Mixturas de Exponenciales Truncadas para modelizar distribuciones continuas

Departamento de Estadística y Matemática Aplicada

Universidad de Almería

Índice

- **Aprendizaje**:
  - *Estimadores kernel.*
  - *Aprendizaje estructural.*

Índice

- **Aprendizaje**:
  - *Estimadores kernel.*
  - *Aprendizaje estructural.*
- **Inferencia**:
  - *Propagación Penniless.*

Índice

- **Aprendizaje**:
  - *Estimadores kernel.*
  - *Aprendizaje estructural.*
- **Inferencia**:
  - *Propagación Penniless.*
- **En curso**.

Estimadores kernel

Introducción

Estimación de densidad univariante MTE

$x_1, \ldots, x_n$

Muestra de X

$\downarrow$

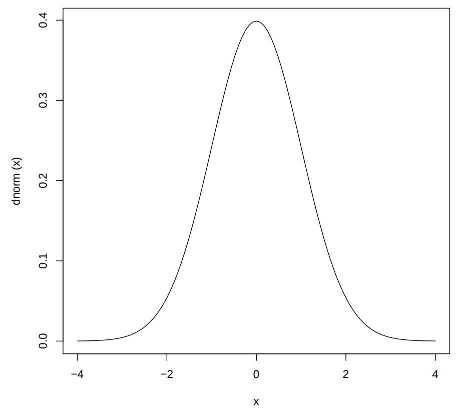
$f(x)$ *Densidad*

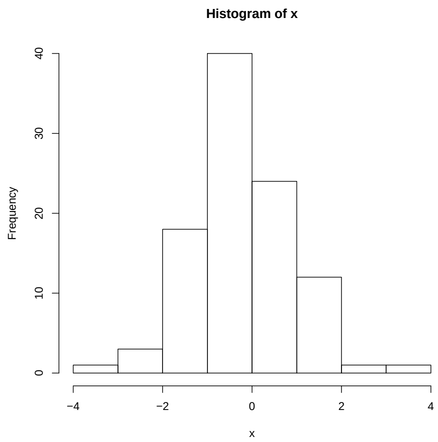
Introducción

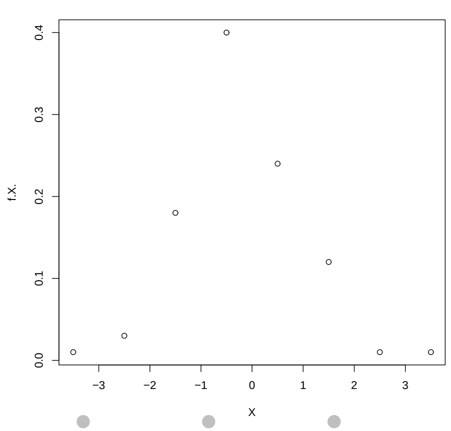
Estimación de densidad univariante MTE

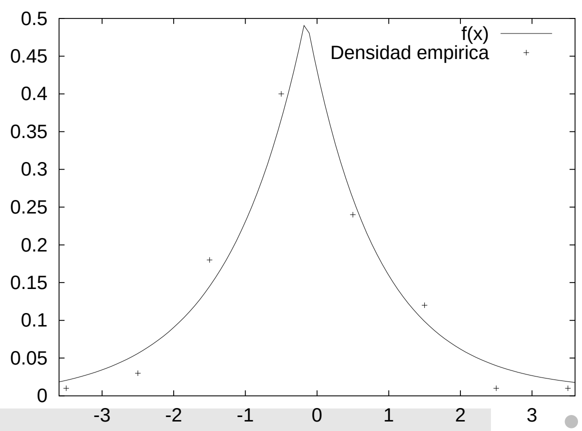
$$\begin{tikzcd}[column sep=huge, row sep=huge]
\begin{array}{c} x_1,\ldots,x_n \\ \text{Muestra de } x \end{array} \arrow[r, "\text{Densidad}", "\text{empírica}"'] & \begin{array}{ll} x_1,\ldots,x_m & \text{valores de } x \\ y_1,\ldots,y_m & \text{valores de } f(x) \end{array} \arrow[d, "\begin{array}{l}\text{Regresión}\\ \text{mínimo}\\ \text{cuadrática}\end{array}"] \\
f(x)\ \text{Densidad} & k+aexp(bx)+cexp(dx) \arrow[l, "\text{Normalizando}"']
\end{tikzcd}$$

Introducción

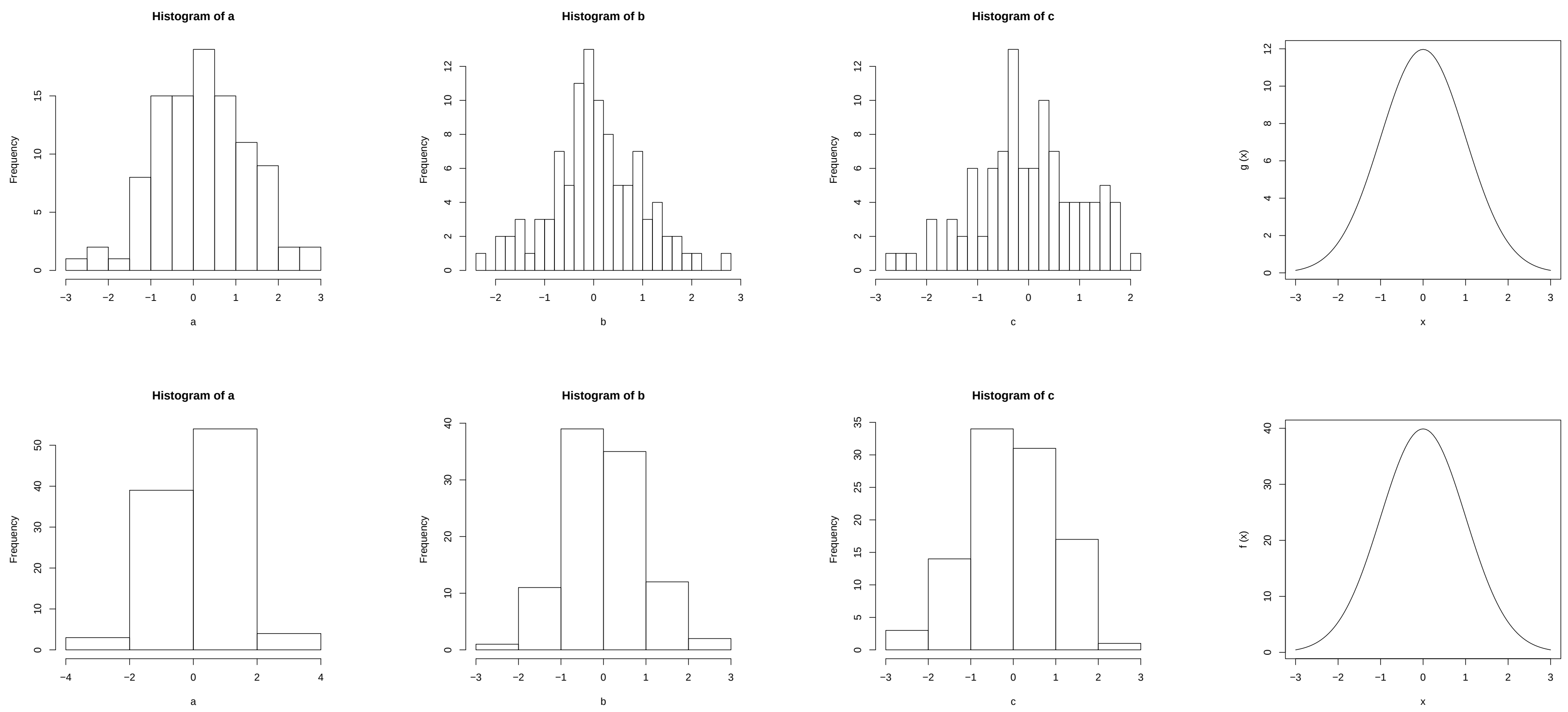








Introducción



Forma del histograma

- *Número de intervalos:*
  - *Pocos: Mala regresión / densidad suave.*
  - *Muchos: Mejor regresión / densidad brusca.*

Forma del histograma

- *Número de intervalos:*
  - *Pocos: Mala regresión / densidad suave.*
  - *Muchos: Mejor regresión / densidad brusca.*
- *Selección de los intervalos*
  - *Igual amplitud.*
  - **Igual frecuencia**.

Kernel

Estimación de densidad univariante MTE

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\begin{array}{c} x_1,\ldots,x_n \\ \text{Muestra de } x \end{array} \arrow[r, "\text{Estimador}", "\text{Kernel}"'] & \begin{array}{ll} x_1,\ldots,x_m & \text{valores de } x \\ y_1,\ldots,y_m & \text{valores de } f(x) \end{array} \arrow[d, "\begin{array}{l}\text{Regresión}\\ \text{mínimo}\\ \text{cuadrática}\end{array}"] \\
f(x) \ \text{Densidad} & k+a\exp(bx)+c\exp(dx) \arrow[l, "\text{Normalizando}"']
\end{tikzcd}$$

Kernel

$\widehat{f}(x)$ *es una* **densidad kernel** *:*

$$\widehat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h})$$

con K una **función kernel** *Propiedades de K :*

- $\int K(u)du = 1$
- $\int uK(u)du = 0$
- $\int u^2 K(u)du = \sigma_K^2$

Datos $\leftrightarrow h \leftrightarrow \widehat{f}$ *kernel (suavizar)*

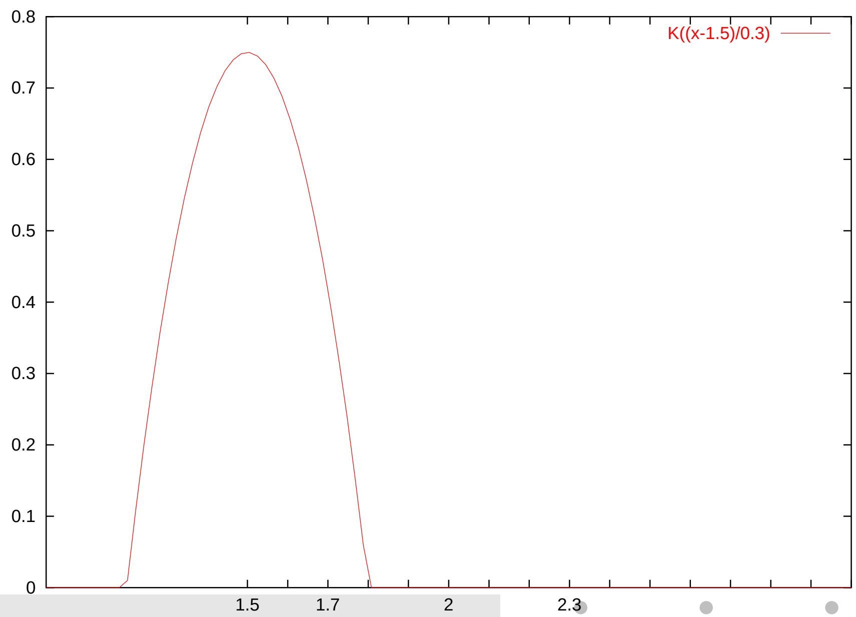
Kernels

Kernel	*Form*
Epanechnikov	$\frac{3}{4}(1-u^2)$
Biweight	$\frac{15}{16}(1-u^2)^2$
Triweight	$\frac{35}{32}(1-u^2)^3$
Gaussian	$(2\pi)^{-1/2}e^{-u^2/2}$
Uniform	$\frac{1}{2}$

