Técnicas para calcular distribuciones de creencias
Estimación de habilidad en la industria del video juego

Gustavo Landfried
@GALandfried

Licenciado en Ciencias Antropológicas
Doctorando en Ciencias de la Computación

DEPARTAMENTO DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Incertidumbre

Sorpresa

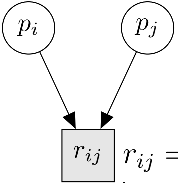
Sorpresa
El punto débil de las creencias

¿Cómo estimar habilidades?

Arpad Elo

Modelo Elo

Rendimiento aleatorio oculto (p) centrado en la habilidad estimada (s)

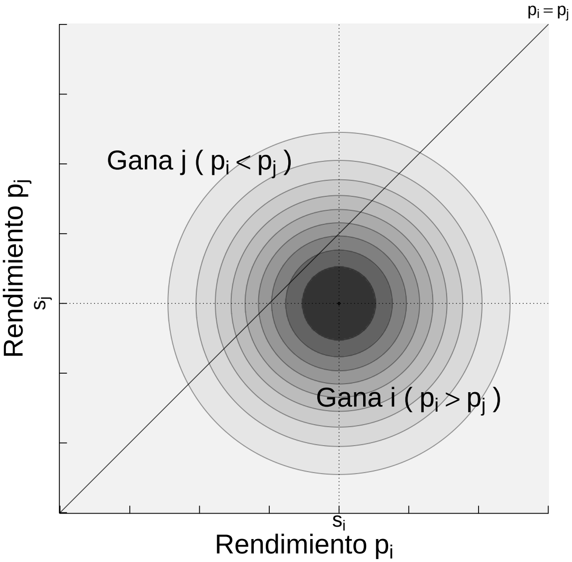


$$p \sim N(s, \beta^2)$$

Resultado observado (r)

$$r_{ij} = \mathbb{I}(p_i > p_j)$$

Probabilidad de ganar (gráfica)



Estimación Elo

$$s_i^{\text{new}} = s_i^{\text{old}} + K\Delta$$

Estimación Elo

$$s_i^{\text{new}} = s_i^{\text{old}} + K\Delta$$

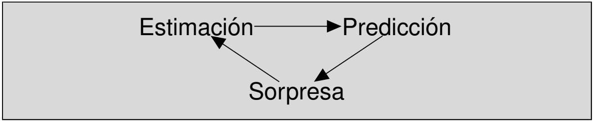
$$\Delta = \underbrace{(2\, r_{ij} - 1)}_{\substack{\textbf{Signo} \\ \text{del resultado}}} \underbrace{(1 - P(r_{ij}|s_i, s_j))}_{\substack{\textbf{Sorpresa} \\ \text{del resultado}}}$$

Estimación Elo

$$s_i^{\text{new}} = s_i^{\text{old}} + K\Delta$$

$$\Delta = \underbrace{(2\, r_{ij} - 1)}_{\substack{\textbf{Signo} \\ \text{del resultado}}} \; \underbrace{(1 - P(r_{ij}|s_i, s_j))}_{\substack{\textbf{Sorpresa} \\ \text{del resultado}}}$$

Modelo de solución



Hoy
Estimación de habilidad en la industria del video juego

Hoy

Estimación de habilidad en la industria del video juego

Inferencia Bayesiana

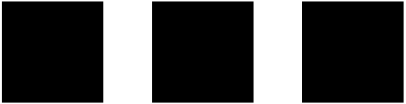
¿Por qué inferencia Bayesiana?

¿Por qué inferencia Bayesiana?

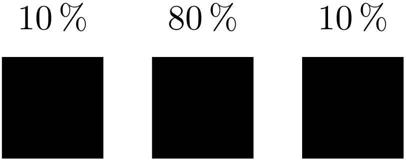
Permite computar creencias óptimas
dadas restricciones: modelos y datos

Books: Jaynes. Probability theor, the logic of science – Bishop. Pattern recognition and machine learning

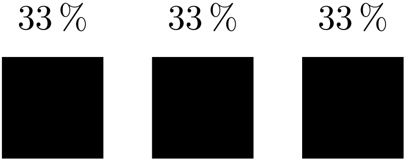
Principio de máxima incertidumbre



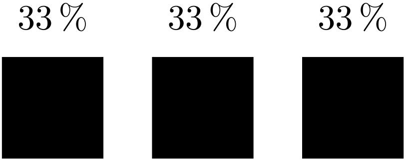
Principio de máxima incertidumbre



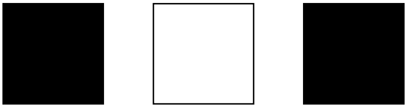
Principio de máxima incertidumbre



Principio de máxima incertidumbre
Honestidad

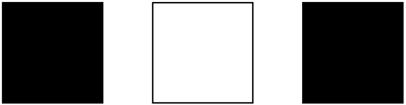


Principio de máxima incertidumbre
Honestidad



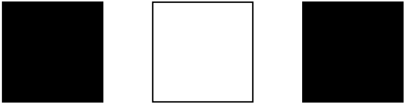
Principio de máxima incertidumbre
Honestidad

$0\,\%$

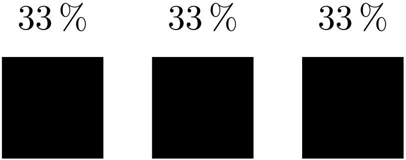


Principio de máxima incertidumbre
Honestidad

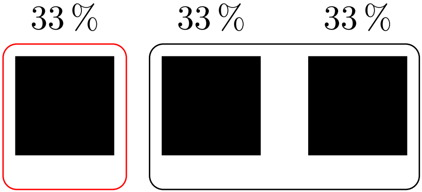
$50\,\%$ $\quad$ $0\,\%$ $\quad$ $50\,\%$



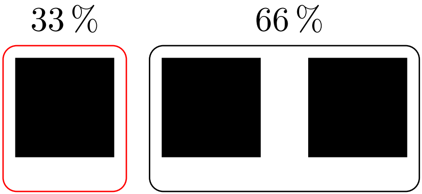
Principio de máxima incertidumbre
Honestidad



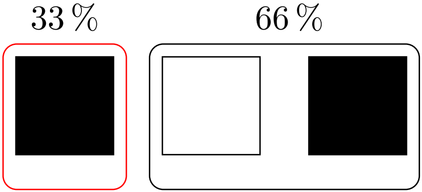
Principio de máxima incertidumbre
Honestidad



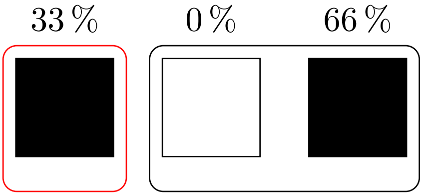
Principio de máxima incertidumbre
Honestidad



