Appeared in: P. Lucas (ed.), Proceedings of the Second European Workshop on Probabilistic Graphical Models, 2004, pp. 57--64, Leiden, Netherlands.

Inference in Hybrid Bayesian Networks with Deterministic Variables

Barry R. Cobb and Prakash P. Shenoy
brcobb@ku.edu, pshenoy@ku.edu
School of Business, University of Kansas
1300 Sunnyside Ave., Summerfield Hall, Lawrence, KS, 66045-7585, U.S.A.

Abstract

An important class of hybrid Bayesian networks are those that have conditionally deterministic variables (a variable that is a deterministic function of its parents). In this case, if some of the parents are continuous, then the joint density function does not exist. Conditional linear Gaussian (CLG) distributions can handle such cases when the deterministic function is linear and continuous variables are normally distributed. In this paper, we develop operations required for performing inference with conditionally deterministic variables using relationships derived from joint cumulative distribution functions (CDF's). These methods allow inference in networks with deterministic variables where continuous variables are non-Gaussian.

1 Introduction

Bayesian networks model knowledge about propositions in uncertain domains using graphical and numerical representations. At the qualitative level, a Bayesian network is a directed acyclic graph where nodes represent variables and the (missing) edges represent conditional independence relations among the variables. At the numerical level, a Bayesian network consists of a factorization of a joint probability distribution into a set of conditional distributions, one for each variable in the network. Hybrid Bayesian networks contain both discrete and continuous conditional probability distributions as numerical inputs.

An important class of hybrid Bayesian networks are those that have conditionally deterministic variables (a variable that is a deterministic function of its parents). In this case, if some of the parents are continuous, then the joint density function does not exist. Conditional linear Gaussian (CLG) distributions (Cowell *et al.*, 1999) can handle such cases when the deterministic function is linear. However, for models where continuous variables are not normally distributed, methods for carrying out inference in networks with linear deterministic relationships have not been developed.

Approximate inference in hybrid Bayesian networks can be performed using mixtures of truncated exponentials (MTE) potentials (Moral *et al.*, 2001). General formulations of MTE potentials which approximate the normal probability density function (PDF) exist (Cobb and Shenoy, 2003); however, these formulations cannot be used to model a conditional distribution where the variance of a variable given values of its continuous parents is zero. In this paper, we develop inference operations for conditionally deterministic variables using relationships derived from joint cumulative distribution functions (CDF's). This allows MTE potentials to be used for inference in any CLG model, as well as other models that have conditionally deterministic variables but do not fit the CLG restrictions, such as those containing discrete nodes with continuous parents.

The remainder of this paper is organized as follows. Section 2 introduces notation and definitions used throughout the paper. Section 3 introduces techniques for using CDF's to construct PDF's for deterministic variables. Section 4 introduces join tree operations for linearly deterministic variables. Section 5 contains an example of inference in a Bayesian network containing linearly deterministic variables. Section 6 summarizes and states directions for future research. This paper is a preliminary version of a longer, forthcoming working paper which includes proofs of all theorems.

2 Notation and Definitions

This section contains notation and definitions that will be used throughout the remainder of the paper.

2.1 Notation

Random variables in a hybrid Bayesian network will be denoted by capital letters, e.g. A, B, C. Sets of variables will be denoted by boldface capital letters, $\mathbf{Y}$ if all variables are discrete, $\mathbf{Z}$ if all variables are continuous, or $\mathbf{X}$ if some of the components are discrete and some are continuous. If $\mathbf{X}$ is a set of variables, $\mathbf{x}$ is a configuration of specific states of those variables. The discrete, continuous, or mixed state space of $\mathbf{X}$ is denoted by $\Omega_{\mathbf{X}}$. MTE probability potentials and discrete probability potentials are denoted by lower-case greek letters, e.g. α, β, γ.

In graphical representations, continuous nodes in hybrid Bayesian networks are represented by double-border ovals, whereas continuous nodes that are deterministic functions of their parents are represented by triple-border ovals. Shaded nodes are degenerate, indicating that evidence has restricted the variable to one value.