Ejemplo

$$\textit{Muestra:}\ 1.5,\ 1.7,\ 2.0,\ 2.3 \qquad h = 0.3$$

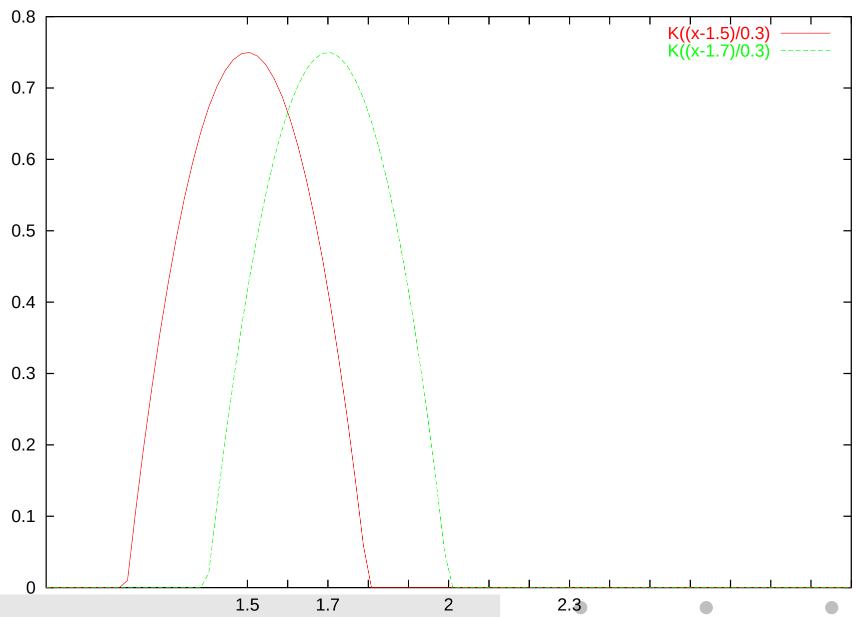
$$K(u) = \frac{3}{4}(1-u^2) \qquad |u| \leq 1$$



Ejemplo

$$\textit{Muestra:}\ 1.5,\ 1.7,\ 2.0,\ 2.3 \qquad h = 0.3$$

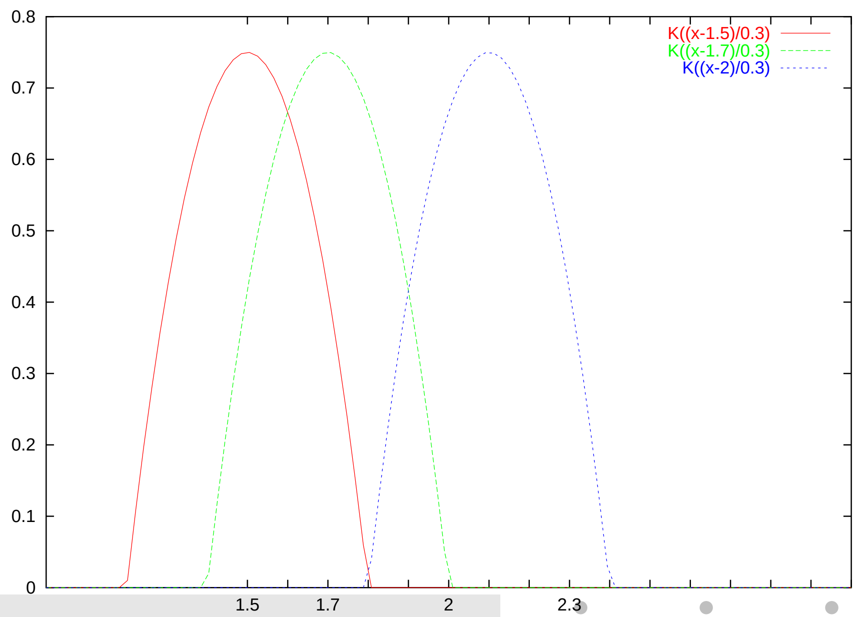
$$K(u) = \frac{3}{4}(1-u^2) \qquad |u| \leq 1$$



Ejemplo

$$\textit{Muestra:}\ 1.5,\ 1.7,\ 2.0,\ 2.3 \qquad h = 0.3$$

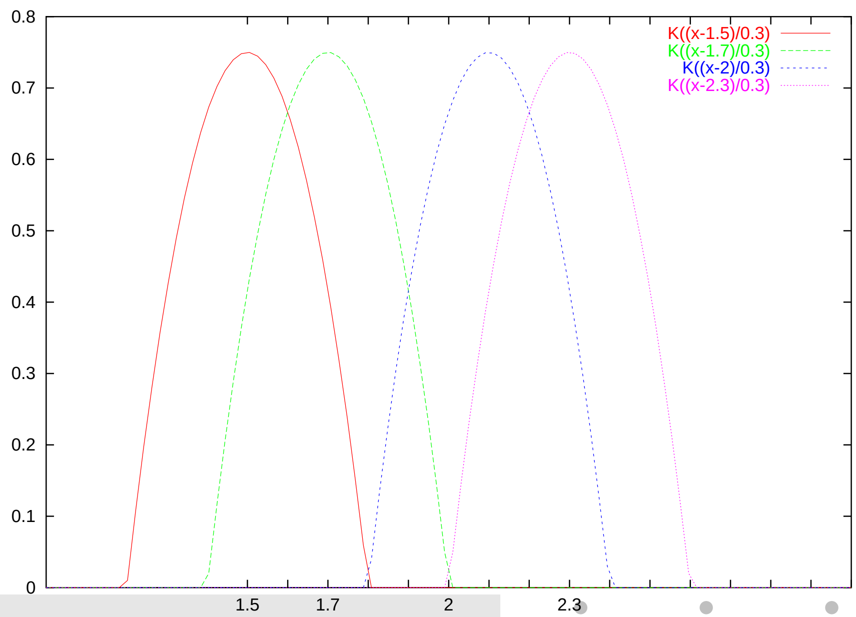
$$K(u) = \frac{3}{4}(1-u^2) \qquad |u| \leq 1$$



Ejemplo

$$\textit{Muestra:}\, 1.5,\ 1.7,\ 2.0,\ 2.3 \qquad h = 0.3$$

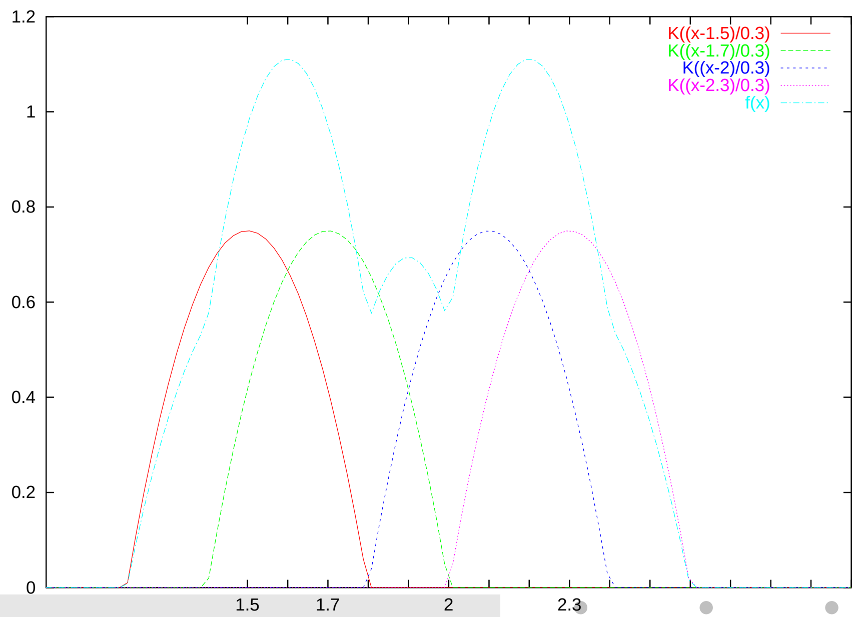
$$K(u) = \frac{3}{4}(1-u^2) \qquad |u| \leq 1$$



Ejemplo

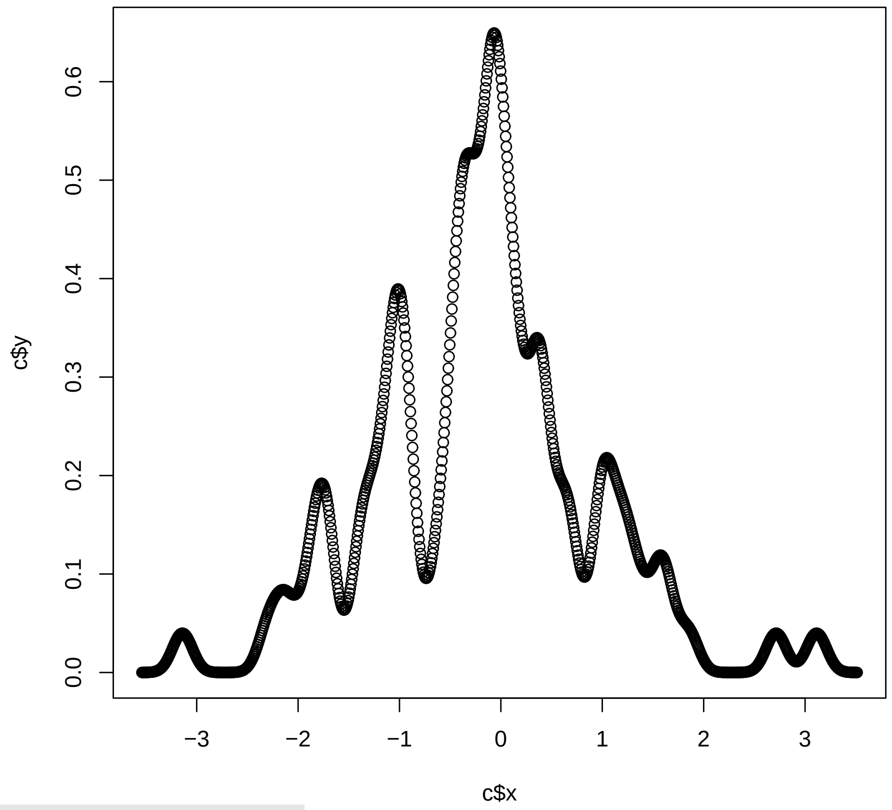
$$\textit{Muestra:}\ 1.5,\ 1.7,\ 2.0,\ 2.3 \qquad h = 0.3$$