Principio de máxima incertidumbre
Honestidad



Principio de máxima incertidumbre
Honestidad



La sorpresa fuente de información
Entropía

La sorpresa fuente de información
Entropía

Información extraida de la sorpresa,

$$h(x) = -\log P(x)$$

La sorpresa fuente de información
Entropía

Información extraida de la sorpresa,

$$h(x) = -\log P(x)$$

Información esperada,

$$H(X) = \sum_x P(x)h(x)$$

La sorpresa fuente de información
Entropía

Información extraida de la sorpresa,

$$h(x) = -\log P(x)$$

Información esperada,

$$H(X) = \sum_{x} P(x) h(x)$$

Honestidad: máxima entropía
Máxima información esperada $\Leftrightarrow$ Máxima incertidumbre

La sorpresa fuente de información
Entropía

Información extraida de la sorpresa,

$$h(x) = -\log P(x)$$

Información esperada,

$$H(X) = \sum_x P(x)h(x)$$

Honestidad: máxima entropía
Máxima información esperada $\Leftrightarrow$ Máxima incertidumbre

Shannon, 1948. A mathematical theory of communication – Jaynes, 1957. Information theory and statistical mechanics

Máxima entropía
ante estados igualmente posibles

Máxima entropía
ante estados igualmente posibles

$$\text{Creencia}(A) = \frac{\text{Cantidad de estados en que } A \text{ es verdadera}}{\text{Cantidad de estados totales}}$$

Máxima entropía
ante estados no igualmente posibles

$$\begin{array}{c|c|c|c|c}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{y_j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-4}
 & \quad & \quad & \quad & \\ \cline{2-4}
x_i & & n_{ij} & & f_i \\ \cline{2-4}
 & & & & \\ \cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{c_j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{N}
\end{array}$$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
 & & y_j & & \\
\cline{2-4}
 & & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
 & & c_j & & N
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) =$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{y_j} & \multicolumn{1}{c}{} & \\
\cline{2-4}
 & & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{c_j} & \multicolumn{1}{c}{} & N
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
 & & y_j & & \\
\hline
 & & & & \\
\hline
x_i & & n_{ij} & & f_i \\
\hline
 & & & & \\
\hline
 & & c_j & & N
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) =$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
\multicolumn{1}{c}{} & \multicolumn{3}{c}{y_j} & \multicolumn{1}{c}{} \\ \cline{2-4}
 & \quad & \quad & \quad & \\ \cline{2-4}
x_i & & n_{ij} & & f_i \\ \cline{2-4}
 & & & & \\ \cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{3}{c}{c_j} & \multicolumn{1}{c}{N}
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N}$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
 & & y_j & & \\
\cline{2-4}
 & \phantom{n_{ij}} & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{c_j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{N}
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \frac{\sum_j n_{ij}}{N}$

Máxima entropía
ante estados no igualmente posibles

		y_j		
x_i		n_{ij}		f_i
		c_j		N

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \sum_j C(X = x_i, Y = y_j)$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
 & & y_j & & \\
\cline{2-4}
 & & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
 & & c_j & & N
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \sum_j C(X = x_i, Y = y_j)$

Creencia condicional: $C(Y = y_j | X = x_i) =$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{y_j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\
\cline{2-4}
 & & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{c_j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{N}
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \sum_j C(X = x_i, Y = y_j)$

Creencia condicional: $C(Y = y_j | X = x_i) = \frac{n_{ij}}{f_i}$

Máxima entropía
ante estados no igualmente posibles

		y_j		
x_i		n_{ij}		f_i
		c_j		N

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \sum_j C(X = x_i, Y = y_j)$

Creencia condicional: $C(Y = y_j | X = x_i) = \frac{n_{ij}}{N} \frac{N}{f_i}$

Máxima entropía
ante estados no igualmente posibles

$$
\begin{array}{c|c|c|c|c}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{y_j} & \multicolumn{1}{c}{} & \\
\cline{2-4}
 & & & & \\
\cline{2-4}
x_i & & n_{ij} & & f_i \\
\cline{2-4}
 & & & & \\
\cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{c_j} & \multicolumn{1}{c}{} & N
\end{array}
$$

Creencia conjunta: $C(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$

Creencia marginal: $C(X = x_i) = \frac{f_i}{N} = \sum_j C(X = x_i, Y = y_j)$

Creencia condicional: $C(Y = y_j | X = x_i) = \frac{n_{ij}}{N} \frac{N}{f_i} = \frac{C(X=x_i, Y=y_j)}{C(X=x_i)}$

Las reglas de la probabilidad

$$\text{Marginal}_i = \sum_j \text{Conjunta}_{ij} \qquad\qquad \text{Condicional}_{j|i} = \frac{\text{Conjunta}_{ij}}{\text{Marginal}_i}$$

Las reglas de la probabilidad

$$\text{Marginal}_i = \sum_j \text{Conjunta}_{ij} \qquad \text{Condicional}_{j|i} = \frac{\text{Conjunta}_{ij}}{\text{Marginal}_i}$$

Regla de la suma

$$P(X) = \sum_Y P(X,Y)$$

Cualquier distribución marginal puede ser obtenida integrando la distribución conjunta

Regla del producto

$$P(X,Y) = P(Y|X)P(X)$$

Cualquier distribución conjunta puede ser expresada como el producto de distribuciones condicionales unidimensionles.

Las reglas de la probabilidad

Teorema de Cox
Son las reglas del razonamiento con incertidumbre

Las únicas que grantizan:

- Representación de las creencias con valores reales
- Consistencia, cualquier camino lleva a la misma conclusión
- Actualización de las creencias en la dirección de la evidencia

Las reglas de la probabilidad

Teorema de Cox
Son las reglas del razonamiento con incertidumbre

Las únicas que grantizan:

- Representación de las creencias con valores reales
- Consistencia, cualquier camino lleva a la misma conclusión
- Actualización de las creencias en la dirección de la evidencia

Cox, R (1946). Probability, frequency and reasonable expectation.

