$p_R(r \mid g)$
d	1	0
d	2	0
d	3	0
d	4	2/3
d	5	1/3
C	1	
C	2	
C	3	
C	4	
C	5	

g	r	$p_R(r \mid g)$
d	1	0
d	2	0
d	3	0
d	4	2/3
d	5	1/3
C	1	
C	2	
C	3	
C	4	
C	5	

data =
$$\{(d, 4), (d, 4), (d, 5), (c, 1), (c, 5)\}$$

g	r	$p_R(r \mid g)$
d	1	0
d	2	0
d	3	0
d	4	2/3
d	5	1/3
C	1	1/2
C	2	0
C	3	0
C	4	0
C	5	1/2

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
d	0	1	
d	0	2	
d	0	3	
d	0	4	
d	0	5	
d	1	1	
d	1	2	
d	1	3	
d	1	4	
d	1	5	
С	0	1	
С	0	2	
С	0	3	
С	0	4	
С	0	5	
С	1	1	
С	1	2	
С	1	3	
С	1	4	
С	1	5	

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
d	0	1	
d	0	2	
d	0	3	
d	0	4	
d	0	5	
d	1	1	
d	1	2	
d	1	3	
d	1	4	
d	1	5	
С	0	1	
С	0	2	
С	0	3	
С	0	4	
С	0	5	
С	1	1	
С	1	2	
С	1	3	
С	1	4	
С	1	5	

data = (d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
		1	
		2	
d	0	3	
•	•	4	
· 	<u> </u>	5	
d	1	1	
d	1	2	
d	1	3	
d	1	4	
d	1	5	
С	0	1	
С	0	2	
С	0	3	
С	0	4	
С	0	5	
С	1	1	
С	1	2	
С	1	3	
С	1	4	
С	1	5	

$p_R(r \mid$	$\boldsymbol{\varrho}$	a
PK(I)	5,	uj

g	а	r	$p_R(r \mid g, a)$
; ;	1		
		2	
d	0	3	
	_	4	
		5	
	:	1	
	_	2	
d	1	3	
	_	4	
	<u>;</u>	5	
С	0	1	
С	0	2	
С	0	3	
c	0	4	
c	0	5	
С	1	1	
С	1	2	
c	1	3	
С	1	4	
c	1	5	

data = (d, 0, 3)(d, 1, 5)(c, 0, 1)(c, 0, 5)(c, 1, 4) $p_R(r \mid g, a)$

g	а	r	$p_R(r \mid g, a)$
	·	1	
		2	
d	0	3	
		4	
· 	- -	5	
		1	
		2	
d	1	3	
		4	
·	:	5	
		1	
		2	
c	0	3	
		4	
		5	
c	1	1	
С	1	2	
c	1	3	
С	1	4	
С	1	5	

	<i>g</i>	a	r	$p_R(r \mid g, a)$
data =		:	1	
(d, 0, 3)			2	
	d	0	3	
(d, 1, 5)			4	
(c, 0, 1)		·- !	5	
(c, 0, 5)			1	
	. 1	4	2	
(c, 1, 4)	d	1	3	
			4	
$p_R(r \mid g, a)$		- -	5	
	c		1	
			2	
) /		
			5	
		- 	1	
			2	
	C	1	3	
			4	
	•	i	5	

data = (d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
		1	
		2	
d	0	3	
		4	
		5	
		1	
		2	
d	1	3	
		4	
	:	5	
		1	
•	•	2	
c	0	3	
	•	4	
		5	
		1	
	•	2	
C	1	3	
	-	4	
		5	

data = (d, 0, 3) (d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
		1	
		2	
d	0	3	
		4	
		5	
		1	
		2	
d	1	3	
•		4	
		5	
		1	
•		2	
C	0	3	
•		4	
		5	
		1	
•		2	
c	1	3	
	-	4	
,		5	

data = (d, 0, 3)(d, 1, 5)

(c, 0, 1)(c, 0, 5)

(c, 1, 4)

<u>g</u>	a	r	$p_R(r \mid g, a)$
		1	0
		2	0
d	0	3	1
		4	0
<u>.</u>		5	0
:		1	
		2	
d	1	3	
		4	
· •		5	
,		1	
		2	
c	0	3	
		4	
		5	
:		1	
•		2	
c	1	3	
		4	
		5	

	$oldsymbol{g}$	а	r	$p_R(r \mid g, a)$
data =		;	1	0
(d, 0, 3)			2	0
	d	0	3	1
(d, 1, 5)			4	0
(c, 0, 1)		-	5	0
			1	
(c, 0, 5)			2	
(c, 1, 4)	d	1	3	
			4	
$p_R(r \mid g, a)$		-	5	
$PK(r, 8, \omega)$			1	
			2	
	\mathbf{c}	c 0	3	
			4	
		 	5	
			1	
			2	
	\mathbf{c}	1	3	
			4	
			5	

data = (d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
•	1	0	
		2	0
d	0	3	1
		4	0
	-	5	0
·		1	
		2	
d	1	3	
		4	
		5	
		1	
		2	
C	0	3	
		4	
·		5	
;		1	
		2	
C	1	3	
	•	4	
i		5	

data = (d, 0, 3) (d, 1, 5) (c, 0, 1)

(c, 0, 5)(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
•	1	0	
		2	0
d	0	3	1
		4	0
	-	5	0
·		1	
		2	
d	1	3	
		4	
		5	
		1	
		2	
C	0	3	
		4	
·		5	
;		1	
		2	
C	1	3	
	•	4	
i		5	

 $p_R(r \mid g, a)$

(c, 1, 4)

<i>g</i>	a	r	$p_R(r \mid g, a)$
	•	1	0
		2	0
d	0	3	1
		4	0
	<u> </u>	5	0
		1	0
		2	0
d	1	3	0
		4	0
		5	1
	i	1	
		2	
c	0	3	
		4	
		5	
		1	
		2	
C	1	3	
		4	
	i	5	

(ا ہـ	~ \
$p_R(r \mid$	g,	<i>a</i>)

(c, 1, 4)

$oldsymbol{g}$	а	r	$p_R(r \mid g, a)$
	:	1	0
		2	0
d	0	3	1
		4	0
		5 0	0
	!	1	0
		2	0
d	1	3	0
		4	0
	:	5	1
		1	
		2	
c	0	3	
		4	
		5	
	i	1	
		2	
C	1	3	
		4	
	ī	5	

$\underline{\hspace{1cm}}$	а	r	$p_R(r \mid g, a)$
	:	1	0
		2	0
d	0	3	1
		4	0
 -		5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	<u>. </u>	5	1
		1	
		2	
c	0	3	
		4	
		5	
		1	
		2	
C	1	3	
		4	
	i	5	

$p_R(r \mid$	$\boldsymbol{\rho}$	a
PK(I)	5,	uj

$\underline{\hspace{1cm}}$	а	r	$p_R(r \mid g, a)$
		1	0
		2	0
d	0	3	1
		4	0
		5	0
		1	0
		2	0
d	1	3	0
		4	0
	<u> </u>	5	1
		1	1/2
		2	0
С	0	3	0
		4	0
		5	1/2
		1	
		2	
c	1	3	
		4	
		5	

g	a	r	$p_R(r \mid g, a)$
	:	1	0
		2	0
d	0	3	1
		4	0
		5	0
	!	1	0
		2	0
d	1	3	0
		4	0
	<u> </u>	5	1
	i	1	1 / 2
		2	0
c	0	3	0
		4	0
	<u>.</u>	5	1 / 2
		1	0
		2	0
c	1	3	0
		4	1
		5	0

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
	•	1	0
		2	0
d	0	3	1
		4	0
	_	5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	;	5	1
	:	1	1 / 2
		2	0
c	0	3	0
		4	0
		5	1 / 2
	:	1	0
		2	0
c	1	3	0
		4	1
		5	0

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

 $p_R(r \mid g, a)$

input variables

g	a	r	$p_R(r \mid g, a)$
	•	1	0
		2	0
d	0	3	1
		4	0
=-		5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	į	5	1
	·	1	1/2
		2	0
c	0	3	0
		4	0
		5	1/2
	:	1	0
		2	0
c	1	3	0
		4	1
		5	0