$$K(u) = \frac{3}{4}(1-u^2) \qquad |u| \leq 1$$



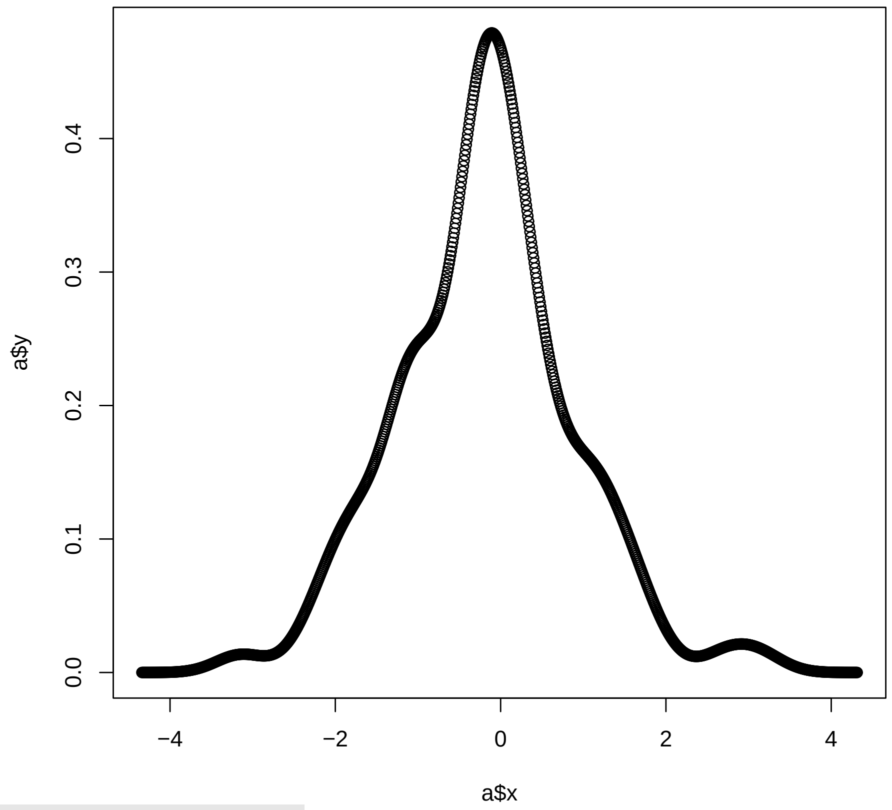
Ejemplo

$$h = 0.1$$



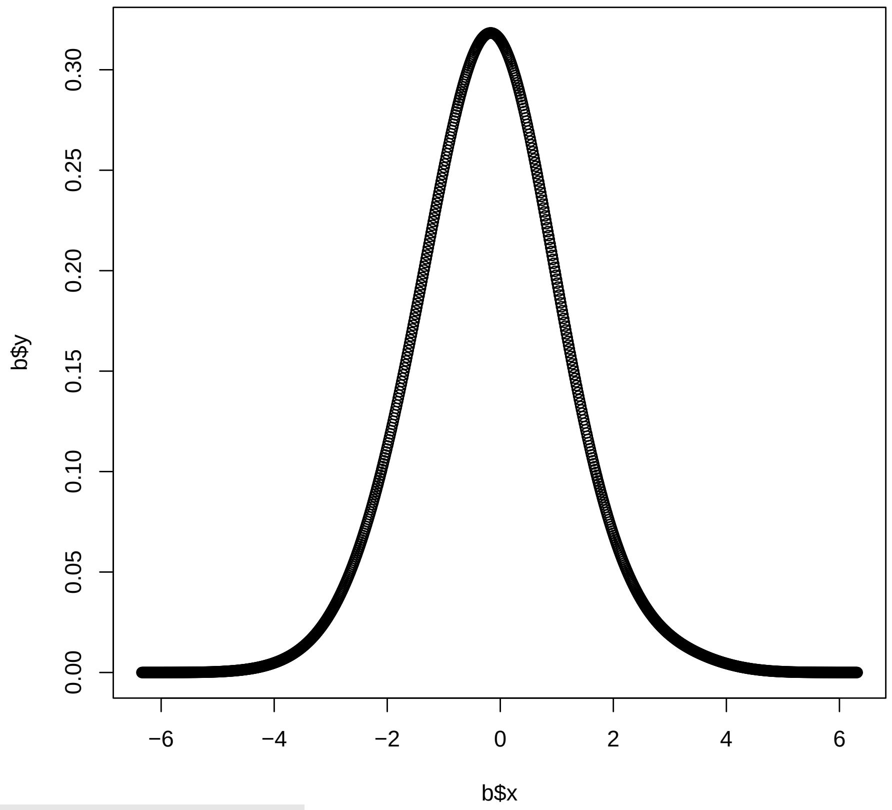
Ejemplo

$$h = 0.3$$



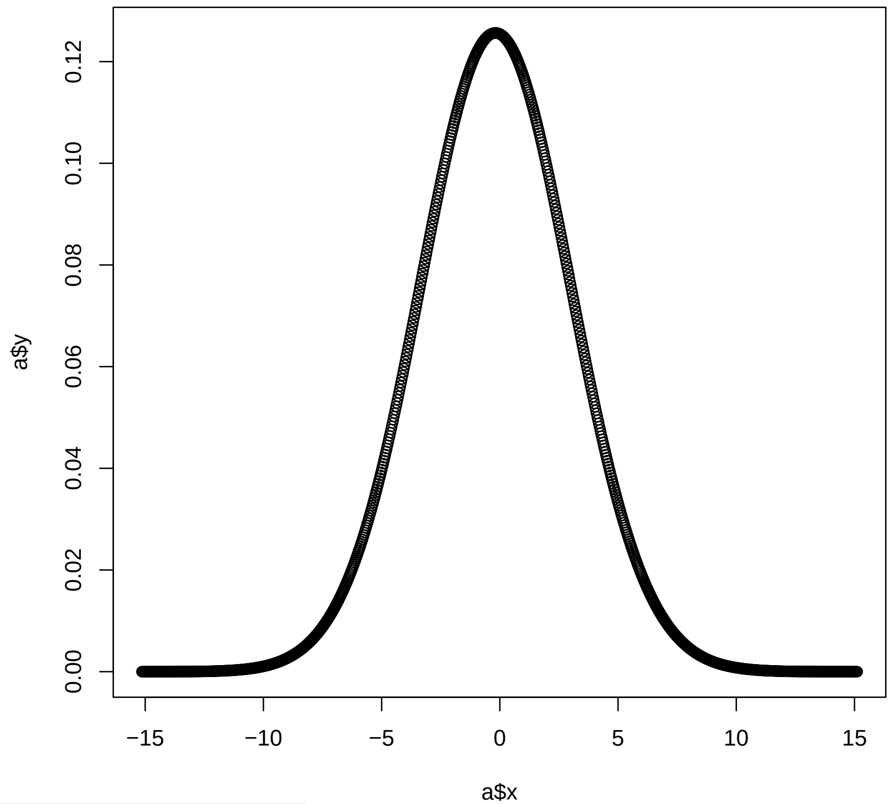
Ejemplo

$$h = 0.8$$



Ejemplo

$$h = 3$$



Error

Medida de error

- $MSE(x) = E_f[\widehat{f}(x) - f(x)]^2.$

Si K es Gaussiano , h minimizando el error es

$$h_0 = 1.059\sigma n^{-1/5}$$

Escoger h

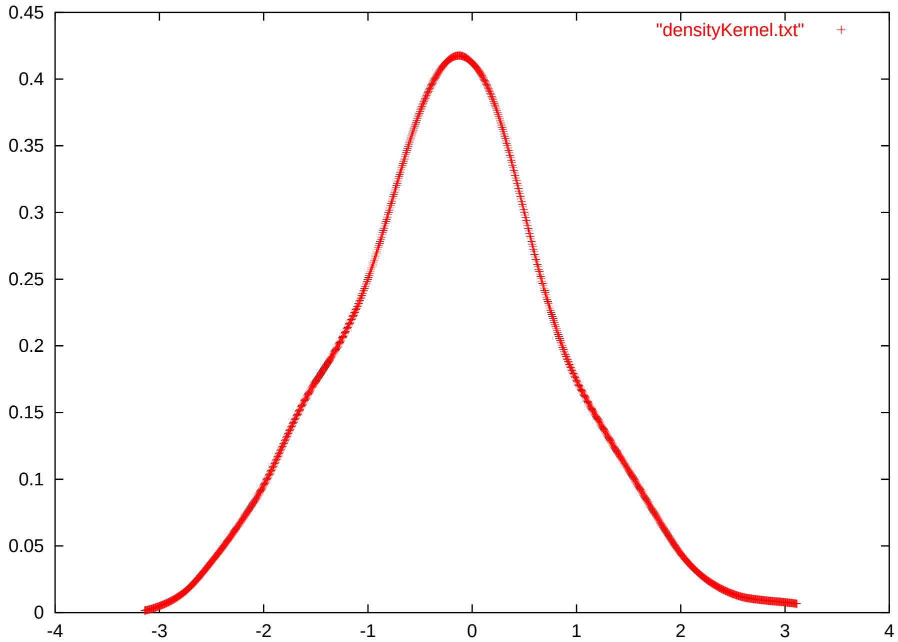
Para cualquier otro kernel K^*, el óptimo h es

$$h_{0,K^*} = C_K * h_{0,G}$$

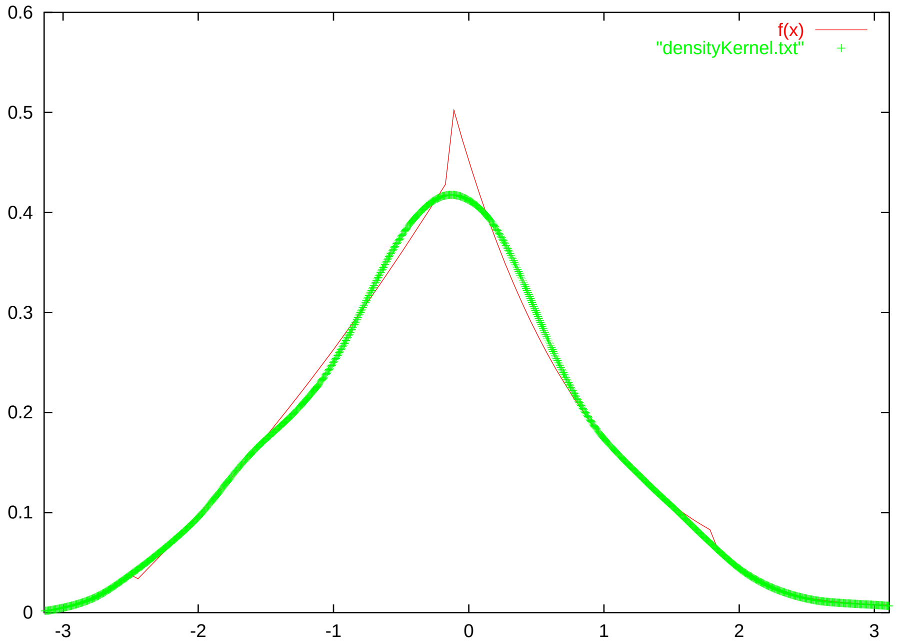
dónde $h_{0,G}$ *es el h del Kernel Gaussiano y* C_K :

Kernel	C_K
Epanechnikov	2.214
Biweight	2.623
Triweight	2.978
Uniform	1.740

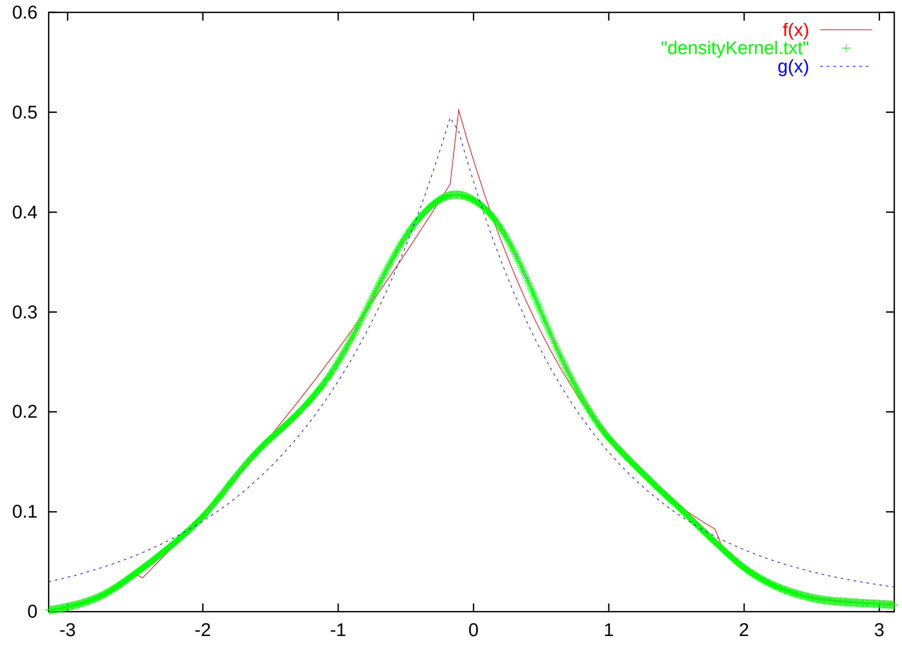
Uso de Kernel



Uso de Kernel



Uso de Kernel



Elvira

KernelDensity (learning/KernelDensity.java)