2.2 Conditional Mass Function (CMF)

When relationships between continuous variables are deterministic, the joint PDF does not exist. We can express a conditional probability mass function as a degenerate function. If Y is a deterministic relationship of variables in $\mathbf{X}$, i.e. $y = g(\mathbf{x})$, the conditional mass function (CMF) for $\{Y \mid \mathbf{x}\}$ is defined as

$$p_{Y|x}(y) = \mathbf{1}\{y = g(\mathbf{x})\}\,, \qquad (1)$$

where $\mathbf{1}\{A\}$ is the indicator of the event A, i.e. $\mathbf{1}\{A\} = 1$ if A occurs and 0 otherwise. Graphically, the conditionally deterministic relationship of Y given $\mathbf{X}$ is represented in a hybrid Bayesian network model as shown in Figure 1.

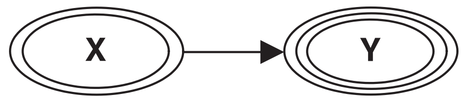

Figure 1: Graphical representation of the conditionally deterministic relationship of Y given $\mathbf{X}$ determined by the CMF $p_{Y|\mathbf{x}}(y)$.

2.3 Mixtures of Truncated Exponentials (MTE) Potentials

A mixture of truncated exponentials (MTE) (Moral *et al.*, 2001) potential has the following definition.

MTE potential. Let $\mathbf{X}$ be a mixed n-dimensional random variable. Let $\mathbf{Y} = (Y_1, \ldots, Y_d)$ and $\mathbf{Z} = (Z_1, \ldots, Z_c)$ be the discrete and continuous parts of $\mathbf{X}$, respectively, with $c + d = n$. A function $\phi : \Omega_{\mathbf{X}} \mapsto \mathbb{R}^+$ is an MTE potential if one of the next two conditions holds:

1. The potential ϕ can be written as

$$\phi(\mathbf{x}) = \phi(\mathbf{y}, \mathbf{z}) = a_0 + \sum_{i=1}^{m} a_i \exp\Big\{ \sum_{j=1}^{d} b_i^{(j)} y_j + \sum_{k=1}^{c} b_i^{(d+k)} z_k \Big\} \qquad (2)$$

   for all $\mathbf{x} \in \Omega_{\mathbf{X}}$, where $a_i, i = 0, \ldots, m$ and $b_i^{(j)}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ are real numbers.

2. There is a partition $\Omega_1, \ldots, \Omega_k$ of $\Omega_{\mathbf{X}}$ verifying that the domain of continuous variables, $\Omega_{\mathbf{Z}}$, is divided into hypercubes, the domain of the discrete variables, $\Omega_{\mathbf{Y}}$, is divided into arbitrary sets, and such that ϕ is defined as

$$\phi(\mathbf{x}) = \phi_i(\mathbf{x}) \qquad \text{if } \mathbf{x} \in \Omega_i, \qquad (3)$$

where each $\phi_i, i = 1, ..., k$ can be written in the form of equation (2) (i.e. each ϕ_i is an MTE potential on Ω_i).

3 Using CDF's to Construct PDF's for Deterministic Variables

This section describes methods of constructing CDF's and their corresponding PDF's for variables that are deterministic functions of their parents.

3.1 Monotonically Increasing Functions

Consider a random variable Y which is a monotonically increasing deterministic function of a random variable X. A Bayesian network representing this relationship is shown in Figure 1. The joint CDF for $\{X, Y\}$ is then given by

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y]$$
$$= \begin{cases} F_X(x) & \text{if } x < g^{-1}(y) \\ F_X(g^{-1}(y)) & \text{if } x \geq g^{-1}(y). \end{cases} \quad (4)$$

Thus, $F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y) = F_X(g^{-1}(y))$, and therefore

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= f_X(g^{-1}(y)) \frac{d}{dy}(g^{-1}(y)) . \end{aligned}$$

Proposition 1. Suppose we have a Bayesian network with two variables X and Y with an arrow from X to Y where Y is a conditionally deterministic, monotonically increasing function of X. Then, the equivalent Bayesian network with an arrow from Y to X where X is a conditionally deterministic function of Y meets the conditions that $f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y))$ and $p_{X|y}(x) = \mathbf{1}\{x = g^{-1}(y)\}$.