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

 $p_R(r \mid g, a)$

input variables

(filter the data)

g	а	r	$p_R(r \mid g, a)$
	•	1	0
		2	0
d	0	3	1
		4	0
		5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	;	5	1
	:	1	1 / 2
		2	0
c	0	3	0
		4	0
		5	1 / 2
	:	1	0
		2	0
c	1	1 2 3 4 5 1 2 3 4 5	0
		4	1
		5	0

data = (d, 0, 3)(d, 1, 5)(c, 0, 1)(c, 0, 5)(c, 1, 4) $p_R(r \mid g, a)$

		distribution in the second said	
g	а	r	$p_R(r \mid g, a)$
	•	1	0
		2	0
d	0	3	1
		. 4	0
	<u>-</u>	5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	;	5	1
	:	1	1 / 2
		2	0
c	0	3	0
		4	0
		5	1 / 2
		1	0
		4 5 1 2 3 4 5 1 2 3 4 5	0
c	1	3	0
		4	1
		5	0

data = (d, 0, 3)(d, 1, 5)(c, 0, 1)(c, 0, 5)(c, 1, 4) $p_R(r \mid g, a)$

$oldsymbol{g}$	а	r	$p_R(r \mid g, a)$
	:	. 1	0
		2	0
d	0	3	1
		4	0
		5	0
	:	1	0
		2	0
d	1	3	0
		4	0
	output	5	1
·	ariable	1	1/2
		2	0
c	0	3	0
		4	0
	<u>.</u>	5	1/2
	;	1	0
		2	0
c	1	3	0
		4	1
	i	5	0

(d, 0, 3)

(d, 1, 5)

(c, 0, 1)

(c, 0, 5)

(c, 1, 4)

g	а	r	$p_R(r \mid g, a)$
A SAME TO SHE WELL AND THE SAME SAME SAME SAME SAME SAME SAME SAM		1	0
		2	0
d	0	3	1
		4	0
	an digital digital di salah digital di salah digital digital di salah di salah di salah digital di salah di sa Salah digital di salah	5	0
		1	0
		2	0
d	1	3	0
		4	0
		5	1
a in the second section of the second section is a second section of the second section in the second section of the second section is a second section of the second section in the second section is a second section of the second section in the second section is a second section of the second section in the second section is a second section of the second section of the second section is a second section of the second section of the section		n in the second	1 / 2
		2	0
c	0	3	0
		4	0
ر المرافق المر المرافق المرافق	and the state of the	5	1/2
and the second s		1	0
		2	0
c	1	3	0
		4	1
		5	0

$p_R(r \mid$	g	a)
$P^{I(I)}$	O ,	

each block sums to 1

\boldsymbol{g}	а	r	$p_R(r \mid g, a)$
d	0	1	0
		2	0
		3	1
		4	0
		5	0
d	1	1	0
		2	0
		3	0
		4	0
		5	1
C	0	1	1 / 2
		2	0
		3	0
		4	0
		5	1/2
C	1	1	0
		2	0
		3	0
	1	1	
	1	4	1

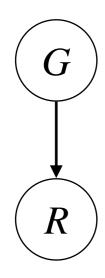
data = (d, 0, 3)(d, 1, 5)(c, 0, 1)(c, 0, 5)(c, 1, 4) $p_R(r \mid g, a)$

\underline{g}	а	r	$p_R(r \mid g, a)$
d	0	1	0
		2	0
		3	1
		4	0
· · · · · · · · · · · · · · · · · · ·		5	0
	1	1	0
		2	0
d		3	0
		4	0
		5	1
	0	1	1 / 2
		2	0
C		3	0
		4	0
		5	1 / 2
	1	1	0
		2	0
c		3	0
		4	1
		5	0

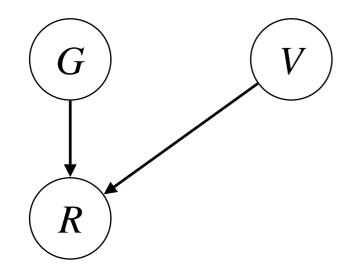
data = (d, 0, 3)	g	a	r	$p_R(r \mid g, a)$
(d, 1, 5) $(c, 0, 1)$	d	0	3	1
(c, 0, 5) (c, 1, 4)	d	1	5	1
$p_R(r \mid g, a)$	c	0	1	1 / 2
	c	0	5	1 / 2
	c	1	4	1

Handling vector-valued variables in a Bayesian network

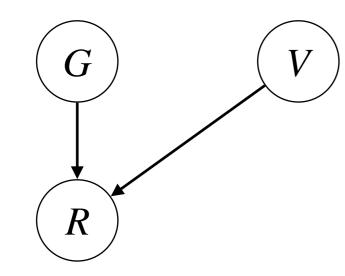
{drama, comedy}

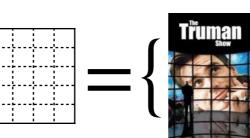


{drama, comedy}



{drama, comedy}



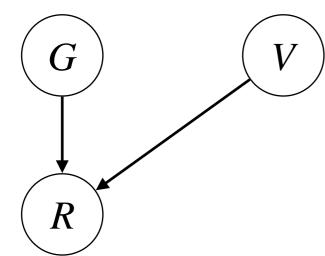


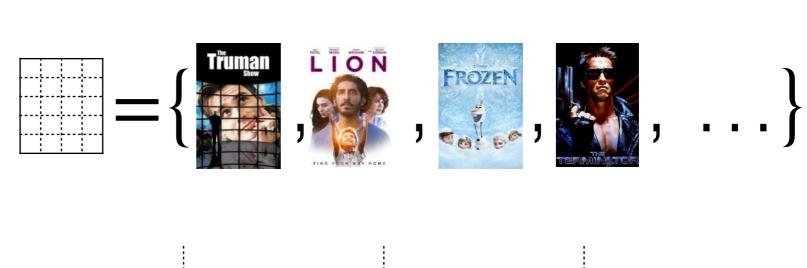






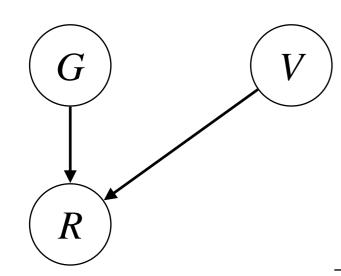
{drama, comedy}





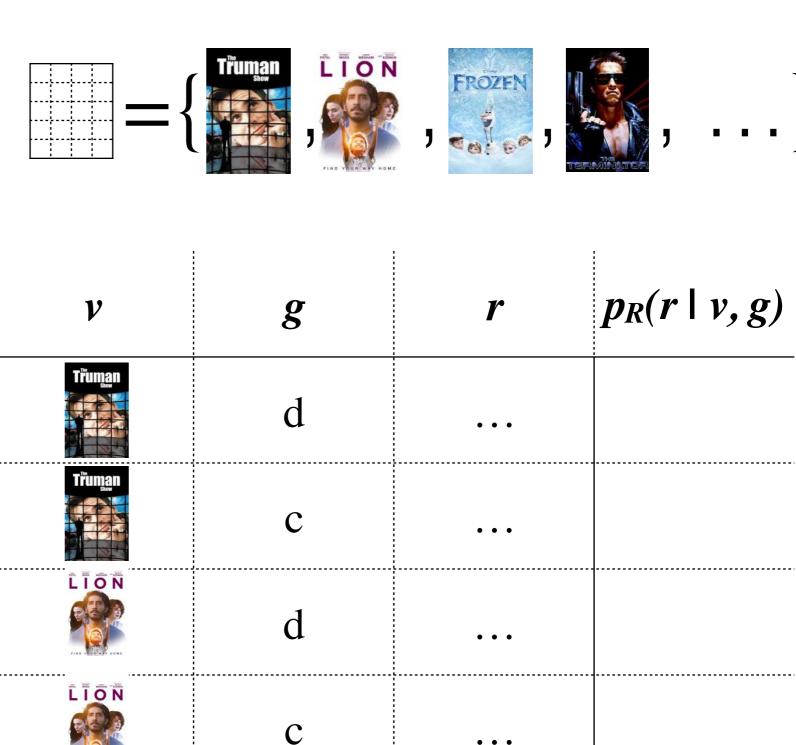
v	g	r	$p_R(r \mid v, g)$
Truman	d	• • •	
Truman Show	c	• • •	
LION	d	• • •	
LION	c	• • •	
• • •	• • •	• • •	• • •

