Modelos de historia completa

Unidad 7

Redes bayesianas de historia completa.

El problema de usar el posterior como prior del siguiente evento

El algoritmo de inferencia por loopy belief propagation.

Consideraciones de inferencia causal en series temporales.

Bibliografía Unidad 7

- Dangauthier P, Herbrich R, Minka T, Graepel T. *Trueskill through time: Revisiting the history of chess*. NeurIPS; 2008 (Descargar). (completo)
- Bishop, C. *Pattern recognition and machine learning*. Springer; 2006 (Descargar). (lecturas 13.2.3-13.2.4, 13.3)

Otros:

- Brodersen KH, Gallusser F, Koehler J, Remy N, Scott SL. *Inferring causal impact using Bayesian structural time-series models*. The Annals of Applied Statistics. 2015 (Paper).
- Hernán. *Causal inference: What if*. CRC Boca Raton, FL. 2020. (Parte III).

La función de costo epistémica

$$\underbrace{P(\text{Hipótesis}, \text{Datos})}_{\substack{\text{Creencia compatible}\\ \text{con los datos}}} = \underbrace{P(\text{Hipótesis})}_{\substack{\text{Acuerdo intersubjetivo}\\ \text{inicial}}} \underbrace{P(\text{dato}_1|\text{Hipótesis})}_{\text{Predicción 1}} \underbrace{P(\text{dato}_2|\text{dato}_1, \text{Hipótesis})}_{\text{Predicción 2}} \dots$$

Series de tiempo
Creencias adaptativas

La función de costo epistémica

$$\underbrace{P(\text{Hipótesis}, \text{Datos})}_{\substack{\text{Creencia compatible}\\ \text{con los datos}}} = \underbrace{P(\text{Hipótesis})}_{\substack{\text{Acuerdo intersubjetivo}\\ \text{inicial}}} \underbrace{P(\text{dato}_1|\text{Hipótesis})}_{\text{Predicción 1}} \underbrace{P(\text{dato}_2|\text{dato}_1, \text{Hipótesis})}_{\text{Predicción 2}} \dots$$

Un único 0 en la secuencia de predicciones
hace falsa la hipótesis para siempre.

Series de tiempo
Creencias adaptativas

La función de costo epistémica

$$\underbrace{P(\text{Hipótesis}, \text{Datos})}_{\substack{\text{Creencia compatible} \\ \text{con los datos}}} = \underbrace{P(\text{Hipótesis})}_{\substack{\text{Acuerdo intersubjetivo} \\ \text{inicial}}} \underbrace{P(\text{dato}_1|\text{Hipótesis})}_{\text{Predicción 1}} \underbrace{P(\text{dato}_2|\text{dato}_1, \text{Hipótesis})}_{\text{Predicción 2}} \dots$$

Ejemplo

Esa persona no está apta para realizar esa tarea.

Series de tiempo
Creencias adaptativas

La función de costo epistémica

$$\underbrace{P(\text{Hipótesis}, \text{Datos})}_{\substack{\text{Creencia compatible} \\ \text{con los datos}}} = \underbrace{P(\text{Hipótesis})}_{\substack{\text{Acuerdo intersubjetivo} \\ \text{inicial}}} \underbrace{P(\text{dato}_1|\text{Hipótesis})}_{\text{Predicción 1}} \underbrace{P(\text{dato}_2|\text{dato}_1, \text{Hipótesis})}_{\text{Predicción 2}} \dots$$

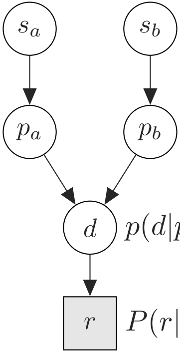
Ejemplo

Esa persona no está apta para realizar esa tarea.

¿Para siempre?!

Estimación de habilidad

en la industria del videojuego



Habilidad

$$p(s_i) = \mathcal{N}(s_i|\mu_i, \sigma_i^2)$$

Desempeño

$$p(p_i|s_i) = \mathcal{N}(p_i|s_i, \beta^2)$$

Diferencia

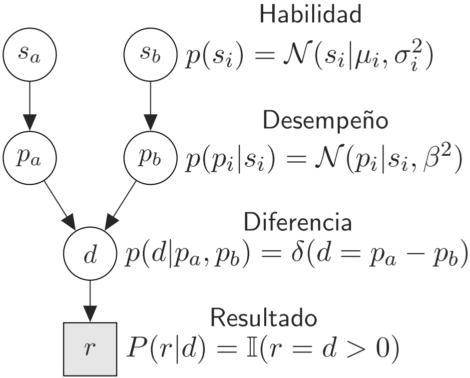
$$p(d|p_a, p_b) = \delta(d = p_a - p_b)$$

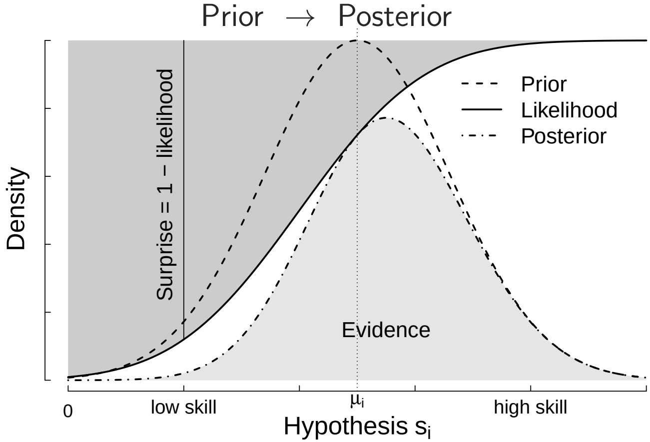
Resultado

$$P(r|d) = \mathbb{I}(r = d > 0)$$

Estimación de habilidad

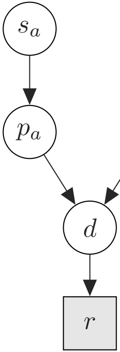
en la industria del videojuego





Estimación de habilidad

en la industria del videojuego



Habilidad

$p(s_i) = \mathcal{N}(s_i|\mu_i, \sigma_i^2)$

Desempeño

$p(p_i|s_i) = \mathcal{N}(p_i|s_i, \beta^2)$

Diferencia

$p(d|p_a, p_b) = \delta(d = p_a - p_b)$

Resultado

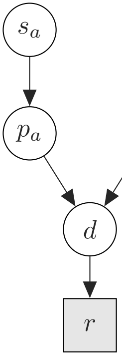
$P(r|d) = \mathbb{I}(r = d > 0)$

Prior $\rightarrow$ Posterior

¿Cómo estimamos una habilidad en el tiempo?

Estimación de habilidad

en la industria del videojuego



Habilidad

$p(s_i) = \mathcal{N}(s_i|\mu_i, \sigma_i^2)$

Desempeño

$p(p_i|s_i) = \mathcal{N}(p_i|s_i, \beta^2)$

Diferencia

$p(d|p_a, p_b) = \delta(d = p_a - p_b)$

Resultado

$P(r|d) = \mathbb{I}(r = d > 0)$

Prior $\rightarrow$ Posterior

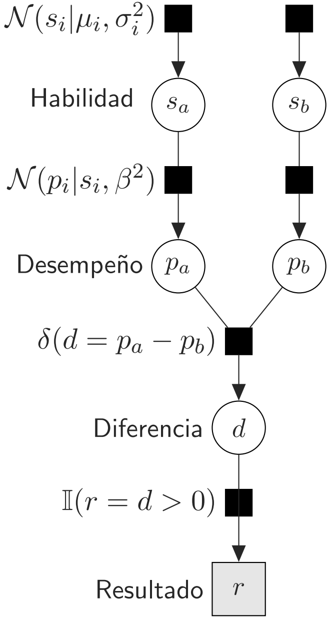
¿Cómo estimamos una habilidad en el tiempo?

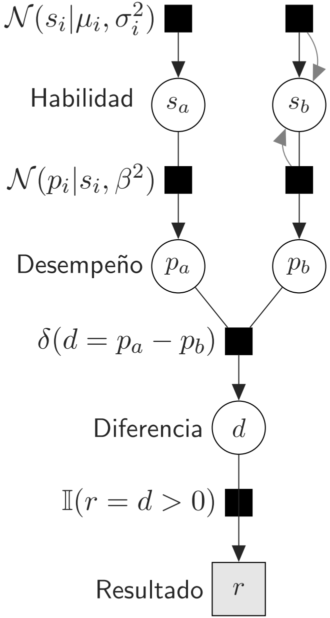
¿Si usamos el último posterior como prior del siguiente evento?

$\text{Posterior}_t \rightarrow \text{Prior}_{t+1}$

Algoritmo suma-producto

Las reglas de la probabilidad por pasaje de mensajes

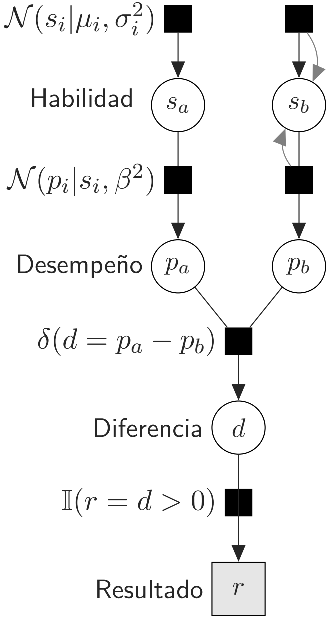




Algoritmo suma-producto

Las reglas de la probabilidad por pasaje de mensajes

$$p(s_b, r) = \underbrace{m_{f_{s_b}\rightarrow s_b}(s_b)}_{\text{Prior}} \cdot \underbrace{m_{f_{p_b}\rightarrow s_b}(s_b)}_{\text{Likelihood}}$$

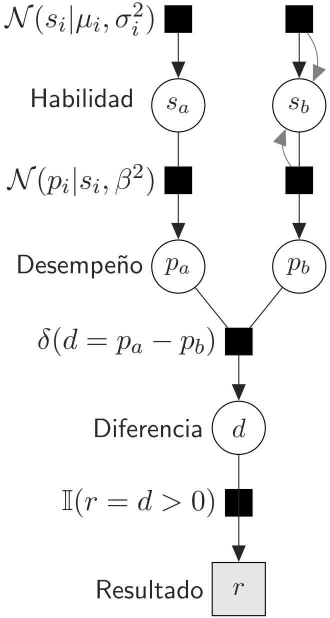


Algoritmo suma-producto

Las reglas de la probabilidad por pasaje de mensajes

$$p(s_b, r) = \underbrace{m_{f_{s_b}\rightarrow s_b}(s_b)}_{\text{Prior}} \cdot \underbrace{m_{f_{p_b}\rightarrow s_b}(s_b)}_{\text{Likelihood}}$$

$$\underbrace{m_{x\rightarrow f}(x)}_{\text{Mensaje de las \textbf{variables}}} = \underbrace{\prod_{g\in v(x)\setminus\{f\}} m_{g\rightarrow x}(x)}_{\text{El producto de lo que recibe de atrás}}$$



Algoritmo suma-producto

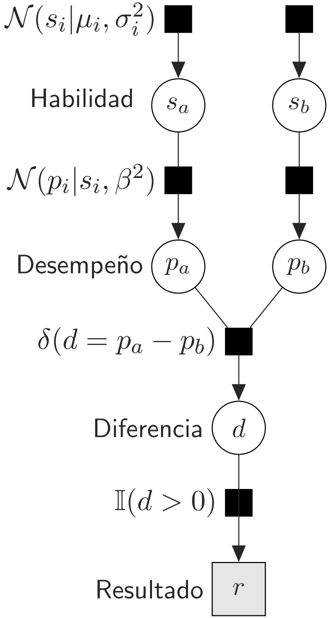
Las reglas de la probabilidad por pasaje de mensajes

$$p(s_b, r) = \underbrace{m_{f_{s_b}\to s_b}(s_b)}_{\text{Prior}} \cdot \underbrace{m_{f_{p_b}\to s_b}(s_b)}_{\text{Likelihood}}$$

$$\underbrace{m_{x\to f}(x)}_{\substack{\text{Mensaje de}\\ \text{las }\textbf{variables}}} = \underbrace{\prod_{g\in v(x)\setminus\{f\}} m_{g\to x}(x)}_{\substack{\text{El producto de}\\ \text{lo que recibe de atrás}}}$$

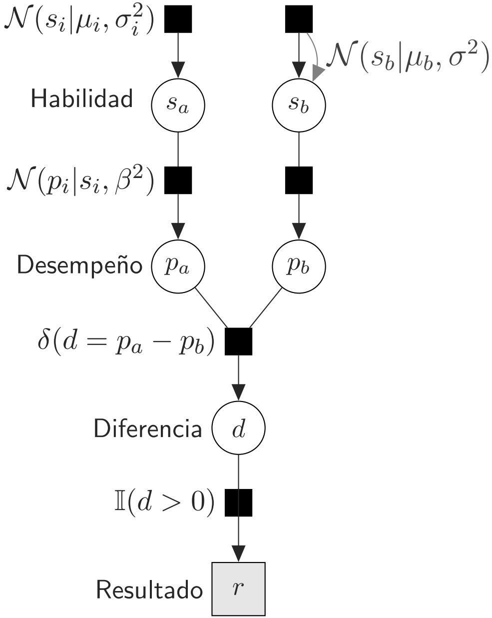
$$\underbrace{m_{f\to x}(x)}_{\substack{\text{Mensaje de}\\ \text{los }\textbf{factores}}} = \underbrace{\sum_{\boldsymbol{y}} \Big(f(\boldsymbol{y}, x) \prod_{y\in v(f)\setminus\{x\}} m_{y\to f}(y) \Big)}_{\substack{\text{Lo que recibe de atrás por el factor,}\\ \text{integrando todas las variables de atrás}}}$$

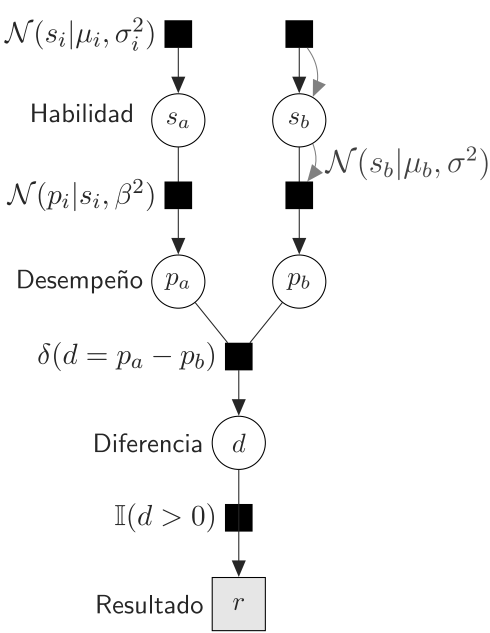
Estimación de habilidad
en la industria del videojuego



Estimación de habilidad

en la industria del videojuego



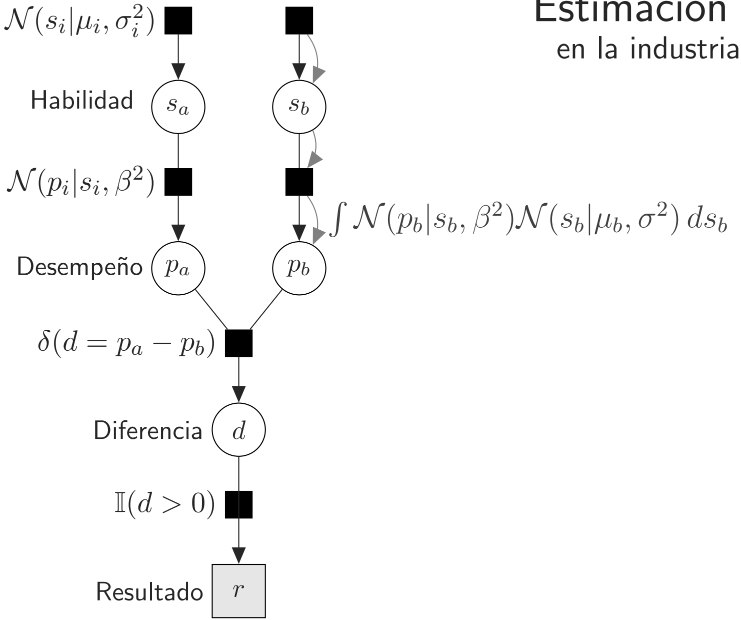


Estimación de habilidad

en la industria del videojuego

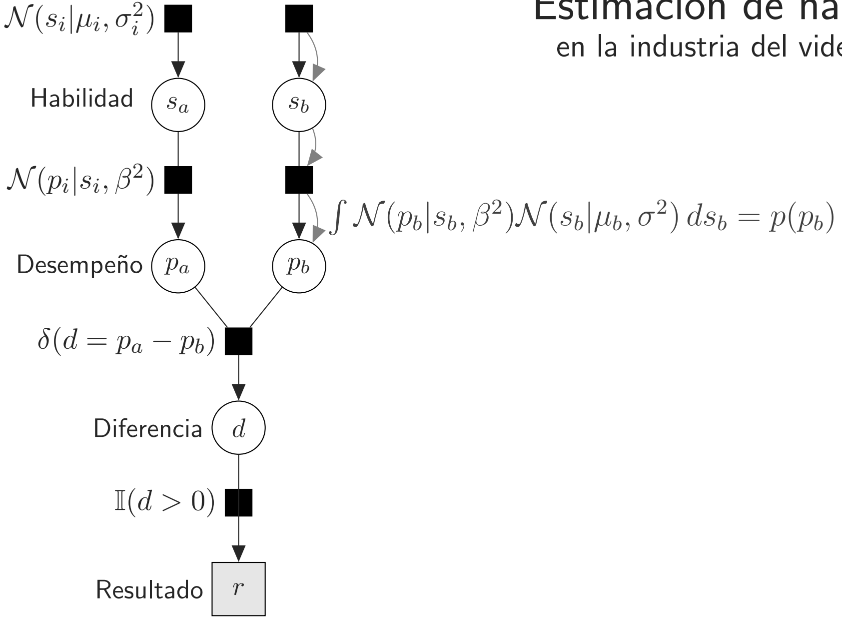
Estimación de habilidad

en la industria del videojuego



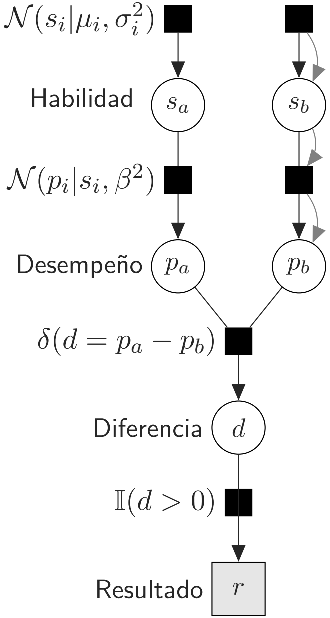
Estimación de habilidad

en la industria del videojuego



Estimación de habilidad

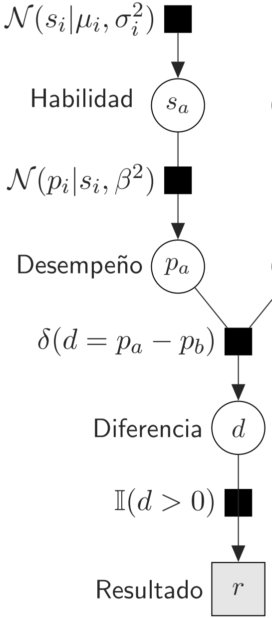
en la industria del videojuego



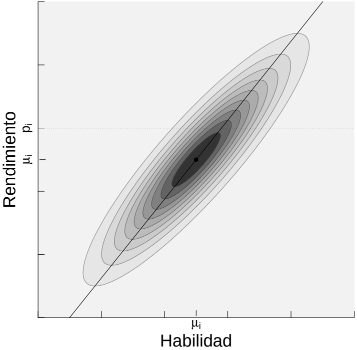
$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$