When Y is a monotonically increasing (and therefore invertible) deterministic function of X, Proposition 1 gives a shortcut to finding the PDF of Y from the PDF of X that does not require the CDF of Y to be computed. We refer to using the operation in Proposition 1 as performing an "arc reversal" on the Bayesian network. After the operation is performed, the Bayesian network appears as in Figure 2.

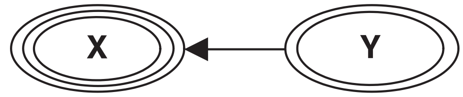

Figure 2: Graphical representation of the conditionally deterministic relationship of X on Y after performing an "arc reversal" on the Bayesian network of Figure 1.

Example 1.

Suppose that a random variable X has PDF

$$f_X(x) = \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere}, \end{cases}$$

and we want to find $f_Y(y)$ if $Y = g(x) = 4x^2$.

Note that $f_X(g^{-1}(y)) = 3y/4$ and $\frac{d}{dy}(g^{-1}(y)) = 1/(4\sqrt{y})$. Using Proposition 1, we compute

$$f_Y(y) = \frac{3y}{4} \cdot \frac{1}{4\sqrt{y}} = \frac{3\sqrt{y}}{16} .$$

3.2 Monotonically Decreasing Functions

Consider a random variable Y which is a monotonically decreasing deterministic function of a random variable X.

We can compute the joint CDF as follows when Y is a monotonically decreasing function of X

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y]$$
$$= \begin{cases} 0 & \text{if } x < g^{-1}(y) \\ F_X(x) - F_X(g^{-1}(y)) & \text{if } x \geq g^{-1}(y). \end{cases} \quad (5)$$

Thus, $F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y) = 1 - F_X(g^{-1}(y))$, and therefore

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) \\ &= -f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y)) . \end{aligned}$$

Proposition 2. Suppose we have a Bayesian network with two variables X and Y with an arrow from X to Y where Y is a conditionally deterministic, monotonically decreasing function of X. Then, the equivalent Bayesian network with an arrow from Y to X where X is a conditionally deterministic function of Y meets the conditions that $f_Y(y) = -f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y))$ and and $p_{X|y}(x) = \mathbf{1}\{x = g^{-1}(y)\}$.

When Y is a monotonically decreasing (and therefore invertible) deterministic function of X, Proposition 2 gives a shortcut to finding the PDF of Y from the PDF of X that does not require the CDF of Y to be computed. As in the monotonically increasing case, we refer to use of the operation in Proposition 2 as an arc reversal.

Example 2.

Let X have the uniform PDF over the unit interval, i.e. $X \sim U(0,1)$. Find $f_Y(y)$ if $Y = g(x) = \frac{-\ln x}{\lambda}$.

Note that $f_X(g^{-1}(y)) = 1$ and $\frac{d}{dy}(g^{-1}(y)) = -\lambda e^{-\lambda y}$. Using Proposition 2, we compute

$$f_Y(y) = (-1) \cdot -\lambda e^{-\lambda y} = \lambda e^{-\lambda y} .$$

3.3 Linear CDF Marginalization Operator

Suppose Y is a conditionally deterministic linear function of X, i.e. $Y = g(x) = ax+b$, $a \neq 0$. The following definition will be used to determine the marginal PDF for Y:

$$f_Y(y) = (f_X(x) \otimes p_{Y|x}(y))^{\downarrow Y}$$

$$\triangleq \begin{cases} \dfrac{1}{a} \cdot f_X\left(\dfrac{y-b}{a}\right) & \text{if } a > 0 \\ \dfrac{-1}{a} \cdot f_X\left(\dfrac{y-b}{a}\right) & \text{if } a < 0 . \end{cases} \qquad (6)$$

The definition of the combination of a deterministic function followed by marginalization follows directly from the expressions in Propositions 1 and 2.

Example 3.

Suppose that a random variable X has PDF

$$f_X(x) = \begin{cases} 6x(1-x) & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Find $f_Y(y)$ if $Y = g(x) = 2x + 1$.