{drama, comedy}

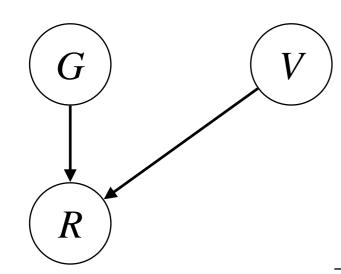


 $\{1, 2, 3, 4, 5\}$

256² pixels



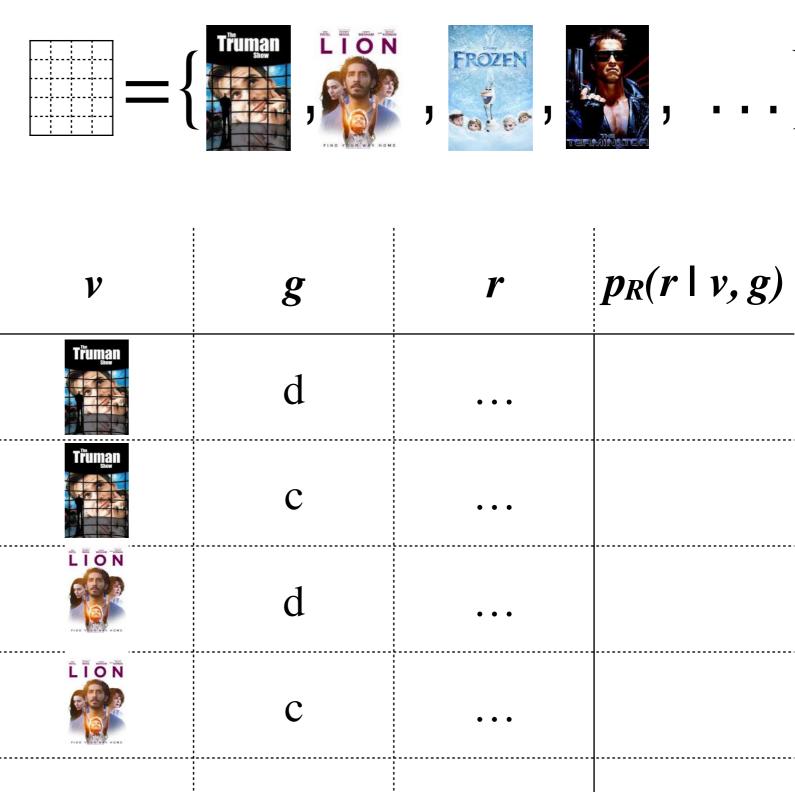
{drama, comedy}



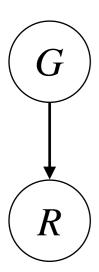
{1, 2, 3, 4, 5}

256² pixels

265536 possible images

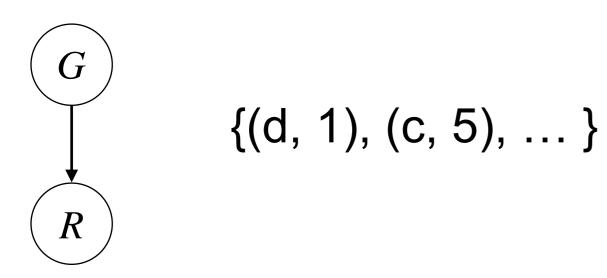


Model



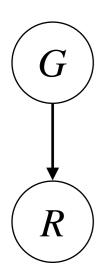
Model

Data



Model

Data



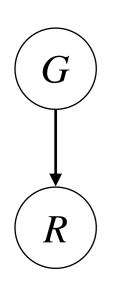
{(d, 1), (c, 5), ...}



Model

Data

Parameters



 $\{(d, 1), (c, 5), \dots\}$

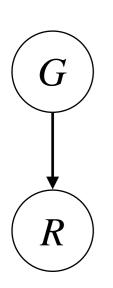


$$heta = egin{array}{c|c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{array}$$

Model

Data

Parameters



 $\{(d, 1), (c, 5), \dots\}$

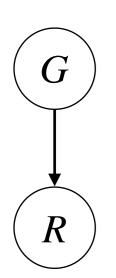


$$heta = egin{array}{c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{pmatrix}$$

Model

Data

Parameters





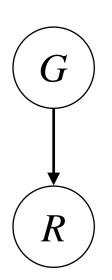
$$heta = egin{array}{c|c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{array}$$

More general principle:

Model

Data

Parameters





$$heta = egin{array}{c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{pmatrix}$$

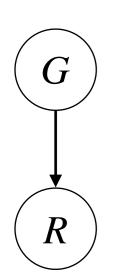
More general principle:

$$\prod_{(g,r)\in\mathcal{D}_{\text{train}}} p_G\left(g\right) p_R\left(r\mid g\right)$$

Model

Data

Parameters





$$heta = egin{array}{c|c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{array}$$

More general principle:

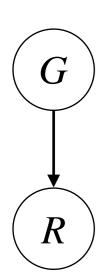
$$\prod_{(g,r)\in\mathcal{D}_{\text{train}}} p_G\left(g\right) p_R\left(r\mid g\right)$$

likelihood of the data

Model

Data

Parameters



{(d, 1), (c, 5), ...}



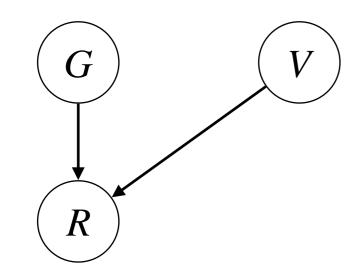
$$heta = egin{array}{c|c} g & p_G(g) \\ d & 3/5 \\ c & 2/5 \\ \hline g & r & p_R(r \mid g) \\ d & 4 & 2/3 \\ d & 5 & 1/3 \\ c & 1 & 1/2 \\ c & 5 & 1/2 \\ \hline \end{array}$$

More general principle:

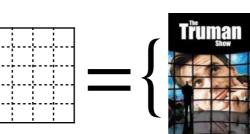
$$\theta = \underset{p_G, \ p_R}{\operatorname{argmax}} \prod_{(g,r) \in \mathcal{D}_{\text{train}}} p_G(g) p_R(r \mid g)$$

likelihood of the data

{drama, comedy}



{1, 2, 3, 4, 5}



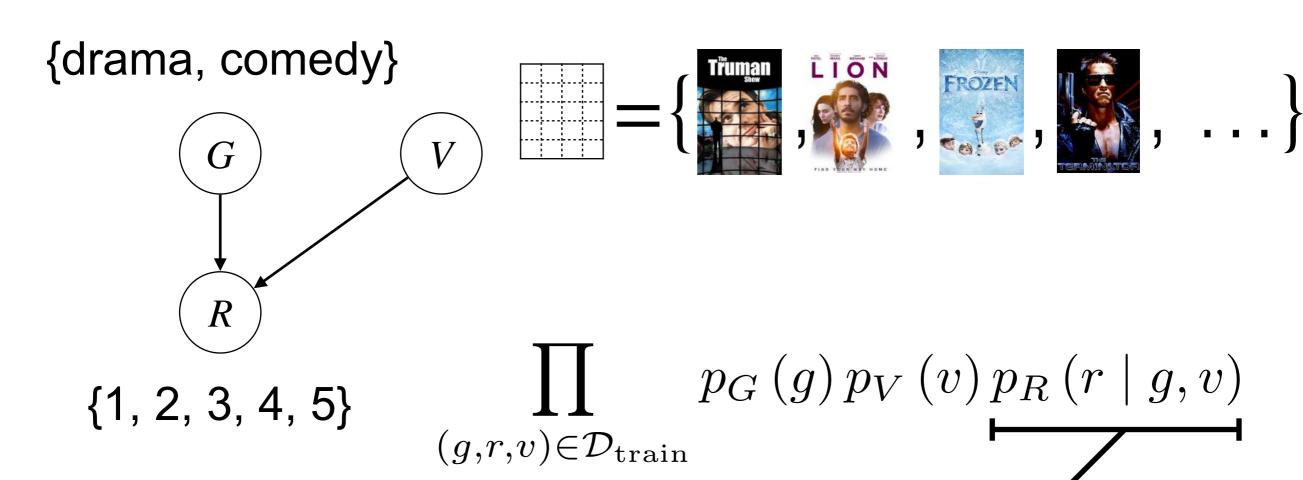






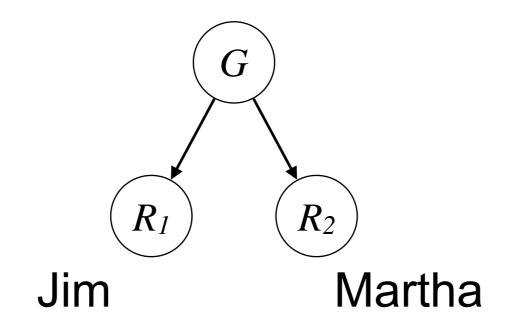
, . . .