Teorema de Bayes

$$P(X,Y) = P(Y,X)$$

Teorema de Bayes

$$P(Y|X)P(X) = P(X,Y) = P(Y,X) = P(X|Y)P(Y)$$

Teorema de Bayes

$$P(Y|X)P(X) = P(X|Y)P(Y)$$

Teorema de Bayes

$$P(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

Teorema de Bayes

$$P(\textsf{Hipótesis} \mid \textsf{Datos}) = \frac{P(\textsf{Datos} \mid \textsf{Hipótesis})P(\textsf{Hipótesis})}{P(\textsf{Datos})}$$

Teorema de Bayes

$$\underbrace{P(\textsf{Hipótesis} \mid \textsf{Datos})}_{\textsf{Posteriori}} = \frac{\overbrace{P(\textsf{Datos} \mid \textsf{Hipótesis})}^{\textsf{Verosimilitud}}\,\overbrace{P(\textsf{Hipótesis})}^{\textsf{Priori}}}{\underbrace{P(\textsf{Datos})}_{\textsf{Evidencia}}}$$

Teorema de Bayes

$$\underbrace{P(\text{Hipótesis} \mid \text{Datos, Modelo})}_{\text{Posteriori}} = \frac{\overbrace{P(\text{Datos} \mid \text{Hipótesis, Modelo})}^{\text{Verosimilitud}}\overbrace{P(\text{Hipótesis} \mid \text{Modelo})}^{\text{Priori}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

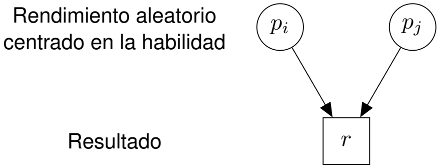
Teorema de Bayes

$$\underbrace{P(\text{Hipótesis} \mid \text{Datos, Modelo})}_{\text{Posteriori}} = \frac{\overbrace{P(\text{Datos} \mid \text{Hipótesis, Modelo})}^{\text{Verosimilitud}}\,\overbrace{P(\text{Hipótesis} \mid \text{Modelo})}^{\text{Priori}}}{\underbrace{P(\text{Datos} \mid \text{Modelo})}_{\text{Evidencia}}}$$

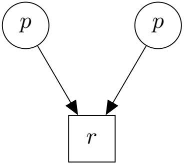
El **modelo** es lo que permite relacionar
los **datos** con nuestras **hipótesis**!

Modelos gráficos

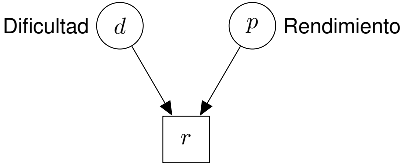
Modelos gráficos



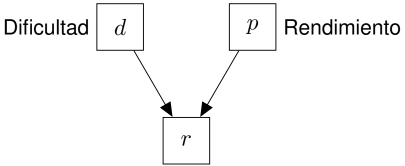
Modelos gráficos



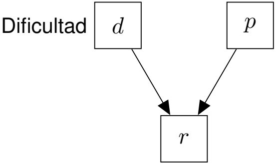
Modelos gráficos



Modelos gráficos



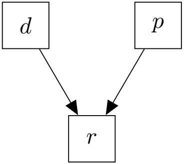
Modelos gráficos



p^0	p^1
0,70	0,30

Modelos gráficos

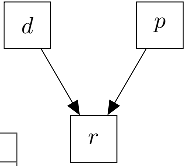
d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

Modelos gráficos

d^0	d^1
0,15	0,85

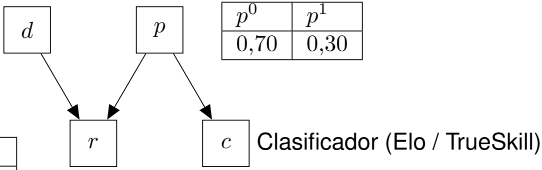


p^0	p^1
0,70	0,30

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

Modelos gráficos

d^0	d^1
0,15	0,85

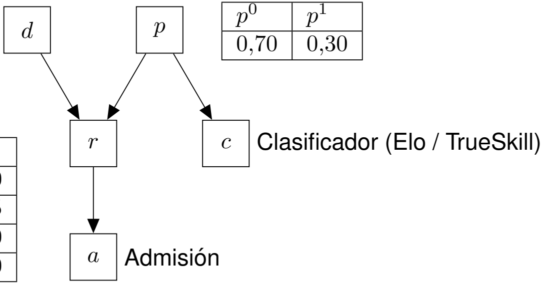


p^0	p^1
0,70	0,30

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

Modelos gráficos

d^0	d^1
0,15	0,85

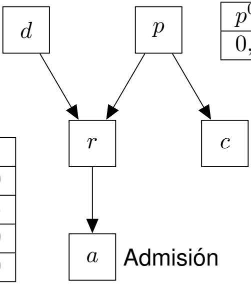


p^0	p^1
0,70	0,30

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

Modelos gráficos

d^0	d^1
0,15	0,85



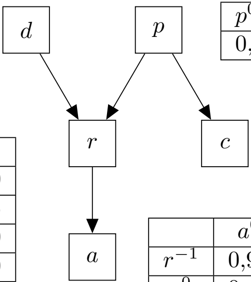
p^0	p^1
0,70	0,30

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

Modelos gráficos

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

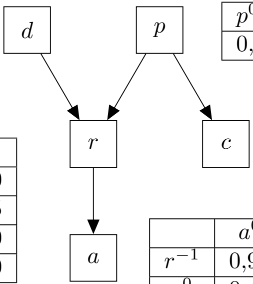
	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

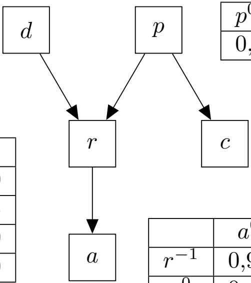
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

: $P(a|r)$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(p)$:

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$P(c|p)$:

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$P(a|r)$:

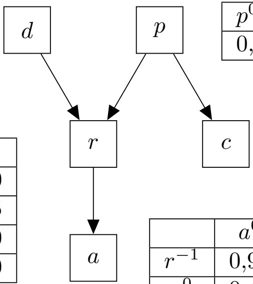
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

Los factores son útiles para definir distribuciones de probabilidad en espacios de alta dimensión.

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(p)$:

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$P(c|p)$:

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$P(a|r)$:

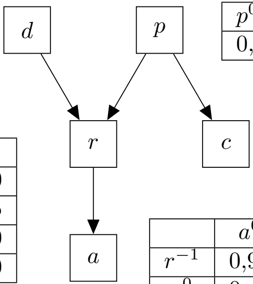
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(d, p, r, c, a) =$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

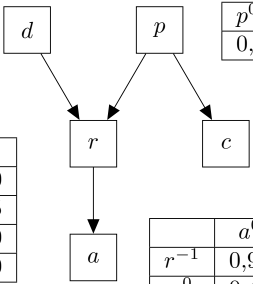
: $P(a|r)$

$$P(d,p,r,c,a) = P(d)P(p|d)P(r|d,p)P(c|d,p,r)P(a|d,p,r,c)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

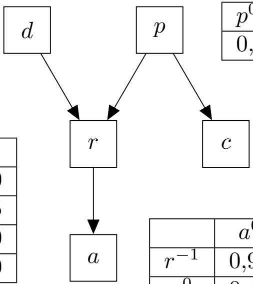
: $P(a|r)$

$$P(d,p,r,c,a) = P(d)P(p|\cancel{d})P(r|d,p)P(c|\cancel{d},p,\cancel{r})P(a|\cancel{d},\cancel{p},r,\cancel{c})$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(p)$:

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$P(c|p)$:

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$P(a|r)$:

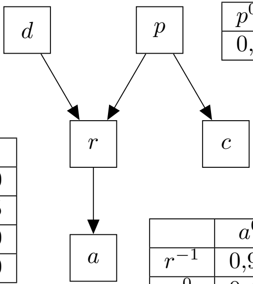
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(d,p,r,c,a) = P(d)P(p)P(r|d,p)P(a|r)P(c|p)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$: P(p)$

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$: P(c|p)$

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$: P(a|r)$

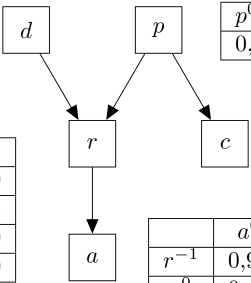
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(d^1, p^1, r^1, c^1, a^1) = P(d^1)P(p^1)P(r^1|d^1, p^1)P(a^1|r^1)P(c^1|p^1)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

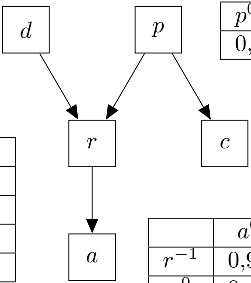
: $P(a|r)$

$$P(d^1, p^1, r^1, c^1, a^1) = 0.85 \cdot P(p^1)P(r^1|d^1, p^1)P(a^1|r^1)P(c^1|p^1)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(p)$:

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$P(c|p)$:

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$P(a|r)$:

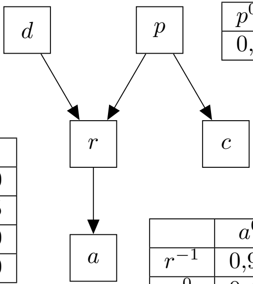
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(d^1, p^1, r^1, c^1, a^1) = 0.85 \cdot 0.30 \cdot 0.50 \cdot 0.80 \cdot 0.90$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



: $P(p)$

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0,d^0)	0,30	0,40	0,30
(p^0,d^1)	0,70	0,25	0,05
(p^1,d^0)	0,02	0,08	0,90
(p^1,d^1)	0,20	0,30	0,50

: $P(c|p)$

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(a|r)$

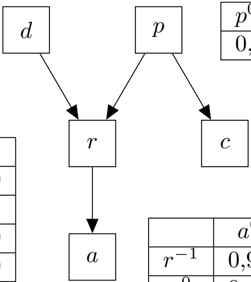
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(d^1,p^1,r^1,c^1,a^1) = 0.85 \cdot 0.30 \cdot 0.50 \cdot 0.80 \cdot 0.90 \approx 0.09$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0,d^0)	0,30	0,40	0,30
(p^0,d^1)	0,70	0,25	0,05
(p^1,d^0)	0,02	0,08	0,90
(p^1,d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

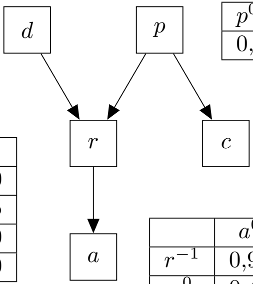
: $P(a|r)$

$$P(r^1) =$$

Modelos gráficos

$P(d):$

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

$: P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$: P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

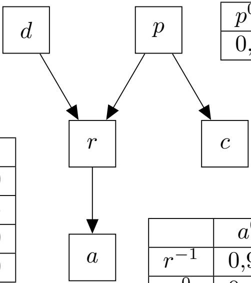
$: P(a|r)$

$$P(r^1) = \sum_i \sum_j \sum_l \sum_m P(d^i, p^j, r^1, c^l, a^m)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

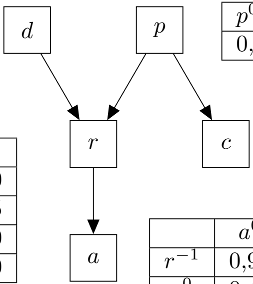
: $P(a|r)$

$$P(r^1) = \sum_{i,j,l,m} P(d^i, p^j, r^1, c^l, a^m)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0,d^0)	0,30	0,40	0,30
(p^0,d^1)	0,70	0,25	0,05
(p^1,d^0)	0,02	0,08	0,90
(p^1,d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

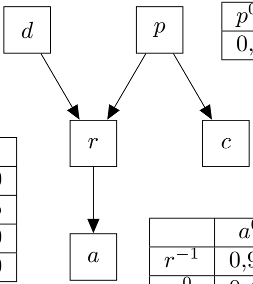
: $P(a|r)$

$$P(r^1) = \sum_{i,j,l,m} P(d^i)P(p^j)P(r^1|d^i,p^j)P(c^l|p^j)P(a^m|r^1)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

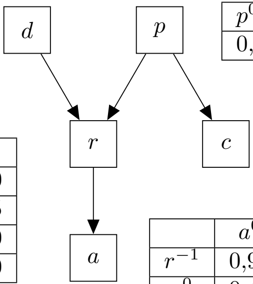
: $P(a|r)$

$$P(r^1) = \left(\sum_m P(a^m|r^1) \right) \left(\sum_{i,j,l} P(d^i)P(p^j)P(r^1|d^i,p^j)P(c^l|p^j) \right)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$: P(p)$

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$: P(c|p)$

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$: P(a|r)$

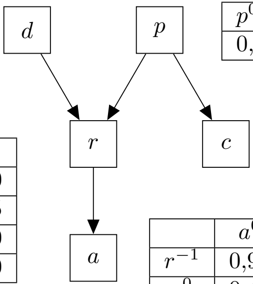
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(r^1) = \left(\sum_m P(a^m|r^1)\right)\left(\sum_{i,j} P(d^i)P(p^j)P(r^1|d^i,p^j)\Big(\sum_l P(c^l|p^j)\Big)\right)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(p)$:

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$P(c|p)$:

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$P(a|r)$:

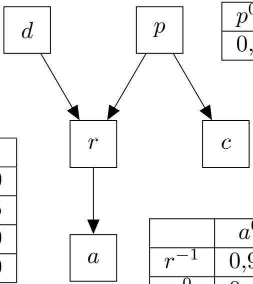
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(r^1) = \left(\cancel{\sum_m P(a^m|r^1)} \right) \left(\sum_{i,j} P(d^i)P(p^j)P(r^1|d^i,p^j) \Big(\sum_l P(c^l|p^j) \Big) \right)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$: P(p)$

p^0	p^1
0,70	0,30

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

$: P(c|p)$

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$: P(a|r)$

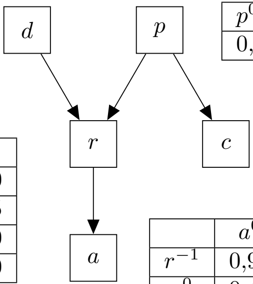
	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$$P(r^1) = \left(\cancel{\sum_m P(a^m | r^1)} \right) \left(\sum_{i,j} P(d^i) P(p^j) P(r^1 | d^i, p^j) \Big(\cancel{\sum_l P(c^l | p^j)} \Big) \right)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



p^0	p^1
0,70	0,30

: $P(p)$

$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

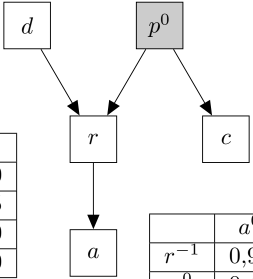
: $P(a|r)$

$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i, p^j)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

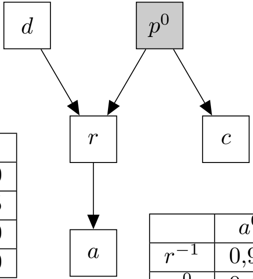
: $P(a|r)$

$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i, p^j) \approx 0.23$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

: $P(a|r)$

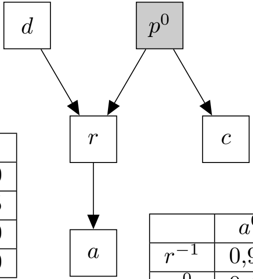
$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i, p^j) \approx 0.23$$

$$P(r^1|p^0) =$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

: $P(a|r)$

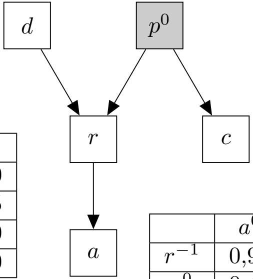
$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i, p^j) \approx 0.23$$

$$P(r^1|p^0) = \frac{P(r^1, p^0)}{P(p^0)}$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

: $P(a|r)$

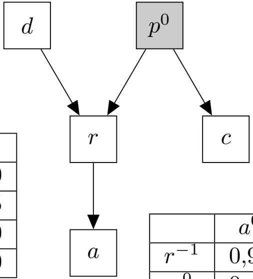
$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i,p^j) \approx 0.23$$

$$P(r^1|p^0) = \frac{1}{P(p^0)} \sum_{i,l,m} P(d^i)P(p^0)P(r^1|d^i,p^0)P(c^l|p^0)P(a^m|r^k)$$

Modelos gráficos

$P(d)$:

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0,d^0)	0,30	0,40	0,30
(p^0,d^1)	0,70	0,25	0,05
(p^1,d^0)	0,02	0,08	0,90
(p^1,d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

: $P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

: $P(a|r)$

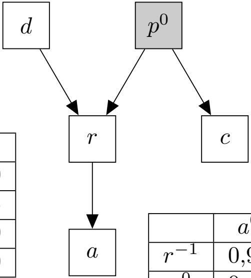
$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i,p^j) \approx 0.23$$

$$P(r^1|p^0) = \frac{1}{P(p^0)} \sum_i P(d^i)P(p^0)P(r^1|d^i,p^0)$$

Modelos gráficos

$P(d):$

d^0	d^1
0,15	0,85



$P(r|d,p)$

	r^{-1}	r^0	r^1
(p^0, d^0)	0,30	0,40	0,30
(p^0, d^1)	0,70	0,25	0,05
(p^1, d^0)	0,02	0,08	0,90
(p^1, d^1)	0,20	0,30	0,50

	c^0	c^1
p^0	0,95	0,05
p^1	0,2	0,8

$: P(c|p)$

	a^0	a^1
r^{-1}	0,99	0,01
r^0	0,40	0,60
r^1	0,10	0,90

$: P(a|r)$

$$P(r^1) = \sum_{i,j} P(d^i)P(p^j)P(r^1|d^i,p^j) \approx 0.23$$

$$P(r^1|p^0) = \frac{1}{P(p^0)} \sum_{i} P(d^i)P(p^0)P(r^1|d^i,p^0) \approx 0.09$$

Flujos de inferencia

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$		
$X \leftarrow V \leftarrow Y$		
$X \leftarrow V \rightarrow Y$		
$X \rightarrow V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \to V \to Y$	Sí	
$X \leftarrow V \leftarrow Y$		
$X \leftarrow V \to Y$		
$X \to V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	
$X \leftarrow V \leftarrow Y$	Sí	
$X \leftarrow V \rightarrow Y$		
$X \rightarrow V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	
$X \leftarrow V \leftarrow Y$	Sí	
$X \leftarrow V \rightarrow Y$	Sí	
$X \rightarrow V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \to V \to Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	
$X \leftarrow V \to Y$	Sí	
$X \to V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \to V \to Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \to Y$	Sí	
$X \to V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \rightarrow Y$	Sí	No
$X \rightarrow V \leftarrow Y$		

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \rightarrow Y$	Sí	No
$X \rightarrow V \leftarrow Y$	No	

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \to V \to Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \to Y$	Sí	No
$X \to V \leftarrow Y$	No	Sí

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \to V \to Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \to Y$	Sí	No
$X \to V \leftarrow Y$	No	$\underbrace{\text{Sí}}_{\text{O algún descendiente observable}}$

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \rightarrow Y$	Sí	No
$X \rightarrow V \leftarrow Y$	$\underbrace{\text{No}}_{\text{Y ningún descendiente observable}}$	$\underbrace{\text{Sí}}_{\text{O algún descendiente observable}}$

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \rightarrow Y$	Sí	No
$X \rightarrow V \leftarrow Y$	$\underbrace{\text{No}}_{\text{Y ningún descendiente observable}}$	$\underbrace{\text{Sí}}_{\text{O algún descendiente observable}}$

Un flujo de inferencia permenece abierto si:

- Toda consecuencia común (o alguno de sus descendientes) es observable
- Ningúna otra variable es observable

Flujos de inferencia

	$V \notin$ Observable	$V \in$ Observable
$X \rightarrow V \rightarrow Y$	Sí	No
$X \leftarrow V \leftarrow Y$	Sí	No
$X \leftarrow V \rightarrow Y$	Sí	No
$X \rightarrow V \leftarrow Y$	$\underbrace{\text{No}}_{\substack{\text{Y ningún descendiente} \\ \text{observable}}}$	$\underbrace{\text{Sí}}_{\substack{\text{O algún descendiente} \\ \text{observable}}}$

Un flujo de inferencia permenece abierto si:

- Toda consecuencia común (o alguno de sus descendientes) es observable
- Ningúna otra variable es observable

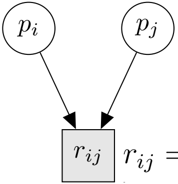
Pearl, J. 2018. The Book of Why: The New Science of Cause and Effect