Note that $f_X(\frac{y-b}{a}) = f_X(\frac{y-1}{2}) = 6(\frac{y-1}{2}) \cdot (1 - (\frac{y-1}{2})) = -\frac{3}{2}y^2 + 6y - \frac{9}{2}$.

Using the operation in (6), we find the PDF for Y as

$$f_Y(y) = (f_X(x) \otimes p_{Y|x}(y))^{\downarrow Y}$$

$$= \begin{cases} -\frac{3}{4}y^2 + 3y - \frac{9}{4} & \text{if } 1 < y < 3 \\ 0 & \text{elsewhere.} \end{cases}$$

The following theorem is required for inference using MTE potentials in Bayesian networks with conditionally deterministic linear variables.

Theorem 3. If $\phi_1(x)$ is an MTE potential for X and Y is a conditionally deterministic linear function of X, then $\phi_2(y) = (\phi_1(x) \otimes p_{Y|x}(y))^{\downarrow Y}$ is an MTE potential.

3.4 Method of Convolutions

The following theorem will be required for join tree operations when a variable is a linearly deterministic function of its parents.

Theorem 4. Let X and Y be continuous, possibly dependent random variables and let $W = a_1 \cdot x + a_2 \cdot y + b$, $a_1 \neq 0$, $a_2 \neq 0$. The PDF for W can be found as

$$f_W(w) = (f_{X,Y}(x,y) \otimes p_{W|x,y}(w))^{\downarrow W}$$

$$\triangleq \tfrac{1}{|a_2|} \int_{-\infty}^{+\infty} f_{X,Y}\left(x, \frac{w - a_1 \cdot x - b}{a_2}\right) \, dx .$$

An integral of the form in Theorem 4 is referred to as a *convolution* of the function $f_{X,Y}(x, y)$ (Larsen and Marx, 2001).

To use the convolution formula in Theorem 4 to find PDF's for linearly deterministic variables in hybrid Bayesian networks with MTE potentials, the following theorem is required.

Theorem 5. If ϕ_1 is a joint MTE potential for $\{X, Y\}$, an MTE potential ϕ for

$$W = a_1 \cdot x + a_2 \cdot y + b$$

can be formed from the convolution of ϕ_1.

4 Join Tree Operations with Linearly Deterministic Variables

Suppose we have a node in a join tree created from a hybrid Bayesian network containing a set of variables $\mathbf{X} = (X_1, ..., X_N)$. Assume a variable $X_i \in \mathbf{X}$ is a linear deterministic function of its continuous parents $(Z_1, ...Z_K) \subset \mathbf{X}$, i.e.

$$X_i = g(z_1, ..., z_K) = w + b\,,$$

where

$$w = \sum_{k=1}^{K} a_k \cdot z_k$$

with $a_1, ..., a_k$ and b defined as real numbers.

The joint PDF for $\{X_i, Z_1, ..., Z_K\}$ does not exist; however, we can find the marginal PDF for X_i by using the operations defined in Section 3 as follows:

$$f_{X_i}(x_i) = (f_W(w) \otimes p_{X_i|w}(x_i))^{\downarrow X_i}\,,$$

where the PDF $f_W(w)$ is obtained by repeated application of the method of convolutions.

Example 5.

Consider the Bayesian network depicted in Figure 3. Suppose $X \sim N(0, 1)$, $Y \sim N(1, 1)$, and Z is a conditionally deterministic function of its parents, $Z \mid x, y \sim N(2 + x - y, 0)$. The objective is to find the marginal PDF for Z.

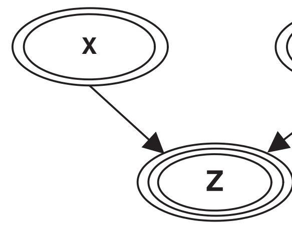

Figure 3: The Bayesian network for Example 5.

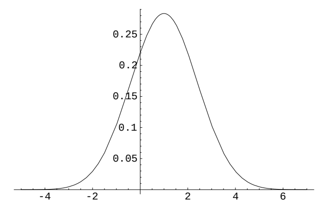

Figure 4: The marginal PDF for Z in Example 5.