- *Campos*
  - *Muestra*
  - *Amplitud h (por defecto la que minimiza el error para el K gaussiano).*
  - *Kernel a usar. (por defecto gaussiano).*

KernelDensity (learning/KernelDensity.java)

- *Campos*
  - *Muestra*
  - *Amplitud h (por defecto la que minimiza el error para el K gaussiano).*
  - *Kernel a usar. (por defecto gaussiano).*
- *Métodos:*
  - *getGaussianH(Vector X): Obtiene el valor de h que minimiza el error para K Gaussiano.*
  - *assignOptimalH(): Asigna el valor de h óptimo para el kernel escogido.*
  - *getValues(Vector A): Obtiene a partir de la muestra los vectores* $x_1, \ldots, x_m$ *e* $y_1, \ldots, y_m$.
  - *Constructores, métodos de acceso, etc.*

Aprendizaje Estructural

Aprendizaje estructural

- $X = \{X_1, \ldots, X_n\}$
- $D = \{x^{(1)}, \ldots, x^{(m)}\}$ *Muestra*
- *Objetivo: Encontrar la red Bayesiana G de variables X que mejor concuerde con D.*

Aprendizaje Estructural

Problemas a resolver:

- *Determinar una estructura candidata.*
- *Estimar las distribuciones condicionadas* $\hat{\theta}$ *para* G.
- *Medir la precisión de* $(G, \hat{\theta})$.

Buscar estructura

- *Buscar en el espacio de redes posibles.*
- *Simulated annealing.*
- *Empezamos con una red vacía.*
- *Operaciones permitidas:*
  - *Inserción de un arco.*
  - *Borrado de un arco.*
  - *Inversión de un arco.*
- *Tras un movimiento se reestiman las distribuciones de probabilidad.*

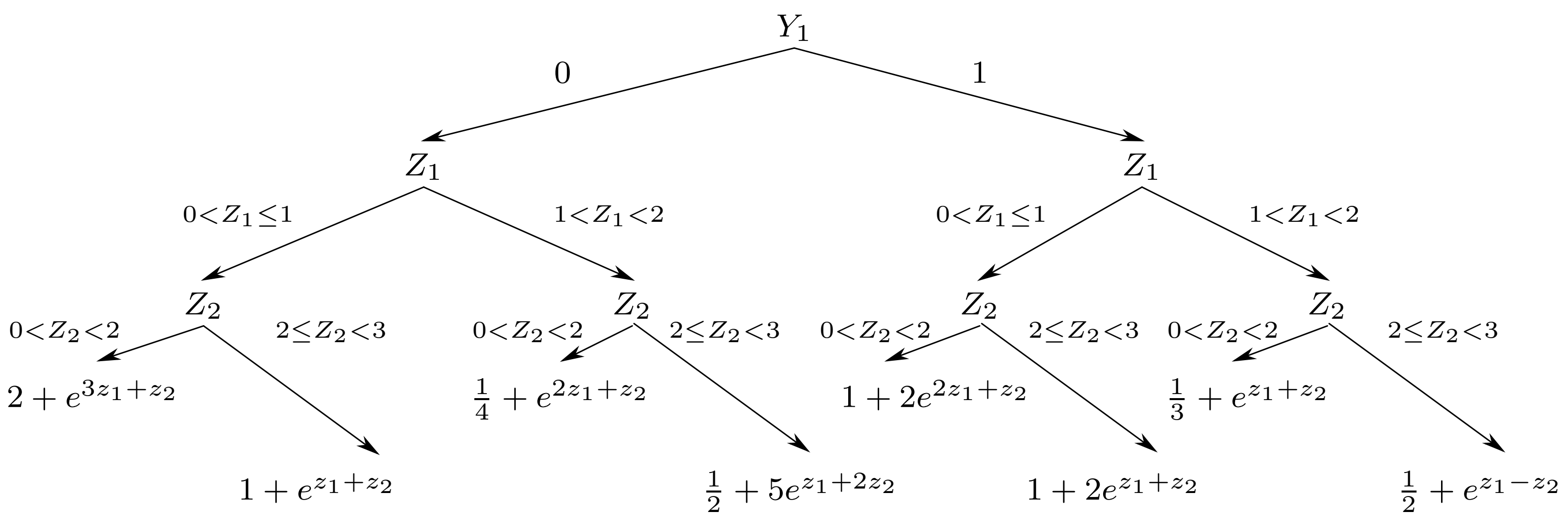
Escoger el mejor movimiento

Maximizar una medida de calidad:

$$Q(G|D,\hat{\theta}) = log\ L(D;G,\hat{\theta}) - \frac{log\ m}{2} Dim(G)$$

- $L(D;G,\hat{\theta})$: *Verosimilitud de los datos dada la red.*
- $Dim(G)$: *Número de parámetros necesarios para especificar la red G.*

Ejemplo


Y_1
0
1
Z_1
Z_1
$0<Z_1\leq 1$
$1<Z_1<2$
$0<Z_1\leq 1$
$1<Z_1<2$
Z_2
Z_2
Z_2
Z_2
$0<Z_2<2$
$2\leq Z_2<3$
$0<Z_2<2$
$2\leq Z_2<3$
$0<Z_2<2$
$2\leq Z_2<3$
$0<Z_2<2$
$2\leq Z_2<3$
$2+e^{3z_1+z_2}$
$1+e^{z_1+z_2}$
$\frac{1}{4}+e^{2z_1+z_2}$
$\frac{1}{2}+5e^{z_1+2z_2}$
$1+2e^{2z_1+z_2}$
$1+2e^{z_1+z_2}$
$\frac{1}{3}+e^{z_1+z_2}$
$\frac{1}{2}+e^{z_1-z_2}$

Parámetros

- *k : Número de términos exponenciales en las hojas (fijo)*
- *$|X|$: Número de estados si X es discreta, número de trozos en los que se divide el dominio si $|X|$ es continua (fijo).*
- $Par(fa(X)) = k \prod_{Y \in fa(X)} |Y|$ *si X es continua.*
- $Par(fa(X)) = |X| \prod_{Y \in pa(X)} |Y|$ *si X es discreta.*

Cortes

- *s : Número de intervalos en los que se divide el dominio de variables continuas (fijo)*
- *s-1: Número de valores a estimar.*
- *Sólo para las variables continuas en* $fa(X)$.
- *Cota superior:* $Spl(fa(X)) = (s-1)^{N_c} \prod_{discretas} d_i$.
  - N_c*: Número de vbles. continuas en* $fa(X)$.
  - d_i *: Número de estados de las vble. discreta i de* $fa(X)$.

$Dim(G)$

- $$Dim(G) = \sum_{i=1}^{n} Dim(fa(X_i))$$

- $$Dim(fa(X)) = Par(fa(X)) + Spl(fa(X))$$

Descomposición de la métrica

$$Q(G|D,\hat{\theta}) = \sum_{j=1}^{n} Q(X_j|D,\hat{\theta})$$

con

$$Q(X_j|D,\hat{\theta}) = \sum_{i=1}^{m} log\ p_j(x_j^{(i)}|pa(x_j^{(i)})) - \frac{log\ m}{2} Dim(fa(X_j))$$

Sólo hay que recalcular $Q(X_j|D,\hat{\theta})$ para la vble. X_j afectada por el movimiento.

Experimentos

Experimentos

Se crearon 4 redes:

	net5_1	net5_2	net10_1	net10_2
No. de enlaces	6	7	13	11
No. de vbles. discretas	1	1	2	
No. de vbles. continuas	4	4	8	8

Se obtuvieron de cada red muestras de tamaño 100, 500 y 1000.

Resultados

Tamaño muestra	*100*	*500*	*1000*
LL	*7*	*8*	*6*
CL	*2*	*2*	*1*
IL	*3*	*3*	*1*
NL	*2*	*3*	*4*

net5_1

Resultados

Tamaño muestra	*100*	*500*	*1000*
LL	*7*	*5*	*8*
CL	*2*	*0*	*4*
IL	*2*	*3*	*1*
NL	*3*	*2*	*3*

net5_2

Resultados

Tamaño muestra	*100*	*500*	*1000*
LL	*21*	*20*	*22*
CL	*3*	*4*	*7*
IL	*4*	*2*	*3*
NL	*14*	*14*	*12*

net10_1

Resultados

Tamaño muestra	100	500	1000
LL	15	25	23
CL	3	3	6
IL	3	5	3
NL	9	7	14

net10_2

Elvira

StructuralMTELearning (learning/)

- *Campos*
  - *Database*
  - *Bnet a aprender.*

StructuralMTELearning (learning/)

- *Campos*
  - *Database*
  - *Bnet a aprender.*
- *Métodos:*
  - *structuralLearning : Aprende la red Bayesiana a partir del DatabaseCases*
  - *semiQuality(Node y, PotentialContinuousPT p) : Calcula la calidad para la vble. y.*
  - *dimension(Node var): Calcula la dimensión para la vble. var.*

Inferencia

Propagación de probabilidades

- *Operaciones de los potenciales:*
  a. *Restricción*
  b. *Multiplicación*
  c. *Marginalización*
  d. *División*

Propagación de probabilidades

- *Operaciones de los potenciales:*
  a. *Restricción*
  b. *Multiplicación*
  c. *Marginalización*
  d. ~~*División*~~

  La clase de los potenciales MTE es cerrada para restricción, multiplicación y marginalización.