TrueSkill

TrueSkill

Rendimiento aleatorio oculto (p)
centrado en la habilidad estimada (s)

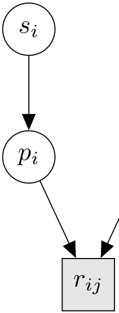
Resultado observado (r)



$$p \sim N(s, \beta^2)$$

$$r_{ij} = \mathbb{I}(p_i > p_j)$$

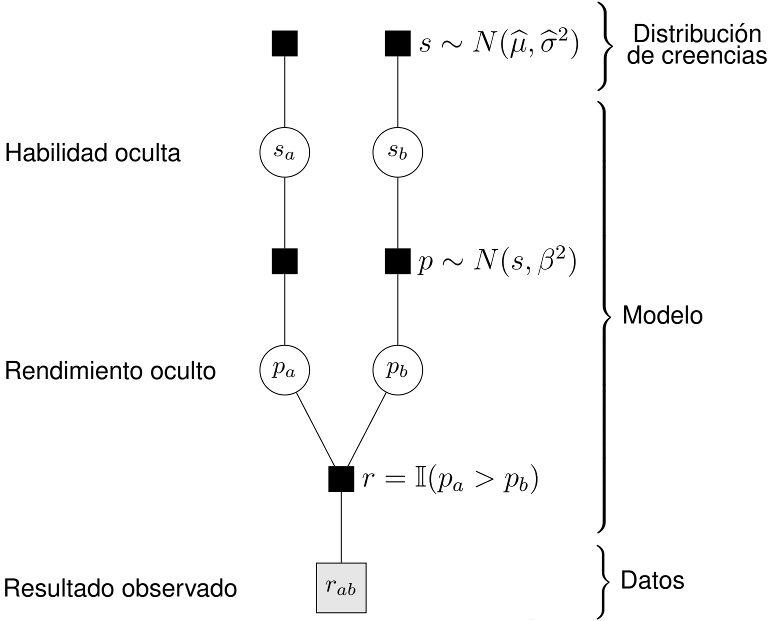
TrueSkill



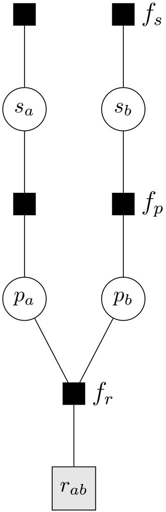
Distribución de creencias sobre la habilidad oculta — $s \sim N(\mu, \sigma^2)$

Rendimiento aleatorio oculto (p) centrado en la habilidad (s) — $p \sim N(s, \beta^2)$

Resultado observado (r) — $r_{ij} = \mathbb{I}(p_i > p_j)$


$s \sim N(\widehat{\mu}, \widehat{\sigma}^2)$
Distribución de creencias
Habilidad oculta
s_a
s_b
$p \sim N(s, \beta^2)$
Modelo
Rendimiento oculto
p_a
p_b
$r = \mathbb{I}(p_a > p_b)$
Resultado observado
r_{ab}
Datos

Sum-product algorithm



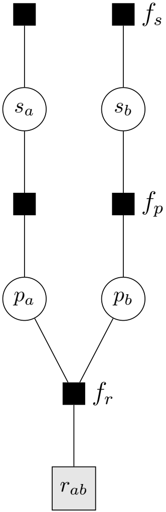
$$P(x) = \prod_{h \in n(x)} m_{h \to x}$$

$m_{x \to f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \to x}(x)$ **:** Mensaje factor f a variable x

$n(v)$ **:** Conjunto de nodos vecinos del nodo v

Sum-product algorithm



$$P(x) = \prod_{h \in n(x)} m_{h \rightarrow x}$$

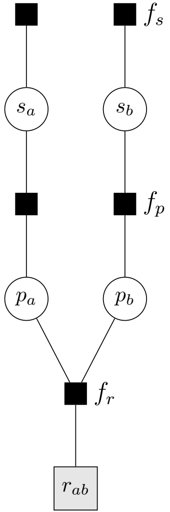
$m_{x \rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \rightarrow x}(x)$ **:** Mensaje factor f a variable x

$n(v)$ **:** Conjunto de nodos vecinos del nodo v

$$m_{x \rightarrow f}(x) = \prod_{h \in n(x) \setminus \{f\}} m_{h \rightarrow x}(x)$$

Sum-product algorithm



$$P(x) = \prod_{h \in n(x)} m_{h \rightarrow x}$$

$m_{x \rightarrow f}(x)$ **:** Mensaje de variable x a factor f

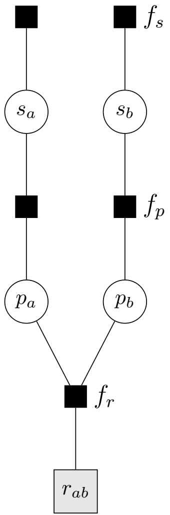
$m_{f \rightarrow x}(x)$ **:** Mensaje factor f a variable x

$n(v)$ **:** Conjunto de nodos vecinos del nodo v

$$m_{x \rightarrow f}(x) = \prod_{h \in n(x) \setminus \{f\}} m_{h \rightarrow x}(x)$$

$$m_{f \rightarrow x}(x) = \sum_{X \setminus \{x\}} \left(f(X) \prod_{h \in n(f) \setminus \{x\}} m_{h \rightarrow f}(h) \right)$$

Sum-product algorithm



$$P(x) = \prod_{h \in n(x)} m_{h \to x}$$

$m_{x \to f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \to x}(x)$ **:** Mensaje factor f a variable x

$n(v)$ **:** Conjunto de nodos vecinos del nodo v

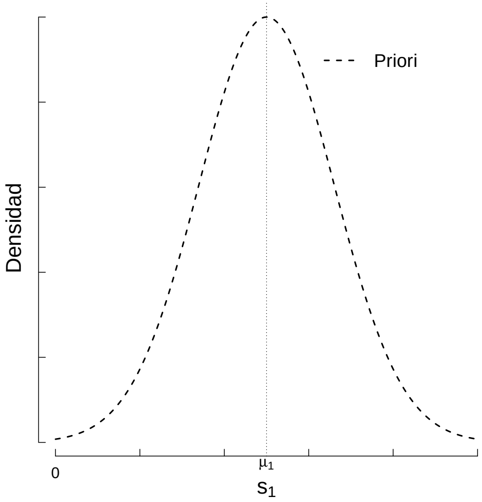
$$m_{x \to f}(x) = \prod_{h \in n(x) \setminus \{f\}} m_{h \to x}(x)$$

$$m_{f \to x}(x) = \sum_{X \setminus \{x\}} \Big(f(X) \prod_{h \in n(f) \setminus \{x\}} m_{h \to f}(h) \Big)$$

