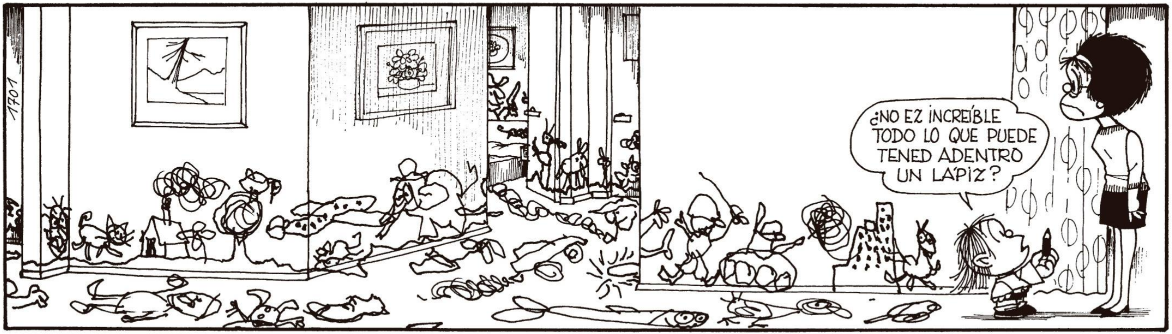
Especificación y evaluación de modelos

Parte 1.

Aprendizaje basado en modelos. Distribuciones de creencias honestas. Las reglas de razonamiento en contextos de incertidumbre. Métodos gráficos de especificación matemática de argumentos causales. Evaluación de modelos causales alternativos.



Verdad

La ciencia tiene pretensión de verdad, de alcanzar
acuerdos intersubjetivos que valgan para todas las personas

Verdad

La ciencia tiene pretensión de verdad, de alcanzar
acuerdos intersubjetivos que valgan para todas las personas

Ciencias formales
Sistemas axiomáticos **cerrados** sin incertidumbre

Verdad

La ciencia tiene pretensión de verdad, de alcanzar
acuerdos intersubjetivos que valgan para todas las personas

Ciencias formales
Sistemas axiomáticos **cerrados** sin incertidumbre

Ciencias con datos
Sistemas naturales **abiertos** con incertidumbre

Verdad

La ciencia tiene pretensión de verdad, de alcanzar
acuerdos intersubjetivos que valgan para todas las personas

Ciencias formales
Sistemas axiomáticos **cerrados** sin incertidumbre

Ciencias con datos
Sistemas naturales **abiertos** con incertidumbre

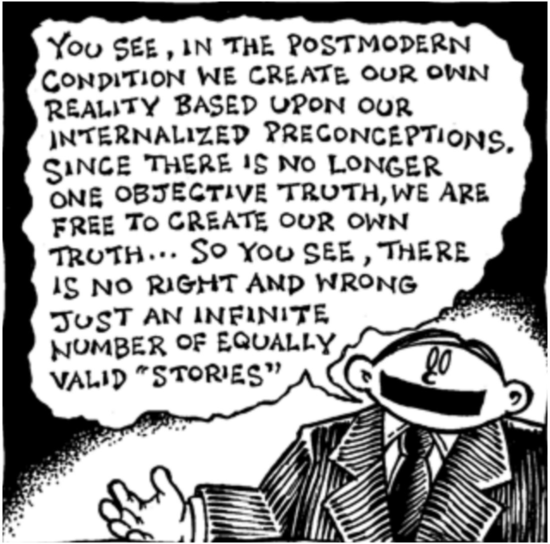
¿Qué es una verdad en
contextos de incertidumbre?

¿Todo vale lo mismo?

Ouch!.. Me estás pisando el cuello!

Bueno, es un punto de vista. Se podría decir que me querés hacer tropezar.

En la posmodernidad ya no hay objetividad, toda historia es igualmente válida



Pero me sigues pisando el cuello! Nunca fuiste a la universidad?

¿Puede haber una verdad si justamente tenemos incertidumbre sobre su valor real?

Al menos sabemos cómo no mentir

Al menos sabemos cómo no mentir

- No afirmar más de lo que se conoce

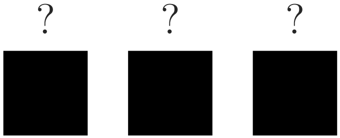
Al menos sabemos cómo no mentir

- No afirmar más de lo que se conoce
- Sin ocultar aquello que sí se conoce

Al menos sabemos cómo no mentir

- No afirmar más de lo que se conoce
- Sin ocultar aquello que sí se conoce

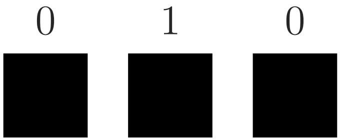
¿Cómo exactamente?



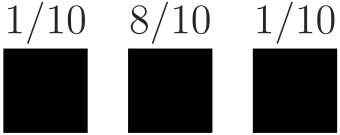
Detrás de una de estas caja hay un regalo.

¿Dónde está el regalo?

Certeza absoluta

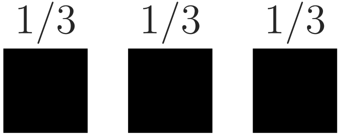


Distribución de creencias



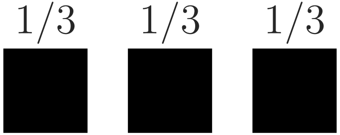
Distribución de creencias

Honesta



Distribución de creencias

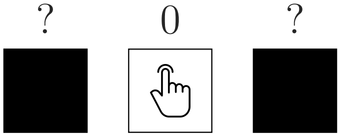
Honesta



Acuerdo intersubjetivo (no mentir)

1. **Maximizar incertidumbre** (dividiendo en partes iguales)
2. **Dada la información disponible** (no asignamos creencia por fuera de las cajas)

¿Cómo preservamos los acuerdos intersubjetivos?

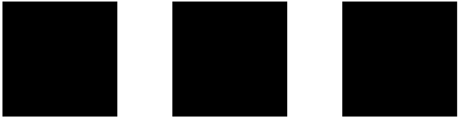


Modelos causales

Modelo gráfico

Regalo: r 

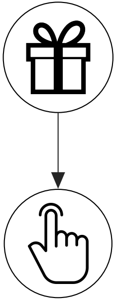
$1/3$ $1/3$ $1/3$



Modelos causales

Modelo gráfico

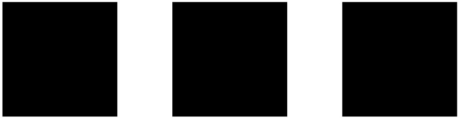
Regalo: r



Pista: s

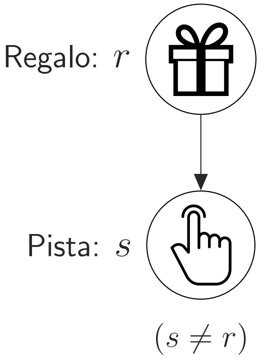
$(s \neq r)$

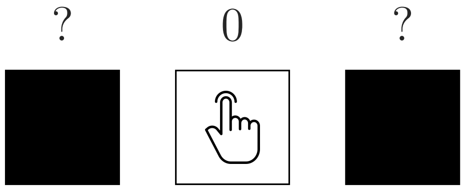
$1/3$ $\quad$ $1/3$ $\quad$ $1/3$



Modelos causales

Modelo gráfico

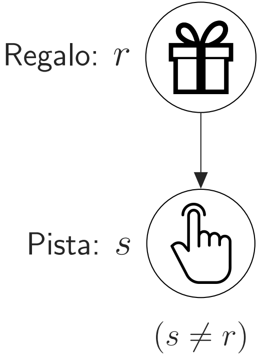


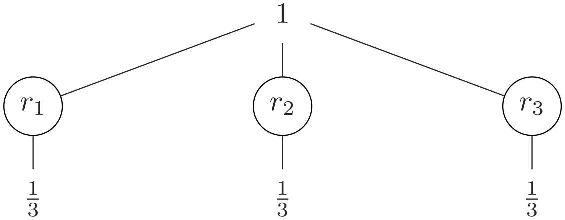


Modelos causales

Máxima incertidumbre dado el modelo

Modelo gráfico





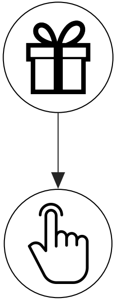
Modelos causales

Máxima incertidumbre dado el modelo

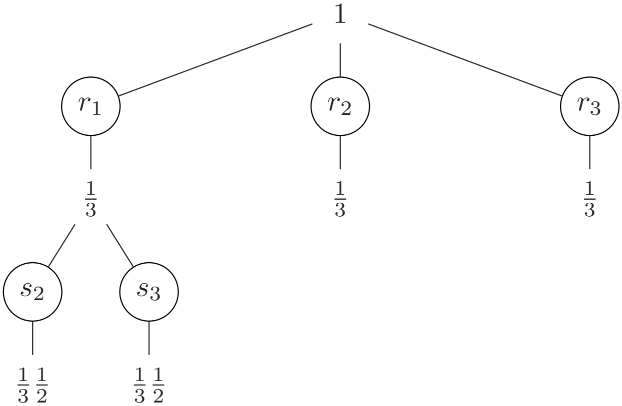
Modelo gráfico

Regalo: r

Pista: s

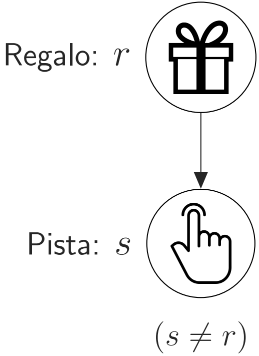


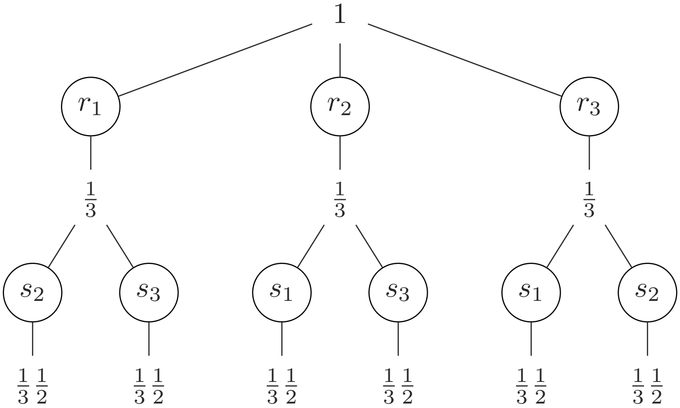
$(s \neq r)$



Modelos causales
Máxima incertidumbre dado el modelo

Modelo gráfico

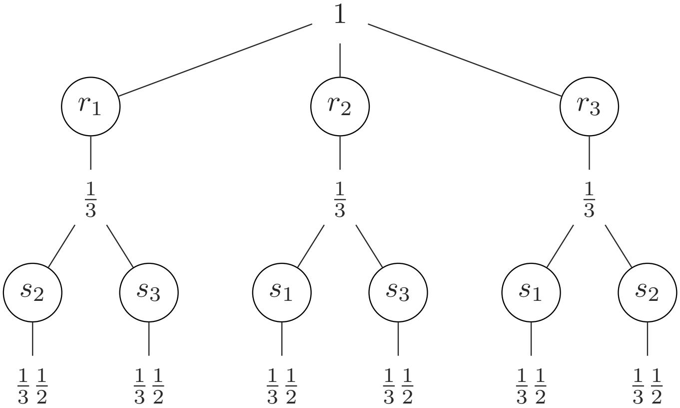




Modelos causales
Máxima incertidumbre dado el modelo

Creencia$(r, s|\text{Modelo})$

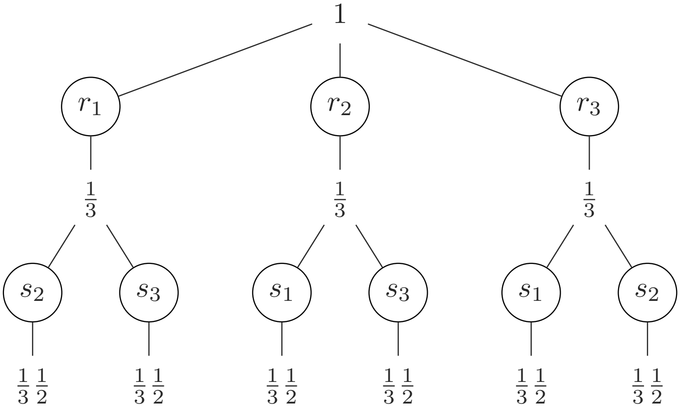
	r_1	r_2	r_3	
s_1				
s_2				
s_3				



Modelos causales
Máxima incertidumbre dado el modelo

Creencia$(r, s|$Modelo$)$

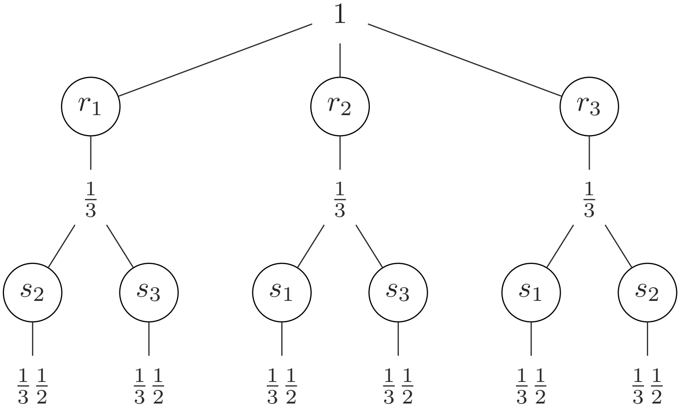
	r_1	r_2	r_3	
s_1	0			
s_2				
s_3				



Modelos causales
Máxima incertidumbre dado el modelo

Creencia$(r, s|$Modelo$)$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2				
s_3				

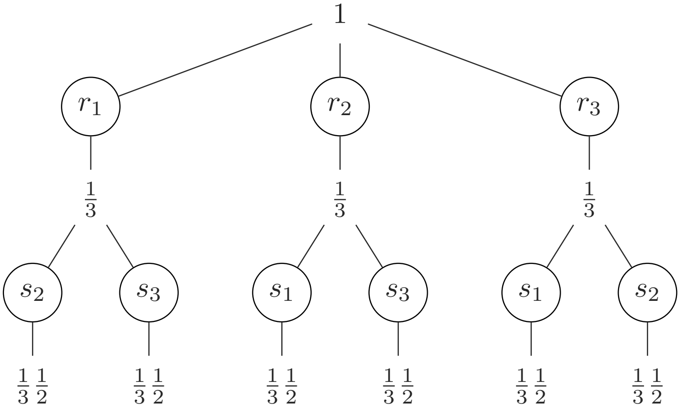


Modelos causales

Máxima incertidumbre dado el modelo

Creencia$(r, s|\text{Modelo})$

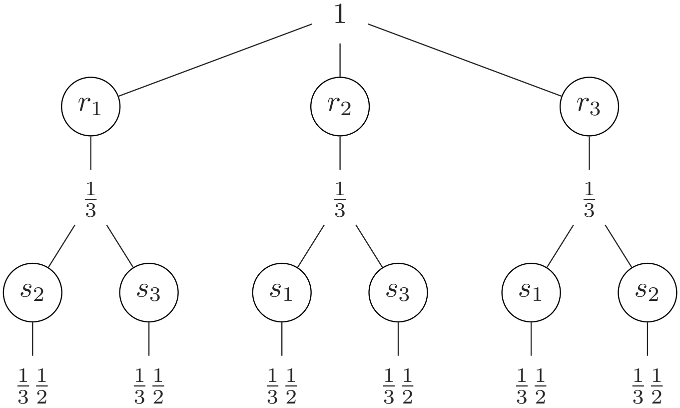
	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3				