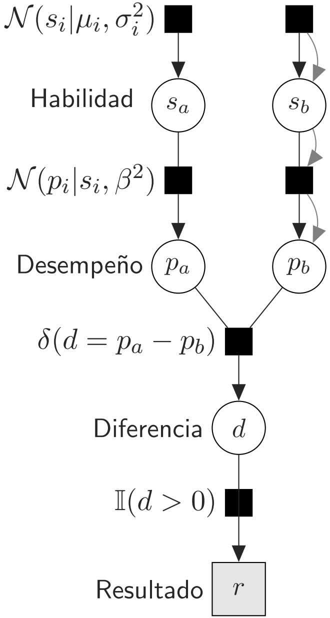
Estimación de habilidad

en la industria del videojuego



$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$

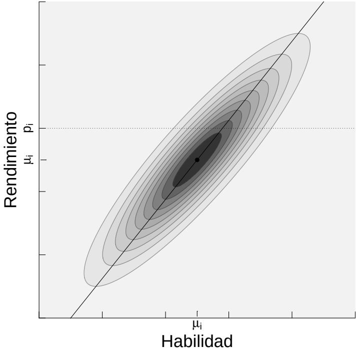


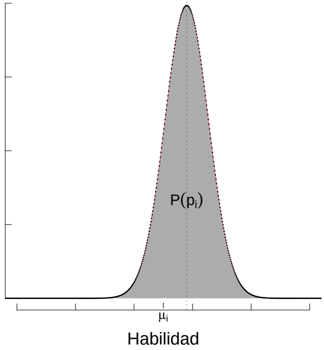


Estimación de habilidad

en la industria del videojuego

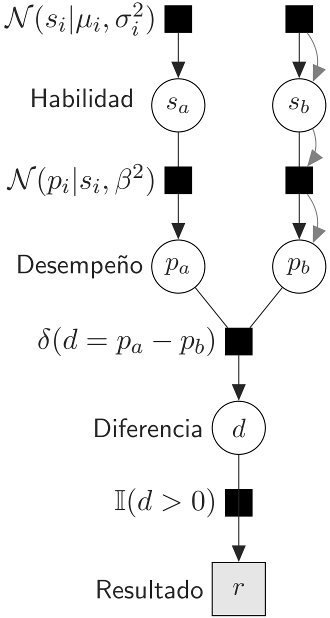
$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$



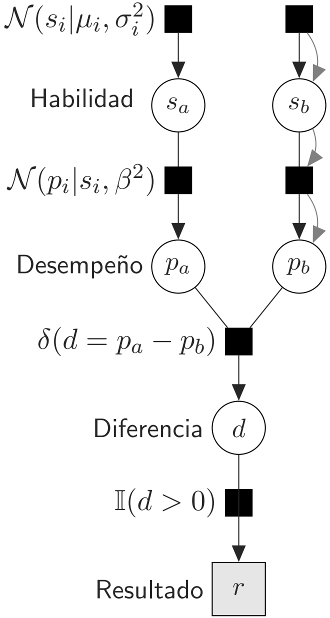


Estimación de habilidad

en la industria del videojuego



$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$
$$\overset{*}{=} \int \mathcal{N}(p_a|\mu_a,\beta^2+\sigma_a^2)\,\mathcal{N}(s_a|\mu_*,\sigma_*^2)ds_a$$

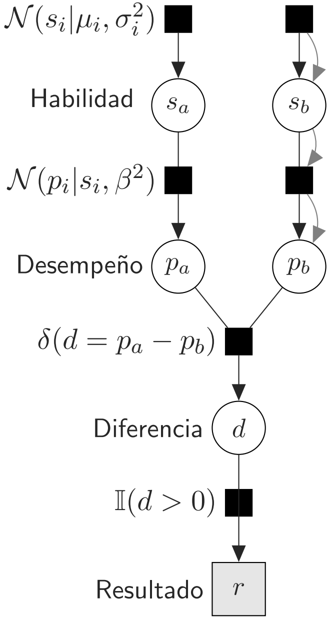


Estimación de habilidad

en la industria del videojuego

$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$

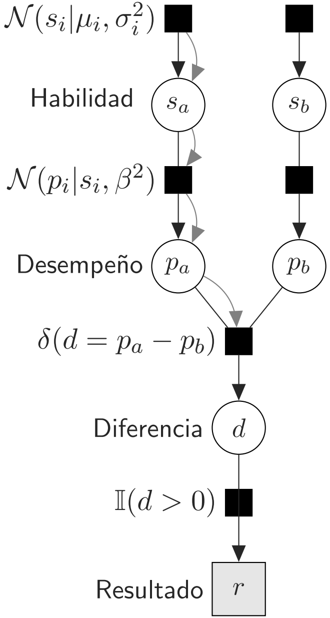
$$\overset{*}{=} \int \underbrace{\mathcal{N}(p_b|\mu_b,\beta^2+\sigma_b^2)}_{\text{const.}}\underbrace{\mathcal{N}(s_a|\mu_*,\sigma_*^2)}_{1}ds_b$$



Estimación de habilidad
en la industria del videojuego

$$p(p_b) = \int \mathcal{N}(p_b|s_b,\beta^2)\mathcal{N}(s_b|\mu_b,\sigma_b^2)ds_b$$

$$\overset{*}{=} \quad \mathcal{N}(p_b|\mu_b,\beta^2+\sigma_b^2)$$



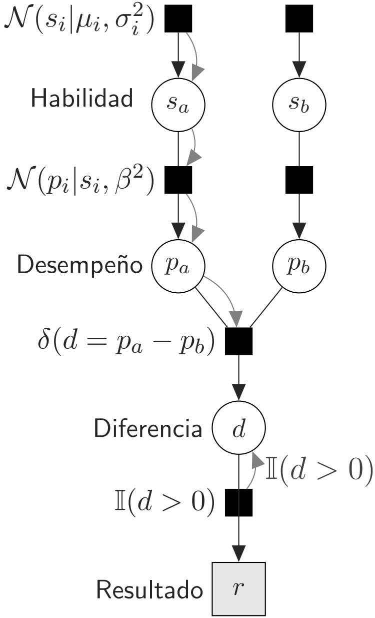
Estimación de habilidad

en la industria del videojuego

$$p(p_a) \overset{*}{=} \quad \mathcal{N}(p_b|\mu_a,\beta^2+\sigma_a^2)$$

Estimación de habilidad

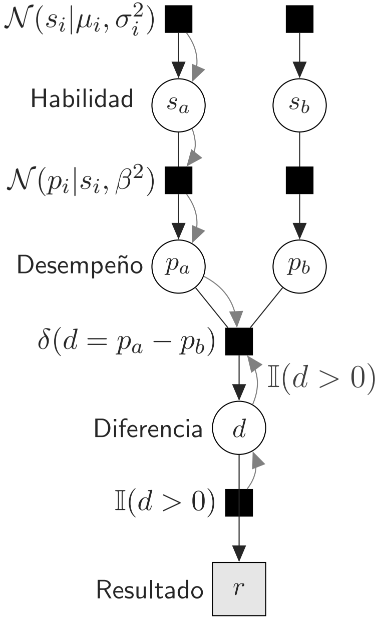
en la industria del videojuego



$$p(p_a) \overset{*}{=} \quad \mathcal{N}(p_b|\mu_a, \beta^2 + \sigma_a^2)$$

Estimación de habilidad

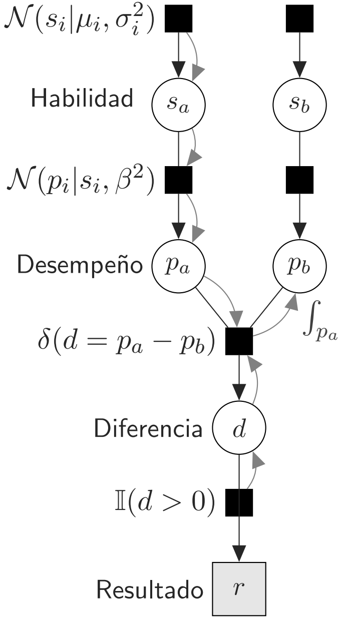
en la industria del videojuego



$$p(p_a) \overset{*}{=} \quad \mathcal{N}(p_b|\mu_a, \beta^2 + \sigma_a^2)$$

Estimación de habilidad

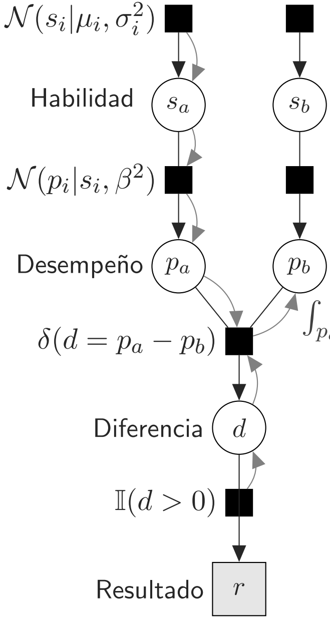
en la industria del videojuego



$$\int_{p_a}\int_d \delta(d = p_a - p_b)\mathbb{I}(d>0)\mathcal{N}(p_a|\mu_a,\beta^2+\sigma_a^2)\,dp_a\,dd$$

Estimación de habilidad

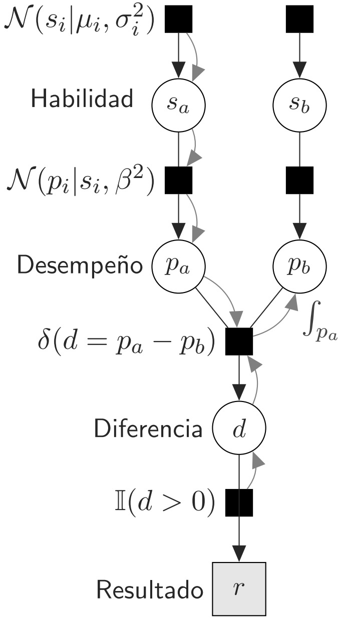
en la industria del videojuego



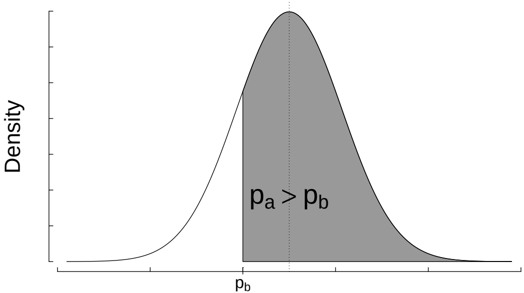
$$\int_{p_a}\int_d \delta(\boldsymbol{d} = \boldsymbol{p_a} - \boldsymbol{p_b})\mathbb{I}(\boldsymbol{d} > 0)\mathcal{N}(p_a|\mu_a, \beta^2 + \sigma_a^2)\, dp_a\, dd$$

Estimación de habilidad

en la industria del videojuego

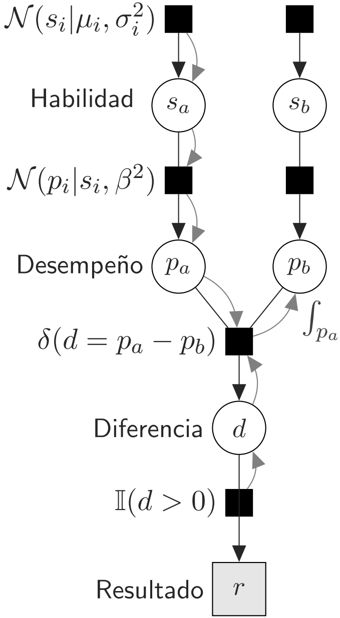


$$\int_{p_a} \mathbb{I}(p_a > p_b)\mathcal{N}(p_a|\mu_a, \beta^2 + \sigma_a^2)\, dp_a$$

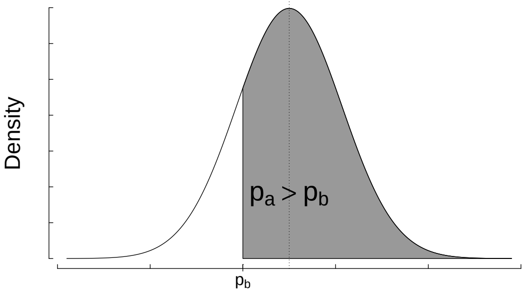


Estimación de habilidad

en la industria del videojuego

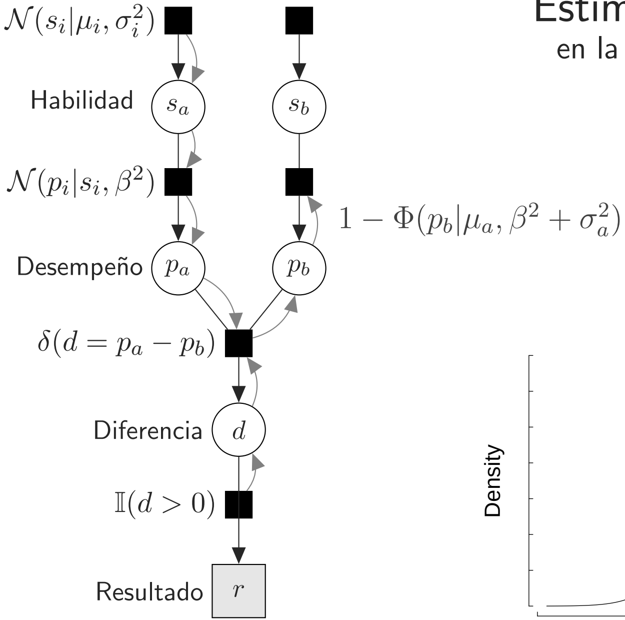


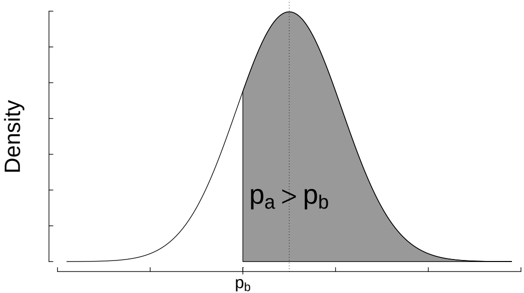
$$\int_{p_a} \mathbb{I}(p_a > p_b)\mathcal{N}(p_a|\mu_a, \beta^2+\sigma_a^2)\, dp_a = 1 - \Phi(p_b|\mu_a,\beta^2+\sigma_a^2)$$



Estimación de habilidad

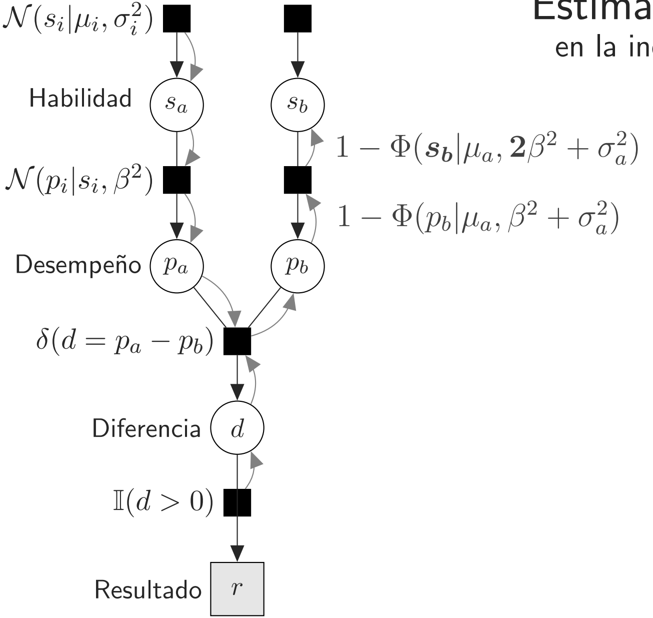
en la industria del videojuego





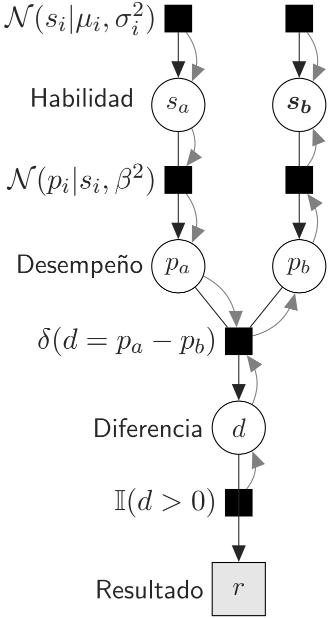
Estimación de habilidad

en la industria del videojuego

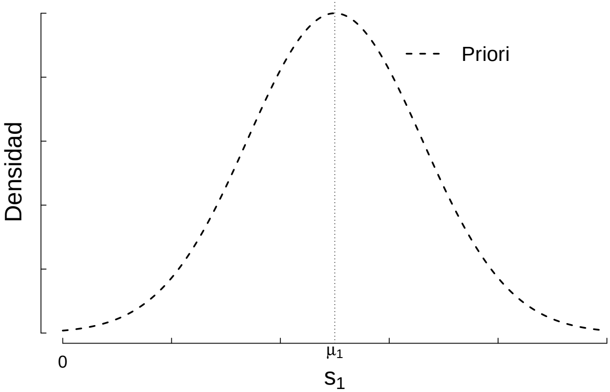


Estimación de habilidad

en la industria del videojuego

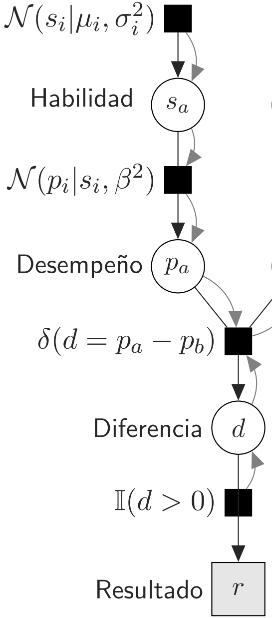


$$p(s_b, r^*) = \overbrace{\mathcal{N}(s_b|\mu_b,\sigma_b^2)}^{\text{Prior}}\,\overbrace{1-\Phi(0|\mu_a - s_b, 2\beta^2+\sigma_a^2)}^{\text{Predicción o verosimilitud}}$$