Propagación de probabilidades

- *Operaciones de los potenciales:*
  a. *Restricción*
  b. *Multiplicación*
  c. *Marginalización*
  d. ~~*División*~~

  La clase de los potenciales MTE es cerrada para restricción, multiplicación y marginalización.
- *Algoritmos que no requieren división:*
  - *Shenoy-Shafer.*
  - *Lazy propagation .*
  - *Eliminación de variables.*

Propagación de probabilidades

- *Operaciones de los potenciales:*
  a. *Restricción*
  b. *Multiplicación*
  c. *Marginalización*
  d. ~~*División*~~

  La clase de los potenciales MTE es cerrada para restricción, multiplicación y marginalización.
- *Algoritmos que no requieren división:*
  - *Shenoy-Shafer: Cobb and Shenoy (2003) y Cobb, Shenoy y Rumí (2004).*

Esquema Shenoy-Shafer

- *Red bayesiana*

Esquema Shenoy-Shafer

- *Red bayesiana*
- *Join Tree*

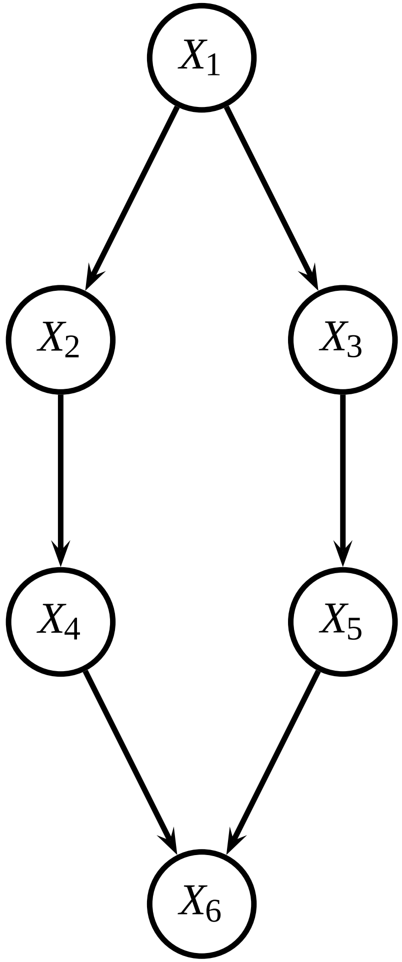
Esquema Shenoy-Shafer

- *Red bayesiana*
- *Join Tree*
- *Mensajes*

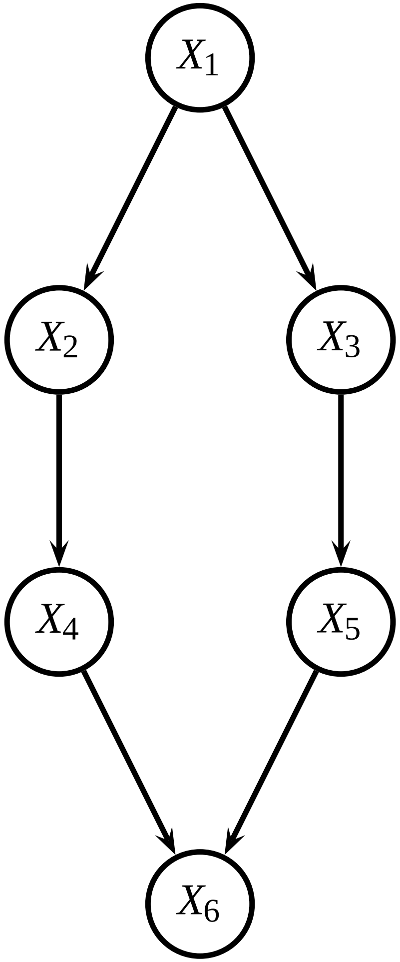
Esquema Shenoy-Shafer

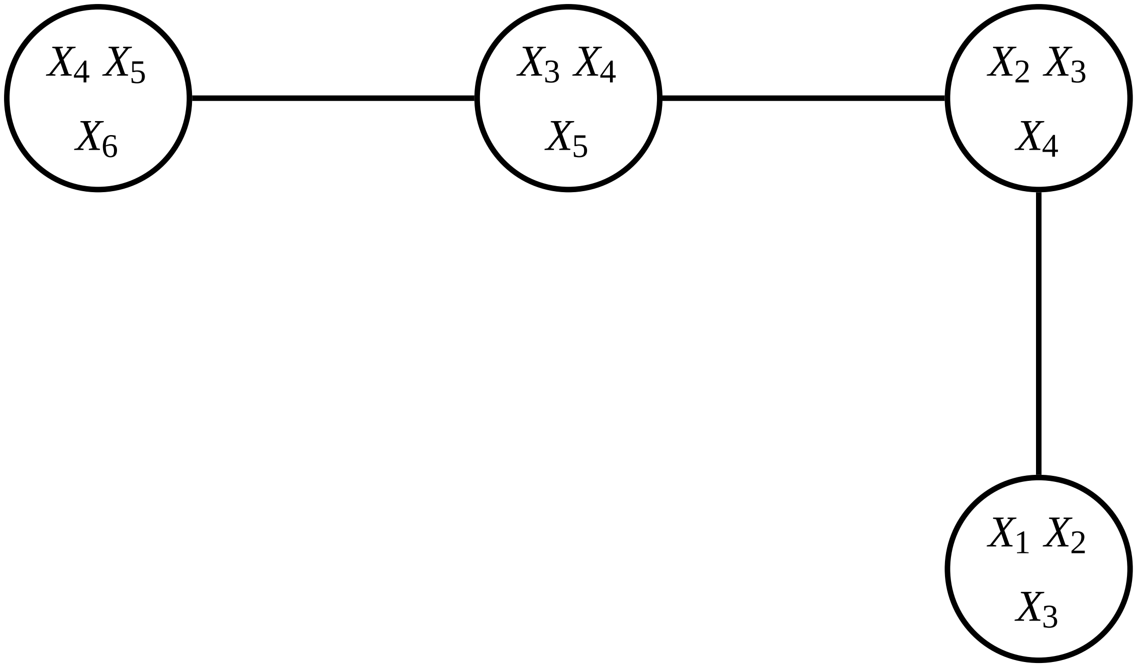
- *Red bayesiana*
- *Join Tree*
- *Mensajes*
- *Distribución marginal para las variables objetivo.*