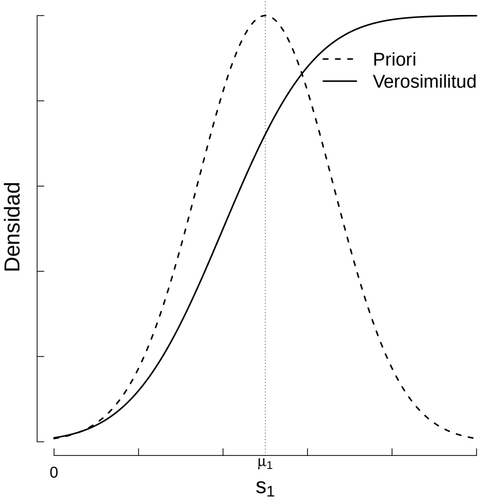
Kschischang FR, Frey BJ, Loeliger HA. Factor graphs and the sum-product algorithm. 2001

$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|\, s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$

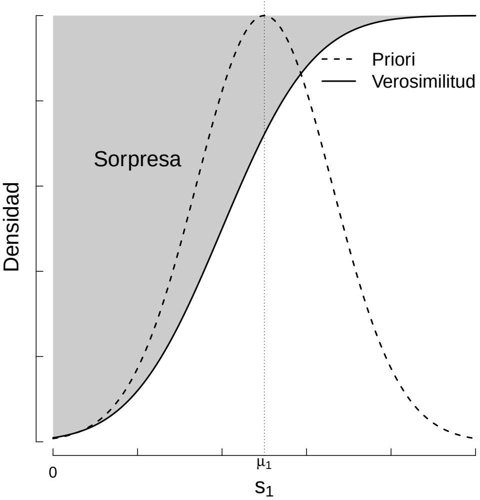
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|\, s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



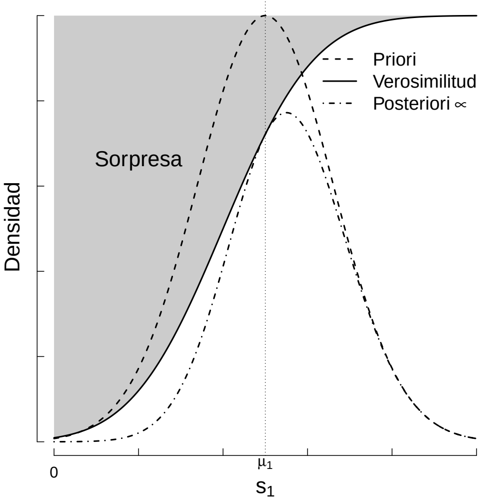
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



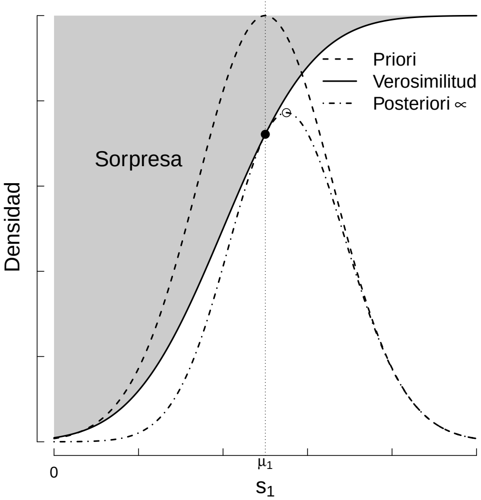
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|\, s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



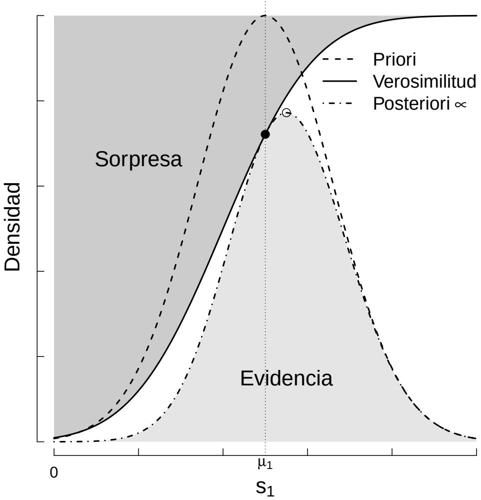
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



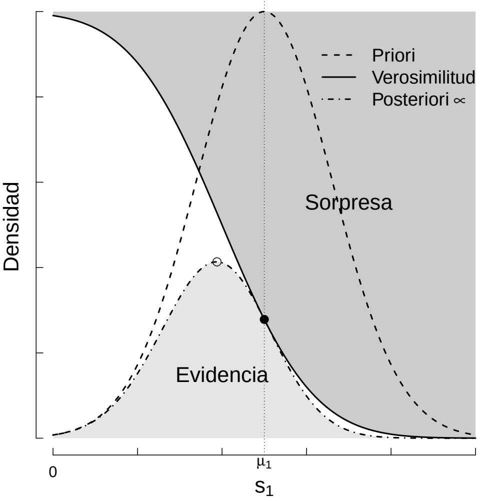
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \ \overbrace{1 - \Phi(0 \,|\, s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



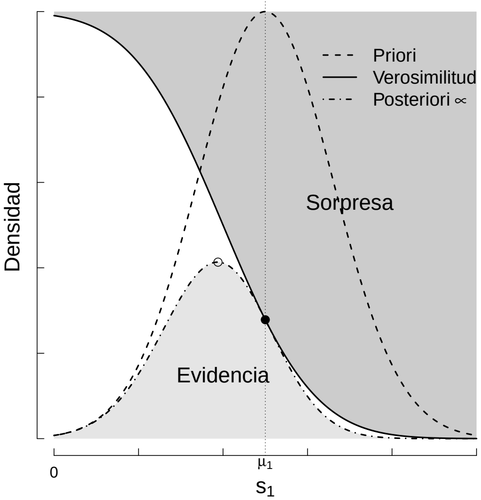
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Priori}} \; \overbrace{1 - \Phi(0 \,|\, s_1 - \mu_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



$$\overbrace{P(s_2 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_2 \,|\, \mu_2, \sigma_2^2)}^{\text{Priori}} \; \overbrace{\Phi(0 \,|\mu_1 - s_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso perdedor}$$



$$\overbrace{P(s_2 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_2 \,|\, \mu_2, \sigma_2^2)}^{\text{Priori}} \overbrace{\Phi(0 \,|\mu_1 - s_2, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso perdedor}$$



Todos los detalles en: Landfried. TrueSkill: Technical Report. 2019

Selección de modelo

¿Y nuestras creencias respecto de modelos alternativos?