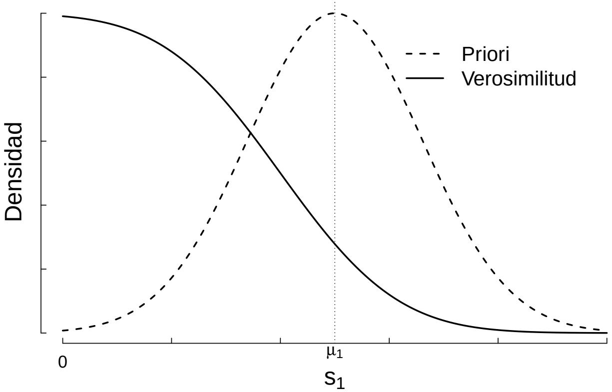


Estimación de habilidad

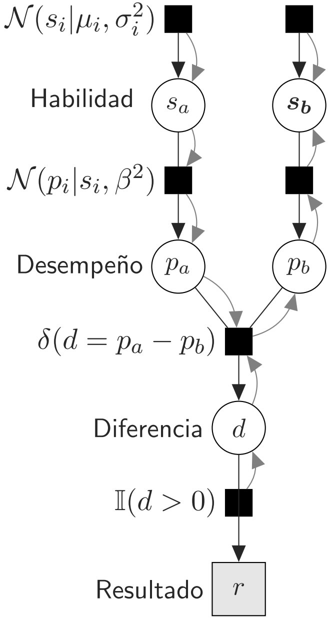
en la industria del videojuego



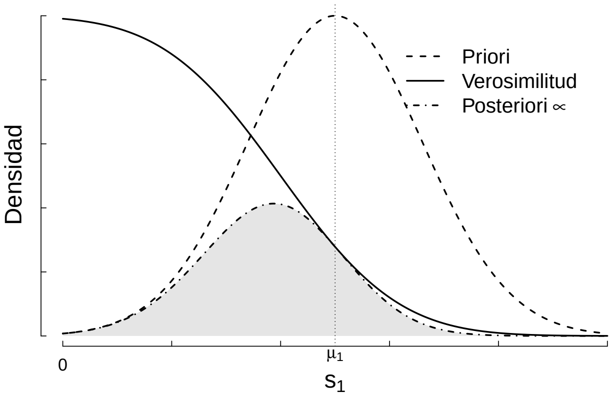
$$p(s_b, r^*) = \underbrace{\mathcal{N}(s_b|\mu_b,\sigma_b^2)}_{\text{Prior}} \underbrace{1-\Phi(0|\mu_a - s_b, 2\beta^2+\sigma_a^2)}_{\text{Predicción o verosimilitud}}$$



Estimación de habilidad
en la industria del videojuego

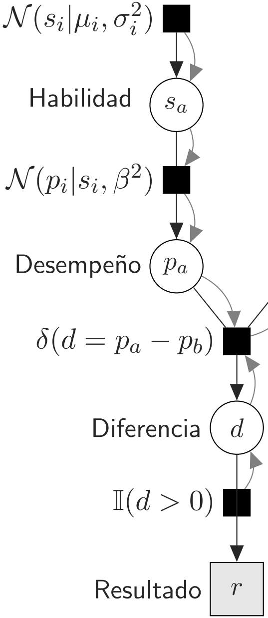


$$p(s_b, r^*) = \underbrace{\mathcal{N}(s_b|\mu_b,\sigma_b^2)}_{\text{Prior}}\ \underbrace{1-\Phi(0|\mu_a - s_b, 2\beta^2+\sigma_a^2)}_{\text{Predicción o verosimilitud}}$$

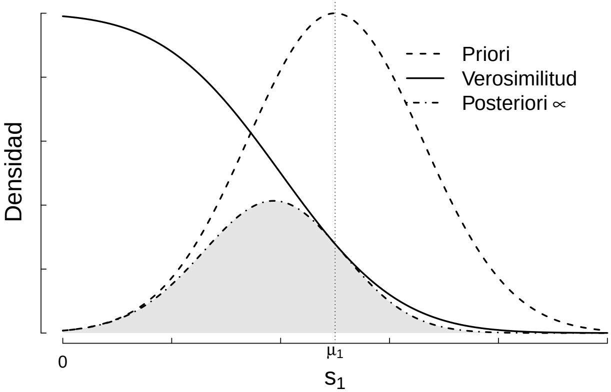


Estimación de habilidad

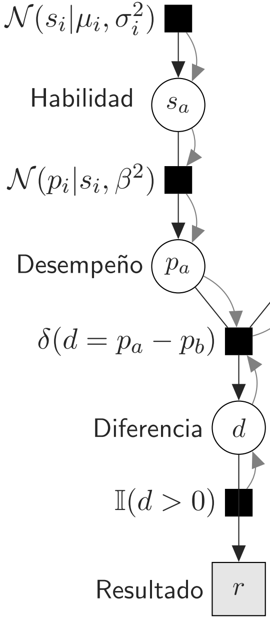
en la industria del videojuego



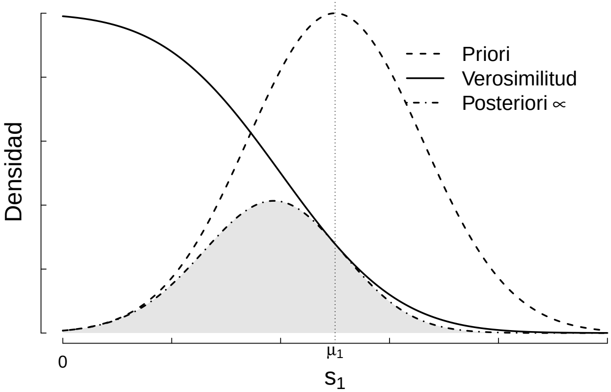
$$\widehat{p}(s_b|r^*) = \underset{\mu,\sigma}{\text{arg min}} \quad \text{KL}(\, p(s_b|r^*) \,||\, \mathcal{N}(s_b|\mu,\sigma^2)\,)$$

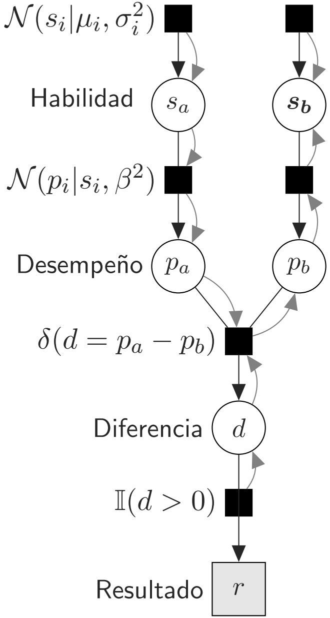


Estimación de habilidad
en la industria del videojuego



$$\widehat{p}(s_b|r^*) = \mathcal{N}(s_b\,|\,\widehat{\mu},\,\widehat{\sigma}^2)$$



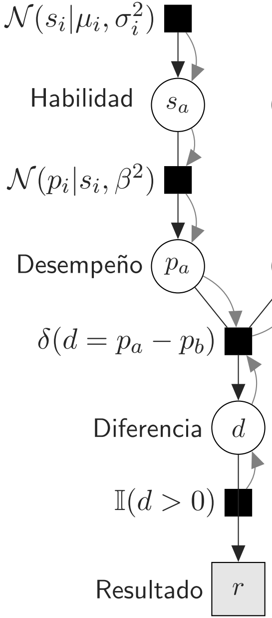


Estimación de habilidad

en la industria del videojuego

$$\underbrace{\widehat{p}(s_b|r^*)}_{\text{Posterior}_1} = \mathcal{N}(s_b\,|\,\widehat{\mu},\,\widehat{\sigma}^2)$$

$$\widehat{\text{Posterior}}_t \rightarrow \text{Prior}_{t+1}$$



Estimación de habilidad

en la industria del videojuego

$$\underbrace{\widehat{p}(s_b|r^*)}_{\text{Posterior}_1} = \mathcal{N}(s_b\,|\,\widehat{\mu},\,\widehat{\sigma}^2)$$

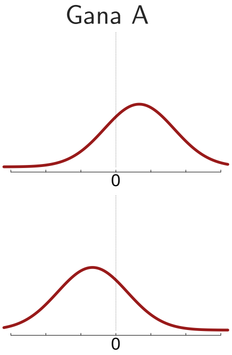
$$\widehat{\text{Posterior}}_t \rightarrow \text{Prior}_{t+1}$$

$$\underbrace{p(s_b)}_{\text{Prior}_2} = \mathcal{N}(s_b\,|\,\widehat{\mu},\,\widehat{\sigma}^2 + \gamma^2)$$

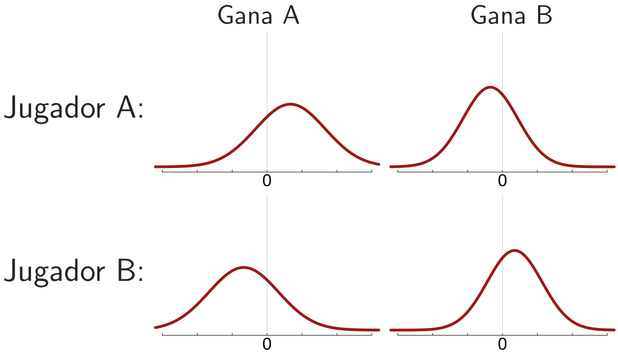
Enfoque Posterior → Prior

Jugador A:

Jugador B:



Enfoque Posterior $\rightarrow$ Prior



Enfoque Posterior $\rightarrow$ Prior

Problemas de un procedimiento ad-hoc

Problemas.

- Ofrece resultados que van en contra de nuestra intuición

Enfoque Posterior $\rightarrow$ Prior
Problemas de un procedimiento ad-hoc

Problemas.

- Ofrece resultados que van en contra de nuestra intuición
- Mucha incertidumbre al inicio de las series temporales

Enfoque Posterior $\rightarrow$ Prior

Problemas de un procedimiento ad-hoc

Problemas.

- Ofrece resultados que van en contra de nuestra intuición
- Mucha incertidumbre al inicio de las series temporales
- No aprovecha la información disponible (entre ramas paralelas)

Enfoque Posterior $\rightarrow$ Prior

Problemas de un procedimiento ad-hoc

Problemas.

- Ofrece resultados que van en contra de nuestra intuición
- Mucha incertidumbre al inicio de las series temporales
- No aprovecha la información disponible (entre ramas paralelas)
- No garantiza comparabilidad entre estimaciones lejanas en el tiempo y el espacio.

Enfoque Posterior $\rightarrow$ Prior

Problemas de un procedimiento ad-hoc

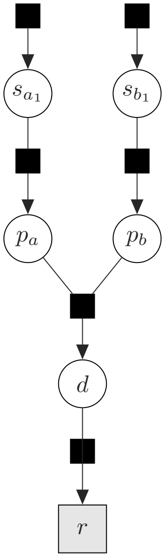
Propaga la información en una sola dirección, del pasado al futuro.

Problemas.

- Ofrece resultados que van en contra de nuestra intuición
- Mucha incertidumbre al inicio de las series temporales
- No aprovecha la información disponible (entre ramas paralelas)
- No garantiza comparabilidad entre estimaciones lejanas en el tiempo y el espacio.

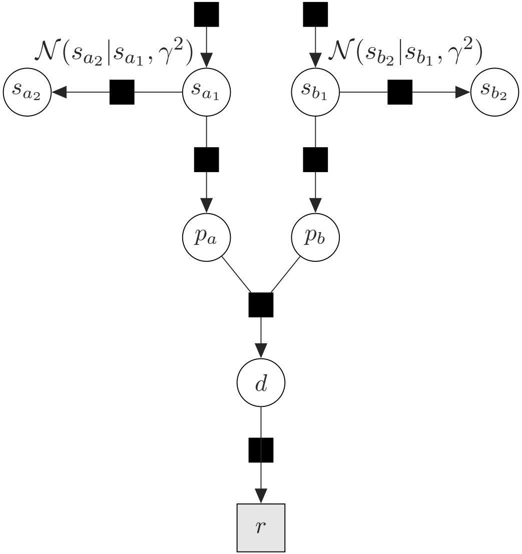
Modelos de historia completa

Estado del arte en la industria del videojuego.



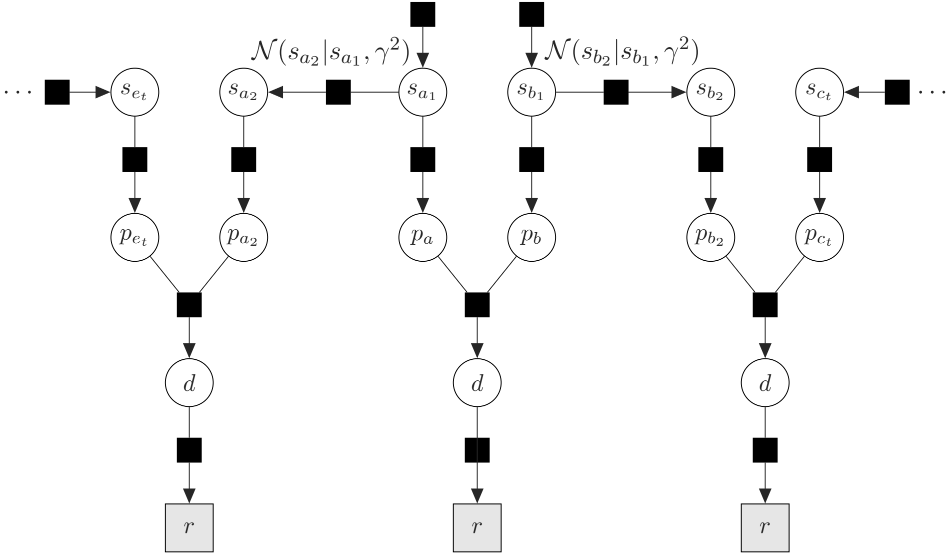
Modelos de historia completa

Estado del arte en la industria del videojuego.



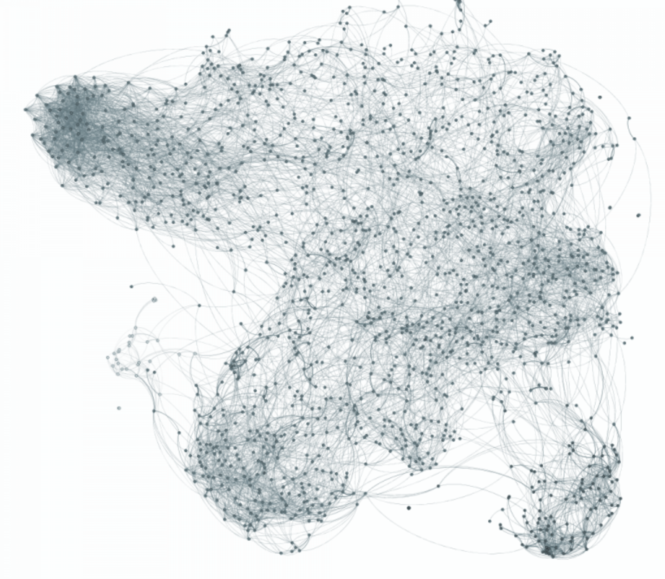
Modelos de historia completa

Estado del arte en la industria del videojuego.



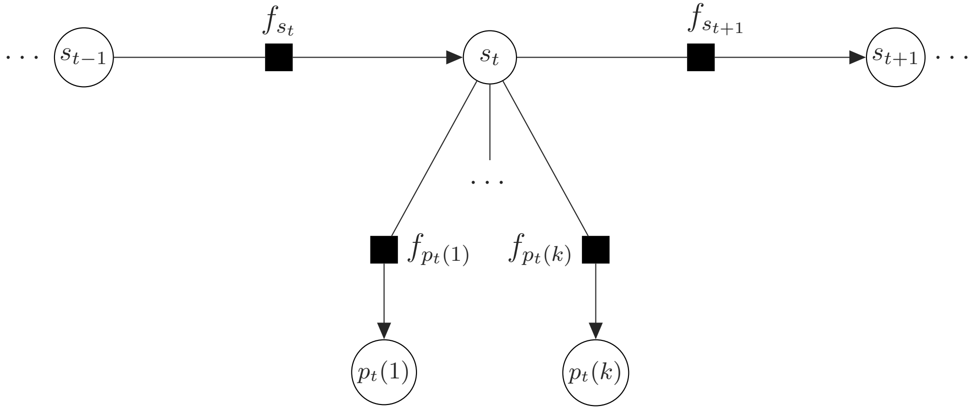
Modelos de historia completa

Estado del arte en la industria del videojuego.