Modelos causales
Máxima incertidumbre dado el modelo

Creencia$(r, s|$Modelo$)$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3	$1/6$	$1/6$	0	

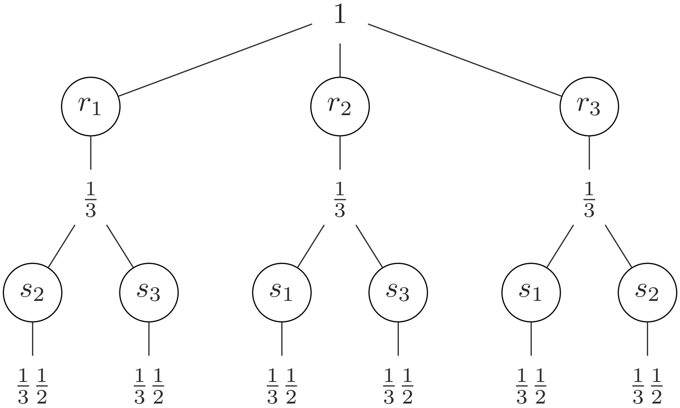


Modelos causales

Máxima incertidumbre dado el modelo

Creencia$(r, s|$Modelo$)$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3	$1/6$	$1/6$	0	



Creencia conjunta
intersubjetiva inicial

Máxima incertidumbre dado el modelo

$$\overbrace{\text{Creencia}(r,s|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3	$1/6$	$1/6$	0	

Máxima incertidumbre dado el modelo

$$\overbrace{\text{Creencia}(r, s|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3	$1/6$	$1/6$	0	

$$\text{Creencia}(r|\text{M}) =$$

Máxima incertidumbre dado el modelo

$$\overbrace{\text{Creencia}(r, s|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	
s_2	$1/6$	0	$1/6$	
s_3	$1/6$	$1/6$	0	
	$1/3$	$1/3$	$1/3$	

$$\text{Creencia}(r|\text{M}) = \sum_s \text{Creencia}(r, s|\text{M})$$

Máxima incertidumbre dado el modelo

$$\overbrace{\text{Creencia}(r,s|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	$1/3$
s_2	$1/6$	0	$1/6$	$1/3$
s_3	$1/6$	$1/6$	0	$1/3$
	$1/3$	$1/3$	$1/3$	

$$\text{Creencia}(r|\text{M}) = \sum_s \text{Creencia}(r,s|\text{M})$$

$$\text{Creencia}(s|\text{M}) = \sum_r \text{Creencia}(r,s|\text{M})$$

Máxima incertidumbre dado el modelo

$$\overbrace{\text{Creencia}(r, s|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	$1/3$
$\boldsymbol{s_2}$	$1/6$	0	$1/6$	$1/3$
s_3	$1/6$	$1/6$	0	$1/3$
	$1/3$	$1/3$	$1/3$	

$$\text{Creencia}(r|\text{M}) = \sum_s \text{Creencia}(r, s|\text{M})$$

$$\text{Creencia}(s|\text{M}) = \sum_r \text{Creencia}(r, s|\text{M})$$

Máxima incertidumbre dado el modelo y el dato

$$\overbrace{\text{Creencia}(r, s_2|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
$\boldsymbol{s_2}$	$1/6$	0	$1/6$	$1/3$

$$\overbrace{\text{Creencia}(r, s_2|\text{M})}^{\text{Creencia compatible}}$$

Máxima incertidumbre dado el modelo y el dato

$$\overbrace{\text{Creencia}(r, s_2|\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
$\boldsymbol{s_2}$	$1/6$	0	$1/6$	$\mathbf{1/3}$

$$\frac{\overbrace{\text{Creencia}(r, s_2|\text{M})}^{\text{Creencia compatible}}}{\underbrace{\text{Creencia}(s_2|\text{M})}_{\text{Creencia total que compatible}}}$$

Máxima incertidumbre dado el modelo y el dato

$$\overbrace{\text{Creencia}(r|s_2,\text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
$\boldsymbol{s_2}$	$\frac{1}{6}/\frac{1}{3}$	0	$\frac{1}{6}/\frac{1}{3}$	$\frac{1}{3}/\frac{1}{3}$

$$\underbrace{\text{Creencia}(r|s_2,\text{M})}_{\text{Nueva creencia}} = \frac{\overbrace{\text{Creencia}(r,s_2|\text{M})}^{\text{Creencia compatible}}}{\underbrace{\text{Creencia}(s_2|\text{M})}_{\text{Creencia total que compatible}}}$$

Máxima incertidumbre dado el modelo y el dato

$$\overbrace{\text{Creencia}(r|s_2,\text{M})}^{\text{De ambas variables}}$$

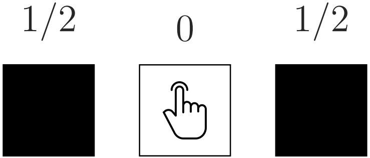
	r_1	r_2	r_3	
$\boldsymbol{s_2}$	$1/2$	0	$1/2$	1

$$\underbrace{\text{Creencia}(r|s_2,\text{M})}_{\text{Nueva creencia}} = \frac{\overbrace{\text{Creencia}(r,s_2|\text{M})}^{\text{Creencia compatible}}}{\underbrace{\text{Creencia}(s_2|\text{M})}_{\text{Creencia total que compatible}}}$$

Máxima incertidumbre dado el modelo y el dato

$$\overbrace{\text{Creencia}(r|s_2, \text{M})}^{\text{De ambas variables}}$$

	r_1	r_2	r_3	
$\boldsymbol{s_2}$	$1/2$	0	$1/2$	1



Las reglas de la probabilidad

Principio de coherencia
(regla del producto)

$$P(\text{Hipótesis}|\text{Dato}) = \frac{P(\text{Hipótesis}, \text{Dato})}{P(\text{Dato})}$$

Preservamos la creencia previa que sigue siendo compatible con el dato

Principio de integridad
(regla de la suma)

$$P(\text{Dato}) = \sum_{\text{Hipótesis}} P(\text{Hipótesis}, \text{Dato})$$

Predecimos con la contribución de todas las hipótesis. No perdemos ni creamos creencia.

Regla del producto

$$P(\ H_1 \mid H_2\ ) = \frac{P(H_1, H_2)}{P(H_2)}$$

Regla del producto

$$P(\ H_1 \mid H_2\ ) = \frac{P(H_1, H_2)}{P(H_2)}$$

$$P(\ H_2 \mid H_1\ ) =$$

Regla del producto

$$P(\ H_1 \mid H_2\ ) = \frac{P(H_1, H_2)}{P(H_2)}$$

$$P(\ H_2 \mid H_1\ ) = \frac{P(H_1, H_2)}{P(H_1)}$$

Regla del producto

$$P(\ H_1 \mid H_2\ ) = \frac{P(H_1, H_2)}{P(H_2)}$$

$$P(H_1)P(\ H_2 \mid H_1\ ) = P(H_1, H_2)$$

Regla del producto

$$P(\ H_1 \mid H_2\ ) = \frac{P(H_1, H_2)}{P(H_2)}$$

$$P(H_1)P(\ H_2 \mid H_1\ ) = \boldsymbol{P(H_1, H_2)}$$

Regla del producto

Teorema de Bayes

$$\overbrace{P(\text{Hipótesis}_i,\, \text{Datos})}^{\text{Creencia que sobrevive}} = \overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i)}^{\text{Predicción (verosimilitud)}} \; \overbrace{P(\text{Hipótesis}_i)}^{\text{Creencia previa}}$$

Regla del producto
Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i | \text{ Datos})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Hipótesis}_i, \text{ Datos})}^{\text{Creencia que sobrevive}}}{\underbrace{P(\text{Datos})}_{\text{Evidencia}}}$$

Regla del producto

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i \,|\, \text{Datos})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Dato} \,|\, \text{Hipótesis}_i)}^{\text{Verosimilitud}} \, \overbrace{P(\text{Hipótesis}_i)}^{\text{Prior}}}{\underbrace{P(\text{Dato})}_{\text{Evidencia}}}$$

Regla del producto

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i |\, \text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i,\, \text{Modelo})}^{\text{Verosimilitud}}\, \overbrace{P(\text{Hipótesis}_i |\, \text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

Y el **modelo**!

que relaciona el **dato** con la **hipótesis**!

Regla del producto

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i \,|\, \text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i\text{, Modelo})}^{\text{Verosimilitud}} \overbrace{P(\text{Hipótesis}_i \,|\, \text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

$$\overbrace{P(\text{Datos} \,|\, \text{Modelo})}^{\text{Evidencia}} = \, ?$$

Regla del producto
Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i \,|\, \text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i\text{, Modelo})}^{\text{Verosimilitud}}\,\overbrace{P(\text{Hipótesis}_i \,|\, \text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

$$\overbrace{P(\text{Datos} \,|\, \text{Modelo})}^{\text{Evidencia}} = \quad \text{Regla de la suma}$$

Regla del producto

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i \,|\, \text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i,\, \text{Modelo})}^{\text{Verosimilitud}}\,\overbrace{P(\text{Hipótesis}_i \,|\, \text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

$$\overbrace{P(\text{Datos} \,|\, \text{Modelo})}^{\text{Evidencia}} = \sum_{\text{Hipótesis}_i} P(\underbrace{\text{Hipótesis}_i,\, \text{Dato}}_{\substack{\text{Creencia} \\ \text{conjunta}}} \,|\, \text{Modelo})$$

Regla del producto

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i \,|\, \text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos} \,|\, \text{Hipótesis}_i\text{, Modelo})}^{\text{Verosimilitud}} \, \overbrace{P(\text{Hipótesis}_i \,|\, \text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

$$P(\text{Modelo}_j|\text{Datos}) =$$

Regla del producto

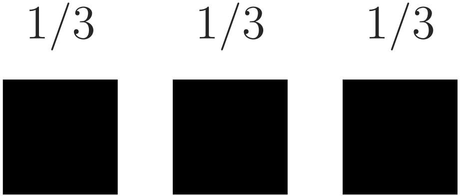
Teorema de Bayes

$$\underbrace{P(\text{Hipótesis}_i|\,\text{Datos, Modelo})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Datos}\,|\,\text{Hipótesis}_i,\,\text{Modelo})}^{\text{Verosimilitud}}\,\overbrace{P(\text{Hipótesis}_i|\,\text{Modelo})}^{\text{Prior}}}{\underbrace{P(\text{Datos}\,|\,\text{Modelo})}_{\text{Evidencia}}}$$

$$P(\text{Modelo}_j|\text{Datos}) = \frac{\overbrace{P(\text{Datos}|\text{Modelo}_j)}^{\text{Evidencia}}\,P(\text{Modelo}_j)}{P(\text{Datos})}$$

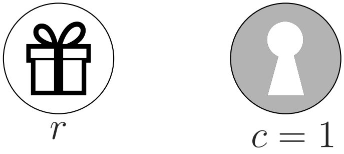
Modelo causal alternativo

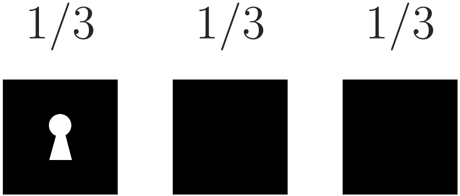
Modelo gráfico:



Modelo causal alternativo

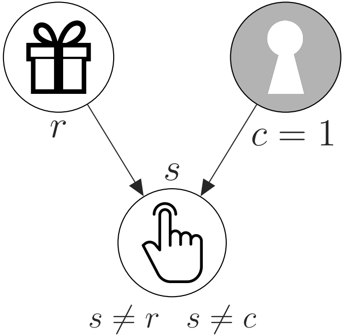
Modelo gráfico:

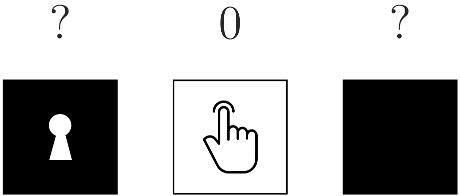




Modelo causal alternativo

Modelo gráfico:

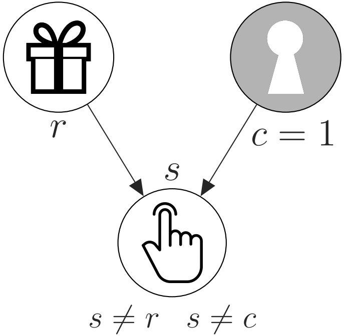




Modelo causal alternativo

Incertidumbre óptima dado el modelo

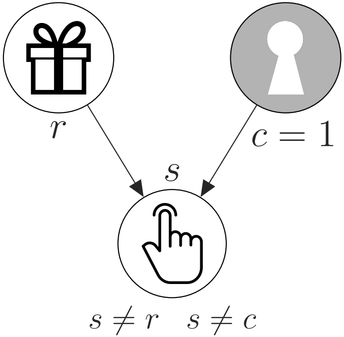
Modelo gráfico:

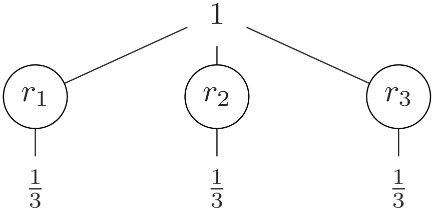


Modelo causal alternativo

Incertidumbre óptima dado el modelo

Modelo gráfico:

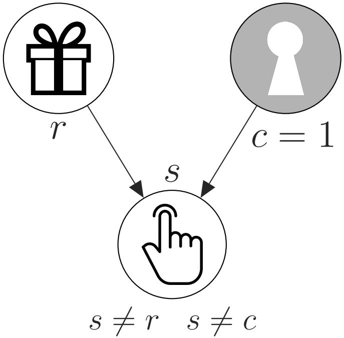


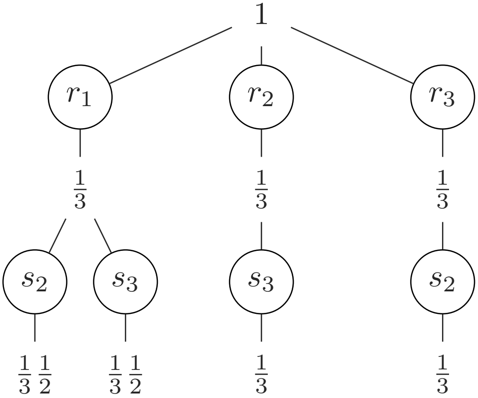


Modelo causal alternativo

Incertidumbre óptima dado el modelo

Modelo gráfico:

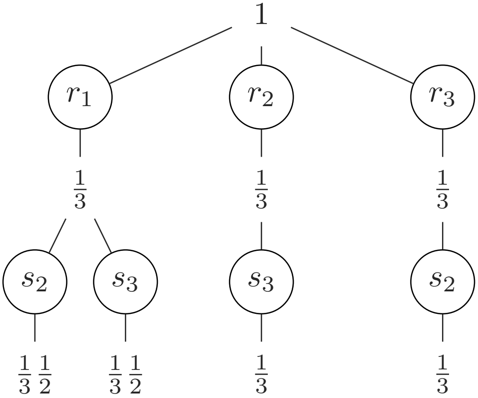




Modelo causal alternativo
Incertidumbre óptima dado el modelo

$P(r, s)$

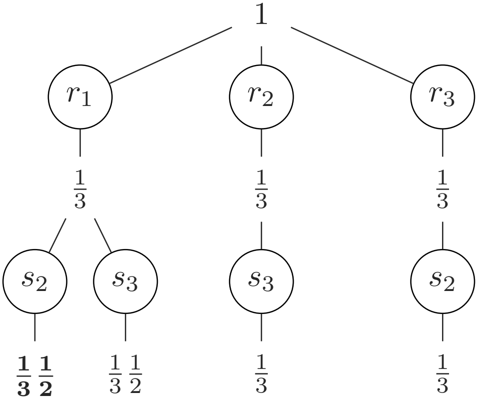
	r_1	r_2	r_3	
s_2				
s_3				



Modelo causal alternativo
Incertidumbre óptima dado el modelo

$P(r, s)$

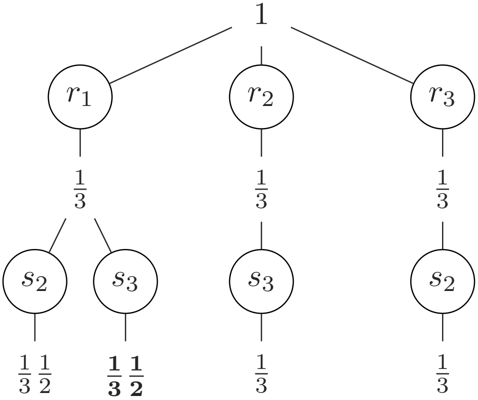
	r_1	r_2	r_3	
s_2	$1/6$			
s_3				



Modelo causal alternativo
Incertidumbre óptima dado el modelo

$P(r,s)$

	r_1	r_2	r_3	
s_2	$1/6$			
s_3	$1/6$			

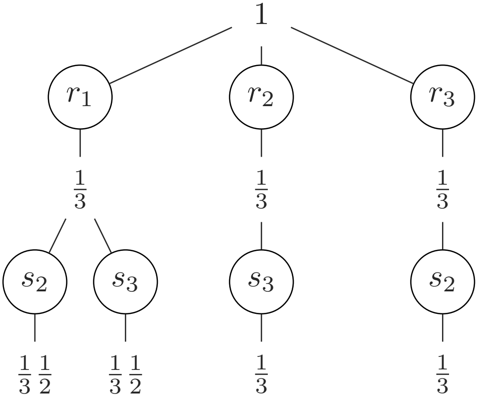


Modelo causal alternativo

Incertidumbre óptima dado el modelo

$$P(r,s)$$

	r_1	r_2	r_3	
s_2	$1/6$	0		
s_3	$1/6$			

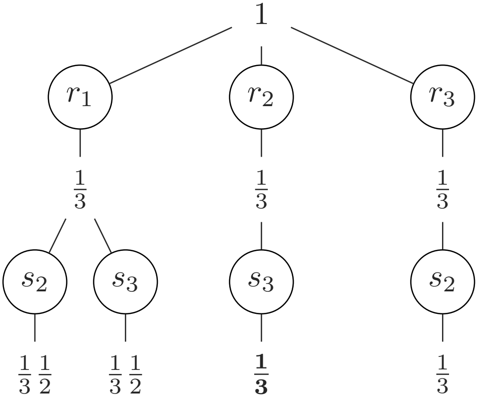


Modelo causal alternativo

Incertidumbre óptima dado el modelo

$P(r, s)$

	r_1	r_2	r_3	
s_2	$1/6$	0		
s_3	$1/6$	$1/3$		

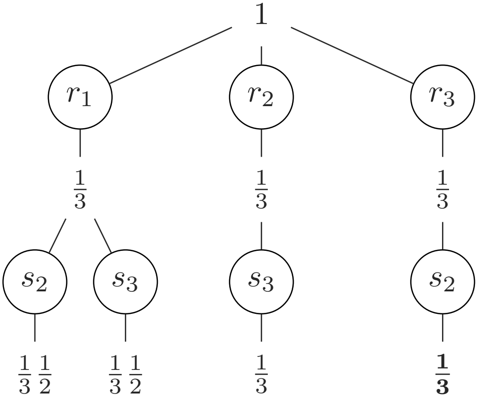


Modelo causal alternativo

Incertidumbre óptima dado el modelo

$P(r, s)$

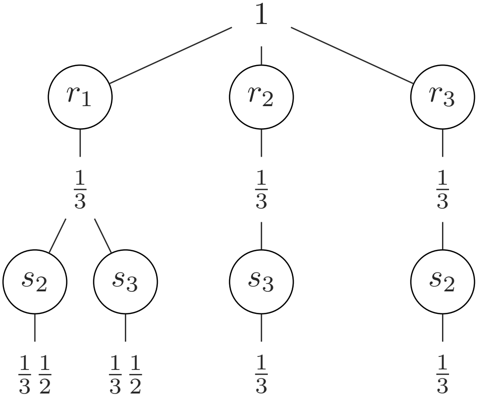
	r_1	r_2	r_3	
s_2	$1/6$	0	$1/3$	
s_3	$1/6$	$1/3$		



Modelo causal alternativo
Incertidumbre óptima dado el modelo

$P(r,s)$

	r_1	r_2	r_3	
s_2	$1/6$	0	$1/3$	
s_3	$1/6$	$1/3$	0	



Modelo causal alternativo

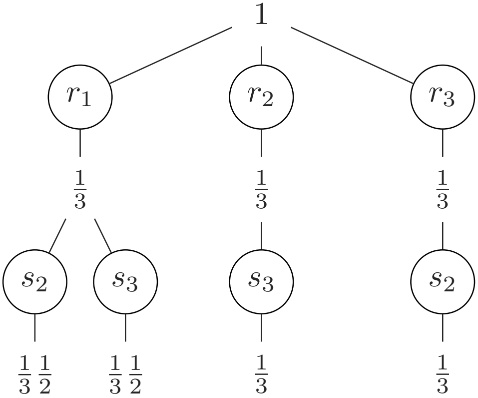
Incertidumbre óptima dado el modelo

$P(r, s)$

	r_1	r_2	r_3	
s_2	$1/6$	0	$1/3$	
s_3	$1/6$	$1/3$	0	

Regla de la suma

$$P(s_i) = \sum_j P(r_j, s_i)$$



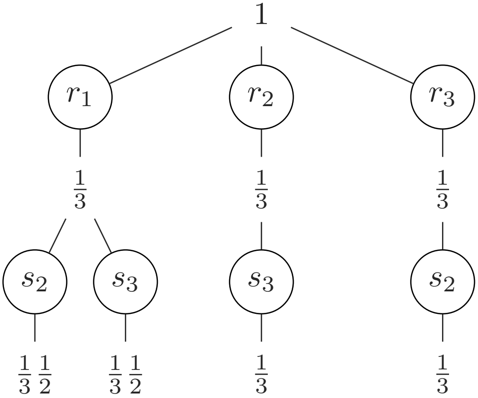
Modelo causal alternativo
Incertidumbre óptima dado el modelo

$P(r,s)$

	r_1	r_2	r_3	
s_2	$1/6$	0	$1/3$	$1/2$
s_3	$1/6$	$1/3$	0	$1/2$
	$1/3$	$1/3$	$1/3$	1

Regla de la suma

$$P(s_i) = \sum_j P(r_j, s_i)$$



Modelo causal alternativo

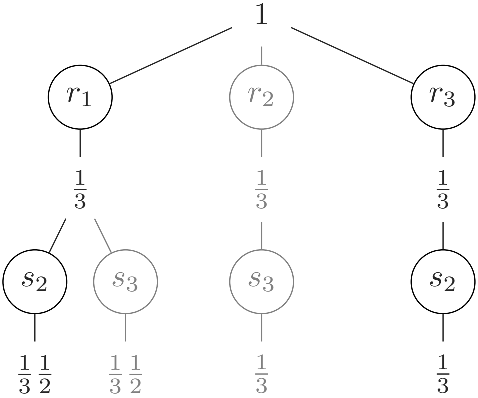
Incertidumbre óptima dado el modelo y el dato

$$P(r, s_2)$$

	r_1	r_2	r_3	
s_2	$1/6$	0	$1/3$	$1/2$

Regla del producto

$$P(r_i|s_2) = \frac{P(r_i, s_2)}{P(s_2)}$$



Modelo causal alternativo

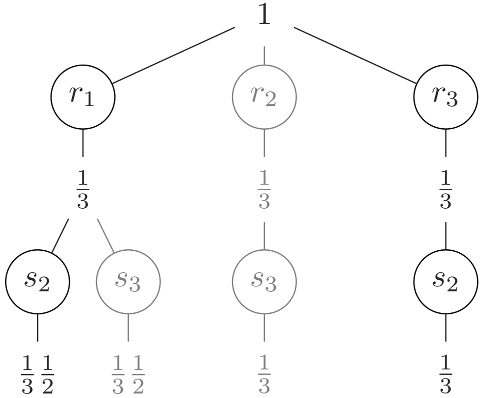
Incertidumbre óptima dado el modelo y el dato

$P(r|s_2)$

	r_1	r_2	r_3	
s_2	$1/3$	0	$2/3$	1

Regla del producto

$$P(r_i|s_2) = \frac{P(r_i, s_2)}{P(s_2)}$$



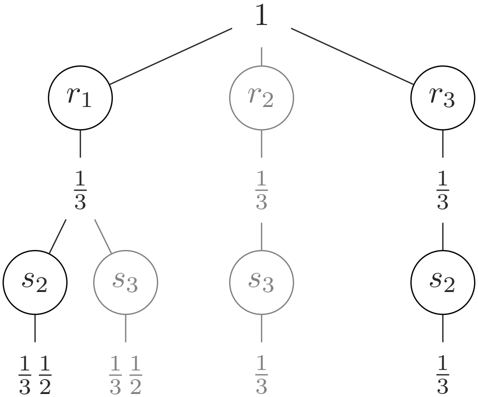
Modelo causal alternativo

Incertidumbre óptima dado el modelo y el dato

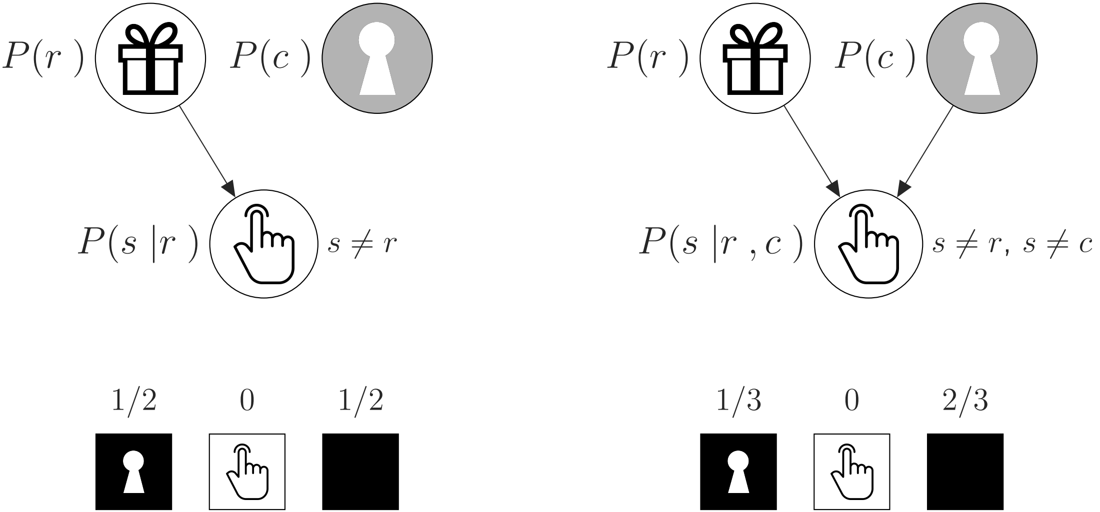
$P(r|s_2)$

	r_1	r_2	r_3	
s_2	$1/3$	0	$2/3$	1

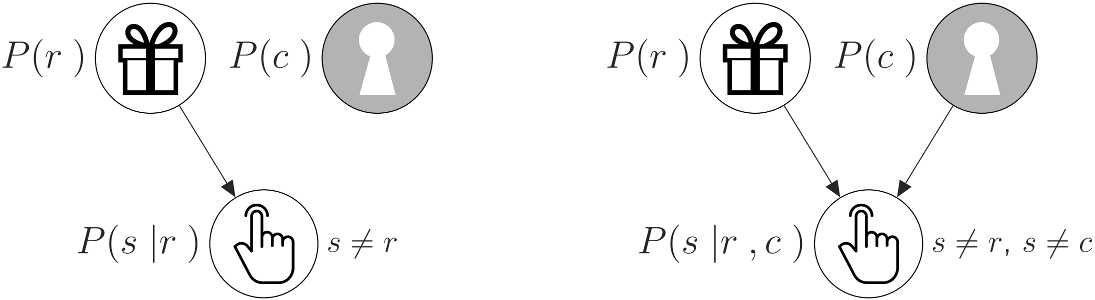
$1/3$ $\quad$ 0 $\quad$ $2/3$



Modelos causales alternativos



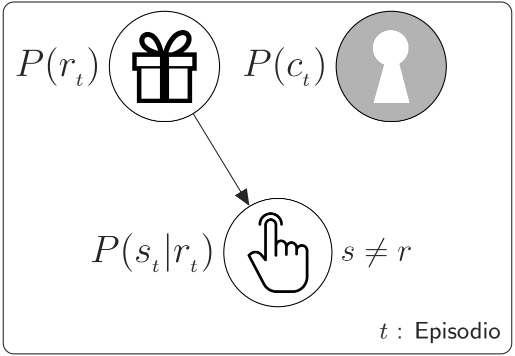
Modelos causales alternativos

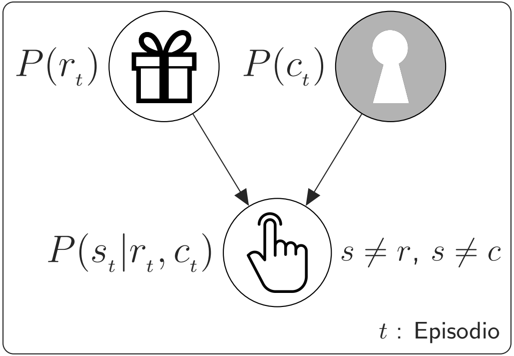


¿Y el acuerdo intersubjetivo respecto de los modelos?