Selección de modelo

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

Selección de modelo

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

$$\frac{P(\mathsf{M}_i|\mathsf{D})}{P(\mathsf{M}_j|\mathsf{D})} = \frac{P(\mathsf{D}|\mathsf{M}_i)\ P(\mathsf{M}_i)}{P(\mathsf{D}|\mathsf{M}_j)\ P(\mathsf{M}_j)}$$

Selección de modelo

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

$$\frac{P(\mathsf{M}_i|\mathsf{D})}{P(\mathsf{M}_j|\mathsf{D})} = \frac{P(\mathsf{D}|\mathsf{M}_i)\ P(\mathsf{M}_i)}{\underbrace{P(\mathsf{D}|\mathsf{M}_j)}_{\text{Evidencia!}}\ P(\mathsf{M}_j)}$$

Selección de modelo

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

$$\frac{P(\mathsf{M}_i|\mathsf{D})}{P(\mathsf{M}_j|\mathsf{D})} = \frac{P(\mathsf{D}|\mathsf{M}_i)\ \cancel{P(\mathsf{M}_i)}}{\underbrace{P(\mathsf{D}|\mathsf{M}_j)}_{\text{Evidencia!}}\ \cancel{P(\mathsf{M}_j)}}$$

Selección de modelo

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

$$\frac{P(\mathsf{M}_i|\mathsf{D})}{P(\mathsf{M}_j|\mathsf{D})} = \frac{P(\mathsf{D}|\mathsf{M}_i)\ \cancel{P(\mathsf{M}_i)}}{\underbrace{P(\mathsf{D}|\mathsf{M}_j)}_{\text{Evidencia!}}\ \cancel{P(\mathsf{M}_j)}}$$

Preferimos modelos con la menor **sorpresa** en la evidencia!

$$\boxed{P(M_q|D) > P(M_r|D) \iff P(D|M_q) > P(D|M_r)}$$

Selección de modelo

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

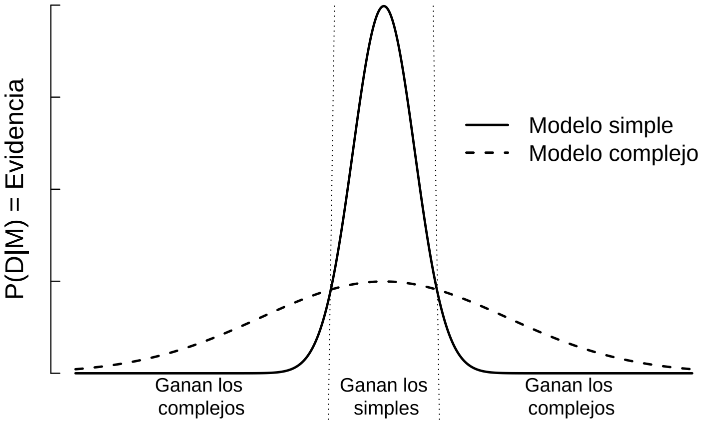
$$\frac{P(\mathsf{M}_i|\mathsf{D})}{P(\mathsf{M}_j|\mathsf{D})} = \frac{P(\mathsf{D}|\mathsf{M}_i)\ \cancel{P(\mathsf{M}_i)}}{\underbrace{P(\mathsf{D}|\mathsf{M}_j)}_{\text{Evidencia!}}\ \cancel{P(\mathsf{M}_j)}}$$

Preferimos modelos con la menor **sorpresa** en la evidencia!

$$\boxed{P(M_q|D) > P(M_r|D) \iff P(D|M_q) > P(D|M_r)}$$

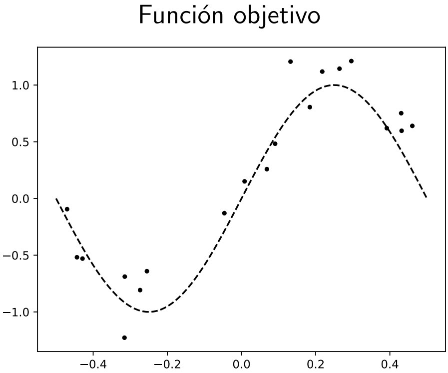
Sobre comparación de modelos ver: Kass & Raftery. Bayes factors. 1995.

Evidencia

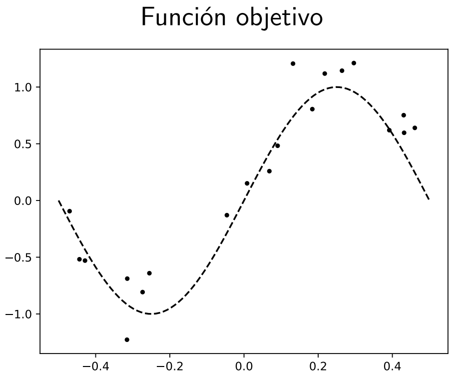


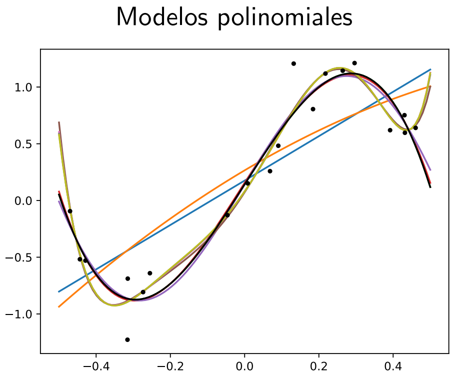
Balance natural entre complejidad y predicción

Regresión lineal Bayesiana

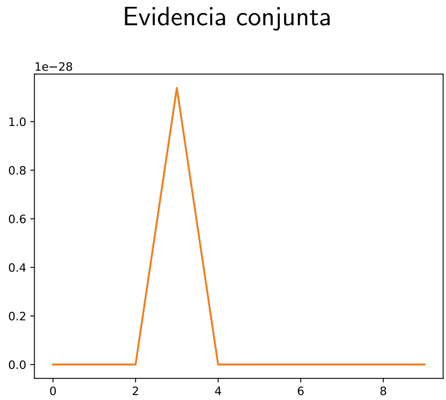


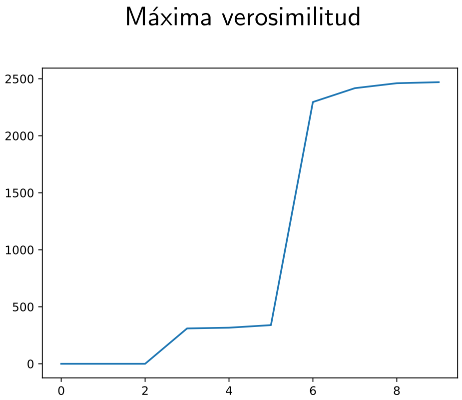
Regresión lineal Bayesiana





Evidencia vs Verosimilitud






quinti cancha