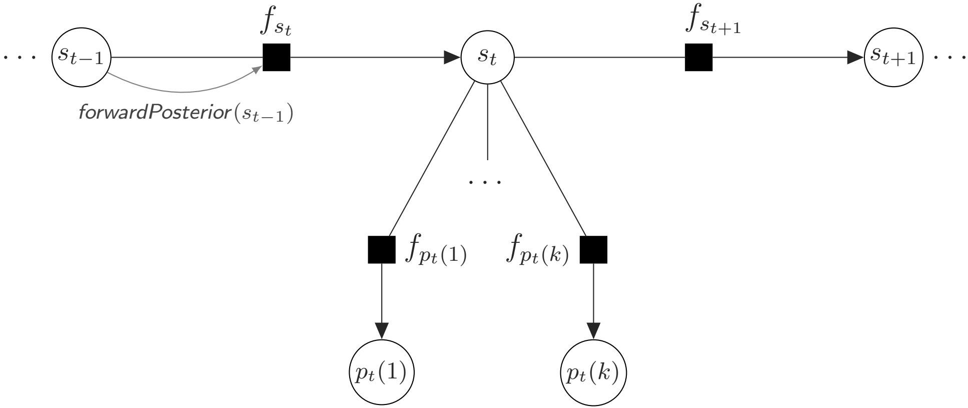
Modelos de historia completa

Smoothing



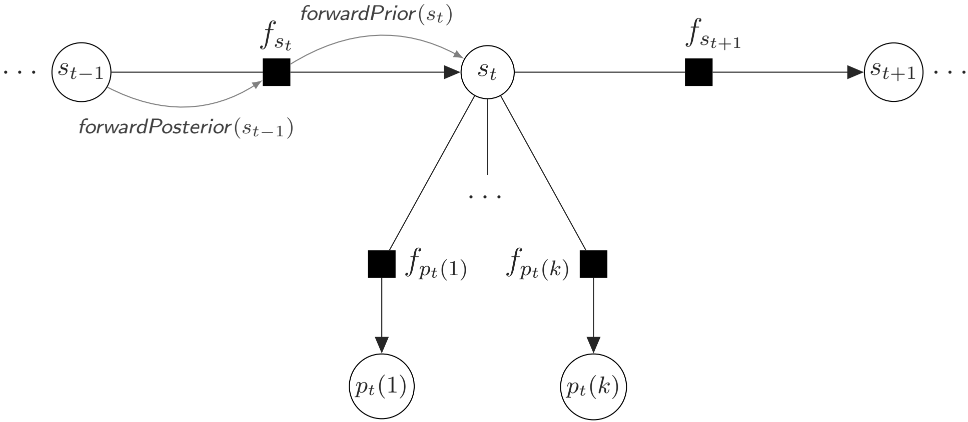
Modelos de historia completa

Smoothing



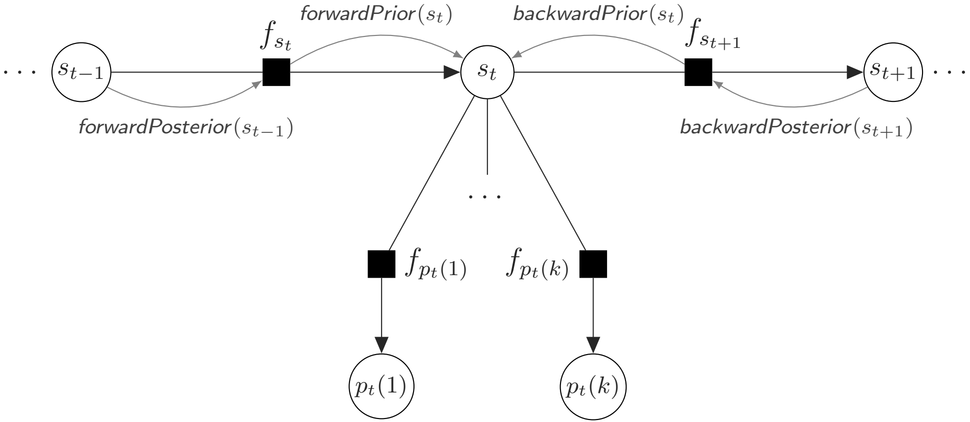
Modelos de historia completa

Smoothing



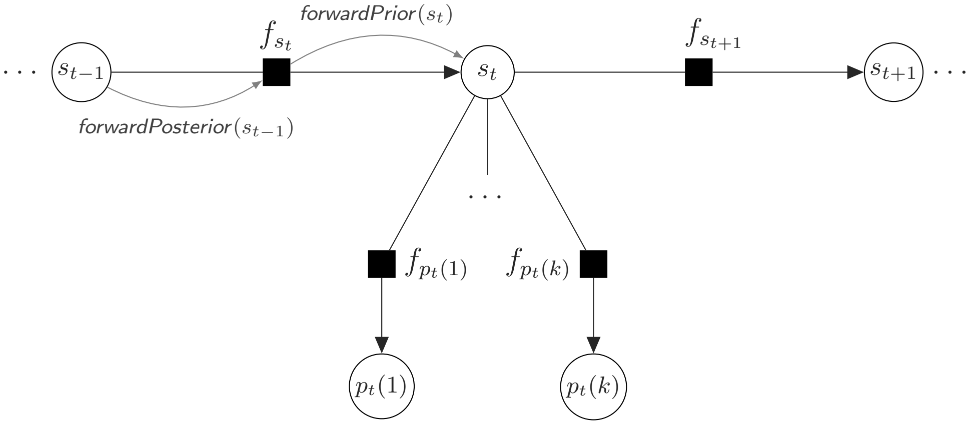
Modelos de historia completa

Smoothing

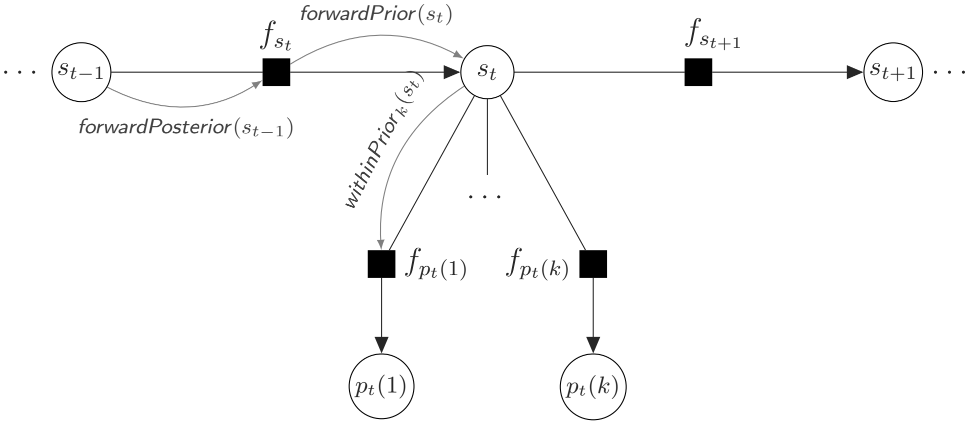


Modelos de historia completa

Smoothing

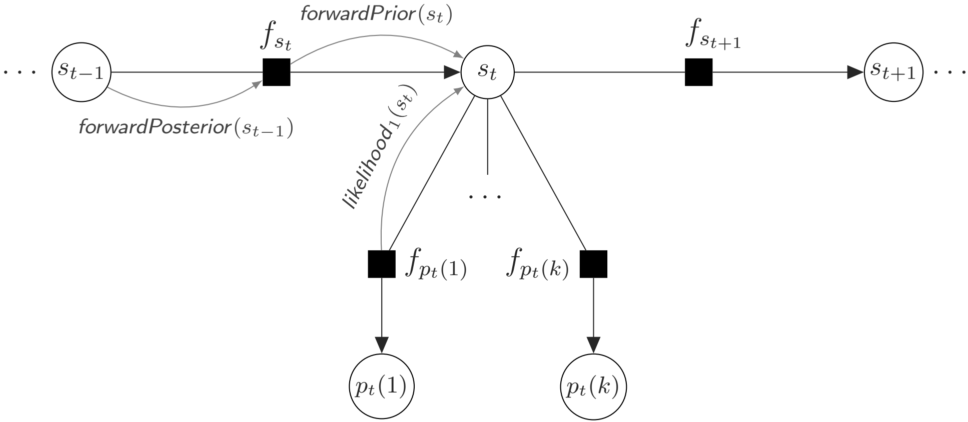


Modelos de historia completa
Smoothing



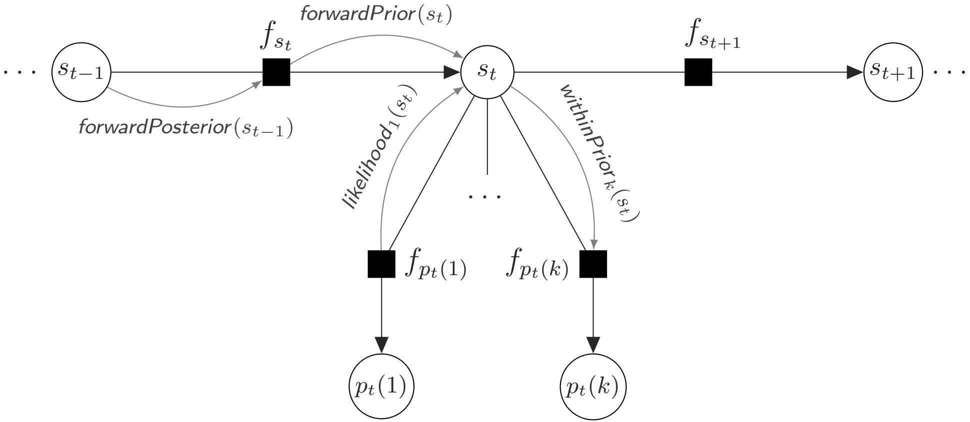
Modelos de historia completa

Smoothing



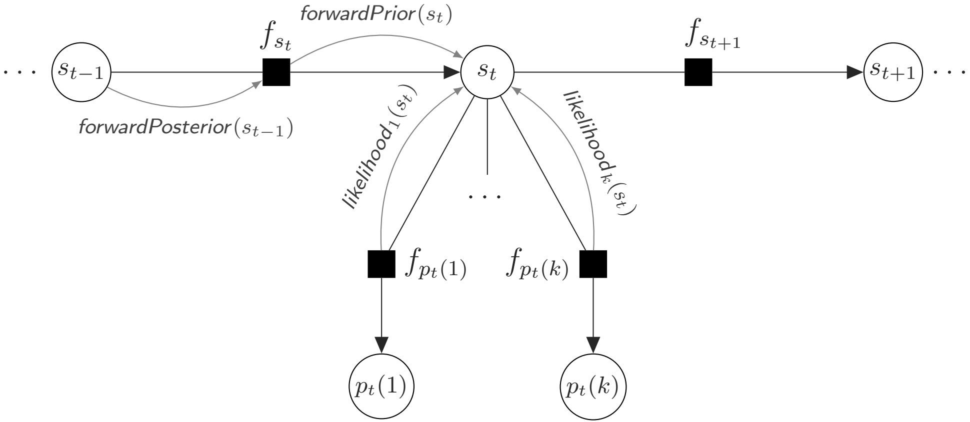
Modelos de historia completa

Smoothing



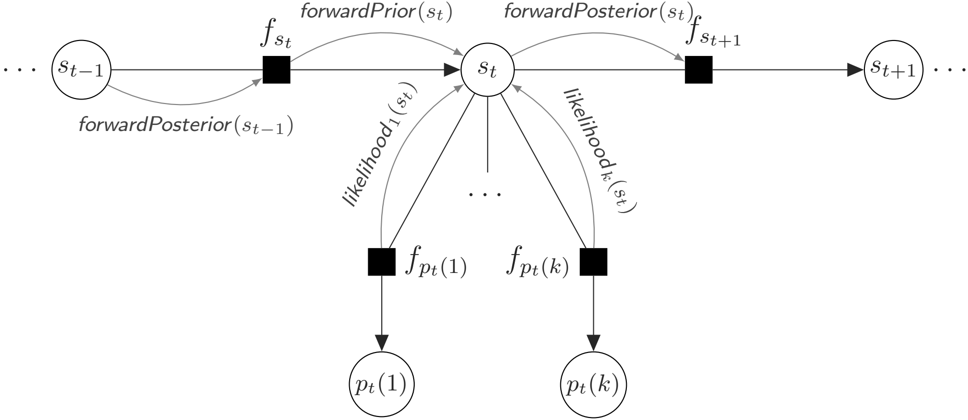
Modelos de historia completa

Smoothing



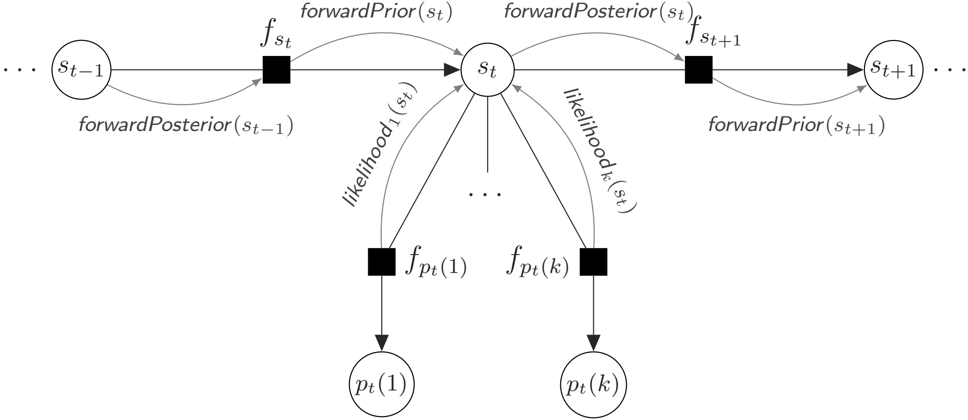
Modelos de historia completa

Smoothing



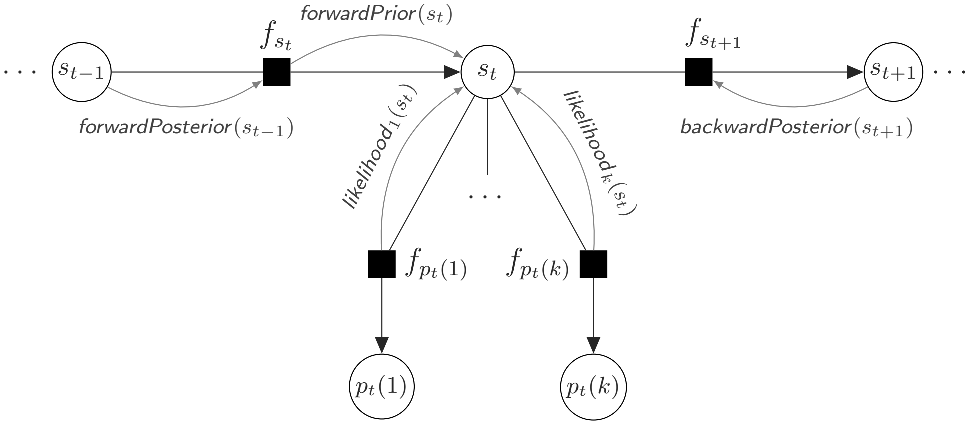
Modelos de historia completa

Smoothing



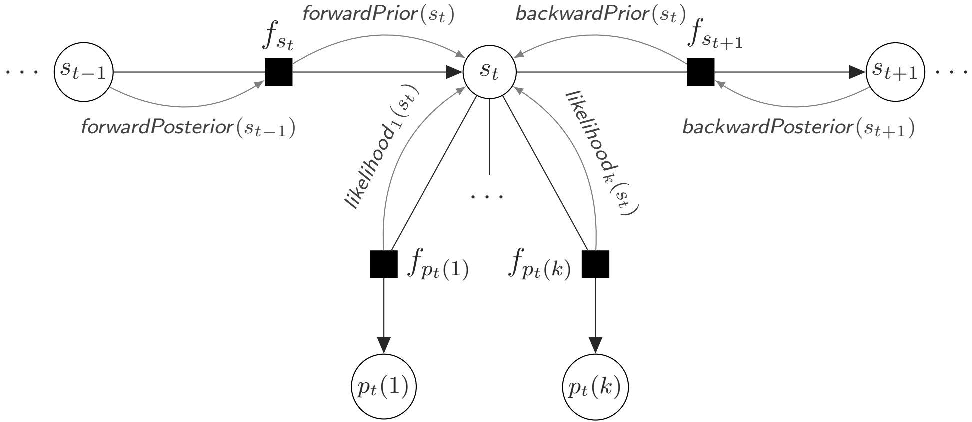
Modelos de historia completa

Smoothing

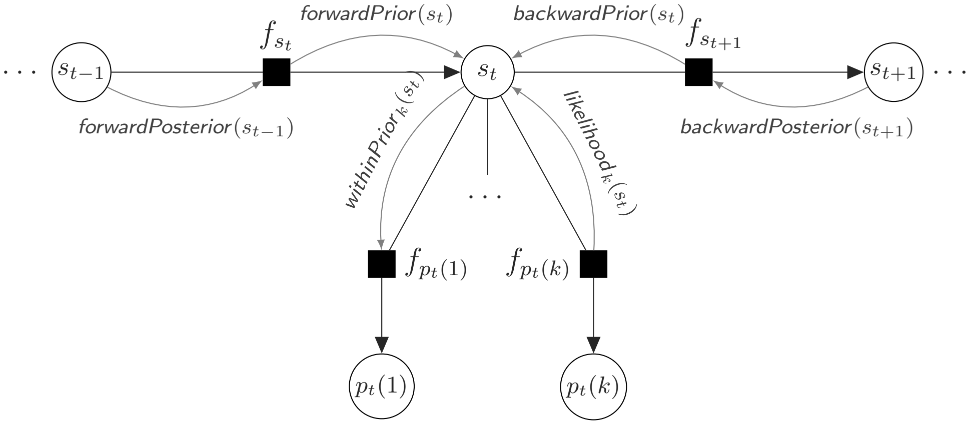


Modelos de historia completa

Smoothing

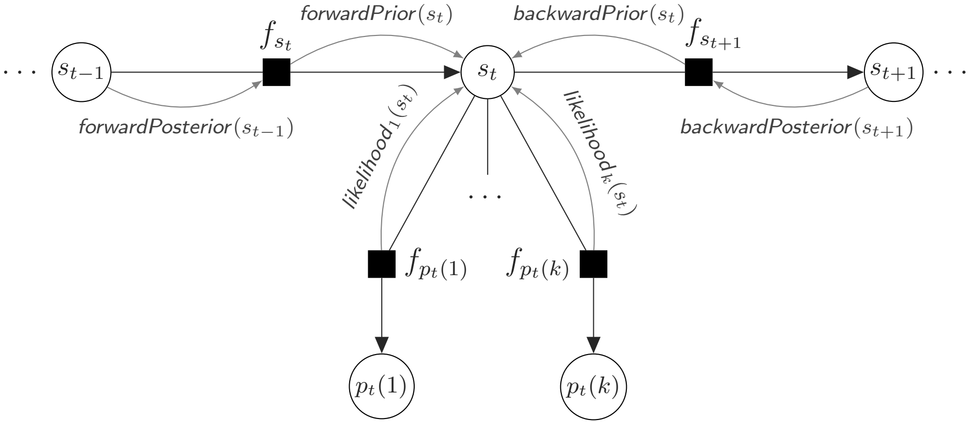


Modelos de historia completa
Smoothing



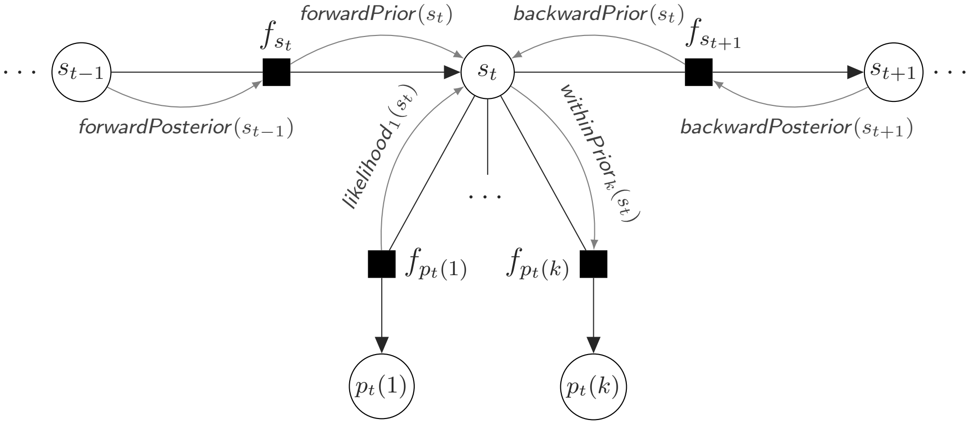
Modelos de historia completa