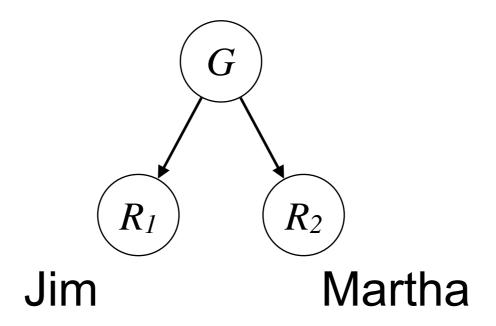
 $(g,r,v)\in\mathcal{D}_{\mathrm{train}}$



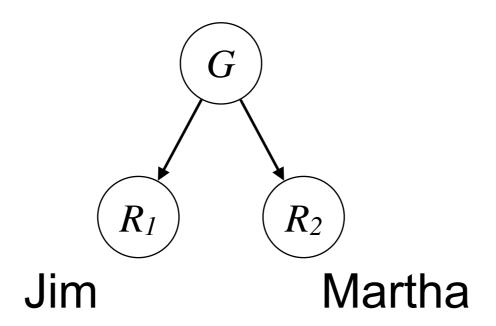
any probabilistic model (e.g. neural network)

Parameter sharing in a Bayesian network



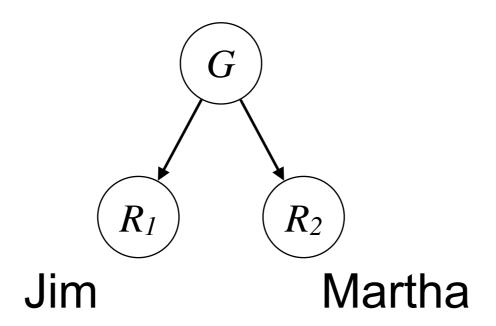


We think Jim and Martha have the exact same preferences



We think Jim and Martha have the exact same preferences





We think Jim and Martha have the exact same preferences



Treat Jim's data and Martha's data as if they came from one person

Jim's data

g	r	count
	1	
	2	
d	3	
	4	2
	5	1
c	1	1
	2	
	3	
	4	
	5	1

Jim's data

Martha's data

g	r	count		g	r	count
	1				1	
	2				2	
d	3			d	3	1
	4	2			4	1
	5	1			5	1
	1	1	+		1	
	2				2	1
c	3			c	3	
	4				4	1
	5	1			5	

Jim's data

Martha's data

Merged data

g	r	count		g	r	count	g	r	count
	1				1			1	
	2				2			2	
d	3			d	3	1	d	3	1
	4	2			4	1		4	3
	5	1			5	1		5	2
	1	1	+ "		1			1	1
	2		2 1		2	1		2	1
c	3			c	3				
	4				4	. 1		4	1
	5	1			5			5	1

Merged data

g	r	count
	1	
	2	
d	3	1
	4	3
	5	2
	1	1
	2	1
c	3	
	4	1
	5	1

Merged data

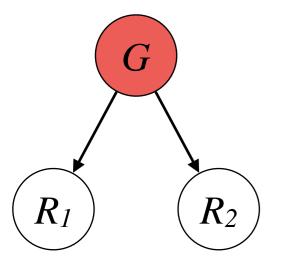
g	r	count
	1	
	2	
d	3	1
	4	3
	5	2
	1	1
	2	1
c	3	
	4	1
	5	1

Normalized

g	r	$p(r \mid g)$
	1	
	2	
d	3	1/6
	4	3/6
	5	2/6
c	1	1/4
	2	1/4
	3	
	4	1/4
	5	1/4

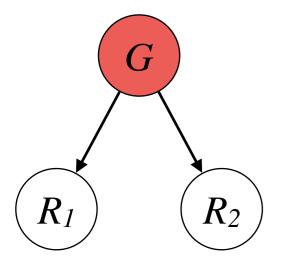
Expectation Maximization

Expectation maximization



data = $\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$

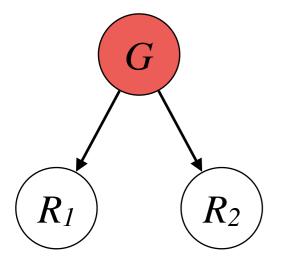
Expectation maximization



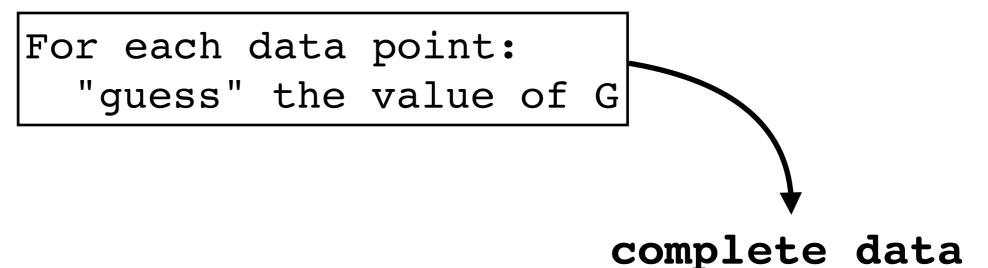
```
data = \{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}
```

```
For each data point:
"guess" the value of G
```

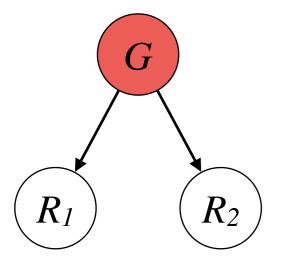
Expectation maximization



```
data = \{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}
```