$$P(\text{Modelo}_i|\text{Datos} = \{\underbrace{(c_0, s_0, r_0)}_{\text{Episodio 0}}, \underbrace{(c_1, s_1, r_1)}_{\text{Episodio 1}}, \dots\})$$

Modelos causales alternativos





¿Y el acuerdo intersubjetivo respecto de los modelos?

$$P(\text{Modelo}_i|\text{Datos} = \{\underbrace{(c_0, s_0, r_0)}_{\text{Episodio 0}}, \underbrace{(c_1, s_1, r_1)}_{\text{Episodio 1}}, \dots\})$$

Evaluación de modelos causales

$$P(\text{Modelo}_i|\text{Datos}) = \frac{\overbrace{P(\text{Datos}|\text{Modelo}_i)}^{\text{Evidencia}}\,P(\text{Modelo}_i)}{P(\text{Datos})}$$

Evaluación de modelos causales

$$P(\mathsf{Modelo}_i|\mathsf{Datos}) = \frac{\overbrace{P(\mathsf{Datos}|\mathsf{Modelo}_i)}^{\mathsf{Evidencia}}\, P(\mathsf{Modelo}_i)}{P(\mathsf{Datos})}$$

$$P(\mathsf{Datos} = \{d_1,\, d_2,\, \dots\}|\, \mathsf{Modelo}) =$$

Evaluación de modelos causales

$$P(\text{Modelo}_i|\text{Datos}) = \frac{\overbrace{P(\text{Datos}|\text{Modelo}_i)}^{\text{Evidencia}} P(\text{Modelo}_i)}{P(\text{Datos})}$$

$P(\text{Datos} = \{d_1, d_2, \dots\}|\,\text{Modelo}) =$ Regla de la suma

Evaluación de modelos causales

$$P(\text{Modelo}_i|\text{Datos}) = \frac{\overbrace{P(\text{Datos}|\text{Modelo}_i)}^{\text{Evidencia}}\,P(\text{Modelo}_i)}{P(\text{Datos})}$$

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\,\text{Modelo}) = \text{ Regla del producto}$$

Evaluación de modelos causales

$$P(\text{Modelo}_i|\text{Datos}) = \frac{\overbrace{P(\text{Datos}|\text{Modelo}_i)}^{\text{Evidencia}}\, P(\text{Modelo}_i)}{P(\text{Datos})}$$

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\,\text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1, \text{Modelo})\dots$$

Evaluación de modelos causales

$$P(\mathsf{Datos} = \{d_1,\, d_2,\, \dots\}|\, \mathsf{Modelo}) = P(d_1|\mathsf{Modelo})P(d_2|d_1, \mathsf{Modelo})\dots$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	$1/3$
s_2	$1/6$	0	$1/6$	$1/3$
s_3	$1/6$	$1/6$	0	$1/3$

	r_1	r_2	r_3	
s_1	0	0	0	0
s_2	$1/6$	0	$1/3$	$1/2$
s_3	$1/6$	$1/3$	0	$1/2$

$P(\mathsf{D}|\mathsf{M})$

$P(\mathsf{D}|\mathsf{M})$

Evaluación de modelos causales

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\} |\, \text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1, \text{Modelo}) \dots$$

	r_1	r_2	r_3	
s_1	0	1/6	1/6	1/3
$\boldsymbol{s_2}$	1/6	0	1/6	**1/3**
s_3	1/6	1/6	0	1/3

$$P(\text{D}|\text{M}) = \frac{1}{3}$$

	r_1	r_2	r_3	
s_1	0	0	0	0
$\boldsymbol{s_2}$	1/6	0	1/3	**1/2**
s_3	1/6	1/3	0	1/2

$$P(\text{D}|\text{M}) = \frac{1}{2}$$

Evaluación de modelos causales

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\,\text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1,\text{Modelo})\dots$$

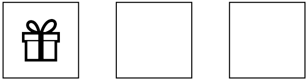
	r_1	r_2	r_3	
s_2	$1/2$	0	$1/2$	1

$$P(\text{D}|\text{M}) = \frac{1}{3}$$

	r_1	r_2	r_3	
s_2	$1/3$	0	$2/3$	1

$$P(\text{D}|\text{M}) = \frac{1}{2}$$

Evaluación de modelos causales



$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\,\text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1,\text{Modelo})\dots$$

	$\boldsymbol{r_1}$	r_2	r_3	
s_2	$\mathbf{1/2}$	0	$1/2$	$\mathbf{1}$

$$P(\text{D}|\text{M}) = \frac{1}{3}\,\frac{1}{2}$$

	$\boldsymbol{r_1}$	r_2	r_3	
s_2	$\mathbf{1/3}$	0	$2/3$	$\mathbf{1}$

$$P(\text{D}|\text{M}) = \frac{1}{2}\,\frac{1}{3}$$

Evaluación de modelos causales

$$P(\mathsf{Datos} = \{d_1,\, d_2,\, \ldots\}|\,\mathsf{Modelo}) = P(d_1|\mathsf{Modelo})P(d_2|d_1, \mathsf{Modelo})\ldots$$

	r_1	r_2	r_3	
s_1	0	$1/6$	$1/6$	$1/3$
s_2	$1/6$	0	$1/6$	$1/3$
s_3	$1/6$	$1/6$	0	$1/3$

	r_1	r_2	r_3	
s_1	0	0	0	0
s_2	$1/6$	0	$1/3$	$1/2$
s_3	$1/6$	$1/3$	0	$1/2$

$$P(\mathsf{D}|\mathsf{M}) = \frac{1}{3}\,\frac{1}{2}$$

$$P(\mathsf{D}|\mathsf{M}) = \frac{1}{2}\,\frac{1}{3}$$

Evaluación de modelos causales

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\,\text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1, \text{Modelo})\dots$$

	r_1	r_2	r_3	
s_1	0	1/6	1/6	1/3
s_2	1/6	0	1/6	1/3
$\boldsymbol{s_3}$	1/6	1/6	0	**1/3**

$$P(\text{D}|\text{M}) = \frac{1}{3}\,\frac{1}{2}\,\frac{1}{3}$$

	r_1	r_2	r_3	
s_1	0	0	0	0
s_2	1/6	0	1/3	1/2
$\boldsymbol{s_3}$	1/6	1/3	0	**1/2**

$$P(\text{D}|\text{M}) = \frac{1}{2}\,\frac{1}{3}\,\frac{1}{2}$$

Evaluación de modelos causales

$$P(\text{Datos} = \{d_1,\, d_2,\, \dots\}|\, \text{Modelo}) = P(d_1|\text{Modelo})P(d_2|d_1, \text{Modelo})\dots$$

	r_1	r_2	r_3	
s_3	$1/2$	$1/2$	0	1

$$P(\text{D}|\text{M}) = \frac{1}{3}\,\frac{1}{2}\,\frac{1}{3}$$

	r_1	r_2	r_3	
s_3	$1/3$	$2/3$	0	1

$$P(\text{D}|\text{M}) = \frac{1}{2}\,\frac{1}{3}\,\frac{1}{2}$$

Evaluación de modelos causales

$$P(\mathsf{Datos} = \{d_1,\, d_2,\, \dots\}|\, \mathsf{Modelo}) = P(d_1|\mathsf{Modelo})P(d_2|d_1, \mathsf{Modelo})\dots$$

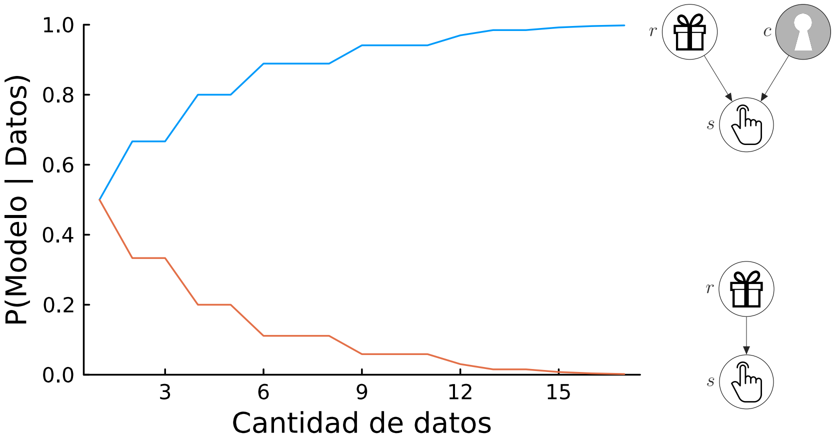
	r_1	$\boldsymbol{r_2}$	r_3	
s_3	$1/2$	$\mathbf{1/2}$	0	$\mathbf{1}$

$$P(\mathsf{D}|\mathsf{M}) = \frac{1}{3}\,\frac{1}{2}\,\frac{1}{3}\,\frac{1}{2}$$

	r_1	$\boldsymbol{r_2}$	r_3	
s_3	$1/3$	$\mathbf{2/3}$	0	$\mathbf{1}$

$$P(\mathsf{D}|\mathsf{M}) = \frac{1}{2}\,\frac{1}{3}\,\frac{1}{2}\,\frac{2}{3}$$

Evaluación de modelos causales



Redes bayesianas
Descomposiciones condicionales

$P(l, e, t, r, a)$

Redes bayesianas
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e, t, r, a|l)$$

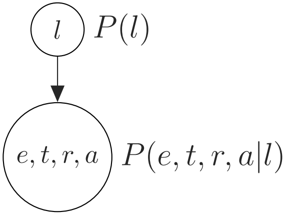
Redes bayesianas
Descomposiciones condicionales

$$P(l,e,t,r,a) = P(l)P(e,t,r,a|l) = \cancel{P(l)}\frac{P(l,e,t,r,a)}{\cancel{P(l)}}$$

Redes bayesianas

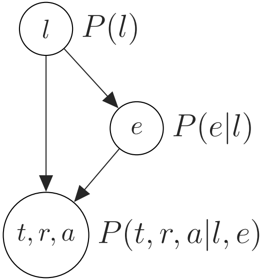
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e, t, r, a|l)$$



Redes bayesianas
Descomposiciones condicionales

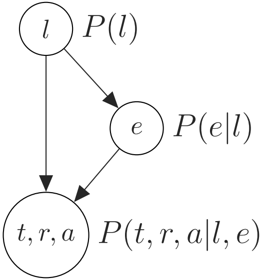
$$P(l, e, t, r, a) = P(l)P(e|l)P(t, r, a|l, e)$$



Redes bayesianas

Descomposiciones condicionales

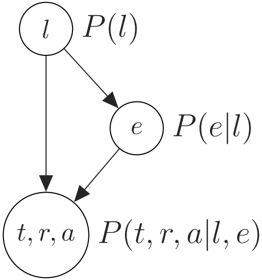
$$P(l,e,t,r,a) = P(l)P(e|l)P(t,r,a|l,e) = \cancel{P(l)}\frac{\cancel{P(l,e)}}{\cancel{P(l)}}\frac{P(l,e,t,r,a)}{\cancel{P(l,e)}}$$



Redes bayesianas

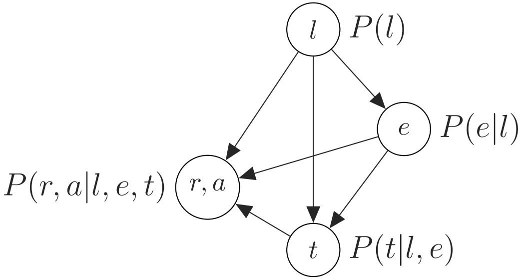
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e|l)P(t, r, a|l, e)$$



Redes bayesianas
Descomposiciones condicionales

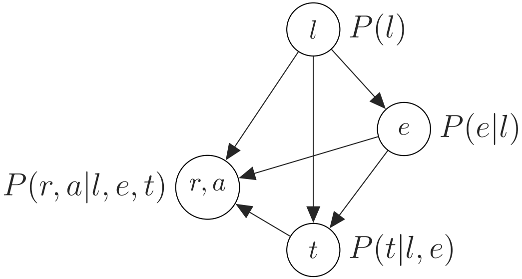
$$P(l, e, t, r, a) = P(l)P(e|l)P(t|l, e)P(r, a|l, e, t)$$



Redes bayesianas

Descomposiciones condicionales

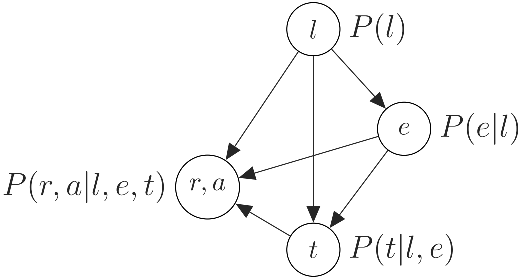
$$P(l,e,t,r,a) = P(l)P(e|l)P(t|l,e)P(r,a|l,e,t) = \cancel{P(l)}\frac{\cancel{P(l,e)}}{\cancel{P(l)}}\frac{\cancel{P(l,e,t)}}{\cancel{P(l,e)}}\frac{P(l,e,t,r,a)}{\cancel{P(l,e,t)}}$$



Redes bayesianas

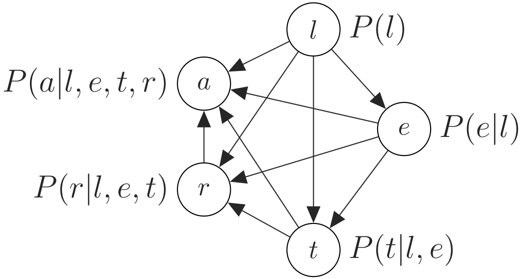
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e|l)P(t|l, e)P(r, a|l, e, t)$$



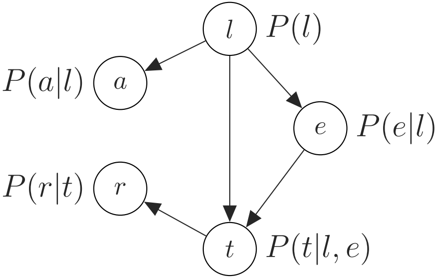
Redes bayesianas
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e|l)P(t|l, e)P(r|l, e, t)P(a|l, e, t, r)$$



Redes bayesianas
Descomposiciones condicionales

$$P(l, e, t, r, a) = P(l)P(e|l)P(t|l, e)P(r|t)P(a|l)$$



Redes bayesianas

Descomposiciones condicionales

$P(l, e, t, r, a) =$ ¿ ?