Smoothing



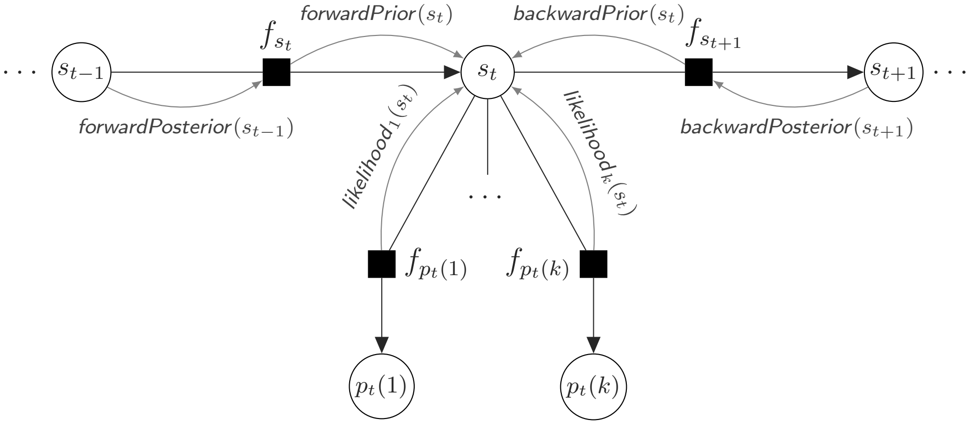
Modelos de historia completa

Smoothing

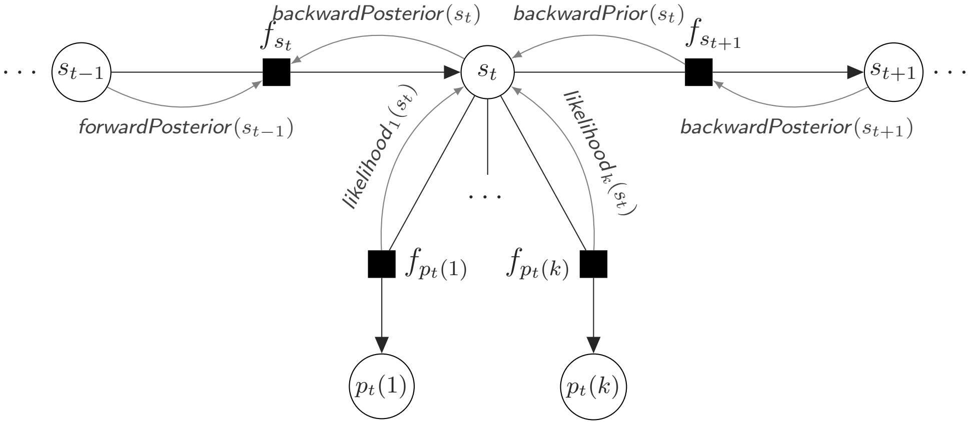


Modelos de historia completa

Smoothing

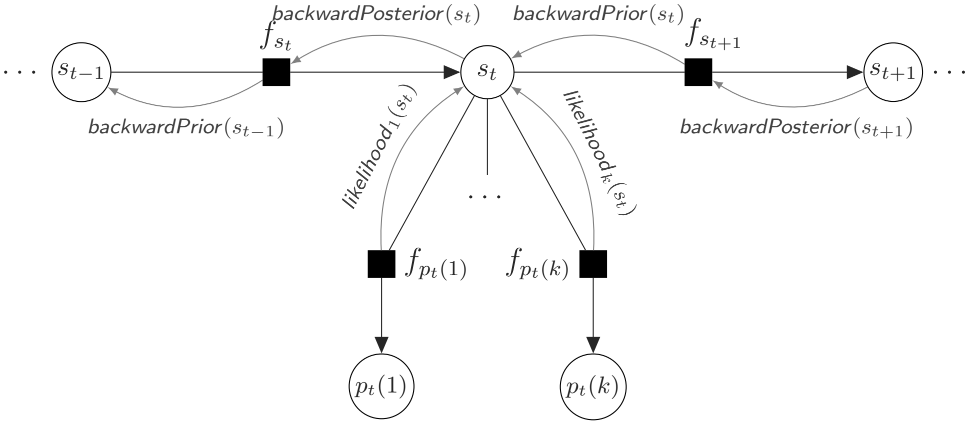


Modelos de historia completa
Smoothing



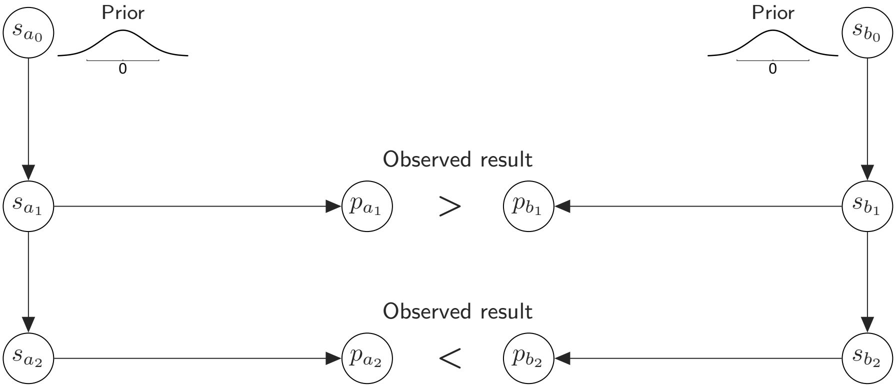
Modelos de historia completa

Smoothing



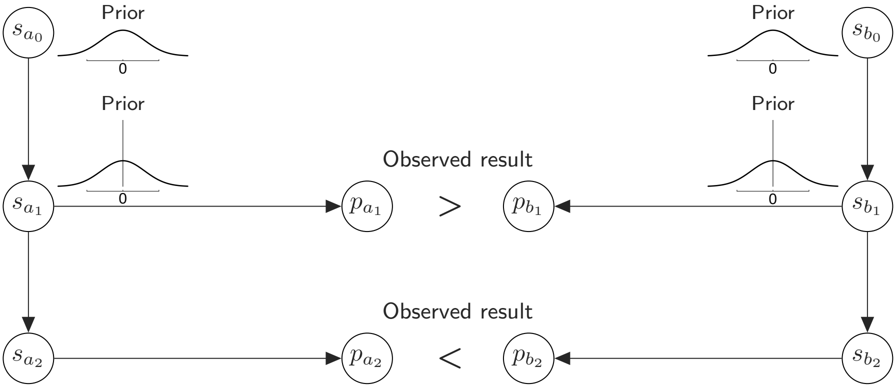
Modelos de historia completa

Smoothing



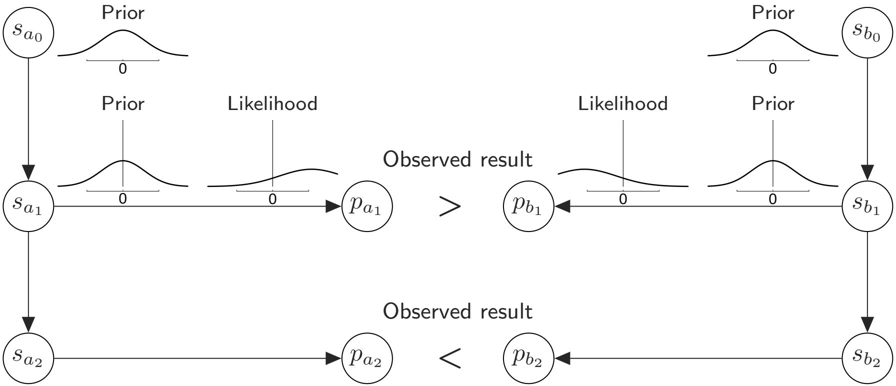
Modelos de historia completa

Smoothing



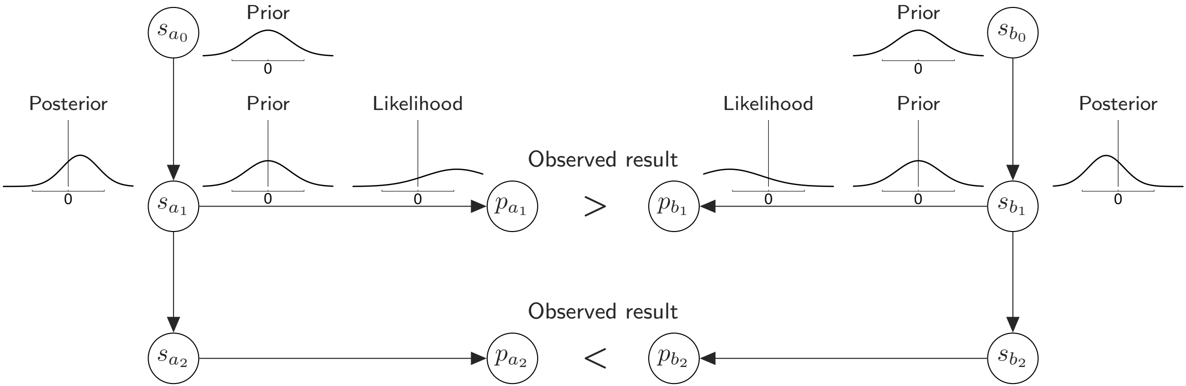
Modelos de historia completa

Smoothing

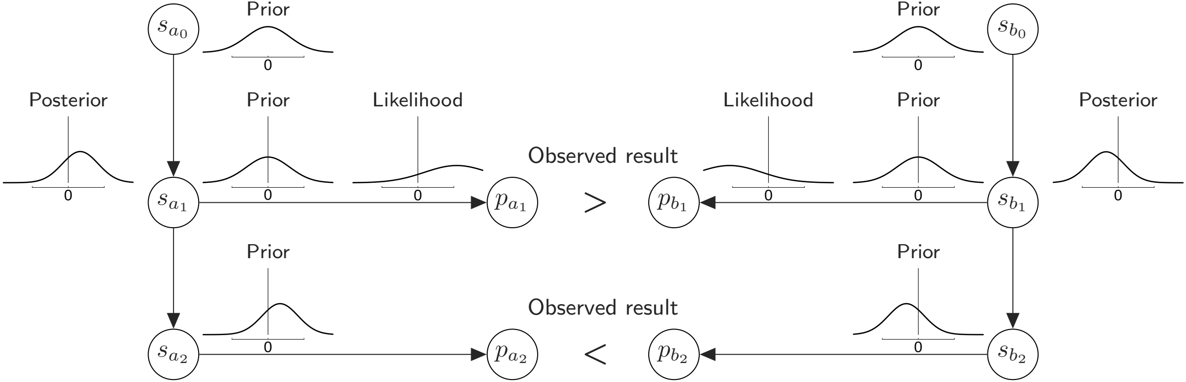


Modelos de historia completa

Smoothing

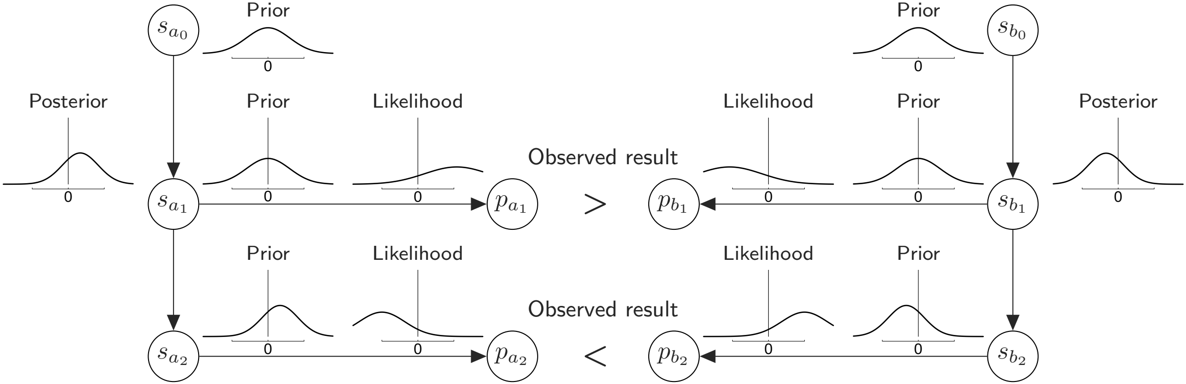


Modelos de historia completa
Smoothing



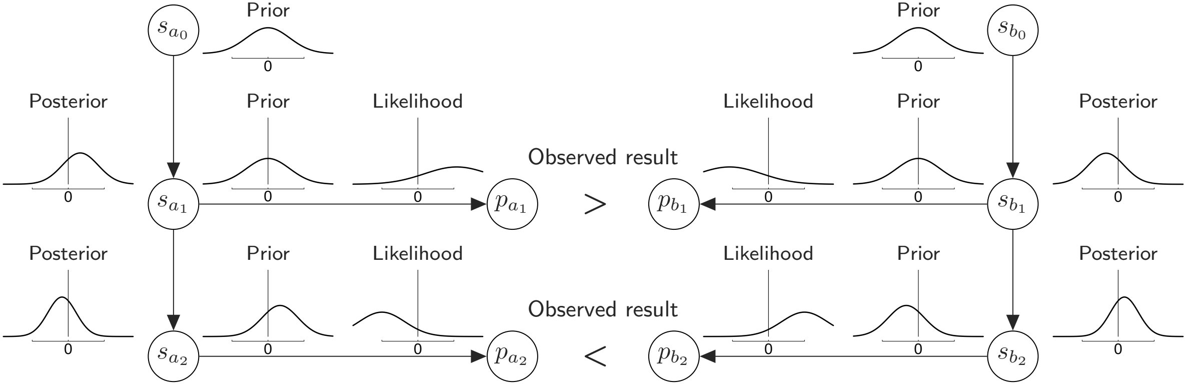
Modelos de historia completa

Smoothing



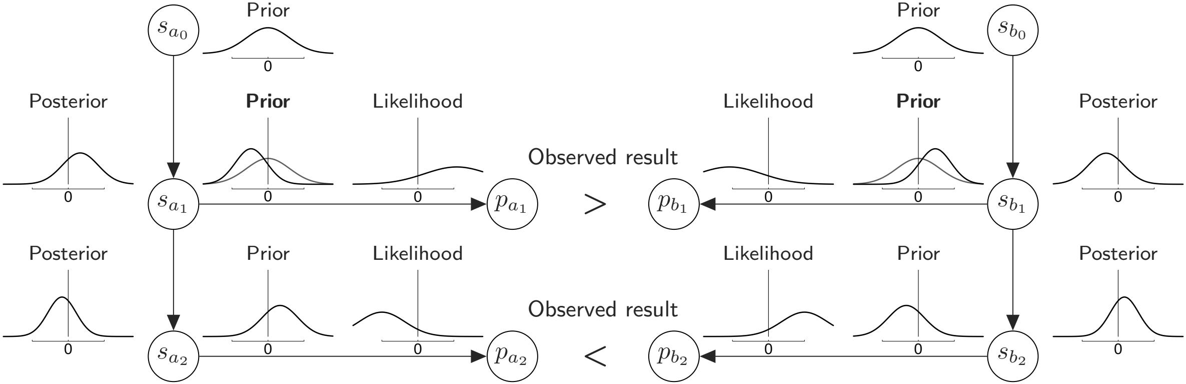
Modelos de historia completa

Smoothing



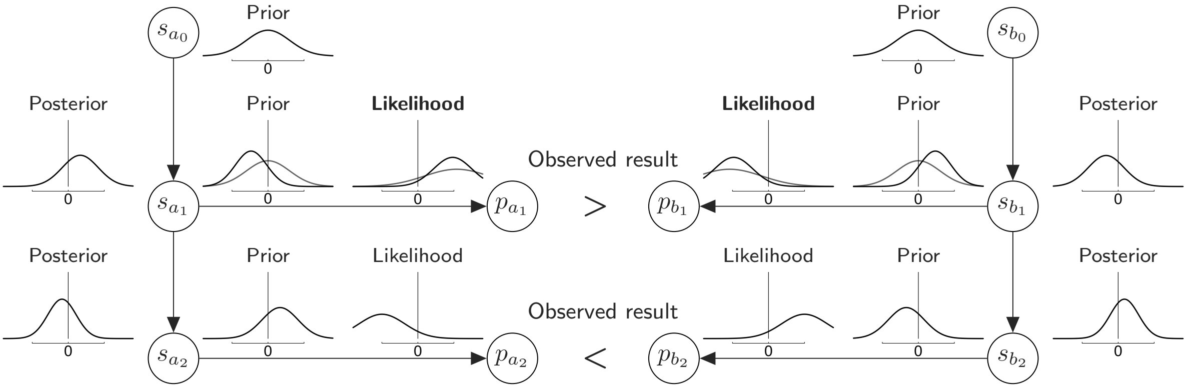
Modelos de historia completa

Smoothing



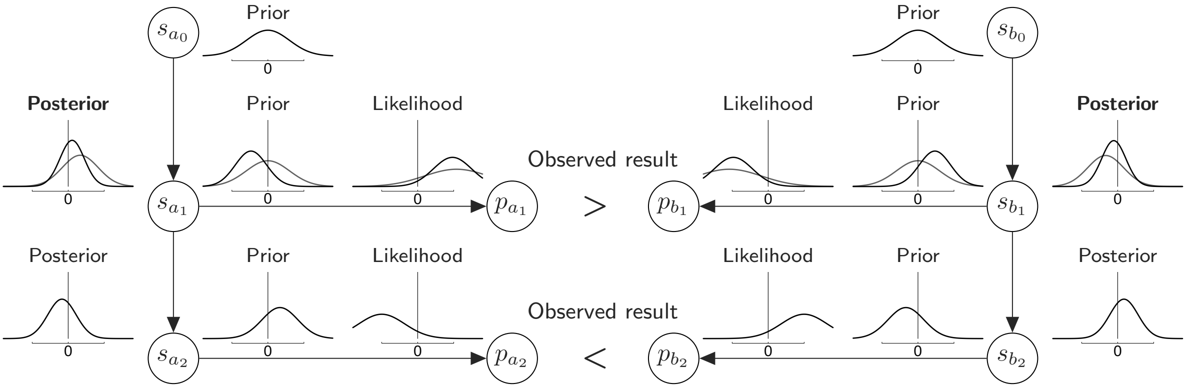
Modelos de historia completa

Smoothing



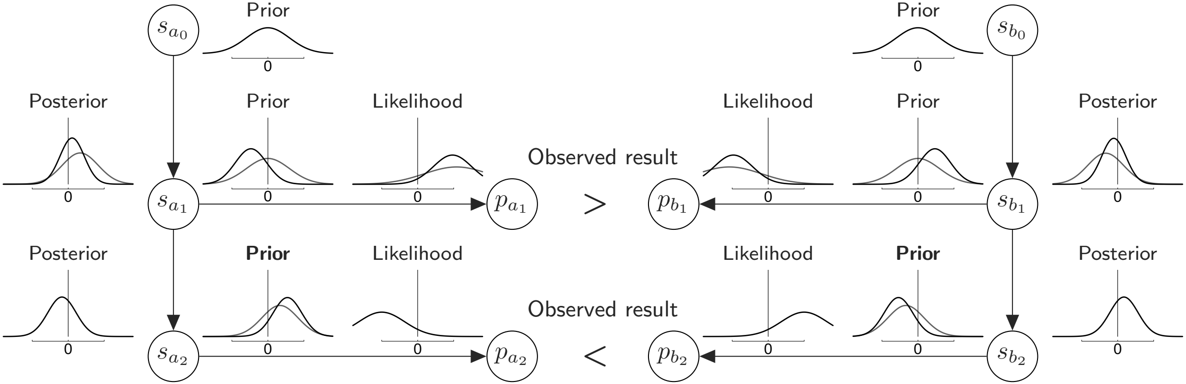
Modelos de historia completa

Smoothing



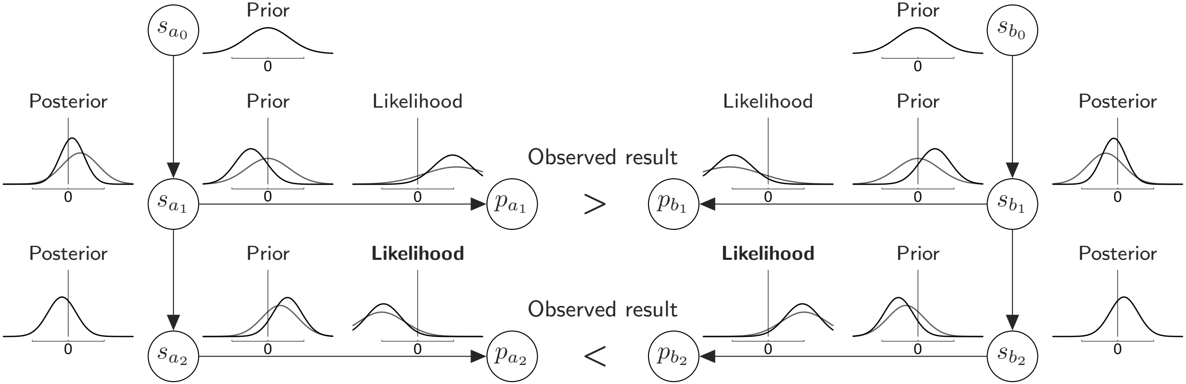