Using the PDF for X and the PDF for Y, we next create the PDF for $Z = x - y + 2$ using the method of convolutions as follows:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, x - z + 2)\, dx\,.$$

Note that in this case, since X and Y are independent, $f_{X,Y}(x, y) = f_X(x)f_Y(y)$. Thus, the calculation above can be simplified to

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(x - z + 2)\, dx\,.$$

The marginal PDF for Z (shown in Figure 4) was created by approximating the normal PDF's in this example with the MTE approximation to the normal PDF presented in Cobb and Shenoy (2003). The mean and variance of this marginal PDF are 1.0000 and 1.9638. All answers presented in this example are comparable to those obtained by solving the problem with Hugin software.

Suppose we obtain evidence that $Z = 3$. Since the existing potential for Z states that $Z = 2 + x - y$, the evidence dictates new deterministic relationships $X - Y = 1$, $X = Y + 1$, and $Y = X - 1$.

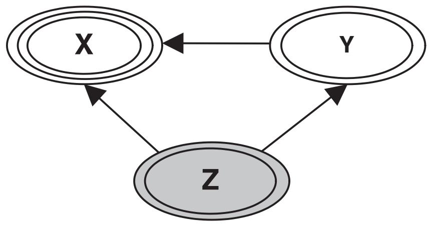

Figure 5: The revised Bayesian network for Example 5 after observing evidence on Z.

The variables X and Y are no longer independent and now have a conditionally deterministic linear relationship. The revised Bayesian network is depicted in Figure 5. To calculate the revised marginal distribution for X—denoted $f_{X_{ev}}(x)$— we combine the prior distribution for X with the distribution created by applying the linear CDF marginalization operator in (6) to the prior distribution for Y. This is done as follows

$$f_{X_{ev}}(x) = K^{-1} f_X(x) \cdot (f_Y(y) \otimes p_{X|y}(x))^{\downarrow X}$$

$$= K^{-1} f_X(x) \cdot f_Y(x-1) \,.$$

The normalization constant, K, is calculated as $K = f_Z(3) = 0.103815$ and represents the likelihood of the observed evidence. The expected value and variance of the posterior marginal PDF for X are calculated as 1.0000 and 0.5004, respectively.

To calculate the revised marginal distribution for Y—denoted $f_{Y_{ev}}(y)$— we combine the prior distribution for Y with the distribution created by applying the linear CDF marginalization operator in (6) to the prior distribution for X. This is done as follows

$$f_{Y_{ev}}(y) = K^{-1} f_Y(y) \cdot (f_X(x) \otimes p_{Y|x}(y))^{\downarrow Y}$$

$$= K^{-1} f_Y(y) \cdot f_X(y+1) \,.$$

The function $(f_X(x) \otimes p_{Y|x}(y))^{\downarrow Y}$ is constructed by reversing the arc between X and Y in Figure 5. The same normalization constant, K, used to calculate the posterior distribution for X remains valid. The expected value and variance of the posterior marginal PDF for Y are calculated as 0.0000 and 0.5004, respectively.

5 Example

The Bayesian network in this example (shown in Figure 7) contains one variable (A) which follows a beta distribution, one variable (C) with a Gaussian potential, and one variable (B) which is a deterministic linear function of its parent. All probability potentials are approximated in the calculations by MTE potentials.

5.1 Representation

The probability distribution for A is a beta distribution, i.e. $\pounds(A) \sim Beta(2.7, 1.3)$. The PDF for A is approximated (using the methods described in (Cobb *et al.*, 2003)) by an MTE potential as follows:

$$\alpha(a) = P(A) =$$

$$\begin{cases} -5.951669 + 5.573316\exp\{0.461388a\} \\ -0.378353\exp\{-6.459391a\} & \text{if } 0 < a < d^- \\ 0.473654 - 6.358483\exp\{-2.639474a\} \\ +2.729395\exp\{-0.331472a\} & \text{if } d^- \le a < m \\ 1.823067 - (5.26E-12)\exp\{26.000041a\} \\ +0.035775\exp\{0.529991a\} & \text{if } m \le a < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

where $m = (1-\alpha)/(2-\alpha-\beta) = 0.85$ and

$$d^- = \frac{(\alpha-1)(\alpha+\beta-3) - \sqrt{(\beta-1)(\alpha-1)(\alpha+\beta-3)}}{(\alpha+\beta-3)(\alpha+\beta-2)} = 0.493.$$

The MTE potential for A is shown graphically in Figure 6, overlayed on the actual $Beta(2.7, 1.3)$ distribution.

