SCHOOL OF BUSINESS WORKING PAPER NO. 324

Representing Piecewise Functions in *Mathematica*[©]

Prakash P. Shenoy

School of Business, University of Kansas,
1300 Sunnyside Ave., Summerfield Hall, Lawrence, KS 66045-7601, USA
pshenoy@ku.edu

March 2011[†]

Representing Piecewise Functions in $Mathematica^{©}$

Prakash P. Shenoy

University of Kansas School of Business

1300 Sunnyside Ave., Summerfield Hall, Lawrence, KS 66045-7601 USA

pshenoy@ku.edu

March 10, 2011

Abstract

There are many ways of representing piecewise functions in $Mathematica^{©}$. We study three different methods. One is by using the function $Piecewise$. Another is by using a difference of two $HeavisideTheta$ functions for each piece. And a third way is by using product to two $HeavisideTheta$ functions for each piece. If efficiency in integration is a criterion, which of these is the best way to do so? Another issue is the limits one should use for integration. Assuming one wishes to integrate over the entire domain, is it better to use $(-\infty, \infty)$ as the limits or is it better to use more specific limits such as $(-3, 3)$ if the piecewise functions are non-zero over $(-3, 3)$? We examine the answers to these questions by doing some experiments and speculate on the reasons for the results.

1 Introduction

A piecewise function is a function that has several (more than one) definitions over different domains. Many probability density functions (PDFs) of continuous random variables in probability theory are piecewise functions. For example, suppose Z is a continuous random variable that has the standard normal distribution. Its PDF is not a piecewise function since the PDF has one definition over the entire domain $(-\infty, \infty)$:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{1.1}$$

However, consider the truncated standard normal distribution PDF on the domain $(-3, 3)$, which is defined as follows:

$$f(z) = \begin{cases} k^{-1}\varphi(z) & \text{if } -3 < x < 3 \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

where k is a normalization constant that ensures that $f(z)$ is a PDF:

$$k = \int_{-3}^{3} \varphi(z) \mathrm{d}z = 0.9973 \tag{1.3}$$

Although the function $f(z)$ has three pieces—a piece from $(-\infty, -3]$, where the function is defined as 0, a piece from $(3, 3)$ where the function is defined as in equation 1.2, and a piece from $[3, \infty)$, where the function is defined to be 0—we will ignore the two pieces where the function is defined to be 0 and consider $f(z)$ to be a 1-piece function.

This note is concerning the most efficient representation of piecewise functions in $Mathematica^{©}$ in the context of inference in hybrid Bayesian networks.

2 Representation of Piecewise Functions

One way of representing a piecewise function in $Mathematica^{©}$ is to use the built-in function called $Piecewise$. Thus the piecewise function $f(z)$ in equation (1.2) can be represented in $Mathematica^{©}$ by $g_1(z)$ as follows:

$$
\begin{aligned}
\varphi[z\_] &= PDF[NormalDistribution[0, 1.], z] \\
k &= \int_{-3}^{3} \varphi[z] \mathrm{d}z \\
g_1[z\_] &= Piecewise[\{\{k^{-1} \varphi[z], -3 < z < 3\}\}]
\end{aligned}
$$

Another way to represent a piecewise function is to use the built-in $HeavisideTheta$ function. Let $\theta(\cdot)$ denote the HeavisideTheta function. Then $\theta(z)$ is defined as follows:

$$\theta(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 & \text{if } z > 0 \\ \text{undefined} & \text{if } z = 0 \end{cases} \tag{2.1}$$

Notice that $\theta(z + 3) - \theta(z - 3)$ is a piecewise function of z that is $= 1$ if $-3 < z < 3$, $= 0$ if $z < 3$ or $z > 3$, and undefined if $z = 3$ or $z = -3$. Thus, we could represent the piecewise function $f(z)$ in equation (1.2) in $Mathematica^{©}$ by $g_2(z)$ as follows:

$$g_2[z\_] = k^{-1} \varphi[z](\theta[z + 3] - \theta[z - 3]) \tag{2.2}$$

A disadvantage of this representation is that we cannot distinguish between intervals that are open/closed at the end-points of the intervals. But since this is not of interest in representing PDFs, it is not material. Thus, $g_2(z)$ as defined in equation (2.2) is a well-defined PDF. We will refer to this representation of a piecewise function as a *difference of HeavisideTheta* (DH) representation. If we had a piecewise function that had more than one piece, we would represent each piece as in equation (2.2) and then sum the pieces.