Modelos de historia completa

Smoothing



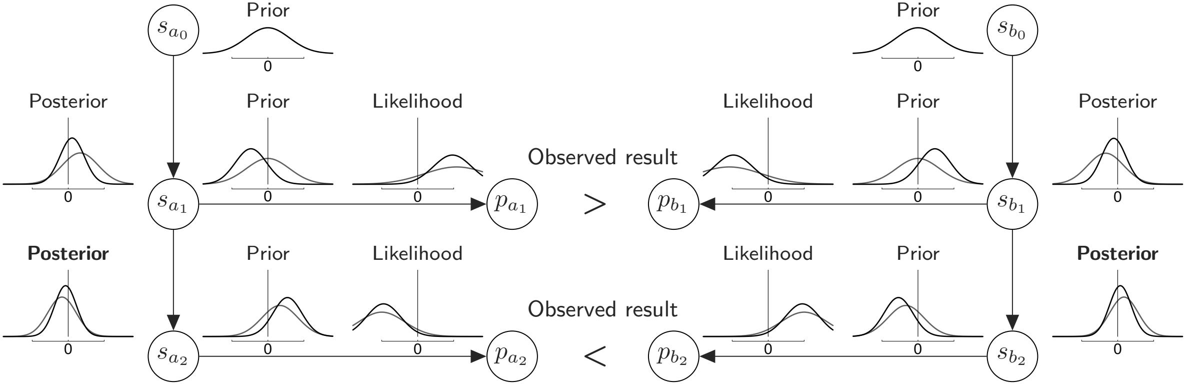
Modelos de historia completa

Smoothing



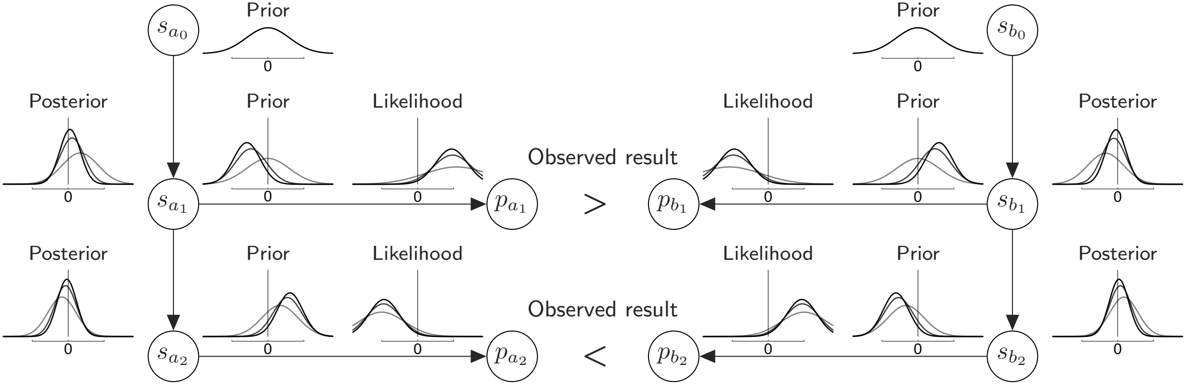
Modelos de historia completa

Smoothing



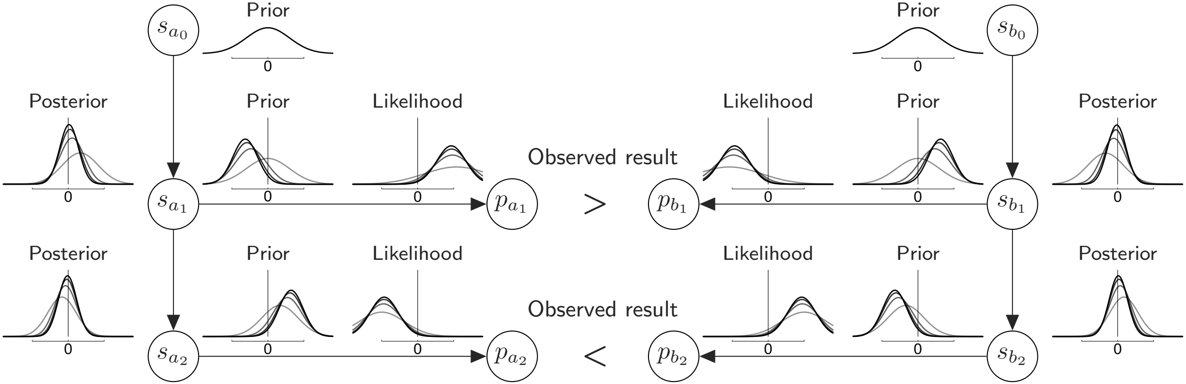
Modelos de historia completa

Smoothing



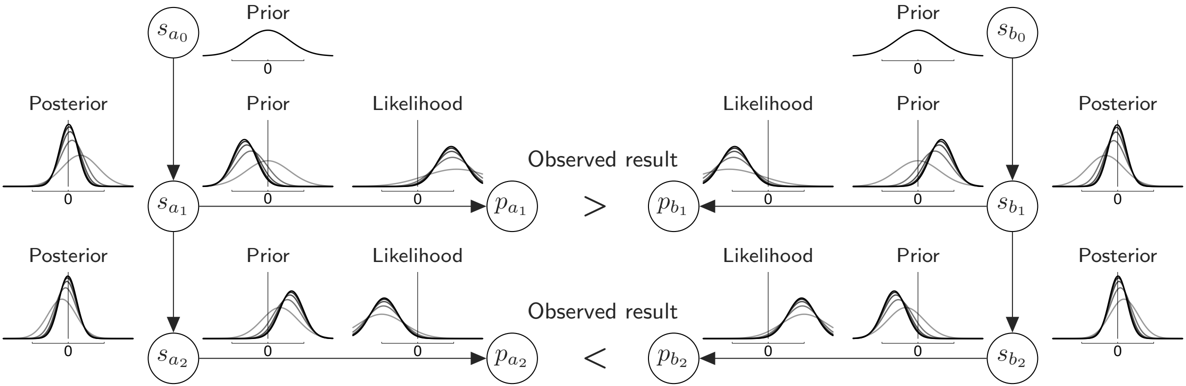
Modelos de historia completa

Smoothing



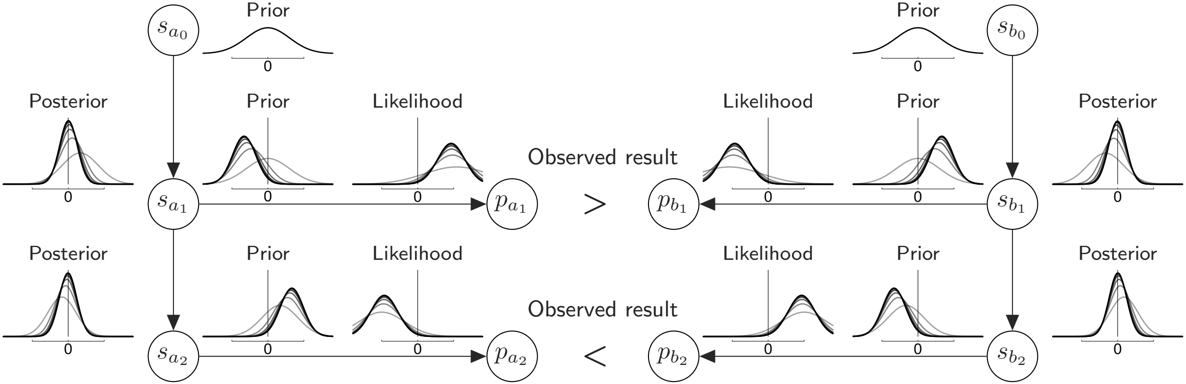
Modelos de historia completa

Smoothing



Modelos de historia completa

Smoothing



TrueSkill Through Time

Estado del arte en la industria del videojuego.

TrueSkill Through Time

Estado del arte en la industria del videojuego.

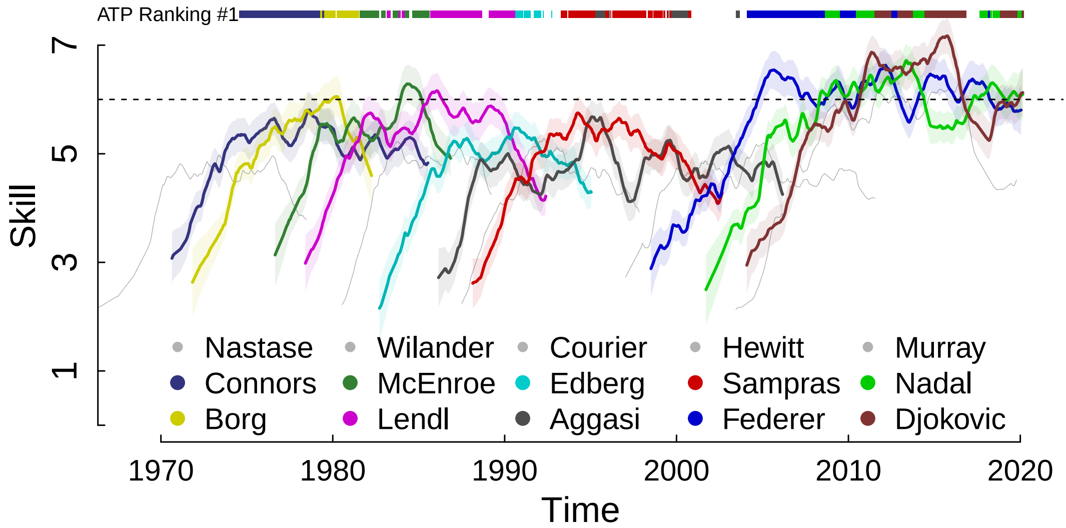
```
using CSV; using Dates
using TrueSkillThroughTime
global const ttt = TrueSkillThroughTime

data = CSV.read("atp.csv")
dates = Dates.value.(data[:,"time_start"] .- Date("1900-1-1"))

matches = [ r.double == "t" ? [[r.w1_id,r.w2_id],[r.l1_id,r.l2_id]] :
    [[r.w1_id],[r.l1_id]] for r in eachrow(data) ]
h = History(composition = matches, times = dates, sigma = 1.6, gamma = 0.036)
convergence(h, epsilon = 0.01, iterations = 10)
```

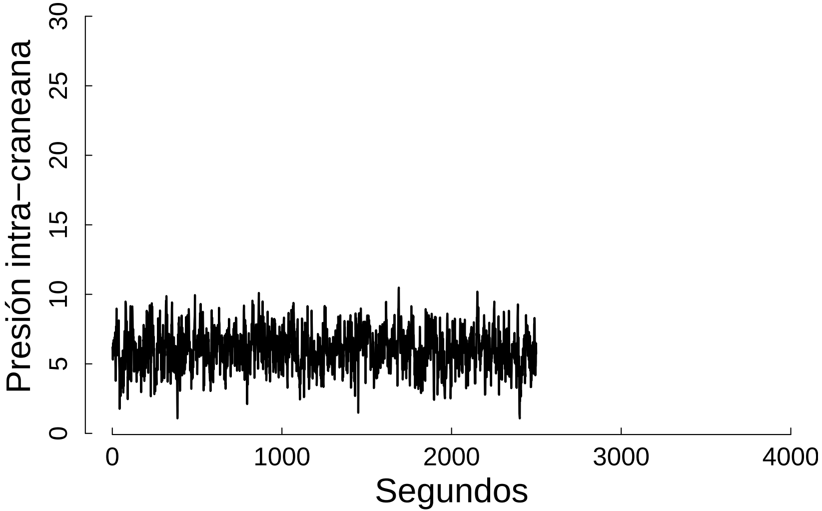
TrueSkill Through Time

Estado del arte en la industria del videojuego.