The probability distribution for B is defined as $\pounds(B \mid a) \sim N(2a+1, 0)$. The conditional distribution for B is represented by a CMF as follows:

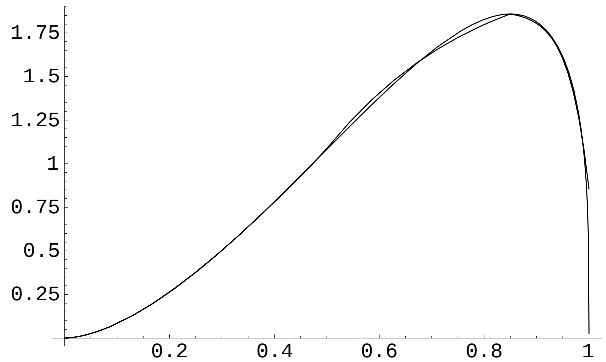

Figure 6: The MTE potential for A overlayed on the actual $Beta(2.7, 1.3)$ distribution.

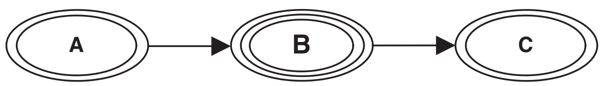

Figure 7: The Bayesian network for the example problem.

$$\beta(a, b) = p_{B|a}(b) = \mathbf{1}\{b = 2a + 1\} .$$

The probability distribution for C is defined as $\pounds(C \mid b) \sim N(2b + 1, 1)$.

5.2 Computing Messages

The join tree for the example problem is shown in Figure 8.

The messages required to calculate posterior marginals for each variable in the network without evidence are as follows:

1) α from $\{A\}$ to $\{A, B\}$
2) $(\alpha \otimes \beta)^{\downarrow B}$ from $\{A, B\}$ to $\{B\}$ and $\{B\}$ to $\{B, C\}$
3) $((\alpha \otimes \beta)^{\downarrow B} \otimes \delta)^{\downarrow C}$ from $\{B, C\}$ to $\{C\}$

5.3 Posterior Marginals

The posterior marginal distribution for B is the message sent from $\{A, B\}$ to $\{B, C\}$. The ex-

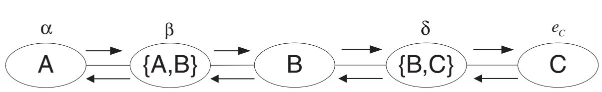

Figure 8: The join tree for the example problem.

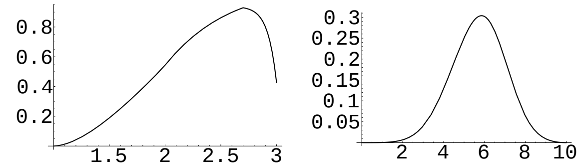

Figure 9: The posterior marginal distributions for B (left) and C (right).

pected value and variance of this distribution are calculated as 2.3488 and 0.1758, respectively. The posterior marginal distribution for C is the message sent from $\{B, C\}$ to $\{C\}$. The expected value and variance of this distribution are calculated as 5.6975 and 1.6851, respectively. The posterior marginal distributions for B and C are shown graphically in Figure 9.

5.4 Entering Evidence

Assume evidence exists that $c = 6$ and define $e_C = 6$. Define $\eta = (\alpha \otimes \beta)^{\downarrow B}$ and $\vartheta(a, b) = \mathbf{1}\{0.5b - 0.5\}$ as the potentials resulting from the reversal of the arc between A and B. The evidence $e_C = 6$ is passed from $\{C\}$ to $\{B, C\}$ in the join tree, where the existing potential is restricted to $\delta(b, 6)$. This likelihood potential is passed from $\{B, C\}$ to $\{B\}$ in the join tree.