Hay $5! = 120$ posibles descomposiciones

Redes bayesianas

Descomposiciones condicionales

$P(l, e, t, r, a)$

Hay $5! = 120$ posibles descomposiciones

¿Cuál de todas ellas usar?

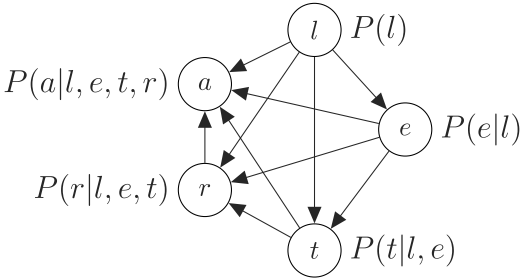
Redes bayesianas

Historias causales

La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.

Redes bayesianas
Historias causales

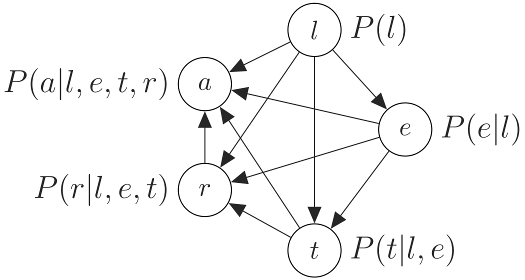
La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.



Redes bayesianas

Historias causales

La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.

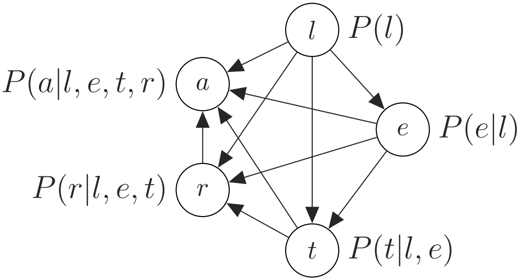


¿Cómo podemos definir la probabilidad conjunta?

Redes bayesianas

Historias causales

La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.



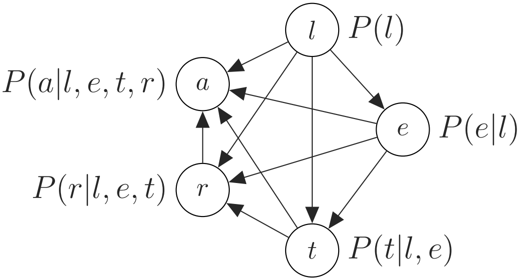
¿Cómo podemos definir la probabilidad conjunta?

$$P(t|l,e) = \,?$$

Redes bayesianas

Historias causales

La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.



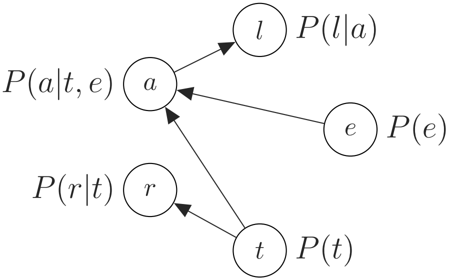
¿Cómo podemos definir la probabilidad conjunta?

$$P(t)\, P(e)\, P(a|e, t)\, P(l|a) P(r|t)$$

Redes bayesianas

Historias causales

La **alarma** (a) de una casa se activa cuando alguien **entra** (e) sin apagarla. Si la alarma se activa, la dueña recibe una **llamada** (l) de la central de seguridad. La alarma se puede activar por otros motivos, como los **terremotos** (t) que no son infrecuentes en la ciudad. Siempre que se produce un terremoto, en las **redes** (r) se habla casi exclusivamente de eso.

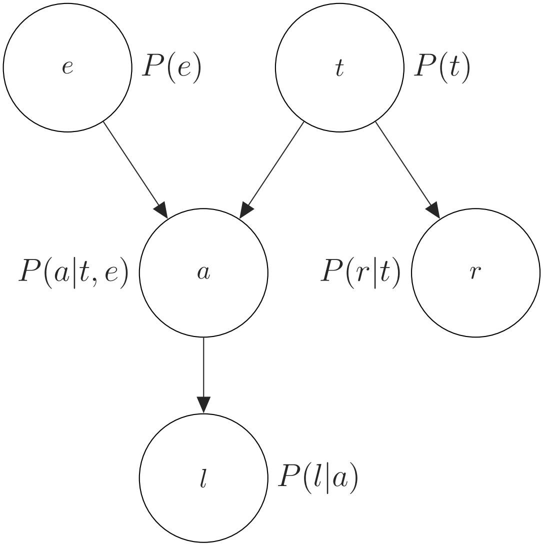


¿Cómo podemos definir la probabilidad conjunta?

$$P(t)\,P(e)\,P(a|e,t)\,P(l|a)P(r|t)$$

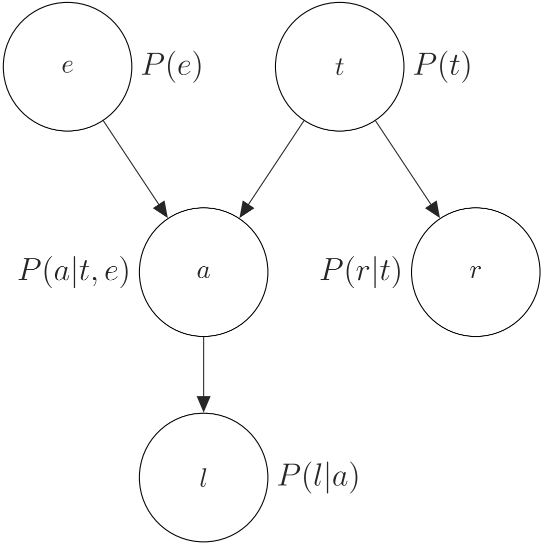
Redes bayesianas

Historias causales



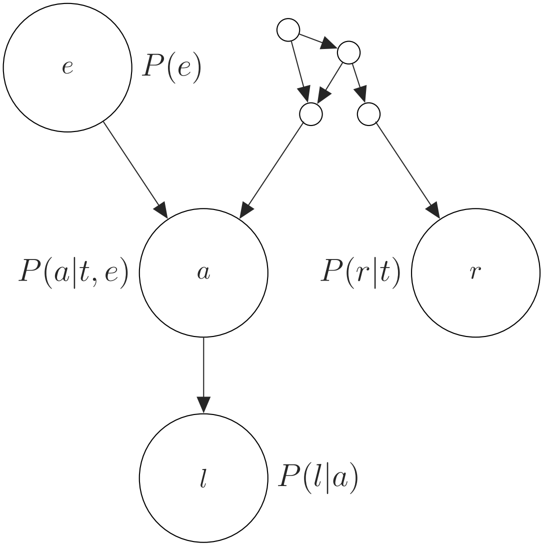
Redes bayesianas

Historias causales



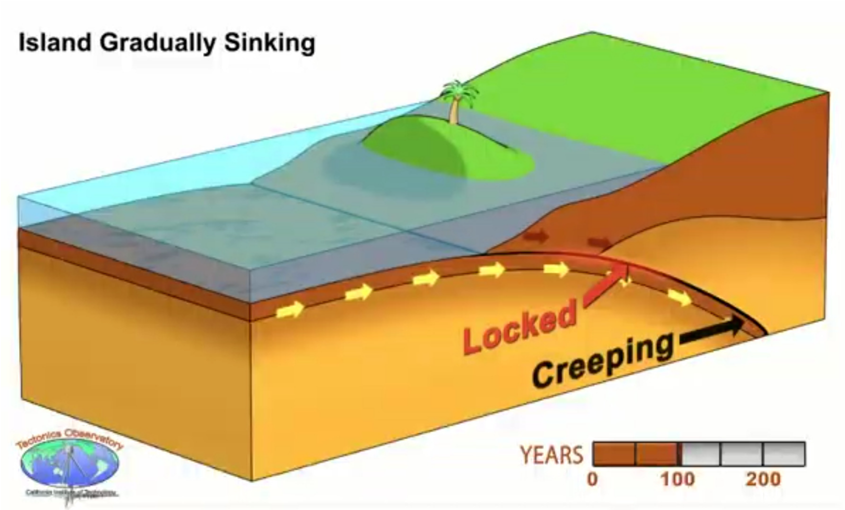
Cada variable resumen el comportamiento de un sistema complejo