A third representation of a piecewise function is by using product of $HeavisideTheta$ functions. Notice that $\theta(z+3) \cdot \theta(-z+3)$ is a piecewise function that $= 1$ if $-3 < z < 3$, $= 0$ if $z < -3$ or $z > 3$, and undefined if $z = -3$ or $z = 3$. Thus, we could represent the piecewise function $f(z)$ in equation (1.2) in $Mathematica^{©}$ by $g_3(z)$ as follows:

$$g_3[z_-] = k^{-1}f[z]\,\theta[z+3]\,\theta[-z+3] \tag{2.3}$$

This representation has the same disadvantages as the DH representation, and $g_3(z)$ is a well-defined PDF. We will refer to this representation of a piecewise as a *product of HeavisideTheta* (PH) representation.

3 Integration of Piecewise Functions

In making inference in hybrid Bayesian networks, marginalization of a continuous variable from a potential involves integration over its entire domain. However, we can choose the limits of integration as either $(-\infty, \infty)$ or we can choose the limits of integration where the function is non-zero, such as $(-3, 3)$ for $f(z)$ in equation (1.2). While both of these give exactly the same result, the time required for integration depends on the representation. An advantage of the $(-\infty, \infty)$ limits is that we do not have to know where the function is non-zero. This is especially so when the integrand is the product of several piecewise functions or the result of marginalizing a product of several piecewise functions.

Integration is also involved in normalizing the marginal distribution of a continuous variable in a hybrid Bayesian network with observations. Also, integration is also involved in finding the mean and variance of a continuous variable.

4 Some Experiments

In this section we will report the time required for a set of tasks for each of the three piecewise representation and for each of the two limits as a 2 by 3 matrix. The time for a task is recorded using the $Timing$ command in $Mathematica^{©}$, version 8.0.1.0, on a laptop MacBook Pro computer (with 2.66 GHz Intel Core i7 processor and 8 GB 1067 MHz DDR3 memory). Since the values recorded by the $Timing$ command are random numbers, we repeated each experiment a total of 10 times and report the mean and its standard error (SE) in parenthesis. Thus, a 95 % confidence interval for the mean time is $mean \pm 2.26 \cdot SE$ (based on the t-distribution with 9 degrees of freedom). The statistics reported in each 2 by 3 matrix are based on experiments done consecutively under identical conditions.

Task 1 is to compute the integral of $f(z)$ over the entire domain of Z, resulting in 1.0. Mean times required for this task are as follows (in seconds).

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.0825 (0.0039)	0.2282 (0.0035)	0.0380 (0.0006)
$(-3, 3)$	0.0169 (0.0007)	0.0102 (0.0003)	0.0083 (0.0004)

For this simple task, using $(-3, 3)$ limits requires less time than using $(-\infty, \infty)$ regardless of the method used. For the DH method, using $(-3, 3)$ required about 4% of the time required by $(-\infty, \infty)$, a dramatic reduction. The best method for this task is the PH method with $(-3, 3)$ limits.

Task 2 is to compute the mean of Z by integrating $z \cdot f(z)$ over the entire domain, resulting in 0.0. Times required for this task are as follows.

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.0309 (0.0004)	0.1927 (0.0015)	0.0472 (0.0007)
$(-3, 3)$	0.0170 (0.0007)	0.0175 (0.0003)	0.0085 (0.0003)

For this task, all methods recorded lower times with $(-3, 3)$ limits compared to $(-\infty, \infty)$ limits. The best method for this task is the PH method with $(-3, 3)$ limits.

Task 3 is to compute the variance of Z by integrating $z^2 \cdot f(z)$ and then subtracting the square of the mean (resulting in 0.973337). Times required for this task are as follows.