Expectation maximization



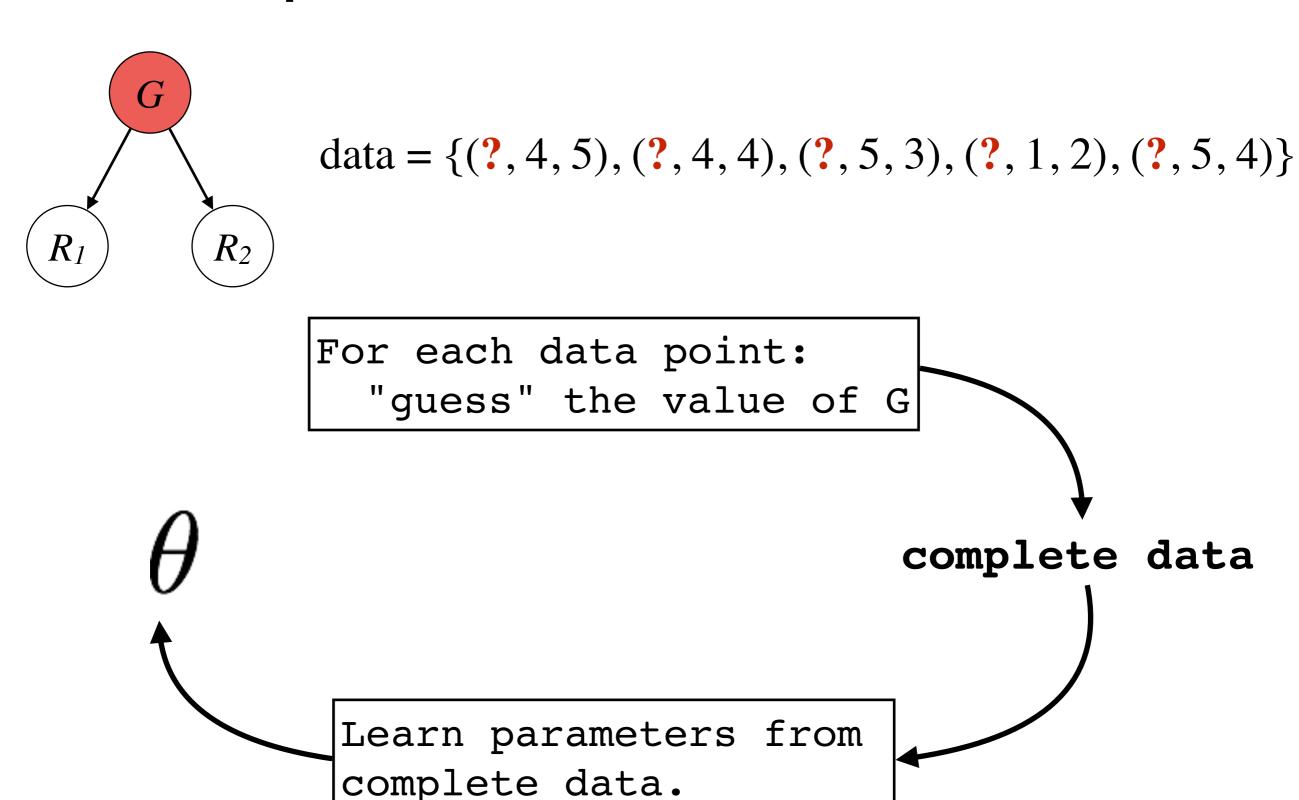
```
data = \{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}
```

For each data point:
"guess" the value of G

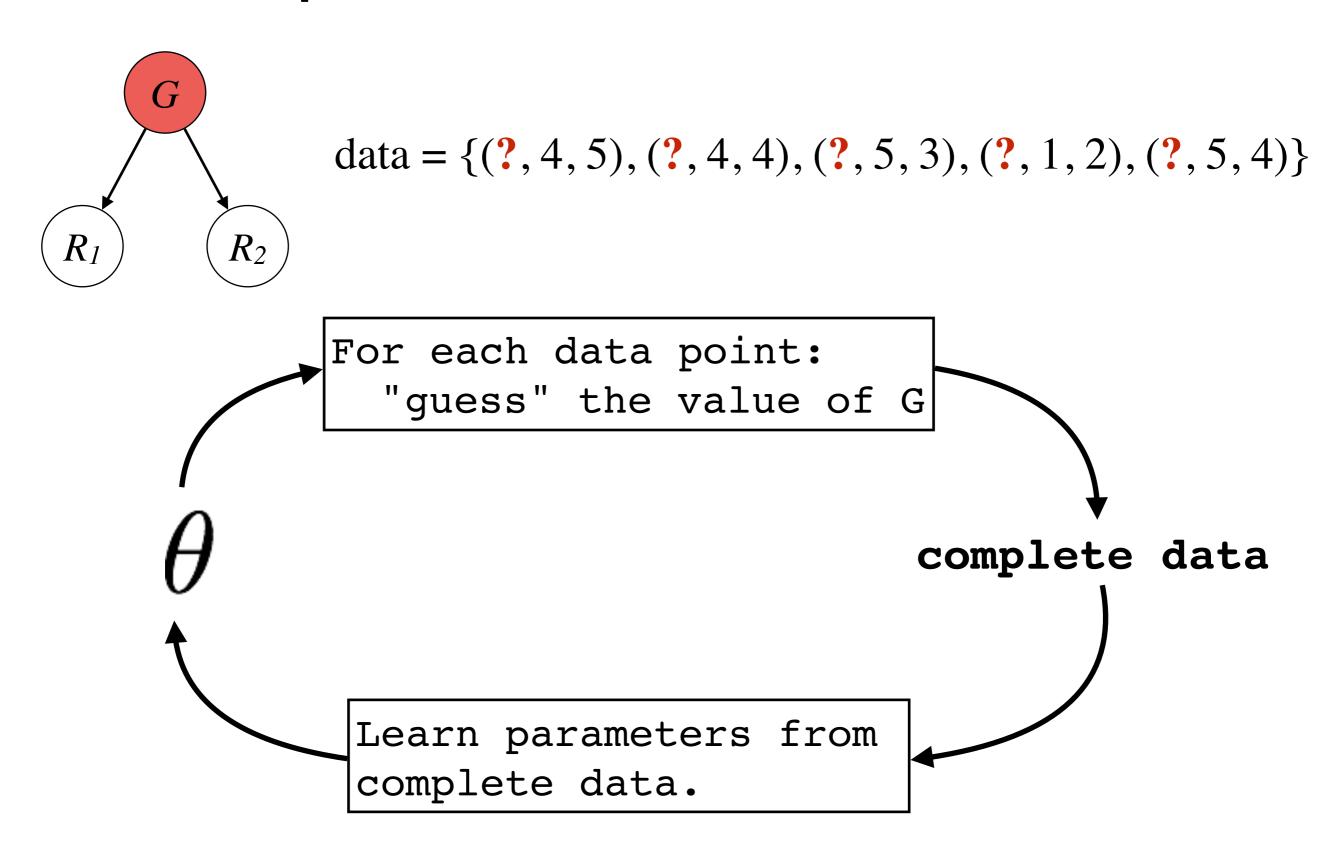
complete data

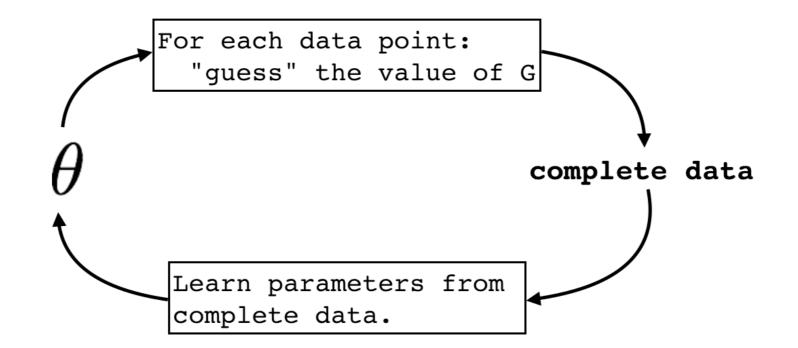
Learn parameters from complete data.