Denote the unnormalized posterior marginal distribution for B as $\xi'(b) = \eta(b) \cdot \delta(b, 6)$. The normalization constant is calculated as $K = \int_b (\eta(b) \cdot \delta(b, 6))\, db = 0.2344$ and represents the probability of the observed evidence. Thus, the normalized marginal distribution for B is found as $\xi(b) = K^{-1} \cdot \xi'(b)$. The expected value and variance of this distribution (which is displayed in Figure 10) are calculated as 2.5049 and 0.0771, respectively.

Using the results of Proposition 1, we determine the posterior marginal distribution for A. Define $\theta = (\xi \otimes \nu)^{\downarrow A}$ as:

$$\theta(a) = \frac{1}{0.5} \xi\,(2a + 1) .$$

The CMF $\nu(a, b) = \mathbf{1}\{2a + 1\}$ is obtained by reversing the arc between A and B in Figure 7. The expected value and variance of this distribution are calculated as 0.7525 and 0.0193, respectively. The posterior marginal distribution

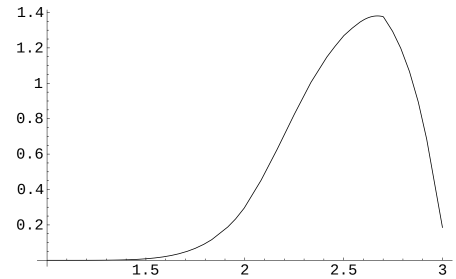

Figure 10: The posterior marginal distribution for B considering the evidence $c = 6$.

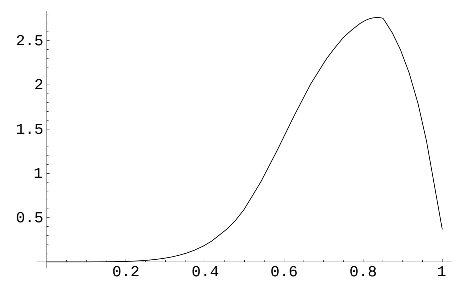

Figure 11: The posterior marginal distribution for A considering the evidence ($c = 6$).

for A considering the evidence is shown graphically in Figure 11.

6 Summary and Conclusions

This paper has described operations required for inference in hybrid Bayesian networks containing variables that are deterministic functions of their parents. Since the joint PDF for a network with deterministic variables does not exist, the operations presented are derived from CDF's. In future work, we plan to create a general inference algorithm using these operations so that Bayesian networks with deterministic variables can be more widely implemented.

References

Cobb, B.R. and P.P. Shenoy (2003), "Inference in hybrid Bayesian networks with mixtures of truncated exponentials," School of Business Working Paper No. 294, University of Kansas. Available for download at: http://www.people.ku.edu/~brcobb/WP294.pdf

Cobb, B.R., Shenoy, P.P. and R. Rumí (2003), "Approximating probability density functions with mixtures of truncated exponentials," Working Paper No. 303, School of Business, University of Kansas. Available for download at: http://www.people.ku.edu/~brcobb/WP303.pdf

Cobb, B.R. and P.P. Shenoy (2004a), "Decision making with hybrid influence diagrams using mixtures of truncated exponentials," School of Business Working Paper No. 304, University of Kansas. Available for download at: http://www.people.ku.edu/~brcobb/WP304.pdf

Cowell, R.G., Dawid, A.P., Lauritzen, S.L. and D.J. Spiegelhalter (1999), *Probabilistic Networks and Expert Systems*, Springer, New York.

Larsen, R.J. and M.L. Marx (2001), *An Introduction to Mathematical Statistics and its Applications*, Prentice Hall, Upper Saddle River, N.J.

Moral, S., Rumí, R. and A. Salmerón (2001), "Mixtures of truncated exponentials in hybrid Bayesian networks," in P. Besnard and S. Benferhart (eds.), *Symbolic and Quantitative Approaches to Reasoning under Uncertainty*, Lecture Notes in Artificial Intelligence, 2143, 156–167.