Redes bayesianas
Sistemas complejos

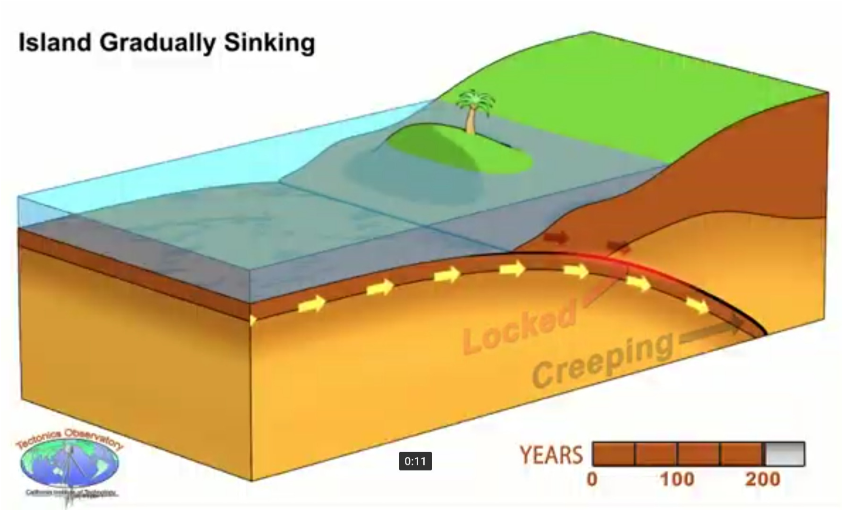


Cada variable resumen el comportamiento de un sistema complejo

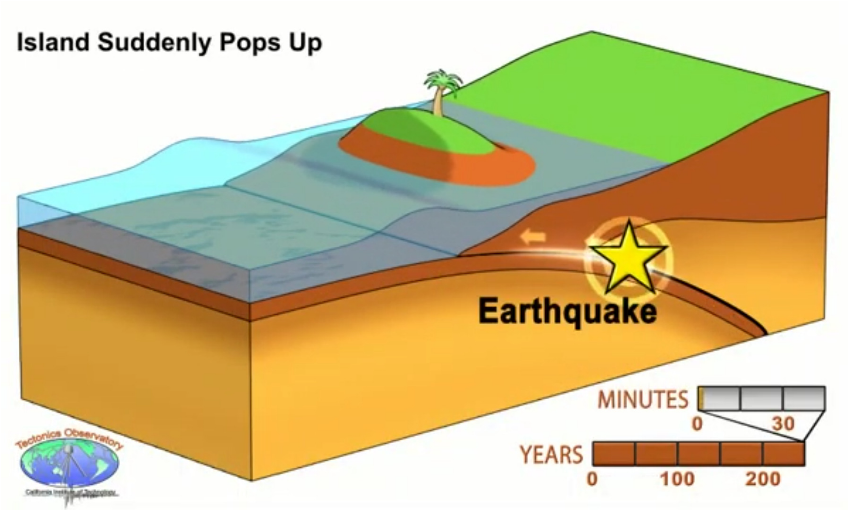
Redes bayesianas
Sistemas complejos



Redes bayesianas
Sistemas complejos

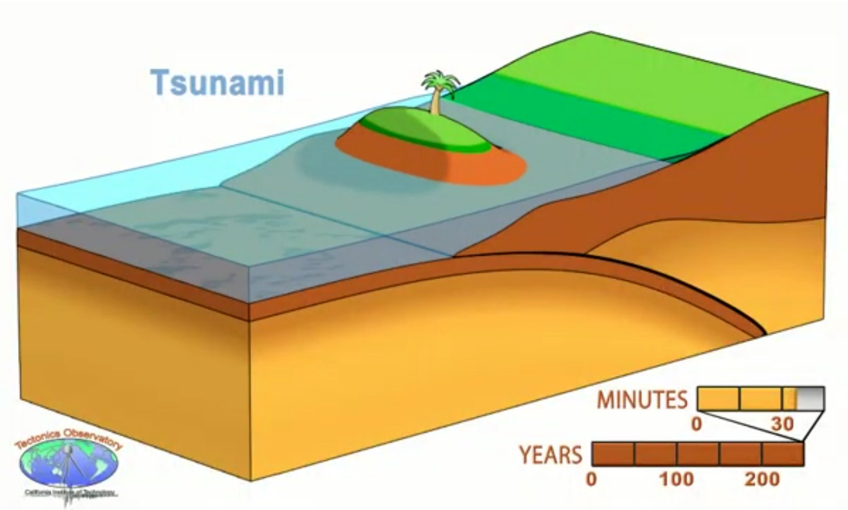


Redes bayesianas
Sistemas complejos

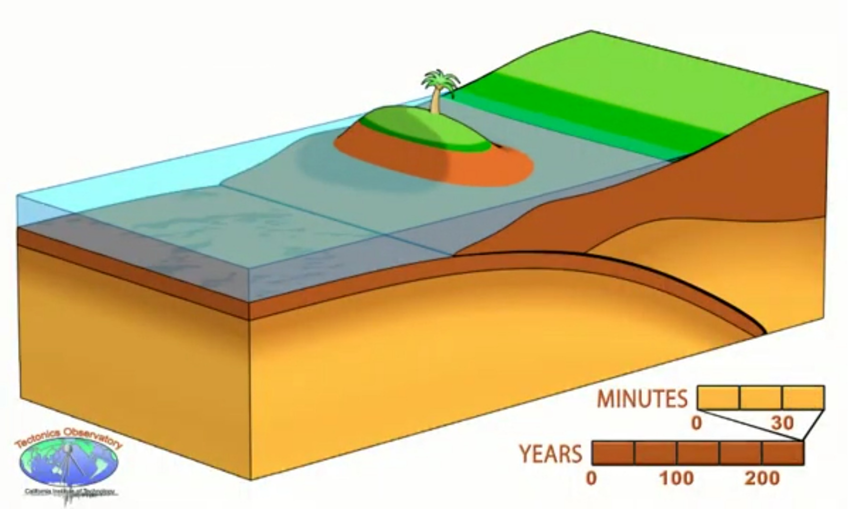


Redes bayesianas

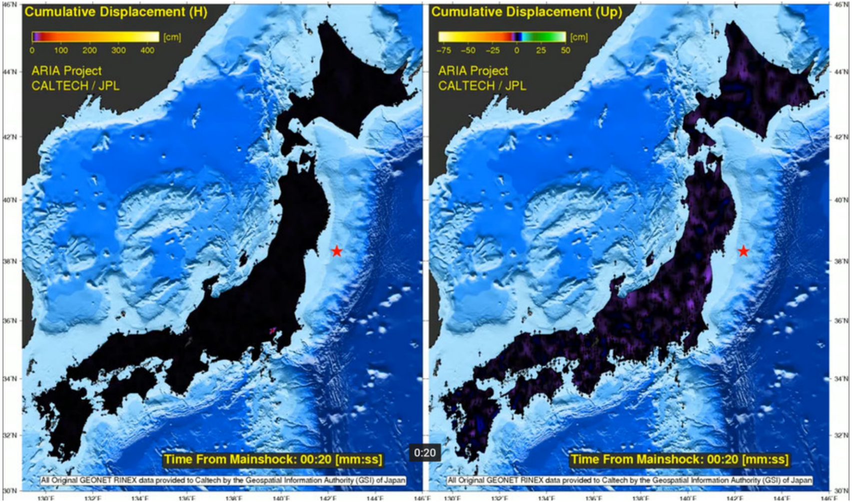
Sistemas complejos



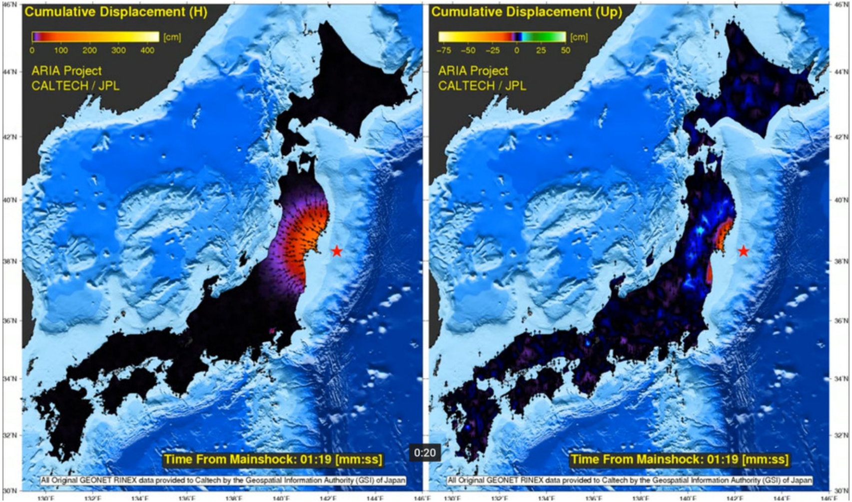
Redes bayesianas
Sistemas complejos



Redes bayesianas
Sistemas complejos

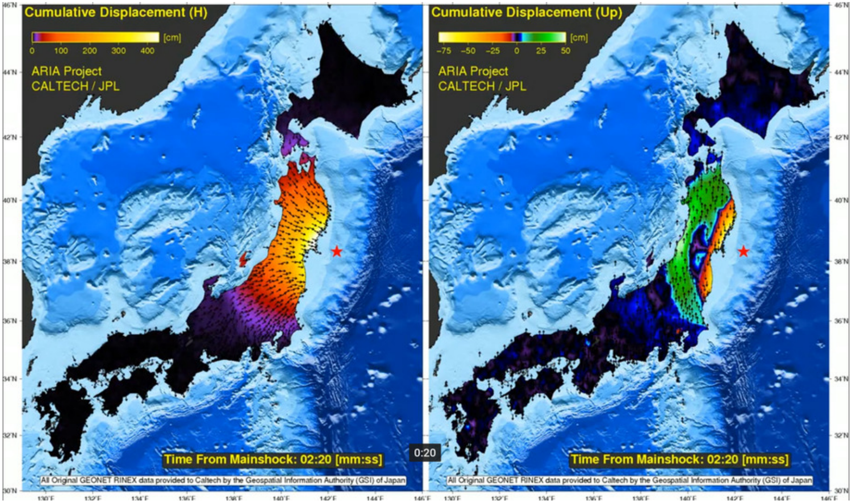


Redes bayesianas
Sistemas complejos

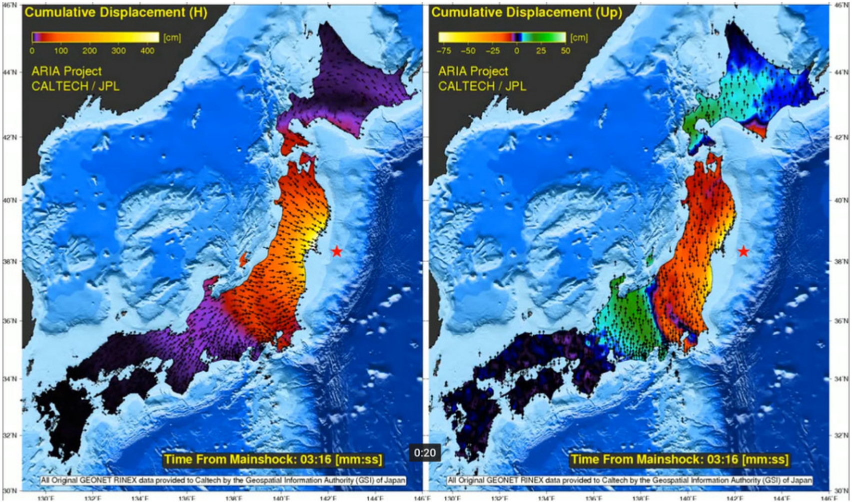


Redes bayesianas

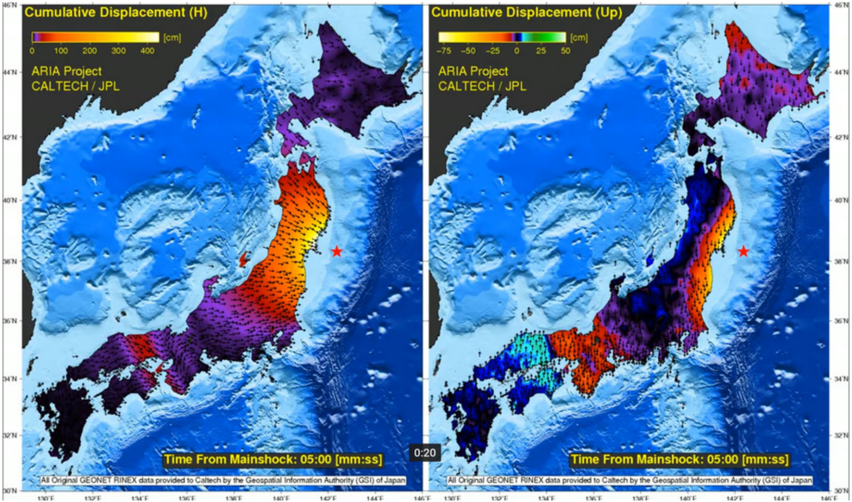
Sistemas complejos



Redes bayesianas
Sistemas complejos

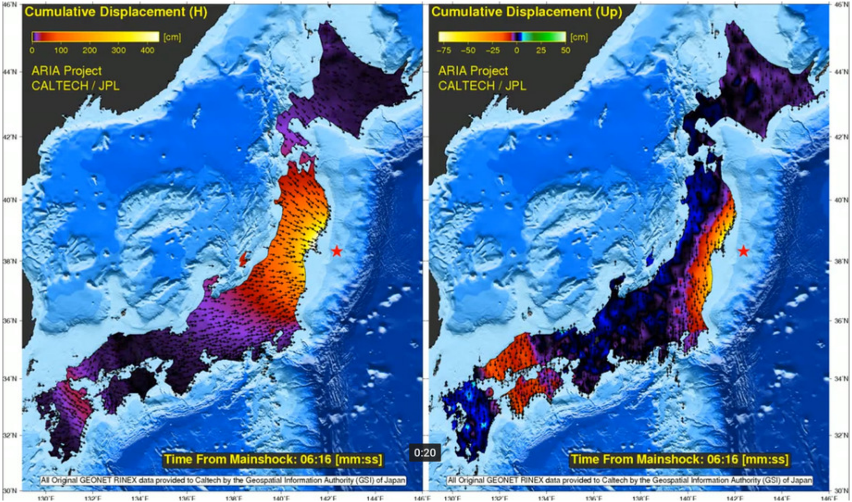


Redes bayesianas
Sistemas complejos

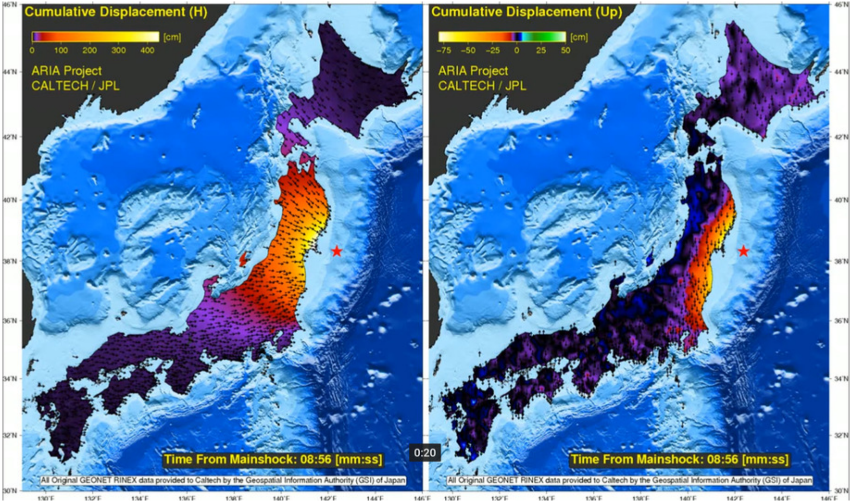


Redes bayesianas

Sistemas complejos

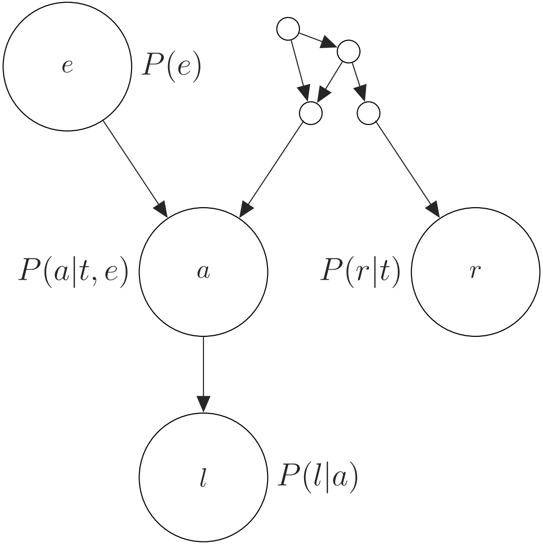


Redes bayesianas
Sistemas complejos



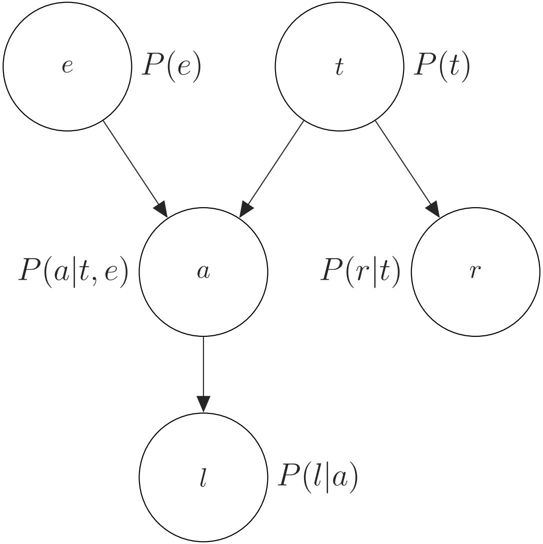
Redes bayesianas

Sistemas complejos



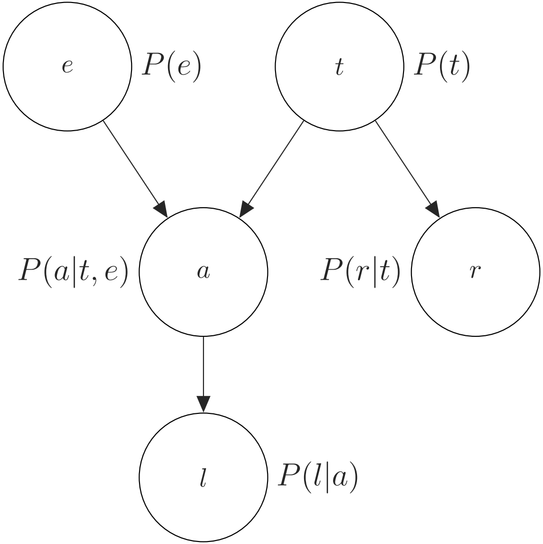
Redes bayesianas

Historias causales



Redes bayesianas

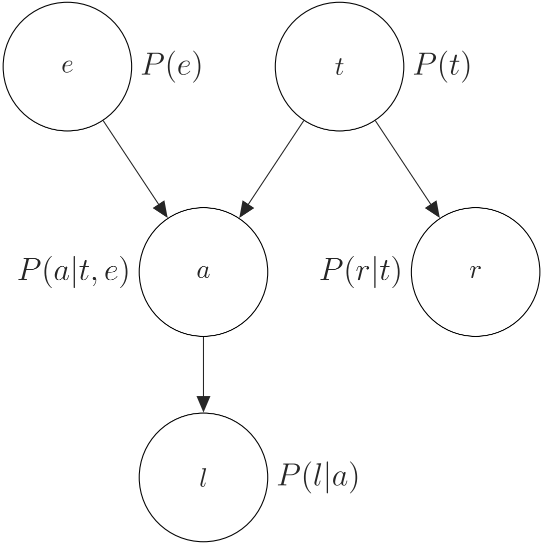
Historias causales



- Define la probabilidad conjunta

Redes bayesianas

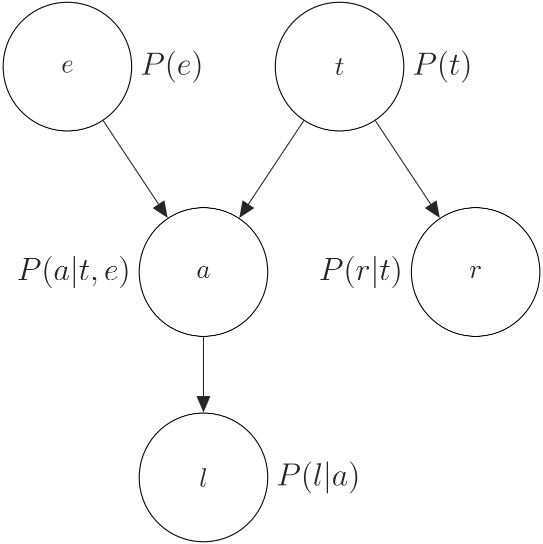
Historias causales



- Define la probabilidad conjunta
- Expresa las hipótesis de forma intuitiva

Redes bayesianas

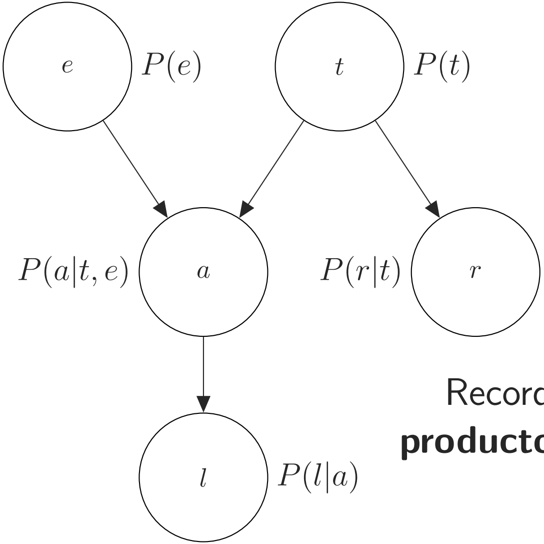
Historias causales



- Define la probabilidad conjunta
- Expresa las hipótesis de forma intuitiva
- Permite hacer inferencia eficientemente