Expectation maximization

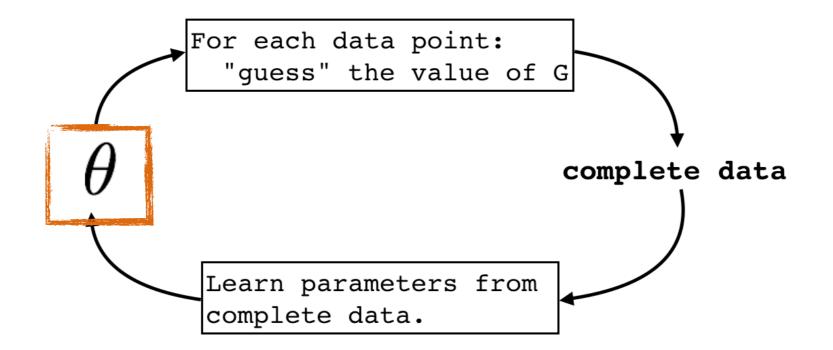


Expectation maximization

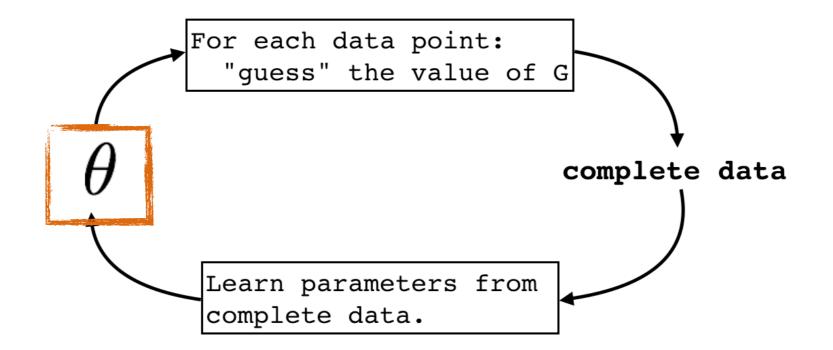




data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

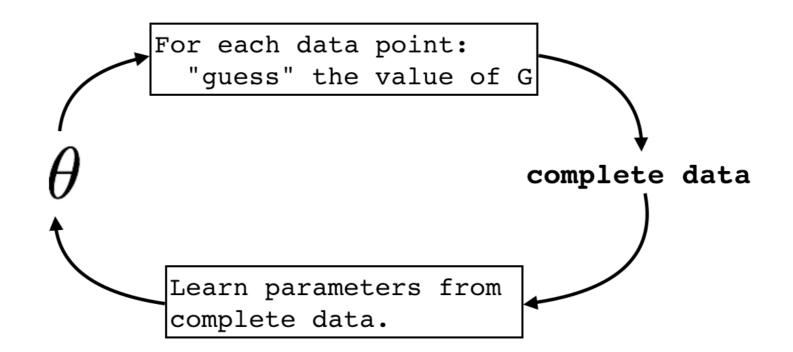


data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$



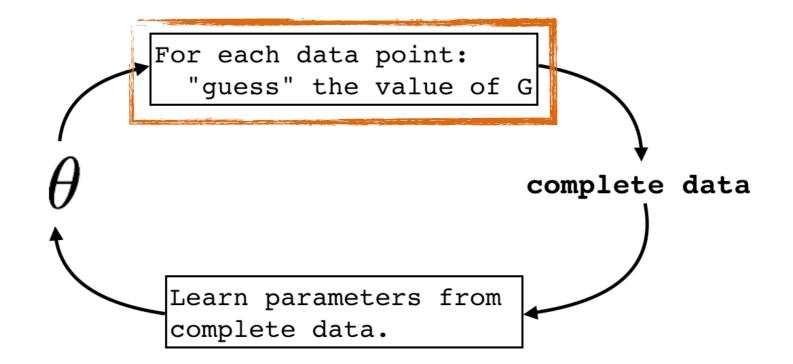
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



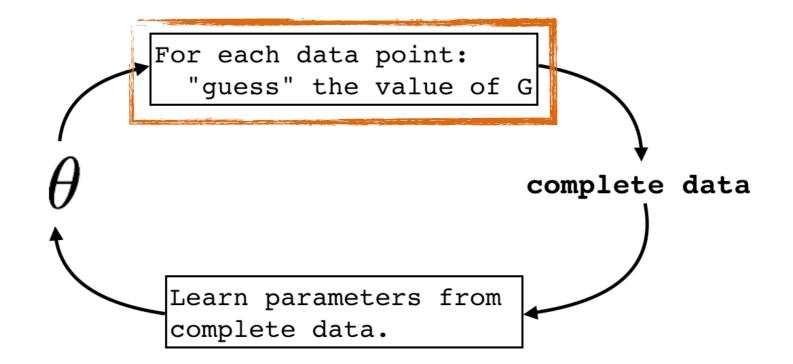
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{ccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \end{array}$



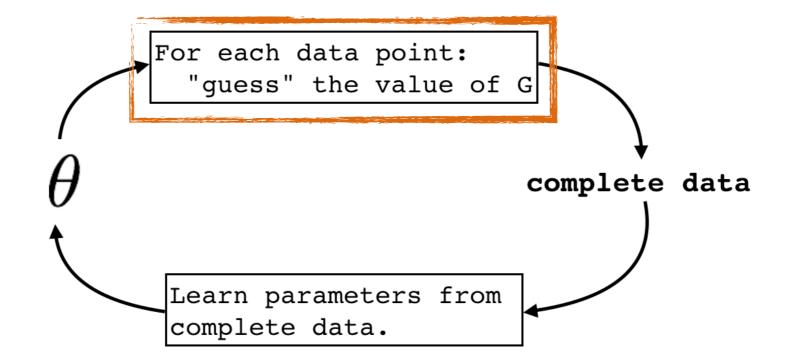
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

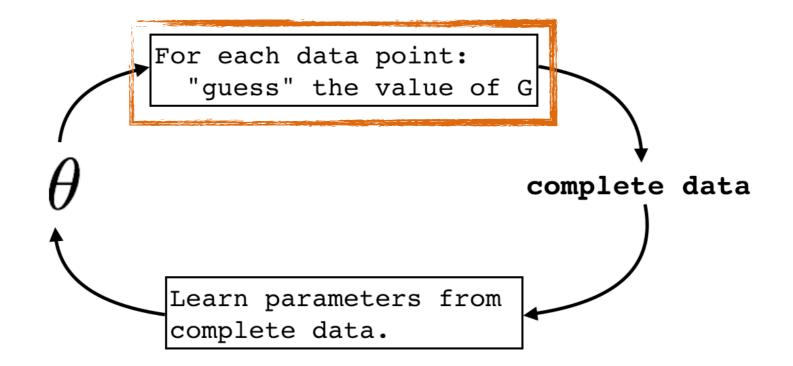
$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(\mathbf{g} \mid r_1 = \mathbf{4}, r_2 = \mathbf{5})$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$

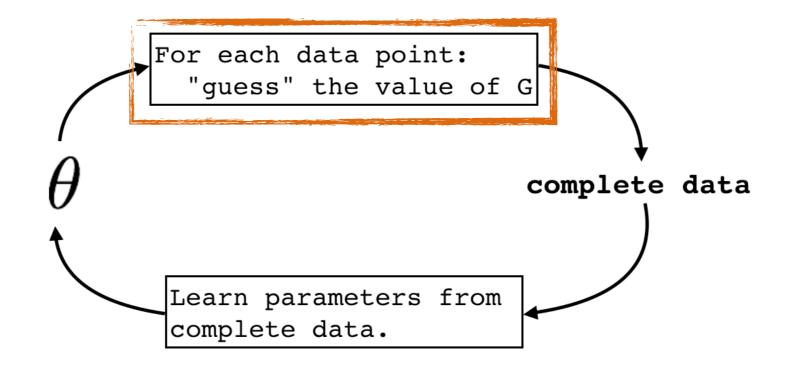


data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(g \mid r_1=4, r_2=5)$$

0.25(c, 4, 5)

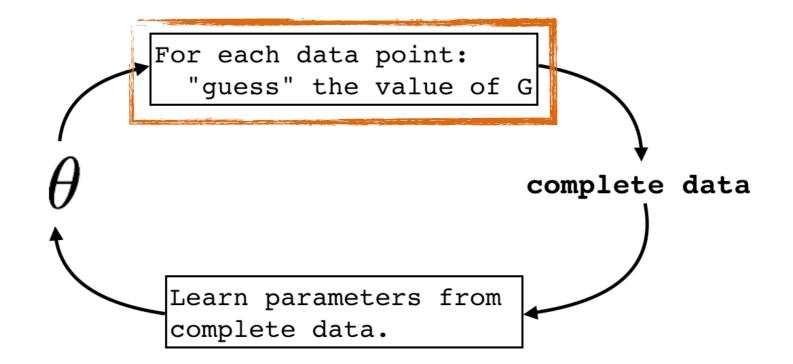
$$heta$$
: $egin{array}{ccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \end{array}$



data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

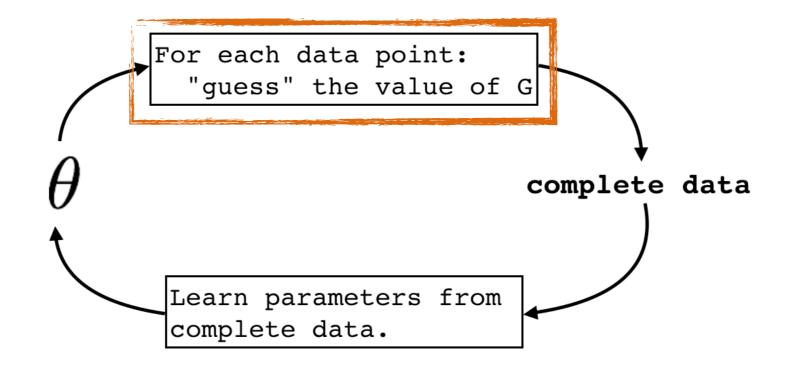
$$p(g \mid r_1=4, r_2=5)$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