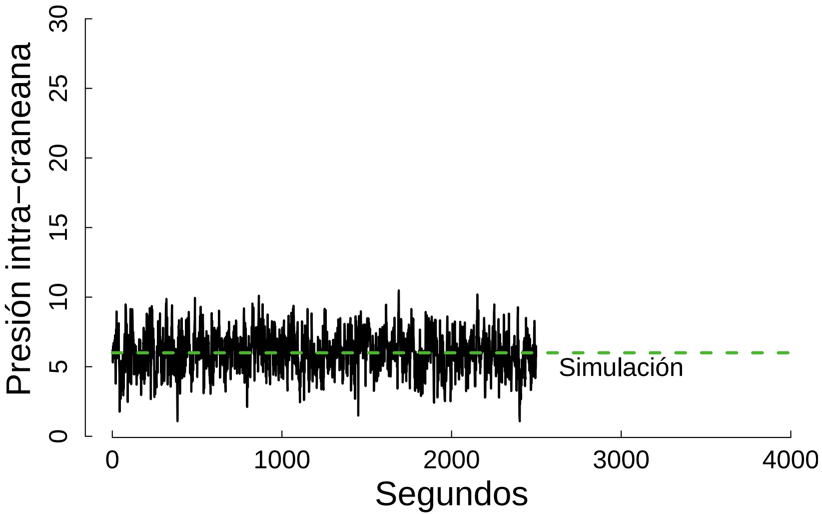


Efecto causal en series temporales

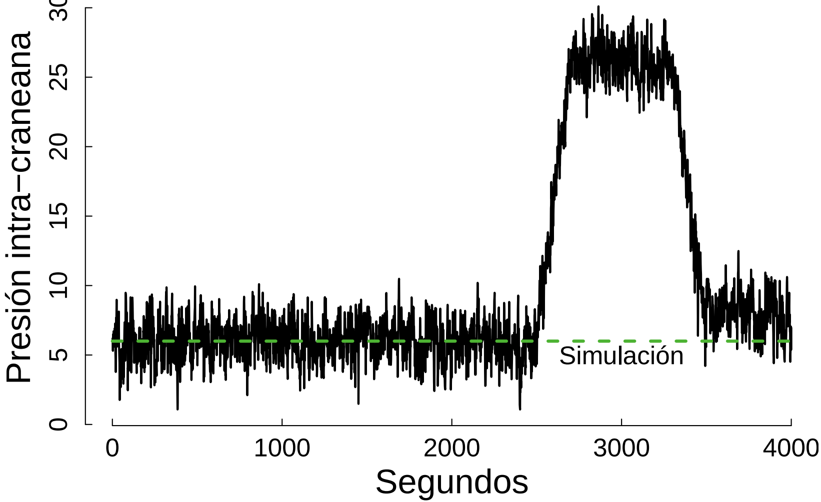
Efecto causal temporal



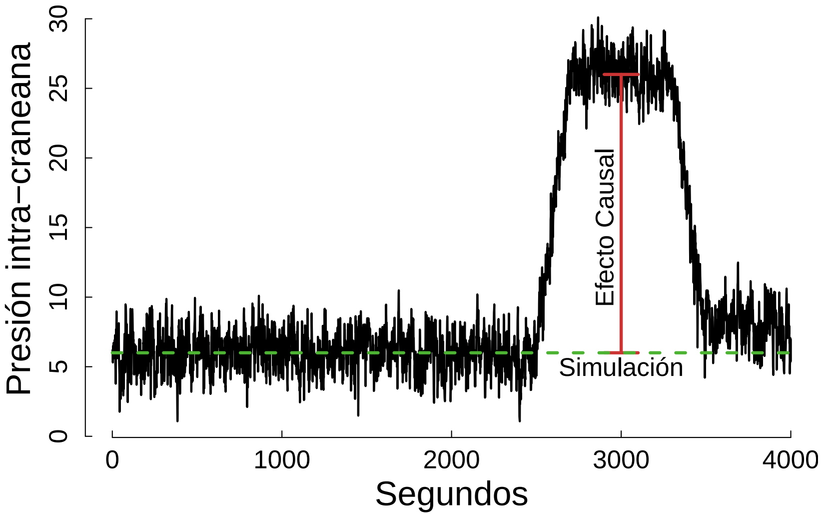
Efecto causal temporal



Efecto causal temporal

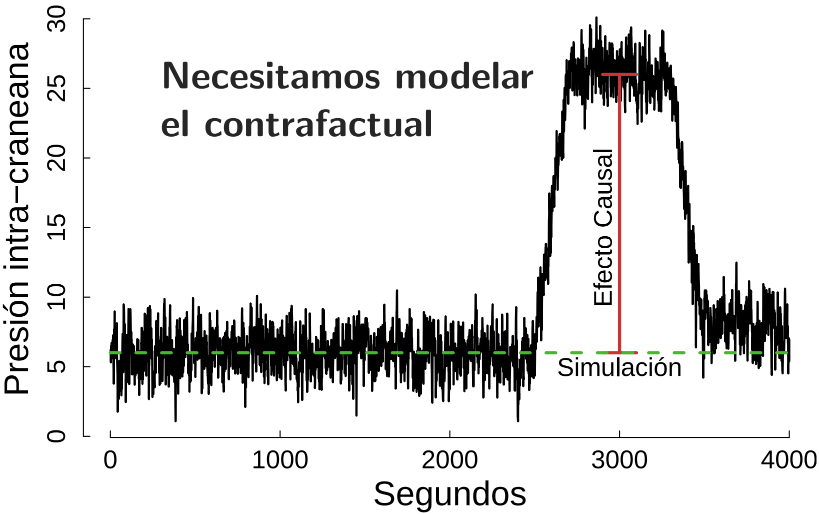


Efecto causal temporal



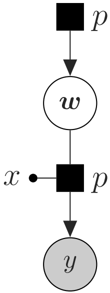
Efecto causal temporal

El control sintético



Modelos lineales para series temporales

Modelo lineal básico

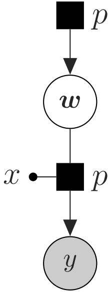


$$p(\boldsymbol{w}) = \prod_i \mathcal{N}(\, w_i \,|\, 0,\, \sigma^2)$$

$$p(y|\boldsymbol{w}, x) = \mathcal{N}(\, y \,|\, w_0 + w_1\, x_1 + w_2\, x_2^2 + w_3\, x_3^3,\, \beta^2)$$

Modelos lineales para series temporales

Modelo lineal básico

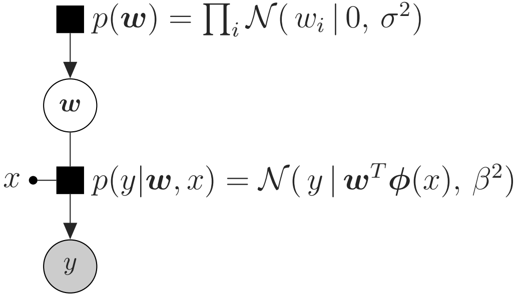


$$p(\boldsymbol{w}) = \prod_i \mathcal{N}(\, w_i \,|\, 0,\ \sigma^2)$$

$$p(y|\boldsymbol{w}, x) = \mathcal{N}(\, y \,|\ \sum_i w_i \phi_i(x),\ \beta^2)$$

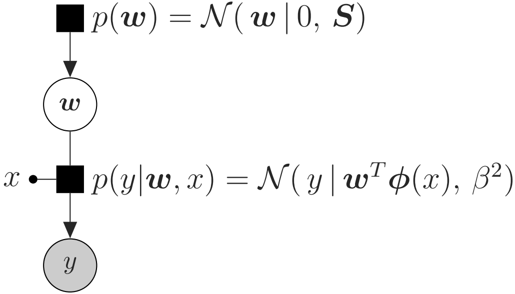
Modelos lineales para series temporales

Modelo lineal básico



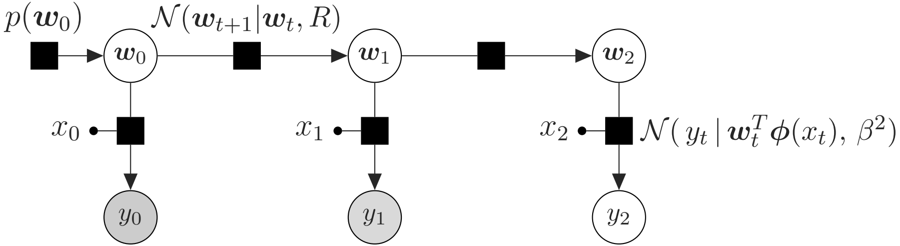
Modelos lineales para series temporales

Modelo lineal básico



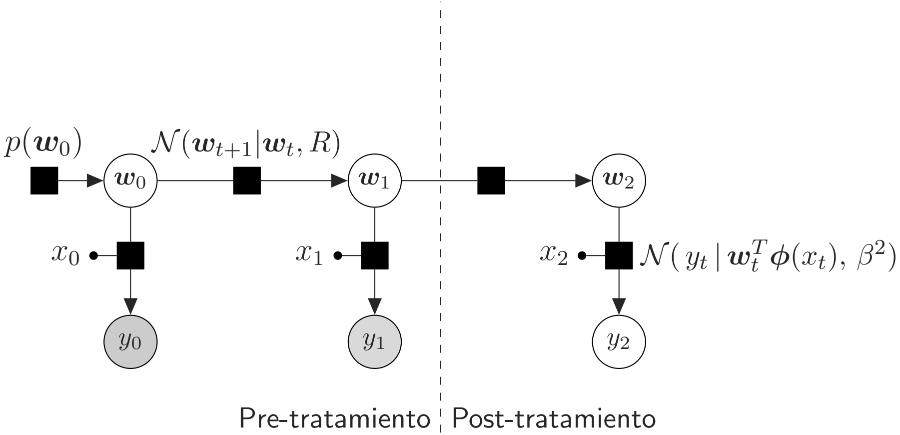
Modelos lineales para series temporales

Structural state-space models



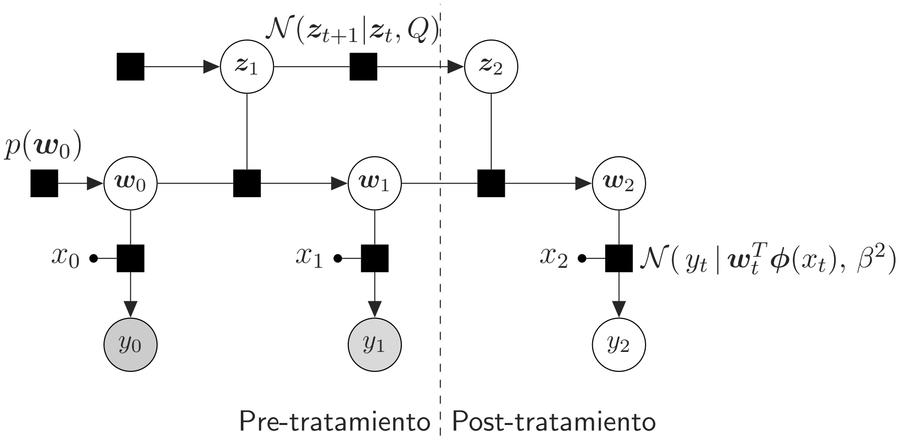
Modelos lineales para series temporales

Structural state-space models

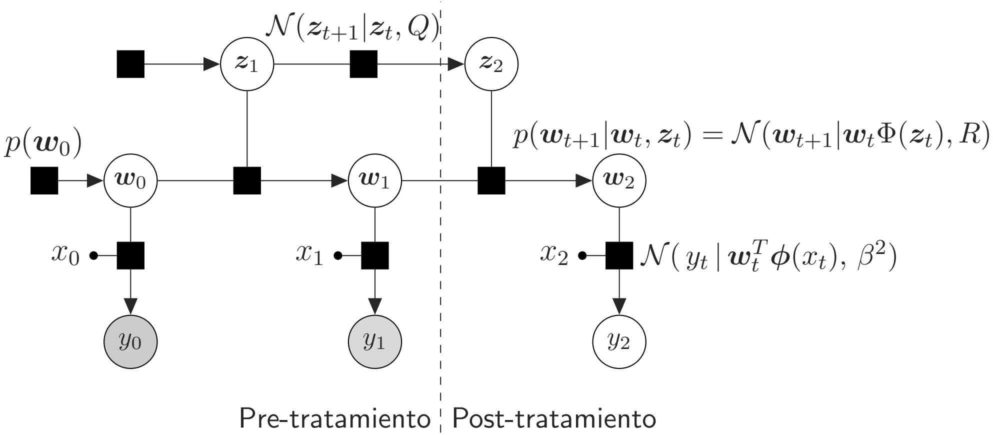


Modelos lineales para series temporales

Structural state-space models

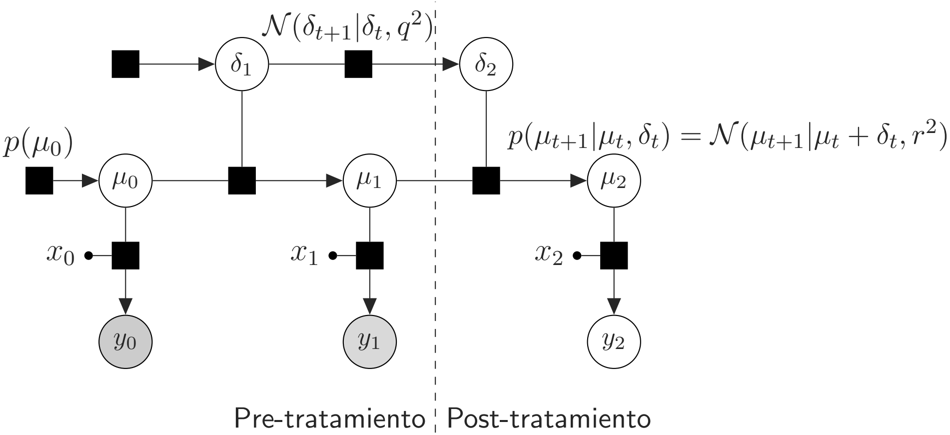


Modelos lineales para series temporales
Structural state-space models



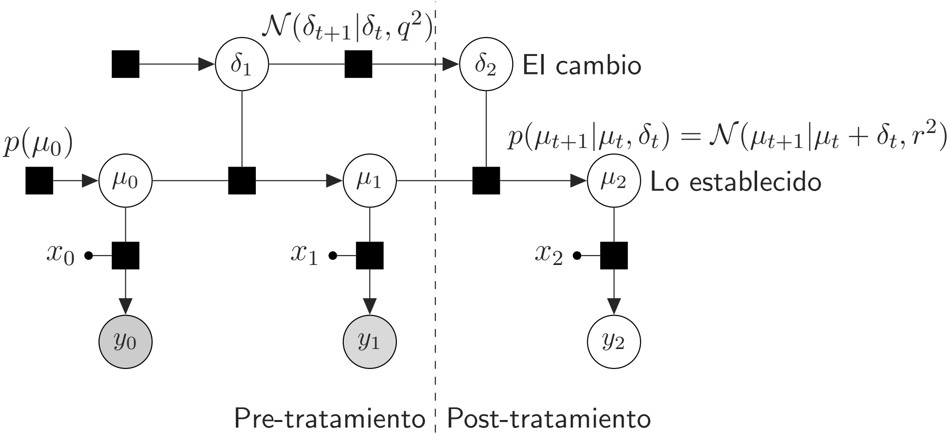
Modelos lineales para series temporales

Tendencia lineal local



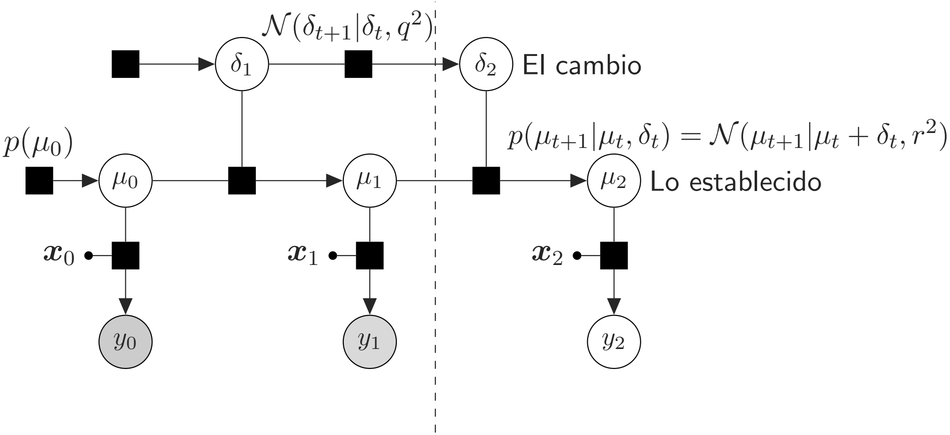
Modelos lineales para series temporales

Tendencia lineal local



Modelos lineales para series temporales

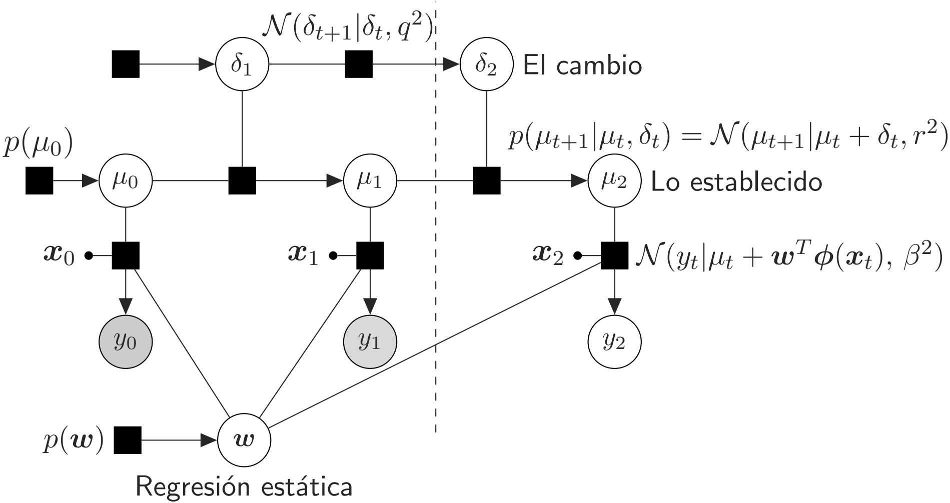
Tendencia lineal local



¿Y la relación con otras series temporales?