Ejemplo

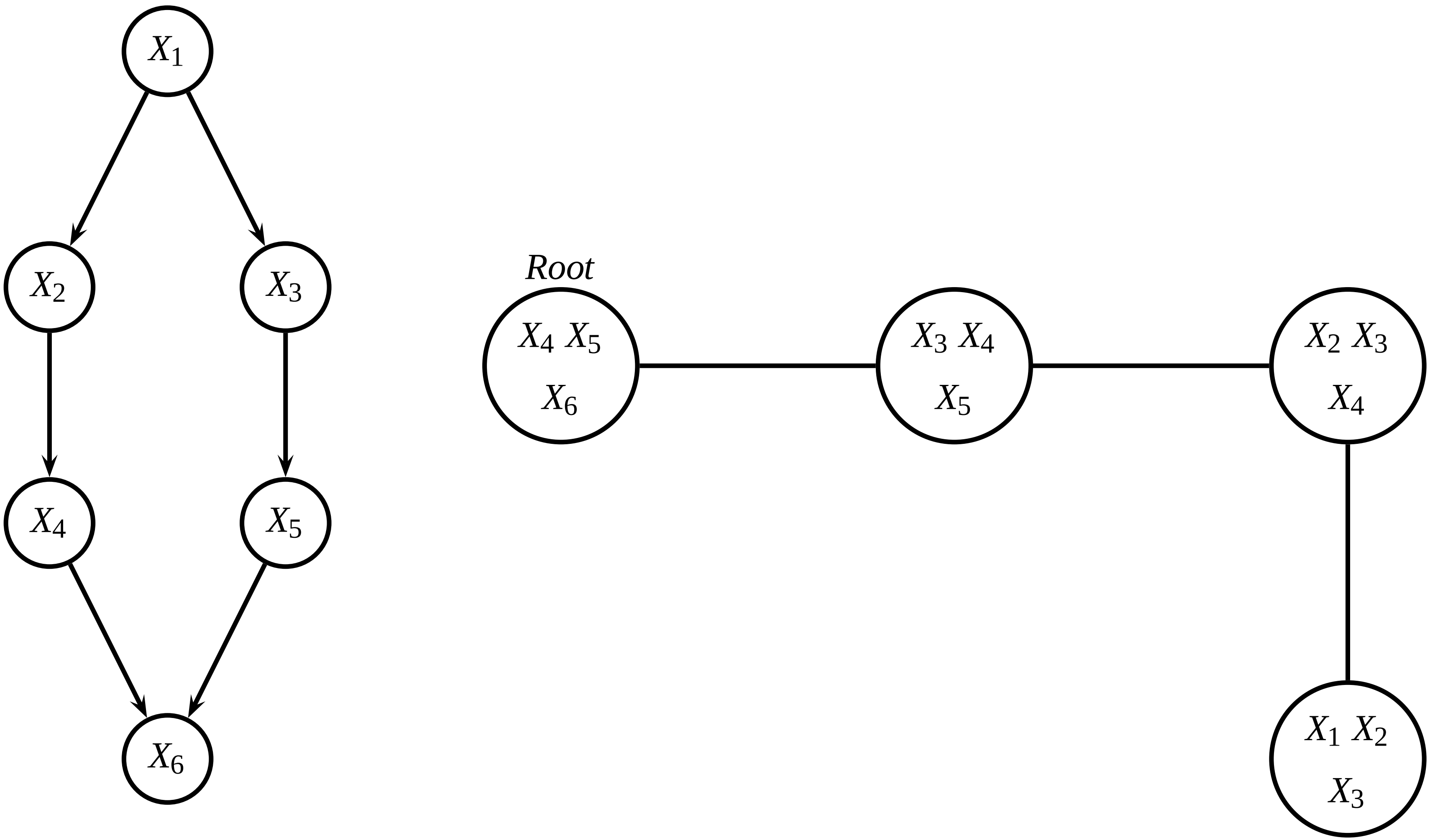


Ejemplo

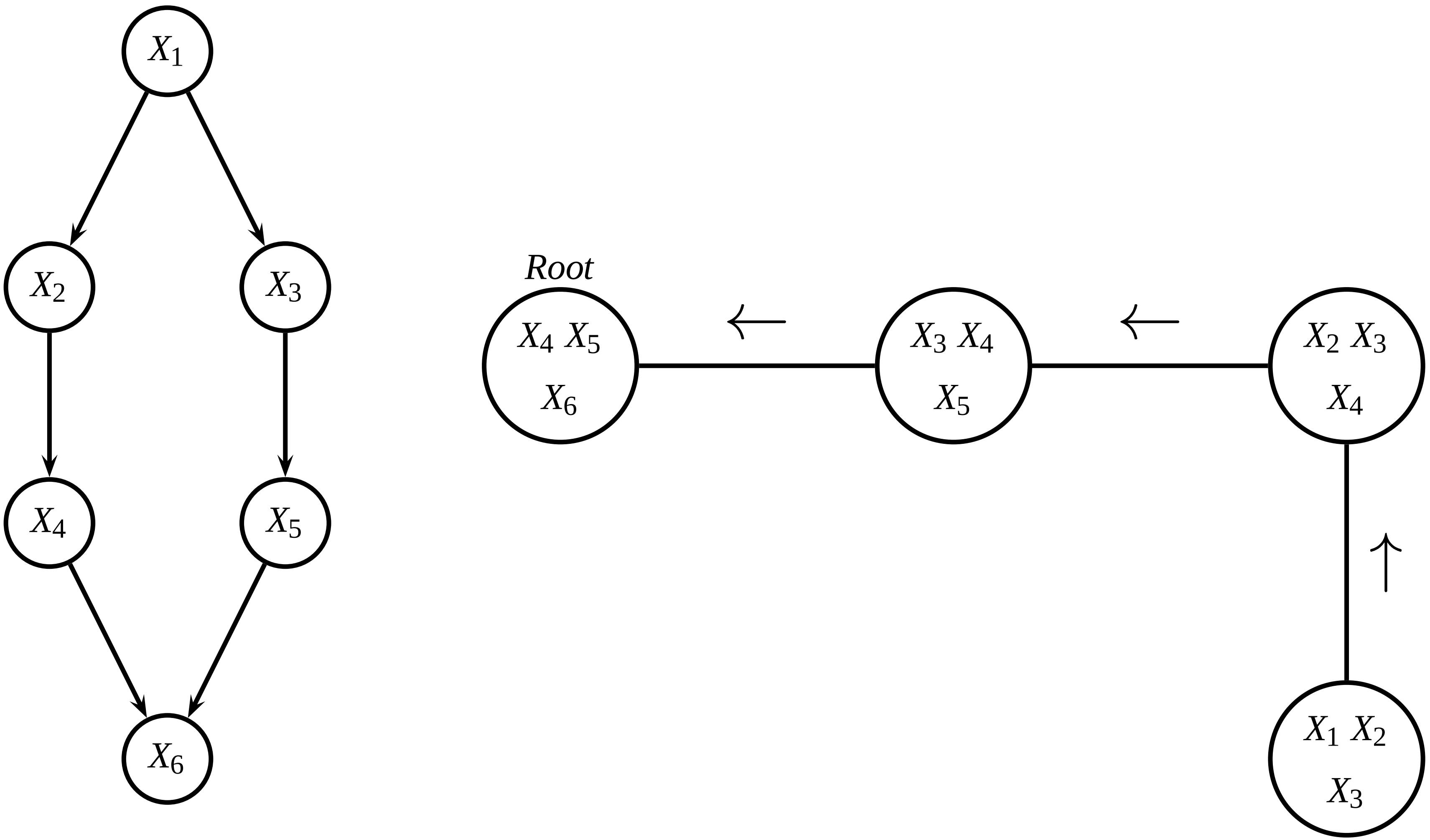




Ejemplo

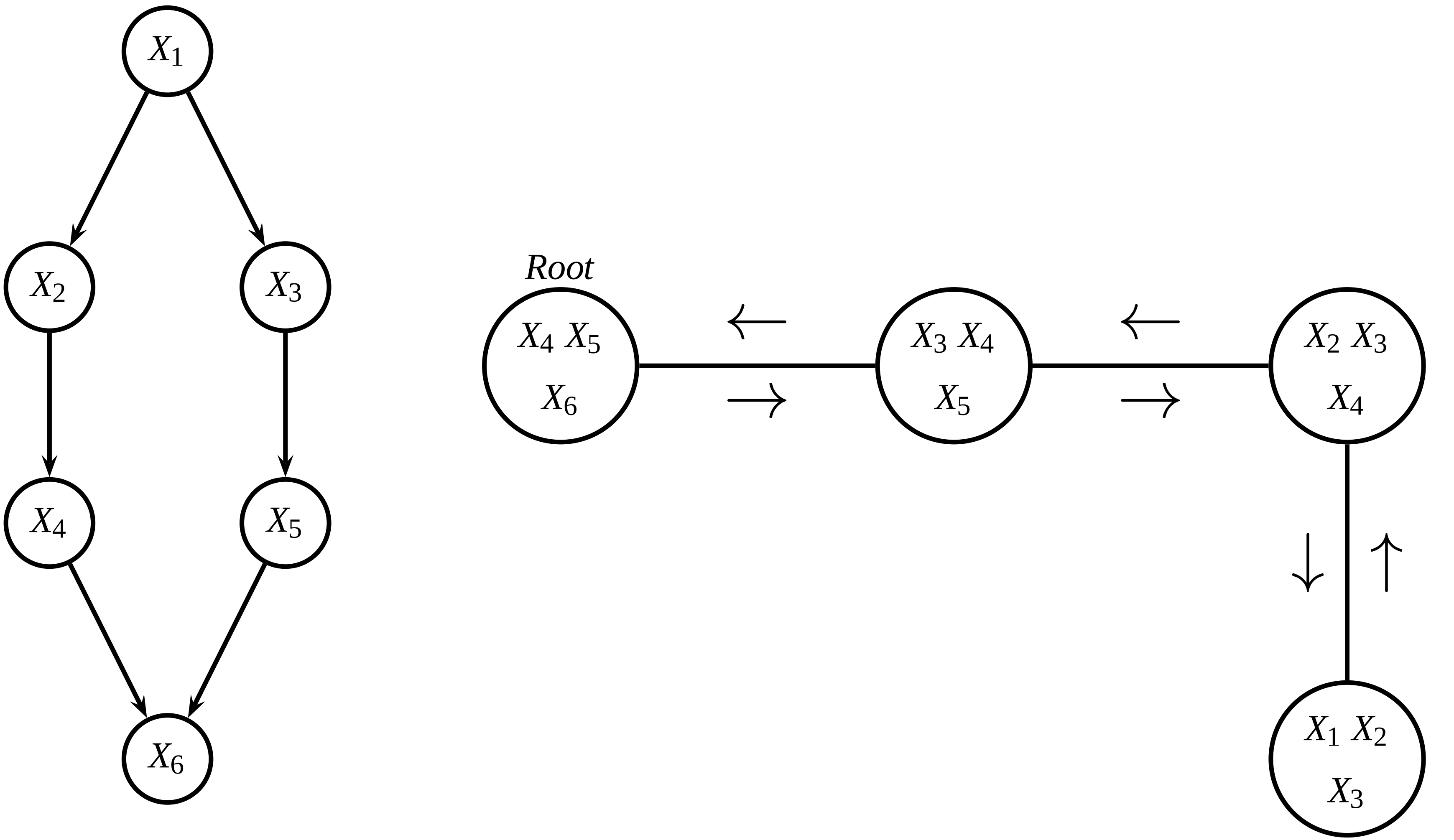


Ejemplo

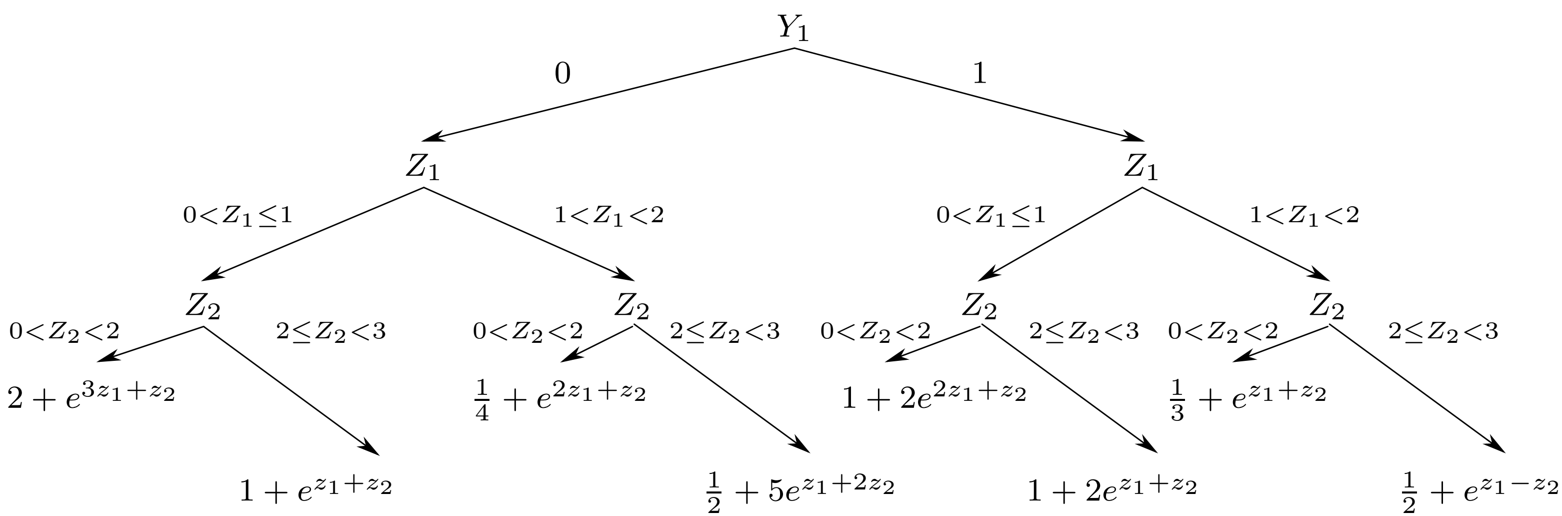


X_1
X_2
X_3
X_4
X_5
X_6
Root
X_4 X_5 X_6
$\leftarrow$
X_3 X_4 X_5
$\leftarrow$
X_2 X_3 X_4
$\uparrow$
X_1 X_2 X_3

Ejemplo



Árboles mixtos



Aproximación de los árboles mixtos

- *Distancia entre árboles:*

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

Aproximación de los árboles mixtos

- *Distancia entre árboles:*

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

- *Disminuir el número de términos exponenciales.*

Aproximación de los árboles mixtos

- *Distancia entre árboles:*

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

- *Disminuir el número de términos exponenciales.*
- *Disminuir el número de hojas:*

Aproximación de los árboles mixtos

- *Distancia entre árboles:*

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

- *Disminuir el número de términos exponenciales.*
- *Disminuir el número de hojas:*
  - *Unir dos hijos (de una vble. continua).*

Aproximación de los árboles mixtos

- *Distancia entre árboles*:

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

- *Disminuir el número de términos exponenciales.*
- *Disminuir el número de hojas:*
    - *Unir dos hijos (de una vble. continua).*
    - *Podar variables discretas.*

Aproximación de los árboles mixtos

- *Distancia entre árboles:*

$$D(\mathcal{T},\mathcal{T}')=E_{\phi_T}[(\phi_T-\phi_{T'})^2]=\int_{\Omega_{\mathbf{Z}}}\phi_T(\mathbf{z})(\phi_T(\mathbf{z})-\phi_{T'}(\mathbf{z}))^2 d\mathbf{z}$$

- *Disminuir el número de términos exponenciales.*
- *Disminuir el número de hojas:*
  - *Unir dos hijos (de una vble. continua.*
  - *Podar variables discretas.*
- *Podamos sólo si* $D(\mathcal{T},\mathcal{T}')<\varepsilon$.

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}}$$

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_{\mathbf{Z}}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_{\mathbf{Z}}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

- k = *Número de términos exponenciales*

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_{\mathbf{z}}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

- *k = Número de términos exponenciales*
- *Mientras k >2 (o parámetro fijado)*

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_{\mathbf{Z}}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

- k = Número de términos exponenciales
- Mientras $k > 2$ (o parámetro fijado)
  - Eliminar $a_i e^{b_i \mathbf{z}}$ si $|p_i| = \min_j |p_i|$

Disminuir el número de términos exponenciales

$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_\mathbf{Z}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

- *k = Número de términos exponenciales*
- *Mientras k >2 (o parámetro fijado)*
    - *Eliminar* $a_i e^{b_i \mathbf{z}}$ *si* $|p_i| = \min_j |p_i|$
    - *Actualizar resto de términos.*

Disminuir el número de términos exponenciales

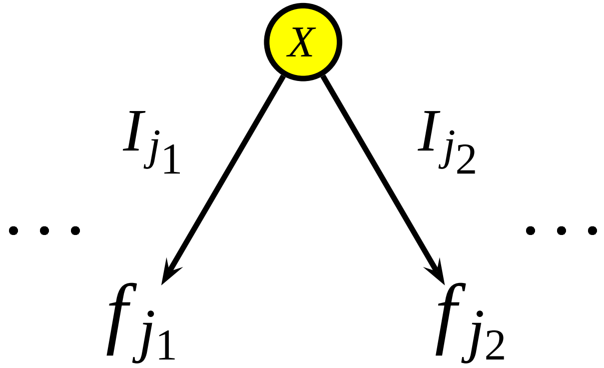
$$f(\mathbf{z}) = k + \sum_{i=1}^{n} a_i e^{b_i \mathbf{z}} \rightarrow p_i = \int_{\Omega_{\mathbf{Z}}} a_i e^{b_i \mathbf{z}} d\mathbf{z}$$

- *k = Número de términos exponenciales*
- *Mientras k >2 (o parámetro fijado)*
  - *Eliminar* $a_i e^{b_i \mathbf{z}}$ *si* $|p_i| = \min_j |p_i|$
  - *Actualizar resto de términos.*
  - *k = k-1.*