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.3187 (0.0015)	0.1977 (0.0016)	0.0775 (0.0007)
$(-3, 3)$	0.0202 (0.0007)	0.0507 (0.0009)	0.0111 (0.0003)

The results are similar to Task 2 results. All methods recorded lower times with $(-3, 3)$ limits compared to $(-\infty, \infty)$ limits. The best method for this task is PH with $(-3, 3)$ limits.

Next, we focus on tasks described by Shenoy [2] in the context of using mixtures of polynomials with the hyper-rhombus condition. Let g_{21} denote a 4-piece, 2-degree MOP approximation of the standard normal PDF using the Piecewise method. Let g_{22} denote the same function represented using the DH method, and let g_{23} denote the same function represented using the PH method. This MOP approximation was found as follows. First, we found a MOP approximation of the cumulative distribution function (CDF) of the standard normal distribution using the Lagrange interpolating polynomial [3] using four Chebyshev points [1] on the intervals $(-3, -1]$, $(-1, 0]$, $(0, 1]$, and $(1, 3)$. Next, we differentiate the MOP approximation of the CDF to obtain an unnormalized PDF. Finally, we normalize the PDF. The $Mathematica^{©}$ commands for the

definition of g_{21}, g_{22}, and g_{23} are as follows.

$$
\begin{aligned}
g_{21u}[z\_] \quad = \quad & Piecewise[\{ \\
& \{0.579878 + 0.410808z + 0.0740051z^2, -3 < z < -1\}, \\
& \{0.403643 + 0.0376952z - 0.128842z^2, -1 \le z < 0\}, \\
& \{0.403643 - 0.0376952z - 0.128842z^2, 0 \le z < 1\}, \\
& \{0.579878 - 0.410808z + 0.0740051z^2, 1 \le z < 3\} \\
& \}]; \\
k_{g_{21}} \quad = \quad & \int_{-3}^{3} g_{21u}[z]\,dz \\
g_{21}[z\_] \quad = \quad & g_{21u}[z]/k_{g_{21}} \\
g_{22u}[z\_] \quad = \quad & (0.579878 + 0.410808z + 0.0740051z^2)(HeavisideTheta[z+3] - HeavisideTheta[z+1]) + \\
& (0.403643 + 0.0376952z - 0.128842z^2)(HeavisideTheta[z+1] - HeavisideTheta[z]) + \\
& (0.403643 - 0.0376952z - 0.128842z^2)(HeavisideTheta[z] - HeavisideTheta[z-1]) + \\
& (0.579878 - 0.410808z + 0.0740051z^2)(HeavisideTheta[z-1] - HeavisideTheta[z-3]) \\
k_{g_{22}} \quad = \quad & \int_{-3}^{3} g_{22u}[z]\,dz \\
g_{22}[z\_] \quad = \quad & g_{22u}[z]/k_{g_{22}} \\
g_{23u}[z\_] \quad = \quad & (0.579878 + 0.410808z + 0.0740051z^2)\,HeavisideTheta[z+3]\,HeavisideTheta[-z-1] + \\
& (0.403643 + 0.0376952z - 0.128842z^2)\,HeavisideTheta[z+1]\,HeavisideTheta[-z] + \\
& (0.403643 - 0.0376952z - 0.128842z^2)\,HeavisideTheta[z]\,HeavisideTheta[-z+1] + \\
& (0.579878 - 0.410808z + 0.0740051z^2)\,HeavisideTheta[z-1]\,HeavisideTheta[-z+3] \\
k_{g_{23}} \quad = \quad & \int_{-3}^{3} g_{23u}[z]\,dz \\
g_{23}[z\_] \quad = \quad & g_{23u}[z]/k_{g_{23}}
\end{aligned}
$$

In all three cases, the normalization constant evaluates to the same constant 0.998427. *Mathematica*© has a built-in function called *LeafCount* that measures the total number of indivisible subexpressions in a function. The leaf counts for the three functions g_{21}, g_{22}, and g_{23} are 72, 87, and 83, respectively.

Task 4 is to compute the integral of $g_{2i}(z)$ over the entire domain of Z resulting in 1. Times required for this task are as follows.