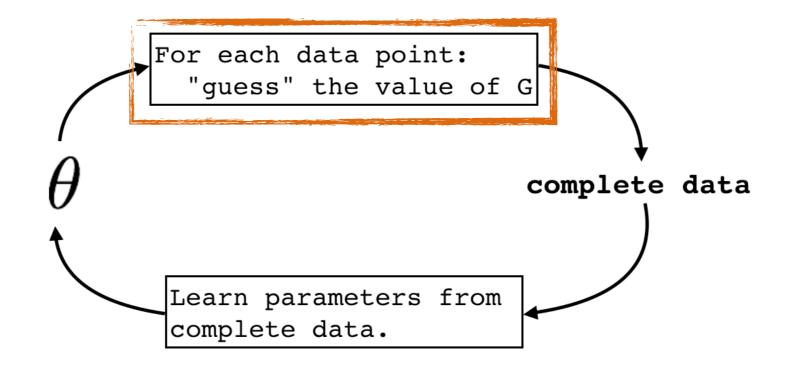
$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(\mathbf{g} \mid r_1 = \mathbf{4}, r_2 = \mathbf{4})$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$

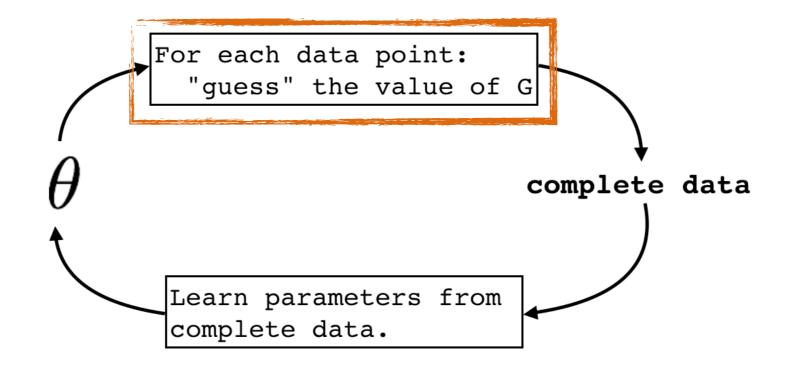


data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(g \mid r_1=4, r_2=4)$$

$$0.25$$
 (c, 4, 5) 0.1 (c, 4, 4)

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$

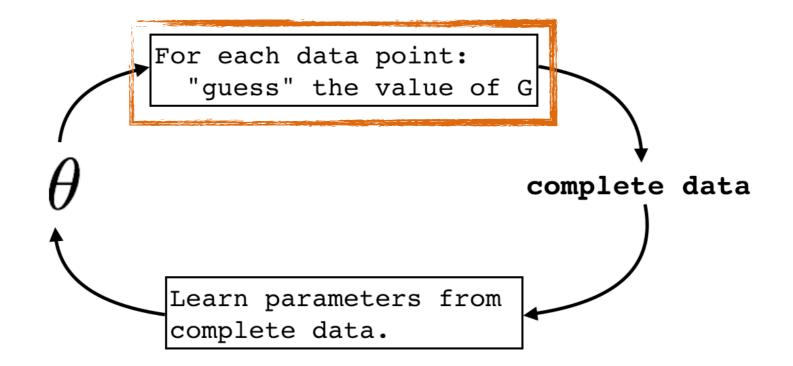


data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(g \mid r_1=4, r_2=4)$$

$$0.25$$
 (c, 4, 5) 0.1 (c, 4, 4)

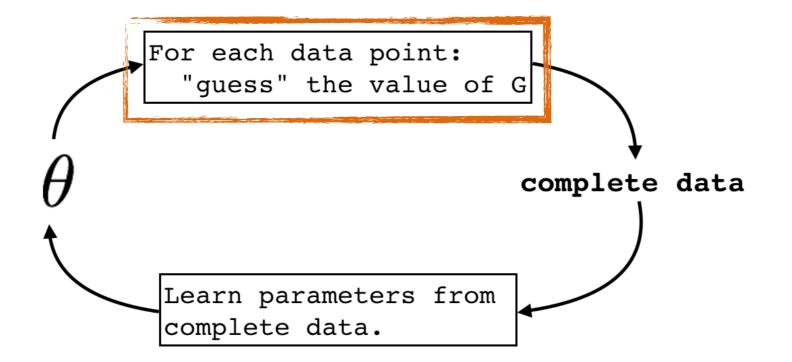
$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$0.25$$
 (c, 4, 5) 0.1 (c, 4, 4)

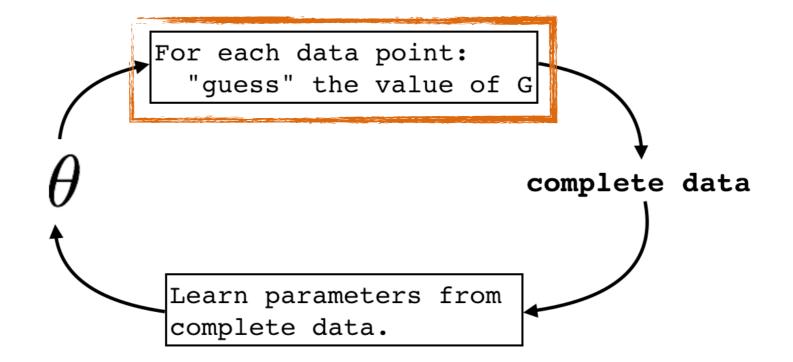
$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}
$$p(\mathbf{g} \mid r_1 = 5, r_2 = 3)$$

$$0.25 (c, 4, 5) \quad 0.1 (c, 4, 4)$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



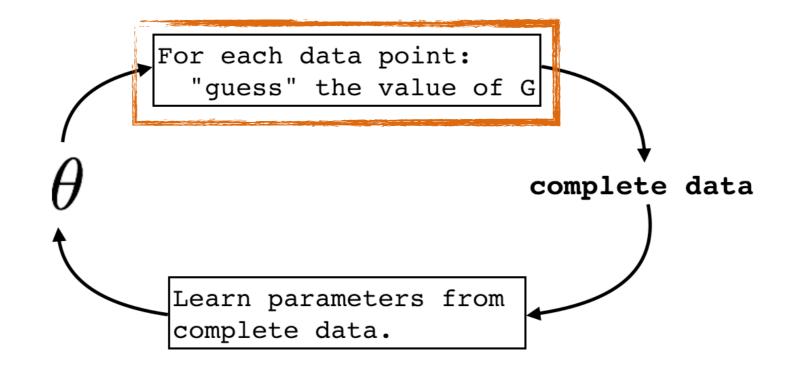
data = {(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)}

$$p(g \mid r_1=5, r_2=3)$$

0.25
$$(\mathbf{c}, 4, 5)$$
 0.1 $(\mathbf{c}, 4, 4)$ **0.8** $(\mathbf{c}, 5, 3)$

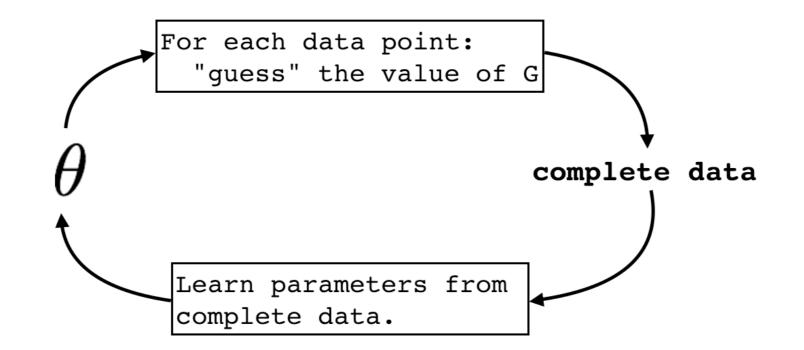
0.75
$$(d, 4, 5)$$
 0.9 $(d, 4, 4)$ **0.2** $(d, 5, 3)$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