Actualizar:

Se calcula el máximo valor del término, $m = \max_{\mathbf{z} \in \mathbf{Z}} \{a_i e^{b_i \mathbf{z}}\}$, *y se suma al término independiente,* $k^* = k + m$, *para posteriormente normalizar.*

Unir dos intervalos



Unir dos intervalos

Unir dos intervalos

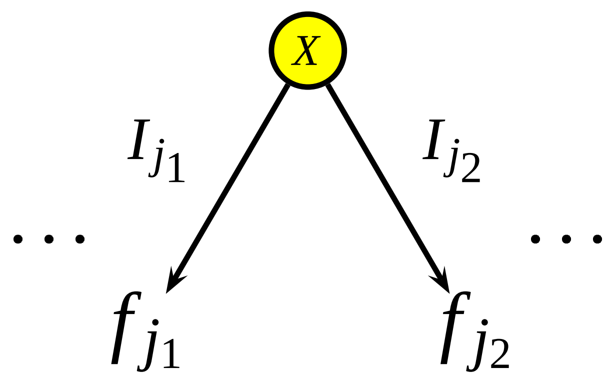
$$p_{j_1} = \int_{\Omega_{\mathbf{Z}}} f_{j_1}(\mathbf{z}) d\mathbf{z} \quad y \quad p_{j_2} = \int_{\Omega_{\mathbf{Z}}} f_{j_2}(\mathbf{z}) d\mathbf{z}$$

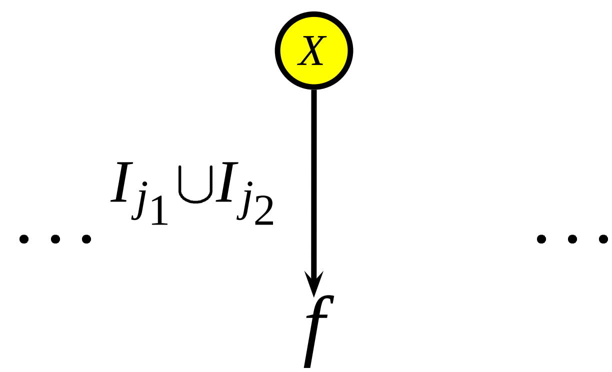
Unir dos intervalos

$$p_{j_1} = \int_{\Omega_{\mathbf{Z}}} f_{j_1}(\mathbf{z}) d\mathbf{z} \quad y \quad p_{j_2} = \int_{\Omega_{\mathbf{Z}}} f_{j_2}(\mathbf{z}) d\mathbf{z}$$

$$f(\mathbf{z}) = \frac{p_{j_1} f_{j_1}(\mathbf{z}) + p_{j_2} f_{j_2}(\mathbf{z})}{p_{j_1} + p_{j_2}}$$

Unir dos intervalos



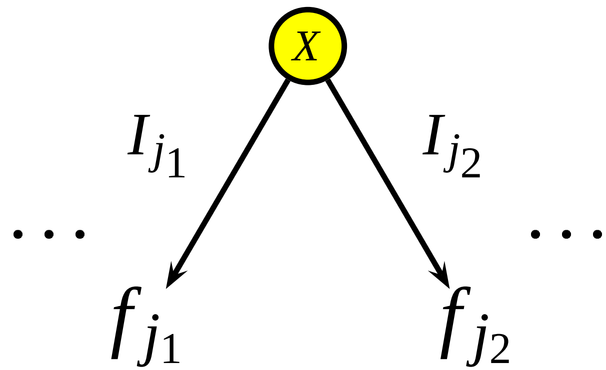


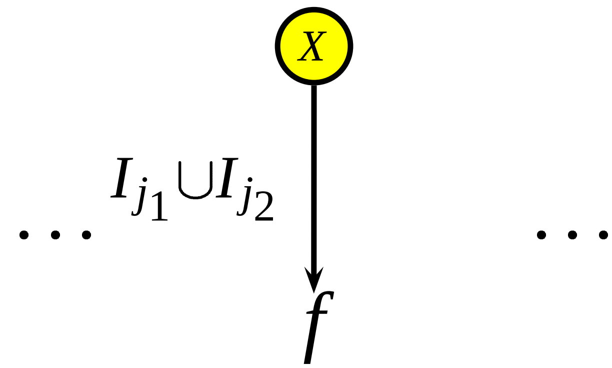
$$p_{j_1} = \int_{\Omega_{\mathbf{Z}}} f_{j_1}(\mathbf{z}) d\mathbf{z} \quad y \quad p_{j_2} = \int_{\Omega_{\mathbf{Z}}} f_{j_2}(\mathbf{z}) d\mathbf{z}$$

$$f(\mathbf{z}) = \frac{p_{j_1} f_{j_1}(\mathbf{z}) + p_{j_2} f_{j_2}(\mathbf{z})}{p_{j_1} + p_{j_2}}$$

$$\int_{\Omega_{\mathbf{Z}}} K f(\mathbf{z}) d\mathbf{z} = p_{j_1} + p_{j_2}$$

Unir dos intervalos



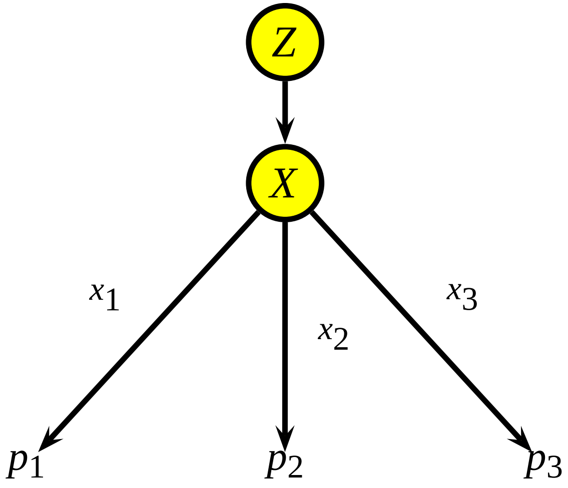


$$p_{j_1} = \int_{\Omega_{\mathbf{Z}}} f_{j_1}(\mathbf{z}) d\mathbf{z} \quad y \quad p_{j_2} = \int_{\Omega_{\mathbf{Z}}} f_{j_2}(\mathbf{z}) d\mathbf{z}$$

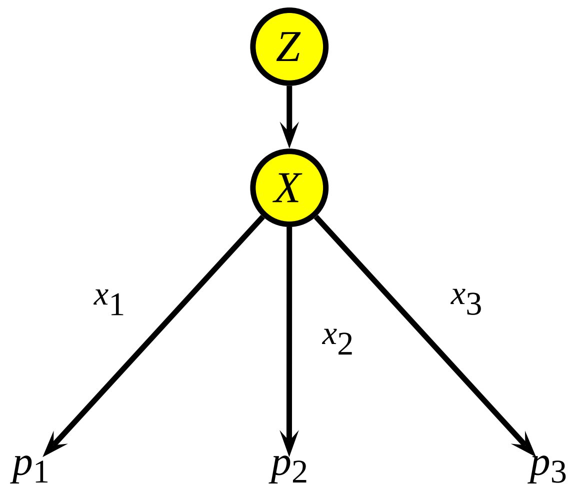
$$f(\mathbf{z}) = \frac{p_{j_1} f_{j_1}(\mathbf{z}) + p_{j_2} f_{j_2}(\mathbf{z})}{p_{j_1} + p_{j_2}}$$

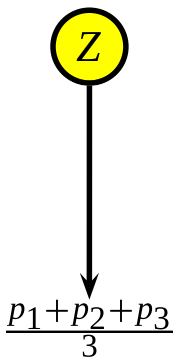
$$\int_{\Omega_{\mathbf{Z}}} K f(\mathbf{z}) d\mathbf{z} = p_{j_1} + p_{j_2} \quad \Rightarrow \quad K = \frac{p_{j_1} + p_{j_2}}{\int_{\Omega_{\mathbf{Z}}} f(\mathbf{z}) d\mathbf{z}}$$

Poda de variables discretas

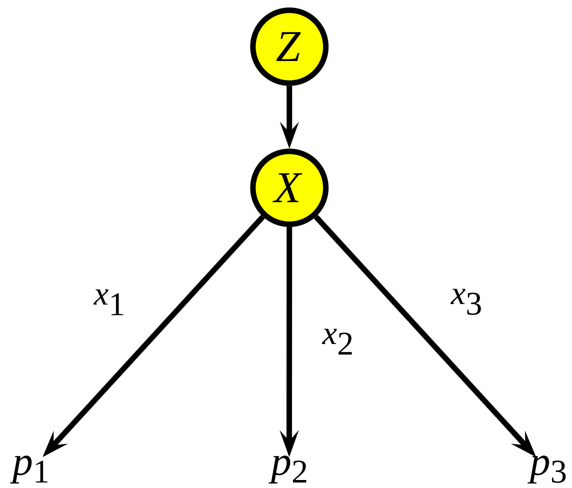


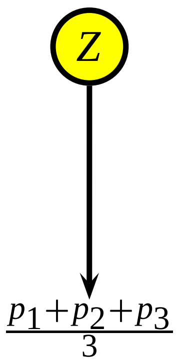
Poda de variables discretas





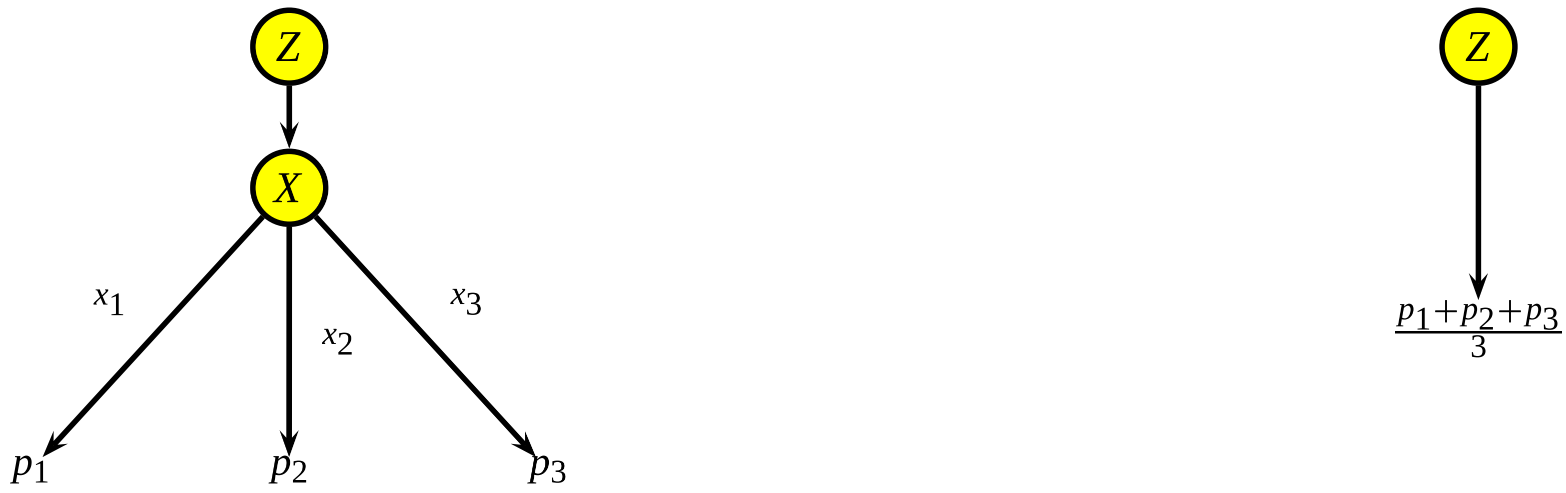
Poda de variables discretas





$$DK(p,q) = \sum_i p_i log(\frac{p_i}{q_i})$$

Poda de variables discretas



$$DK(p,q) = \sum_i p_i log(\frac{p_i}{q_i})$$

$$\xi = -((0.5-\varepsilon)log(0.5-\varepsilon) + (0.5+\varepsilon)log(0.5+\varepsilon))$$

$$Podamos\ si\ DK(\phi, \overline{p}) < \xi$$

Experimentos

Propagación

- *Redes creadas:*

Net	*N. nodos*	*N. nodos discretos*
Net1	*42*	*3*
Net2	*77*	*8*
Net3	*86*	*11*

Propagación

- *Redes creadas:*
- *Evidencia: 30 % de las variables.*

Propagación

- *Redes creadas:*
- *Evidencia: 30 % de las variables.*
- *Comparamos con la discretización. Cada función MTE* $f(\mathbf{z})=k+\sum_{i=1}^{n} a_i e^{b_i \mathbf{z}}$ *se reemplaza por una constante* $f^*(\mathbf{z}) = k^*$ *tal que*

$$\int_{\Omega_{\mathbf{Z}}} f(\mathbf{z}) d\mathbf{z} = \int_{\Omega_{\mathbf{Z}}} f^*(\mathbf{z}) d\mathbf{z} \quad .$$