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.0580 (0.0019)	*	0.2486 (0.0019)
$(-3, 3)$	0.0538 (0.0005)	0.2754 (0.0027)	0.4097 (0.0030)

For the DH method using $(-\infty, \infty)$ limits, *Mathematica*© generates a warning message saying "Indeterminate expression $0 \cdot \infty$ encountered." Thus, we do not report the statistics for this case. Henceforth, we will avoid using the DH method with the $(-\infty, \infty)$ limits. The best method for this task is the Piecewise

method with $(-3, 3)$ limits. The PH method does better with $(-\infty, \infty)$ limits compared to $(-3, 3)$ limits.

Task 5 is to compute the integral of $z \cdot g_2(z)$ over the entire domain of Z. The result of this integral is 0 for the DH representation. However, the Piecewise representation results in the value $-3.59612 \cdot 10^{-16}$ (since there is a slight asymmetry in the definition of this function at the point 0). The PH representation results in $2.22395 * 10^{-16}$ for the case of $(-\infty, \infty)$ limits, and $2.42861 * 10^{-16}$ for the case of $(-3, 3)$ limits. Times required for this task are as follows.

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.0532 (0.0006)	*	0.1692 (0.0011)
$(-3, 3)$	0.0581 (0.0011)	0.2219 (0.0018)	0.3009 (0.0024)

The best method for this task is the Piecewise method with $(-\infty, \infty)$ limits. As for Task 4, the PH method does better with $(-\infty, \infty)$ limits compared to $(-3, 3)$ limits.

Task 6 is to compute the variance of Z by integrating $z^2 \cdot g_2(z)$ and then subtracting the square of the mean resulting in the value 0.982553 for all methods. Times required for this task are as follows.

Time	Piecewise	DH	PH
$(-\infty, \infty)$	0.0541 (0.0004)	*	0.1535 (0.0005)
$(-3, 3)$	0.0574 (0.0008)	0.2175 (0.0021)	0.2666 (0.0009)

The results are similar to Task 5 results. The best methods for this task is the Piecewise method (with $(-\infty, \infty)$ limits). As for Tasks 4 and 5, the PH method does better with $(-\infty, \infty)$ limits compared to $(-3, 3)$ limits.

Task 7 is to compute the marginal distribution of Y in the Bayesian network with two variables: $Z \sim N(0, 1)$ and $Y|z \sim N(z, 1)$. For the PDF of Z, we use the 4-piece, 2-degree MOP approximation g_2 as described earlier. For the conditional PDF of Y given z, we used the MOP approximation $h_1(z, y) = g_2(y - z)$. The marginal distribution of Y is then given as follows.

$$h_2(y) = \int_{-\infty}^{\infty} g_2(z) \, h_1(z, y) \, \mathrm{d}z \qquad (4.1)$$

Notice that we can use limits $(-3, 3)$ in place of $(-\infty, \infty)$ in equation (4.1). The times required for computing the integral in equation (4.1) are as follows (in seconds)

Time	Piecewise	DH	PH
$(-\infty, \infty)$	10.24 (0.0142)	*	*
$(-3, 3)$	5.41 (0.0107)	13.58 (0.0053)	76.5 (0.0761)

For this task using the $(-\infty, \infty)$, like the DH method, the PH method results in the same warning as for the DH method. Thus, we don't report the statistics for this case. Here onwards, we skip the PH method for the case of $(-\infty, \infty)$ limits.

The best method for this task is Piecewise using $(-3, 3)$ limits, and the second best is Piecewise using $(-\infty, \infty)$ limits. One explanation for the difference in times is as follows. Using LeafCount, we measured the size of the expression for h_2 using the different methods. Results are as follows.

Leaf Counts of h_2	Piecewise	DH	PH
$(-\infty, \infty)$	560	*	*
$(-3, 3)$	560	1,449	17,587

The piecewise method results in the same size expression for both limits and has the lowest leaf count, and PH has the highest leaf count. This is one possible explanation for why the time for PH is some much higher then the times for Piecewise and DH methods
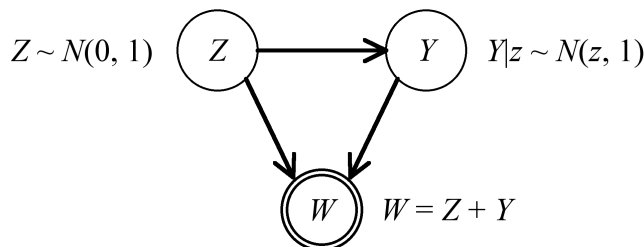