Redes bayesianas

Historias causales

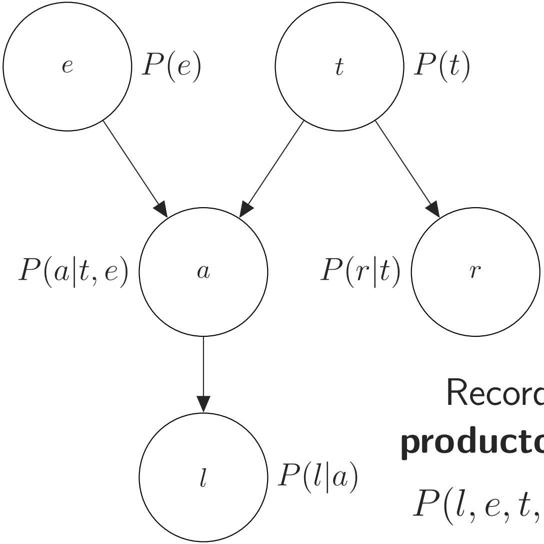


- Define la probabilidad conjunta
- Expresa las hipótesis de forma intuitiva
- Permite hacer inferencia eficientemente

Recordar que la probabilidad conjunta es el **producto de las distribuciones condicionales**

Redes bayesianas

Historias causales



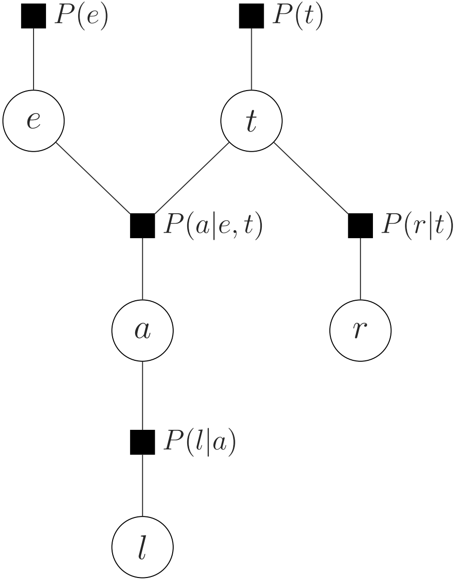
- Define la probabilidad conjunta
- Expresa las hipótesis de forma intuitiva
- Permite hacer inferencia eficientemente

Recordar que la probabilidad conjunta es el **producto de las distribuciones condicionales**

$$P(l,e,t,r,a) = P(e)P(t)P(r|t)P(a|t,e)P(l|a)$$

Método de especificación

Grafo de factorización (factor graph)

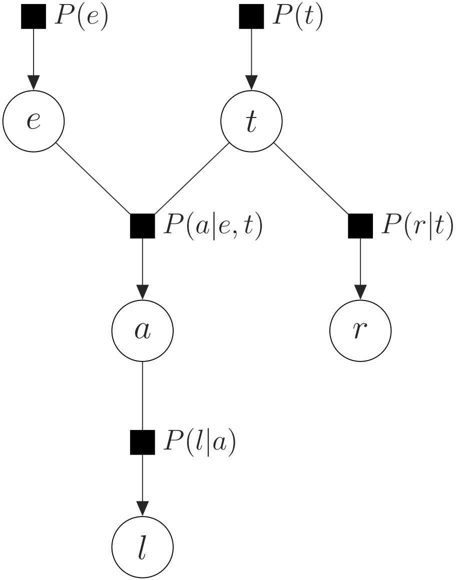


Nodos:
variables y funciones

Ejes:
"la variable v es argumento de la función P"

Método de especificación

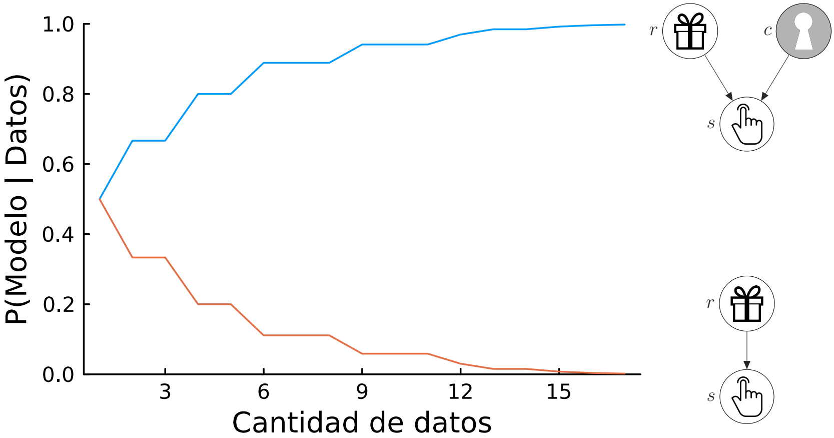
Grafo de factorización (factor graph)



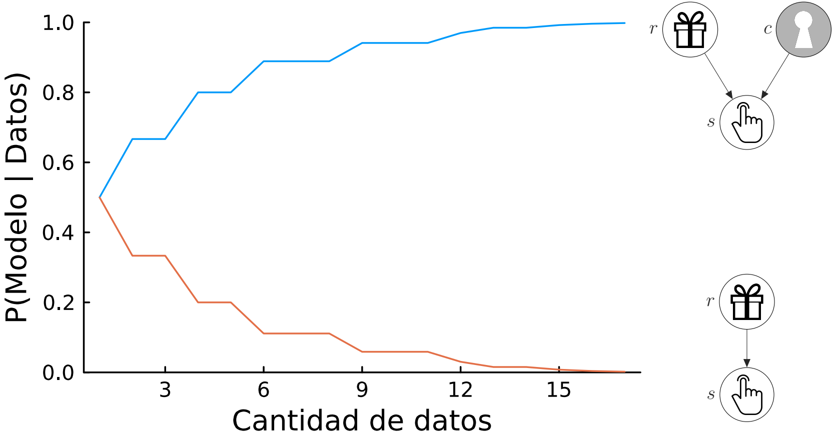
Nodos:
variables y funciones

Ejes:
"la variable v es argumento de la función P"

Evaluación de modelos causales

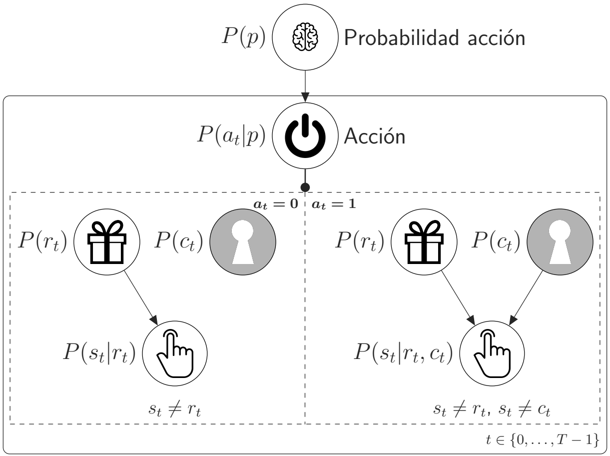


Evaluación de modelos causales

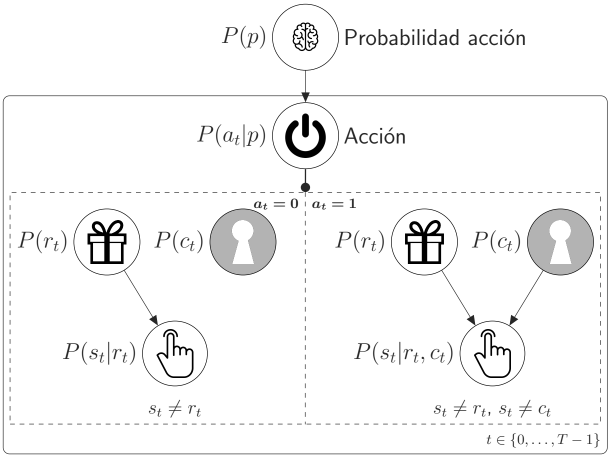


¿Qué pasaría con la evaluación de modelos si la persona que da la pista a veces se olvida de tener en cuenta la caja que elegimos?

Mecanismos causales dinámicos

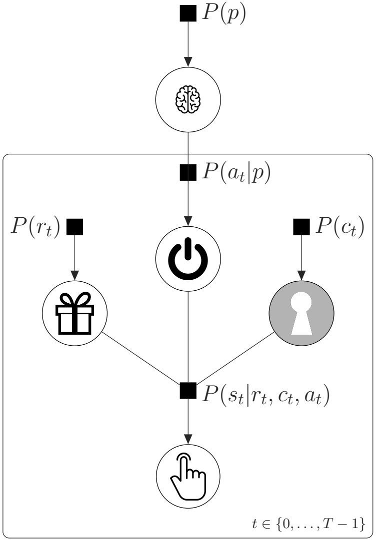


Mecanismos causales dinámicos



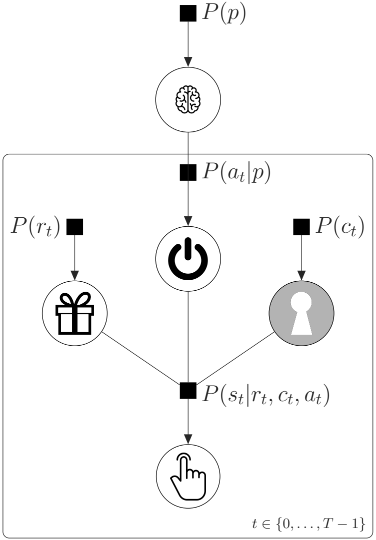
Veamos la especificación precisa usando factor graphs.

Mecanismos causales dinámicos



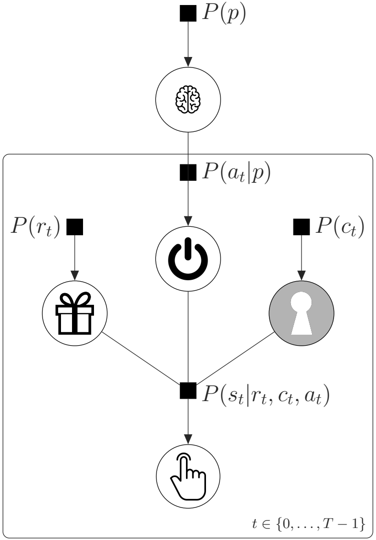
Mecanismos causales dinámicos

$$P(s_t|r_t, c_t, a_t = 0) = P(s_t|r_t)$$



Mecanismos causales dinámicos

$$P(s_t|r_t, c_t, a_t = 0) = P(s_t|r_t)$$



$$P(s_t|r_t, c_t, a_t = 1) = P(s_t|r_t, c_t)$$

Mecanismos causales dinámicos

- Minka, T; Winn, J. *Gates*; Advances in Neural Information Processing Systems. 2008. (Descargar).
- Winn, J. **Causality with Gates**; Proceedings of Machine Learning Research. 2012. (Descargar).

Mecanismos causales dinámicos

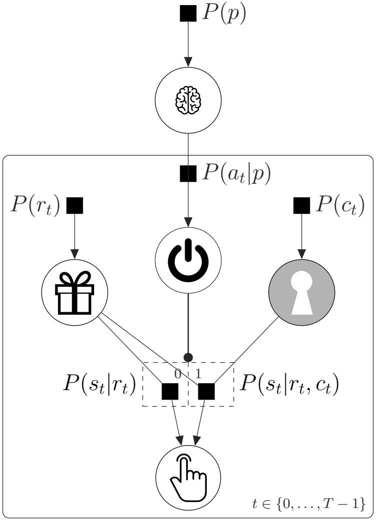
- Minka, T; Winn, J. *Gates*; Advances in Neural Information Processing Systems. 2008. (Descargar).

- Winn, J. **Causality with Gates**; Proceedings of Machine Learning Research. 2012. (Descargar).

La notación de compuertas (gates) permite especificar mecanismos causales dinámicos que dependen del contexto.