Propagación

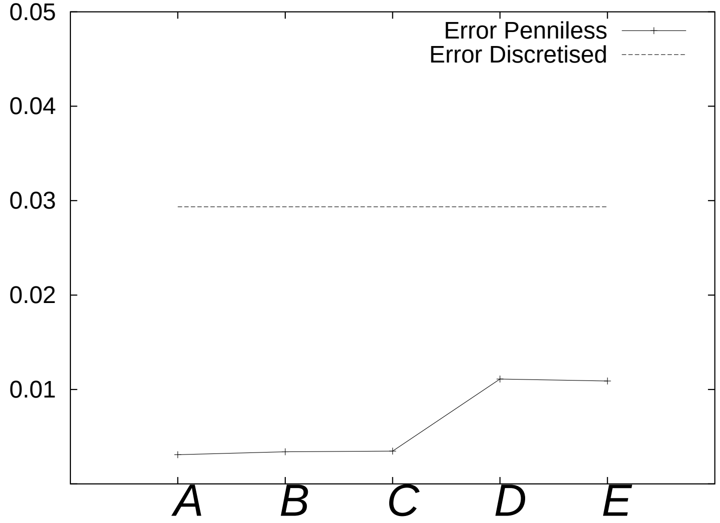
- *Redes creadas:*
- *Evidencia: 30 % de las variables.*
- *Comparamos con la discretización.*
- *Estadísticos de la propagación: Para cada vble. no observada se calcula*
  - *Tamaño del máximo potencial necesario para calcular la distribución marginal.*
  - *Error asociado.*

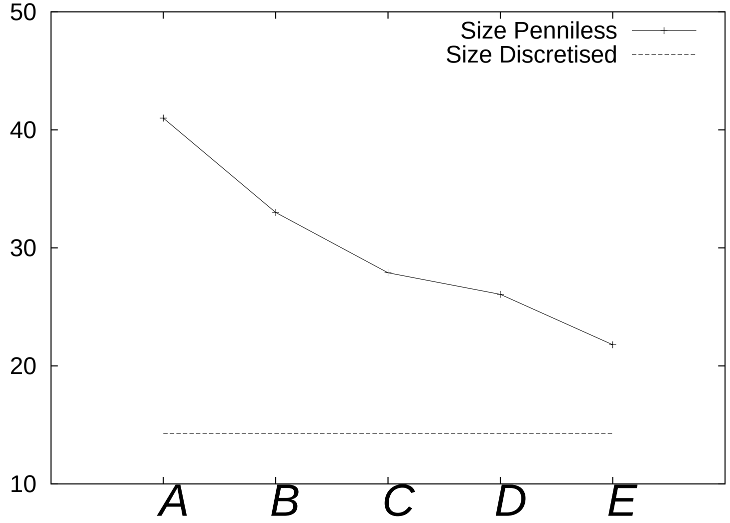
Se presenta la media de estas cantidades.

Parámetros de poda

Poda	*Parámetro unión*	*Parámetro discreto*
A	*0*	*0*
B	*0.005*	*0*
C	*0.005*	*0.01*
D	*0.05*	*0*
E	*0.05*	*0.01*

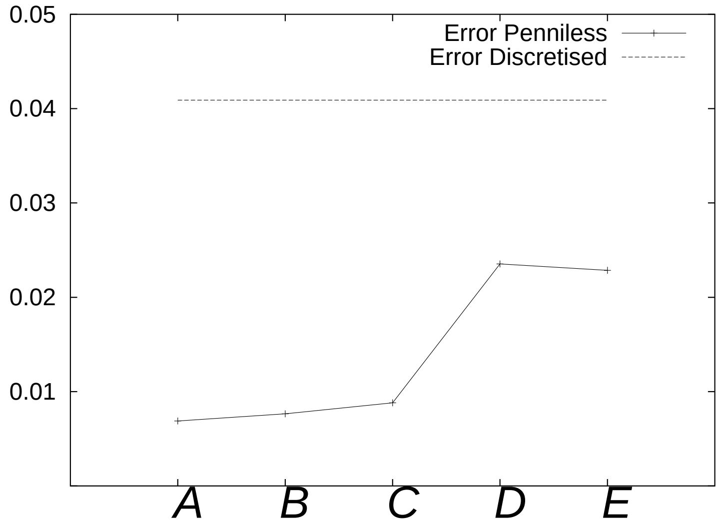
Resultados

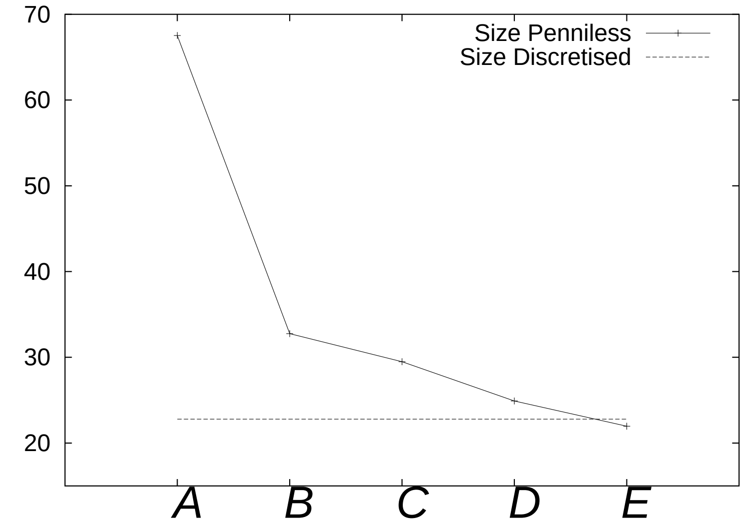




Errores y tamaños para Net1

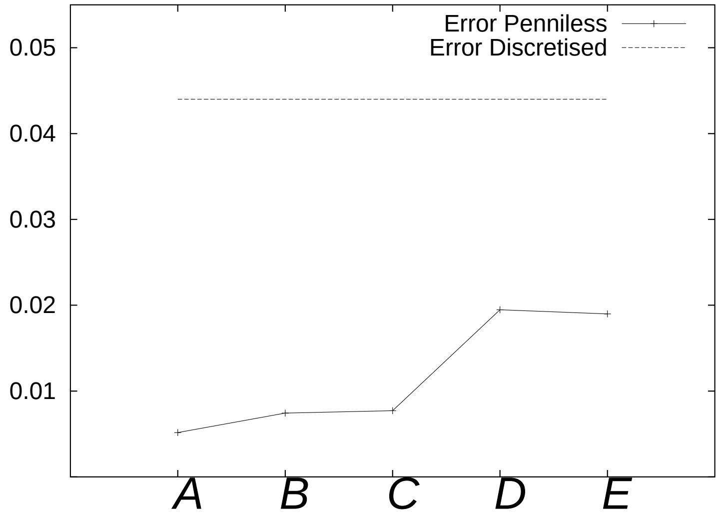
Resultados

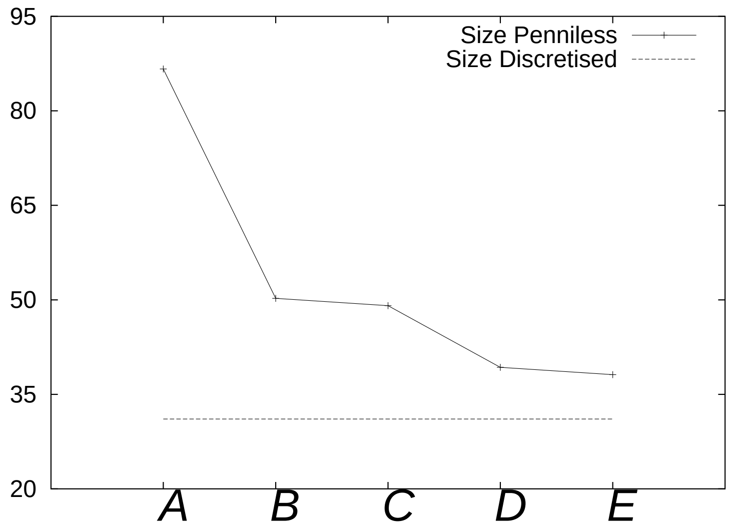




Errores y tamaños para Net2

Resultados





Errores y tamaños para Net3

Elvira

MTESimplePenniless (inference/clustering)

- *Hereda de SimplePenniless.*
  - *propagate(String resultFile, double epsilonJoin, double epsilonDisc, Evidence e, Vector sizeVec, int prune2Int): Realiza la propagación y guarda el resultado en el fichero resultFile como una red Bayesiana.*
  - *Mismos métodos que SimplePenniless pero «actualizados» para trabajar con MTEs.*

PotentialContinuousPT (potential/)

- *Métodos de poda:*
  - *prune*
  - *prune2*

ContinuousProbabilityTree (potential/)

- *Métodos de poda:*
  - *prune*
  - *pruneIterative*
  - *pruneTerms*
  - *pruneJoin*
  - *pruneDiscrete*
  - *prune2*
  - *ErrorPruning*

En curso

Trabajos actuales

- *Mejora del proceso de selección de la partición del dominio de vbles. continuas en el aprendizaje de distribuciones condicionadas.*
- *Algoritmo de muestreo por importancia : ¿$P(X \in (a,b))$?*

Trabajos actuales

- *Mejora del proceso de selección de la partición del dominio de vbles. continuas en el aprendizaje de distribuciones condicionadas.*
- *Algoritmo de muestreo por importancia :* ¿$P(X \in (a,b))$?

Trabajos actuales

- *Mejora del proceso de selección de la partición del dominio de vbles. continuas en el aprendizaje de distribuciones condicionadas.*
- *Algoritmo de muestreo por importancia :* ¿$P(X \in (a,b))$?

Trabajos actuales

- *Mejora del proceso de selección de la partición del dominio de vbles. continuas en el aprendizaje de distribuciones condicionadas.*
- *Algoritmo de muestreo por importancia :* ¿$P(X \in (a,b))$?

Trabajos actuales

- *Mejora del proceso de selección de la partición del dominio de vbles. continuas en el aprendizaje de distribuciones condicionadas.*
- *Algoritmo de muestreo por importancia : ¿$P(X \in (a,b))$?*