Modelos lineales para series temporales

Tendencia lineal local + controles externos



Modelos lineales para series temporales

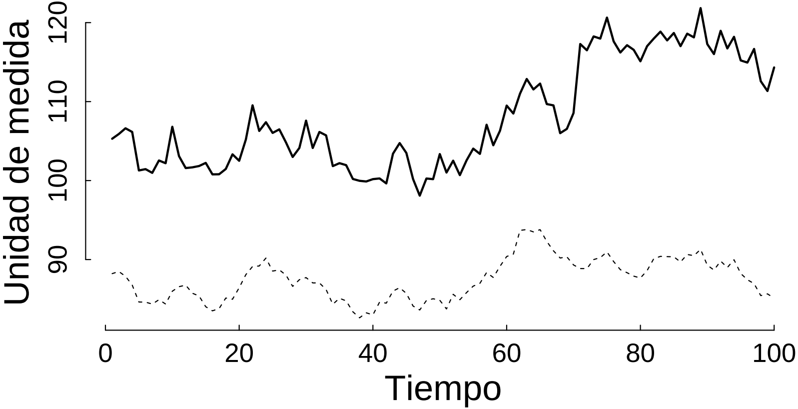
Ejemplo. Generación de datos.

```
set.seed(1)
x <- 100 + arima.sim(model = list(ar = 0.999), n = 100)
y <- 1.2 * x + rnorm(100)
y[71:100] <- y[71:100] + 10 # Intervencion
```

Modelos lineales para series temporales

Ejemplo. Generación de datos.

```
set.seed(1)
x <- 100 + arima.sim(model = list(ar = 0.999), n = 100)
y <- 1.2 * x + rnorm(100)
y[71:100] <- y[71:100] + 10 # Intervencion
```



Modelos lineales para series temporales

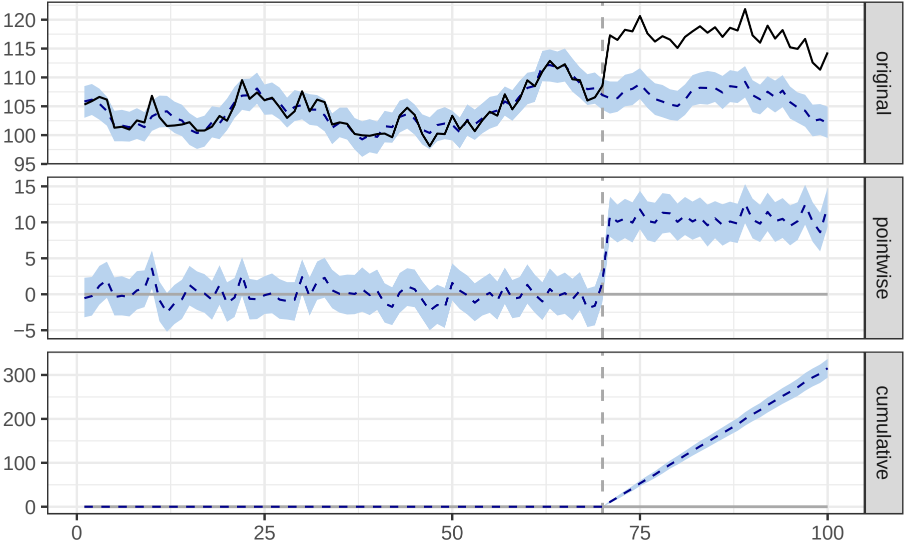
Ejemplo. Inferencia.

```
library("CausalImpact")
# Google "Inferring causal impact using
# bayesian structural time-series models"

data <- cbind(y, x)
pre.period <- c(1, 70)
post.period <- c(71, 100)
impact <- CausalImpact(data, pre.period, post.period)
```

Modelos lineales para series temporales

Ejemplo. Inferencia.



Limitaciones
Modelo lineal

La transformaciones no-lineales $\Phi(x)$ están fijas.

$$f(x, \boldsymbol{w}) = \boldsymbol{w}^T \Phi(x)$$

Limitaciones

Modelo lineal

La transformaciones no-lineales $\Phi(x)$ están fijas.

$$f(x, \boldsymbol{w}) = \boldsymbol{w}^T \Phi(x)$$

Necesitaríamos infinitos modelos
para poder representar cualquier función!

Procesos Gaussianos

Evaluación de infinitos modelos $f(\boldsymbol{x})$

$p(y_i) = \mathcal{N}(y_i | f(\boldsymbol{x}_i), \beta^2)$ donde f es la función objetivo

Prior sobre los modelos:

$$f(\boldsymbol{x}) \sim \mathcal{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}'))$$

Procesos Gaussianos

Evaluación de infinitos modelos $f(\boldsymbol{x})$

$p(y_i) = \mathcal{N}(y_i | f(\boldsymbol{x}_i), \beta^2)$ donde f es la función objetivo

Prior sobre los modelos:

$$f(\boldsymbol{x}) \sim \mathcal{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}'))$$

$$m(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})]$$

$$k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x}') - m(\boldsymbol{x}'))]$$

Procesos Gaussianos

Evaluación de infinitos modelos $f(\boldsymbol{x})$

$p(y_i) = \mathcal{N}(y_i | f(\boldsymbol{x}_i), \beta^2)$ donde f es la función objetivo

Prior sobre los modelos:

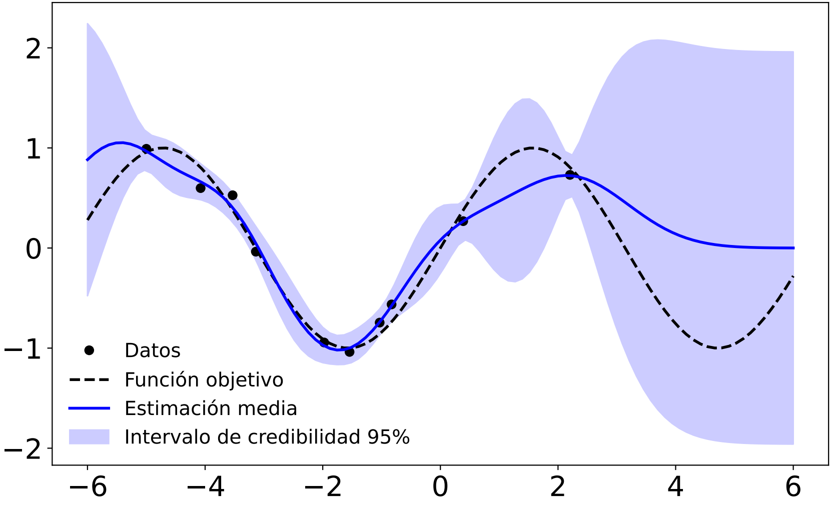
$$f(\boldsymbol{x}) \sim \mathcal{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}'))$$

$$m(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})]$$
$$k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x}') - m(\boldsymbol{x}'))]$$

Posterior sobre los modelos tiene solución analítica!

Procesos Gaussianos

Evaluación de infinitos modelos $f(x)$



Procesos Gaussianos

y Redes Neuronales

Procesos Gaussianos
y Redes Neuronales

Se han demostrado las siguientes equivalencia

Procesos Gaussianos
y Redes Neuronales

Se han demostrado las siguientes equivalencia

- Neal, RM. 1994. *Priors for infinite networks*. Universidad de Toronto.

Equivalencia entre una red neuronal de una sola capa totalmente conectada infinitamente ancha y los proceso gaussiano.

Procesos Gaussianos
y Redes Neuronales

Se han demostrado las siguientes equivalencia

- Neal, RM. 1994. *Priors for infinite networks*. Universidad de Toronto.

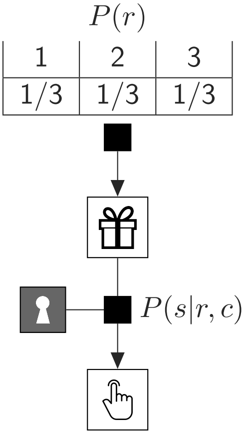
Equivalencia entre una red neuronal de una sola capa totalmente conectada infinitamente ancha y los proceso gaussiano.

- Lee et al (Google Brain). 2018. *Deep Neural Networks as Gaussian Process.*

International Conference on Learning Representations

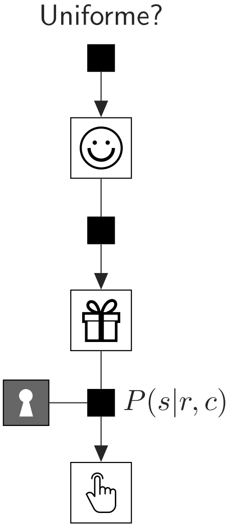
Derivan la equivalencia exacta entre redes infinitamente profundas y GPs.

Monty Hall



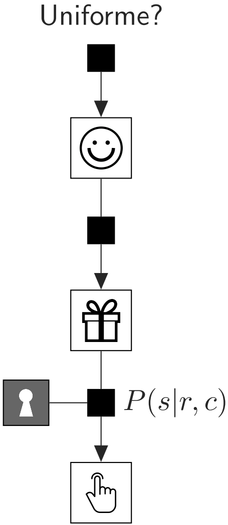
Monty Hall

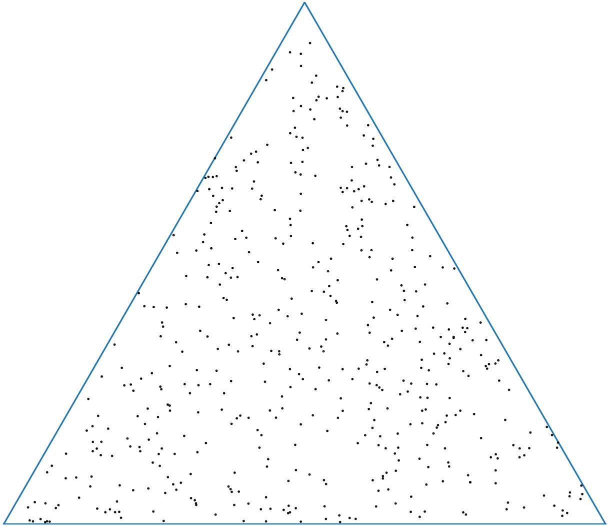
Con sesgo



Monty Hall

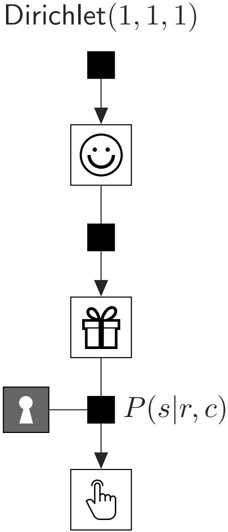
Con sesgo

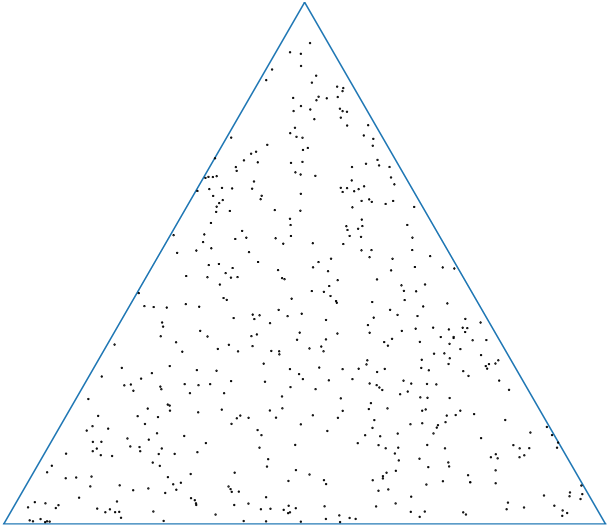




Monty Hall

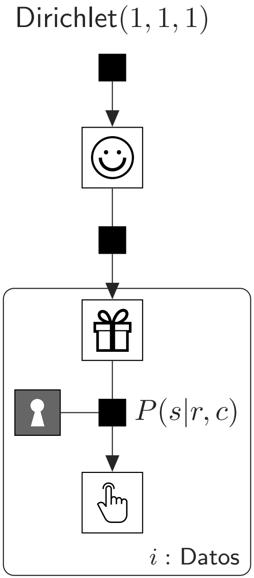
Con sesgo

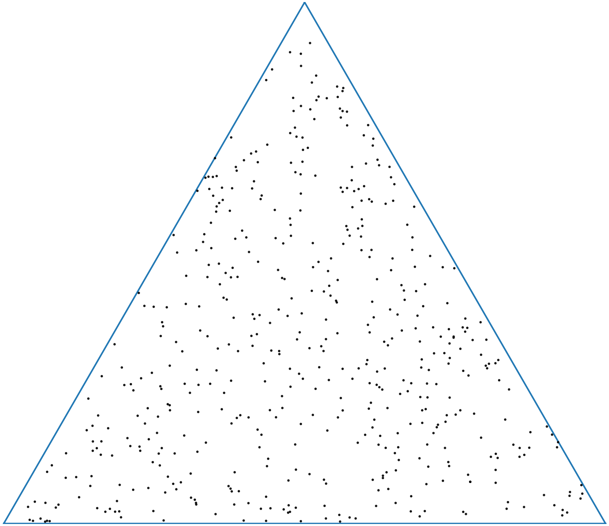




Monty Hall

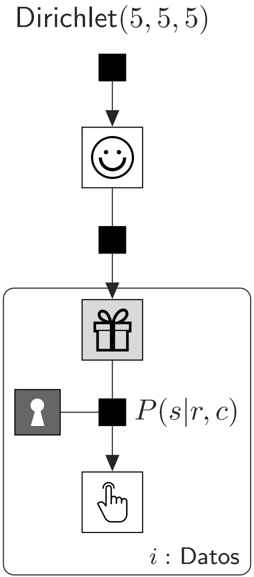
Con sesgo

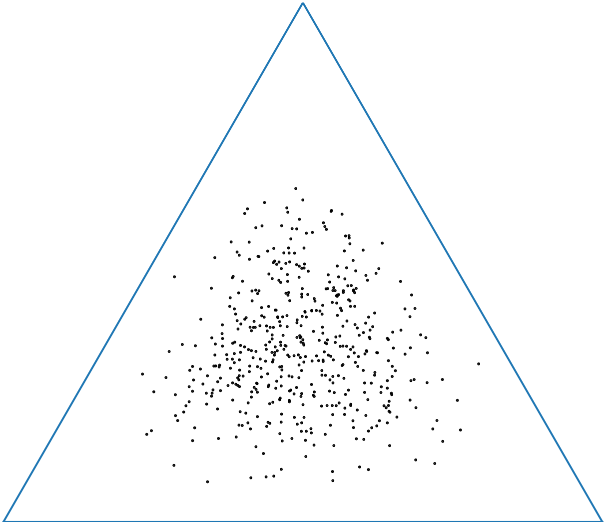




Monty Hall

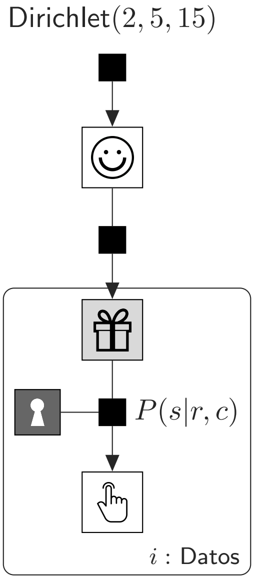
Con sesgo

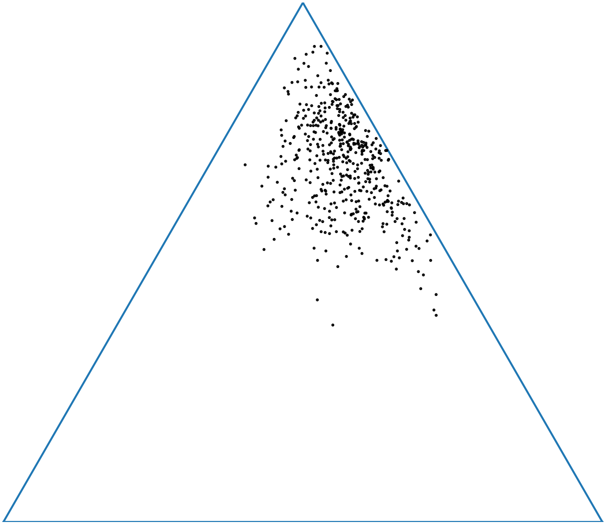




Monty Hall

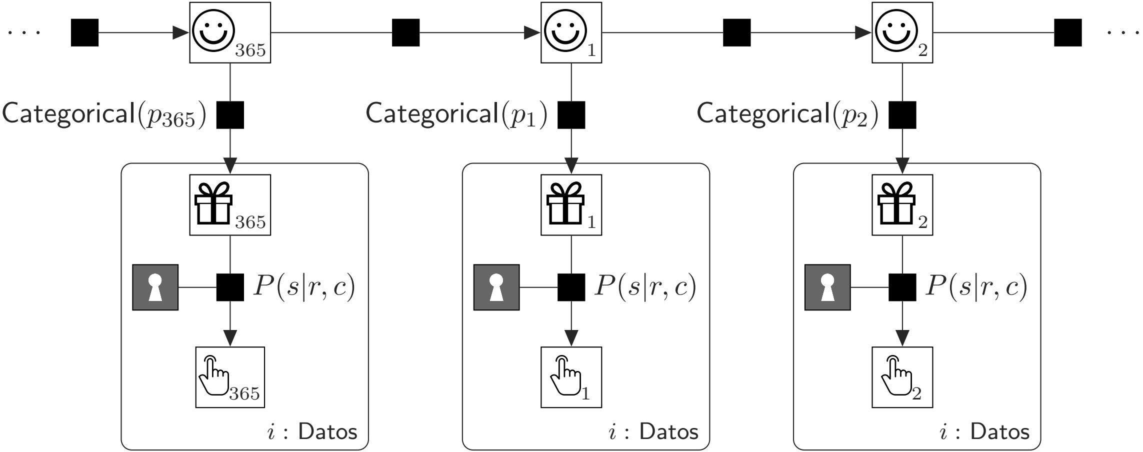
Con sesgo





Monty Hall

Con sesgo dinámico



$p = b$

Laboratorios de

Métodos Bayesianos