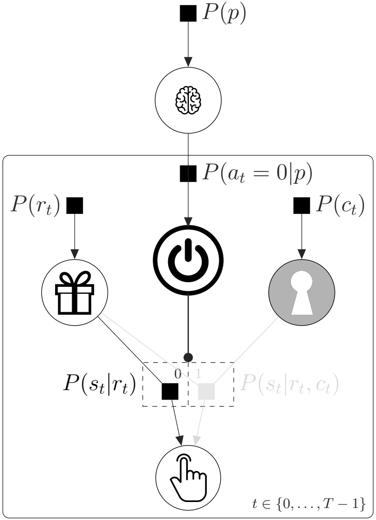
Mecanismos causales dinámicos

Gates



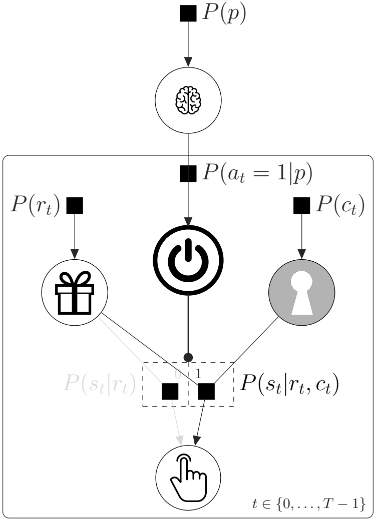
Mecanismos causales dinámicos

Gates



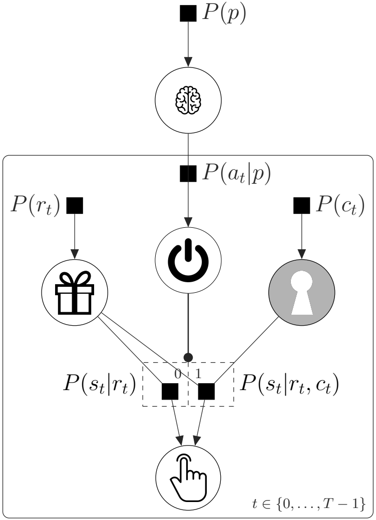
Mecanismos causales dinámicos

Gates



Mecanismos causales dinámicos

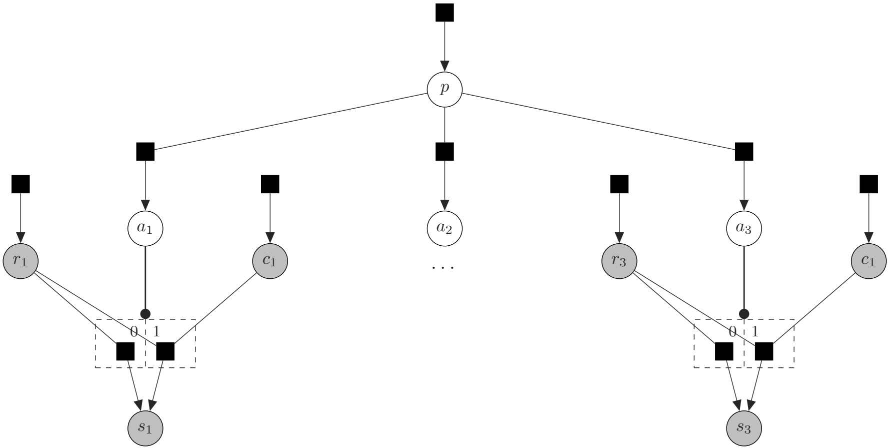
Gates



$$P(s_t|r_t, c_t, a_t) = P(s_t|r_t)^{\mathbb{I}(a_t=0)} P(s_t|r_t, c_t)^{\mathbb{I}(a_t=1)}$$

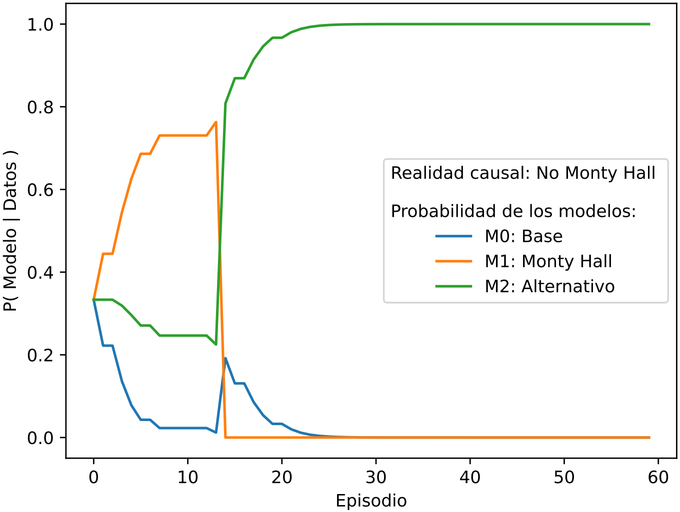
Mecanismos causales dinámicos

Gates



Mecanismos causales dinámicos

Evaluación de modelos



Evaluación de modelos causales alternativos

con misma distribución conjunta

Modelos causales alternativos

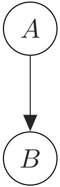
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95



Modelos causales alternativos

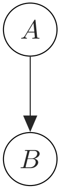
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

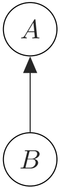


$P(A|B)$

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(B)$

$B=0$	$B=1$
0.5	0.5



Modelos causales alternativos

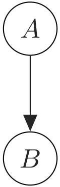
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

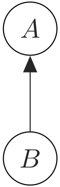


$P(A|B)$

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(B)$

$B=0$	$B=1$
0.5	0.5



$$P(A,B) = P(A)P(B|A) = P(B)P(A|B)$$

Ambos modelos tienen la misma distribución conjunta.

Modelos causales alternativos

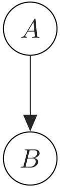
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

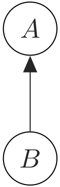


$P(A|B)$

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(B)$

$B=0$	$B=1$
0.5	0.5



$$P(A,B) = P(A)P(B|A) = P(B)P(A|B)$$

Entonces predicen igual!

Modelos causales alternativos

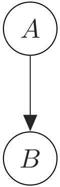
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

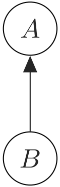


$P(A|B)$

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(B)$

$B=0$	$B=1$
0.5	0.5



$$P(\text{Modelo}_{A\to B}|\text{Datos}) = P(\text{Modelo}_{B\to A}|\text{Datos})$$

Modelos causales alternativos

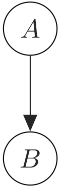
con misma distribución conjunta

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

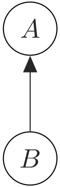


$P(A|B)$

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(B)$

$B=0$	$B=1$
0.5	0.5



$$P(\text{Modelo}_{A\to B}|\text{Datos}) = P(\text{Modelo}_{B\to A}|\text{Datos})$$

No podemos identificar cuál es el modelo causal correcto!


¿El movimiento de las
ramas produce viento,
o
el viento mueve
las ramas?

Modelos causales alternativos

con misma distribución conjunta

Con **intervenciones** (mover las ramas)
podemos identificar el modelo causal correcto

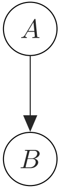
Modelos causales alternativos
Identificación mediante intervenciones

$P(A)$

$A=0$	$A=1$
0.5	0.5

$P(B|A)$

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

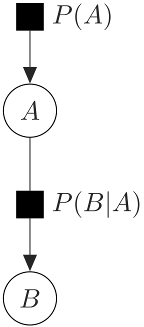


$$P(A,B|\text{Modelo}_{A\to B})$$

Modelos causales alternativos

Identificación mediante intervenciones

$A=0$	$A=1$
0.5	0.5



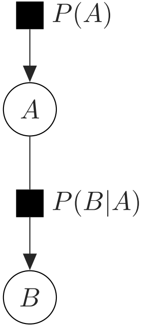
	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95

$$P(A,B|\text{Modelo}_{A\rightarrow B})$$

Modelos causales alternativos
Identificación mediante intervenciones

$A=0$	$A=1$
0.5	0.5

	$B=0$	$B=1$
$A=0$	0.95	0.05
$A=1$	0.05	0.95



Nodos: Variables y Funciones

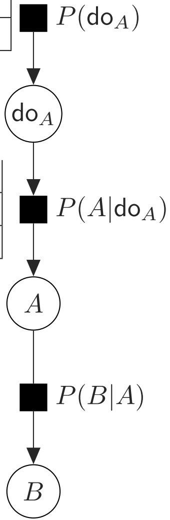
Ejes: Variable v es parámetro de la función f

Modelos causales alternativos
Identificación mediante intervenciones

$\text{do}_A = 0$	$\text{do}_A = 1$
$1 - \delta_A$	δ_A

	$A = 0$	$A = 1$
$\text{do}_A = 0$	0.5	0.5
$\text{do}_A = 1$	$1 - \alpha$	α

	$B = 0$	$B = 1$
$A = 0$	0.95	0.05
$A = 1$	0.05	0.95

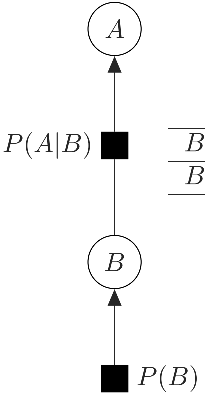


$$P(A, B, \text{do}_A | \text{Modelo}_{A \to B})$$

Modelos causales alternativos

Identificación mediante intervenciones

$$P(A, B, \text{do}_A | \text{Modelo}_{B \to A})$$

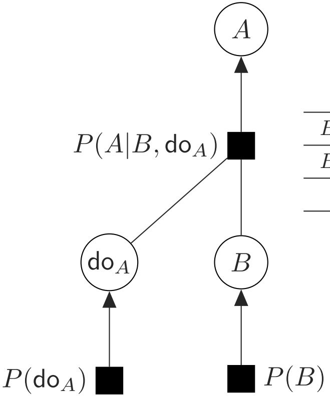


	$A = 0$	$A = 1$
$B = 0$	0.95	0.05
$B = 1$	0.05	0.95

$B = 0$	$B = 1$
0.5	0.5

Modelos causales alternativos

Identificación mediante intervenciones



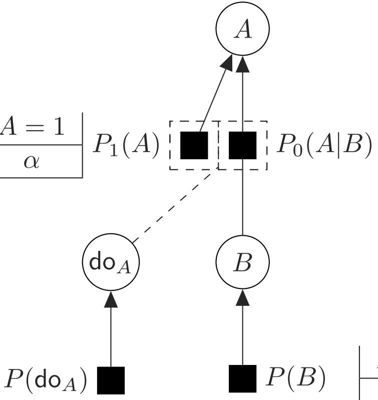
	$A=0$	$A=1$
$B=0$ $\text{do}_A=0$	0.95	0.05
$B=1$ $\text{do}_A=0$	0.05	0.95
$\text{do}_A=1$	$1-\alpha$	α

$\text{do}_A=0$	$\text{do}_A=1$
$1-\delta_A$	δ_A

$B=0$	$B=1$
0.5	0.5

Modelos causales alternativos

Identificación mediante intervenciones



$A=0$	$A=1$
$1-\alpha$	α

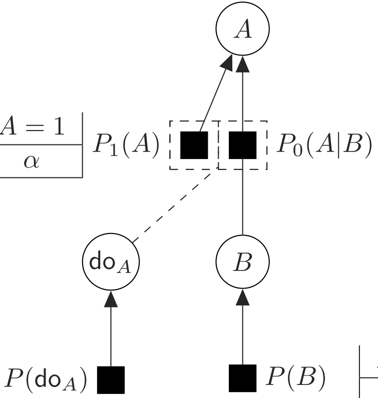
	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$\mathsf{do}_A=0$	$\mathsf{do}_A=1$
$1-\delta_A$	δ_A

$B=0$	$B=1$
0.5	0.5

Modelos causales alternativos

Identificación mediante intervenciones



$P_1(A)$:

$A=0$	$A=1$
$1-\alpha$	α

$P_0(A|B)$:

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$P(\mathsf{do}_A)$:

$\mathsf{do}_A=0$	$\mathsf{do}_A=1$
$1-\delta_A$	δ_A

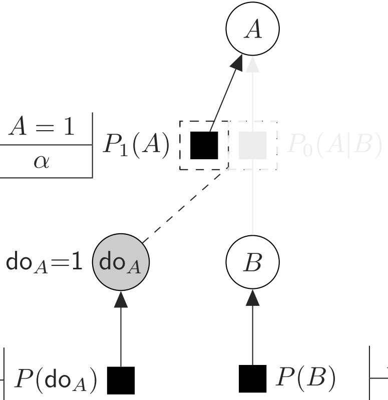
$P(B)$:

$B=0$	$B=1$
0.5	0.5

$$P(A,B,\mathsf{do}_A|\mathsf{Modelo}_{B\to A}) = P(B)\,P_0(A|B)^{1-\mathsf{do}_A}\,P_1(A)^{\mathsf{do}_A}\,P(\mathsf{do}_A)$$

Modelos causales alternativos

Identificación mediante intervenciones



$A=0$	$A=1$
$1-\alpha$	α

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

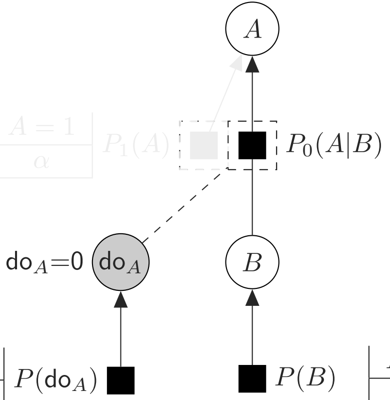
do$_A=0$	do$_A=1$
$1-\delta_A$	δ_A

$B=0$	$B=1$
0.5	0.5

$$P(A,B|\underbrace{\text{do}_A=1,\text{Modelo}_{B\to A}}_{\text{Intervención}}) = P(B)\,P_1(A)$$

Modelos causales alternativos

Identificación mediante intervenciones



$A=0$	$A=1$
$1-\alpha$	α

	$A=0$	$A=1$
$B=0$	0.95	0.05
$B=1$	0.05	0.95

$\text{do}_A=0$	$\text{do}_A=1$
$1-\delta_A$	δ_A

$B=0$	$B=1$
0.5	0.5

$$P(A,B|\underbrace{\text{do}_A=0, \text{Modelo}_{B\to A}}_{\text{Sin intervención}}) = P(B)\,P_0(A|B)$$

Modelos causales alternativos

Identificación mediante intervenciones

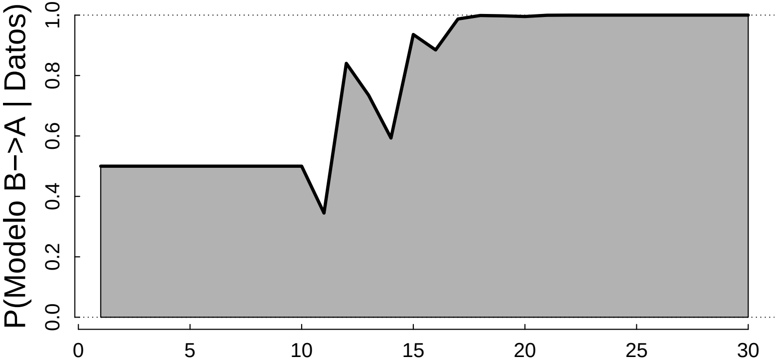
Datos:

i	do_{Ai}	A_i	B_i
1	0	1	1
$\dots$	0	$\dots$	$\dots$
10	0	0	0
11	1	0	1
12	1	1	0
$\dots$	1	$\dots$	$\dots$

Modelos causales alternativos
Identificación mediante intervenciones

Datos:

i	do_{Ai}	A_i	B_i
1	0	1	1
...	0	...	...
10	0	0	0
11	1	0	1
12	1	1	0
...	1	...	...



Identificación de modelo causal
El conocimiento experto

La principal fuente de información para la identificación modelos causales alternativos es el conocimiento experto.

$p = b$

Laboratorios de
Métodos Bayesianos