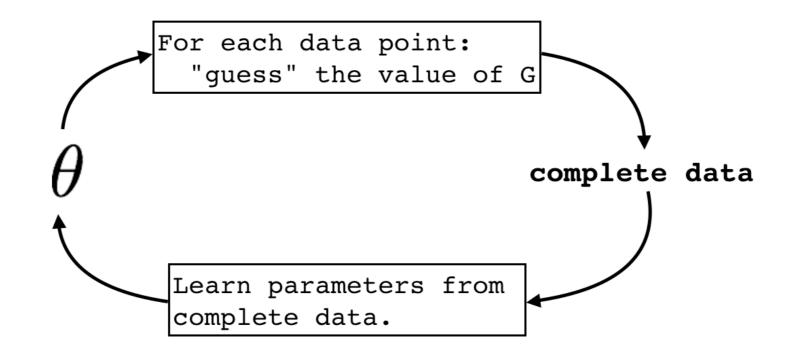
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



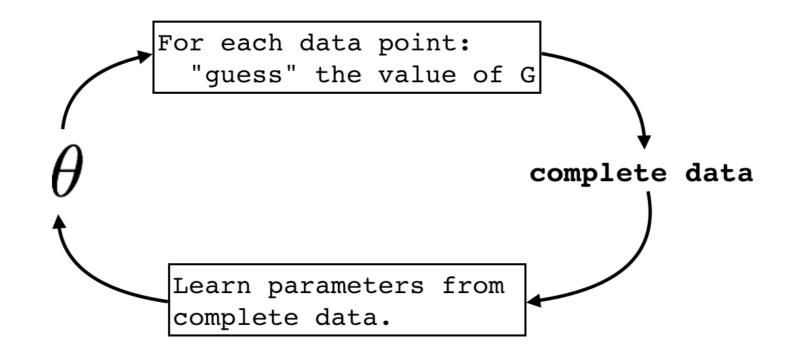
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ {
m d} & 3/5 \ {
m c} & 2/5 \ \end{array}$



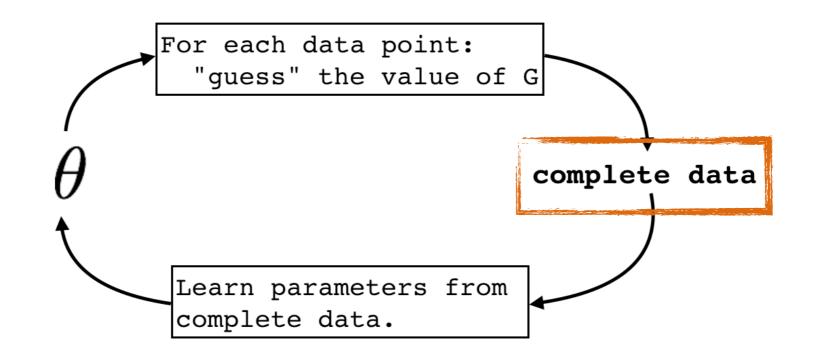
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



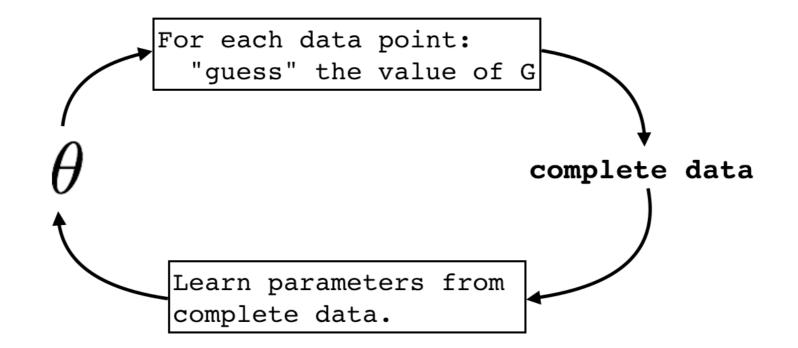
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



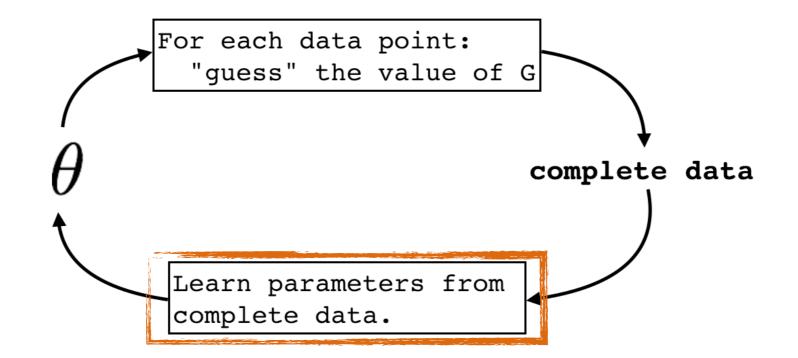
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

$$heta$$
: $egin{array}{cccc} g & p_G(g) \ d & 3/5 \ c & 2/5 \ \end{array}$



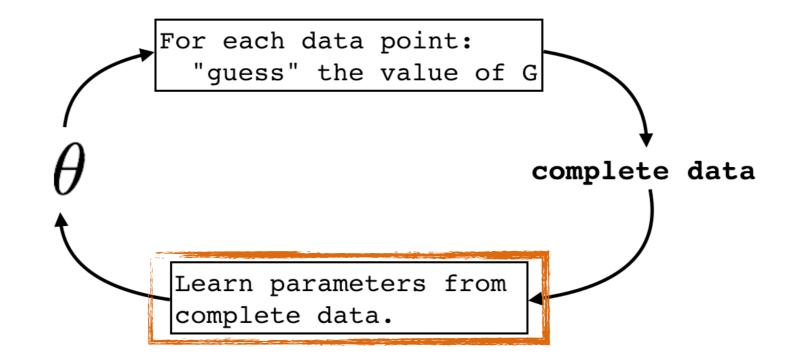
data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

```
0.25 (c, 4, 5) 0.1 (c, 4, 4) 0.8 (c, 5, 3) 0.9 (c, 1, 2) 0.7 (c, 5, 4) 0.75 (d, 4, 5) 0.9 (d, 4, 4) 0.2 (d, 5, 3) 0.1 (d, 1, 2) 0.3 (d, 5, 4)
```



data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$

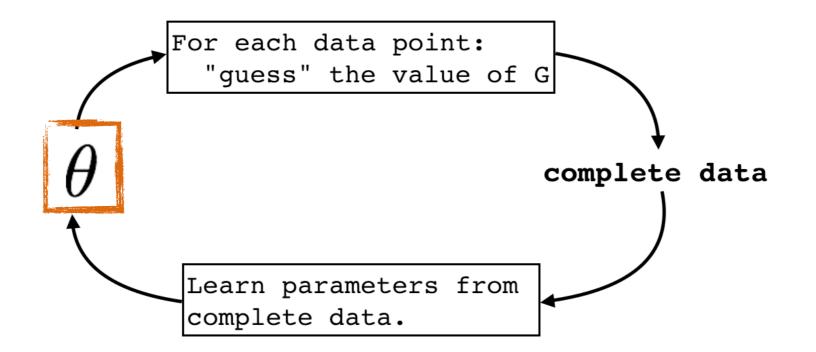
```
0.25 (c, 4, 5) 0.1 (c, 4, 4) 0.8 (c, 5, 3) 0.9 (c, 1, 2) 0.7 (c, 5, 4) 0.75 (d, 4, 5) 0.9 (d, 4, 4) 0.2 (d, 5, 3) 0.1 (d, 1, 2) 0.3 (d, 5, 4)
```



data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$







data =
$$\{(?, 4, 5), (?, 4, 4), (?, 5, 3), (?, 1, 2), (?, 5, 4)\}$$