Figure 1: A Bayesian network with three variables

Task 8 is to compute the marginal distribution of W in the Bayesian network with three variables, $Z \sim N(0, 1)$, $Y|z \sim N(z, 1)$, and $W = Z + Y$ (see Figure 1). The conditional for W is deterministic. As for Task 7, the PDF of Z is given by the 4-piece, 2-degree MOP function $g_2(z)$, and the conditional for Y is denoted by $h_1(z, y) = g_2(y - z)$. The marginal for W is then computed using the convolution formula as follows:

$$h_3(w) = \int_{-\infty}^{\infty} g_2(z)\, h_1(z, w - z)\, \mathrm{d}z \tag{4.2}$$

Notice that we can use limits $(-3, 3)$ in place of $(-\infty, \infty)$ in equation (4.2). The times required for computing the integral in equation (4.2) are as follows (in seconds)

Time	Piecewise	DH	PH
$(-\infty, \infty)$	13.89 (0.0482)	*	*
$(-3, 3)$	9.27 (0.0201)	16.76 (0.0134)	98.9 (0.2828)

The best method for this task is Piecewise with $(-3, 3)$ limits.

5   Summary and Conclusions

The main goal of this note is to describe some different representations of piecewise function in $Mathematica^{\copyright}$ version 8.0.1.0, and study their implications

for the efficiency of computation. We also look at the effect of limits on the time required for doing some integrations.

One conclusion is that most of the time, using exact limits needs less time than using $(-\infty, \infty)$ limits. This seems intuitive as some knowledge is needed to compute exact limits for the marginals, and thus, it should be faster using exact limits. There are some exceptions to this rule (e.g., Task 5), and we do not know why. Perhaps due to a very small sample size of 10, the statistics may not be very reliable.

Another conclusion is that exact limits must be used with DH and PH methods to guarantee correct answers in $Mathematica^{©}$ 8.0.1.0. DH with exact limits usually does better than PH with exact limits.

The results presented here are valid only for $Mathematica^{©}$, version 8.0.1.0. The results for $Mathematica^{©}$, version 7.0.1.0, are quite different. The implementation of *HeavisideTheta* function in Mathematica 8.0.1.0 has apparently changed, and this is probably the reason for the difference in the efficiency of computation. For example, consider Task 8, computing the marginal of W in the Bayesian network of Figure 1. If this task was done in $Mathematica^{©}$ version 7.0.1.0, the results are as follows:

Time	Piecewise	DH	PH
$(-\infty, \infty)$	16.94 (0.0298)	10.19 (0.0300)	36.04 (0.1295)
$(-3, 3)$	17.38 (0.0559)	15.24 (0.0430)	54.91 (0.1517)

First, integrating the DH and PH representations with $(-\infty, \infty)$ limits does not generate any error messages. Second, the best method for this task is the DH method with $(-\infty, \infty)$ limits. The DH method with $(-\infty, \infty)$ limits in $Mathematica^{©}$ v. 7.0.1.0 is almost as fast as the Piecewise method with $(-3, 3)$ in $Mathematica^{©}$ v. 8.0.1.0. Notice that the times for Piecewise methods are lower using $Mathematica^{©}$ v. 8.0.1.0 compared to $Mathematica^{©}$ v. 7.0.1.0. Regardless of the version of $Mathematica^{©}$ used, the PH method is not as fast as either Piecewise or DH. The leaf counts of $h_3(w)$ with the different methods are as follows.

Leaf Counts of h_3	Piecewise	DH	PH
$(-\infty, \infty)$	759	669	15,583
$(-3, 3)$	759	3,798	26,622

There is a high correlation between the leaf counts of the computed function and the time required for computation.

So, assuming one has access to both versions of $Mathematica^{©}$, versions 7.0.1.0 and 8.0.1.0, should we use $Mathematica^{©}$, v. 8.0.1.0 with Piecewise or $Mathematica^{©}$, v. 7.0.1.0 with DH? To resolve this question, consider the Bayesian network shown in Figure 2. This Bayesian network has a 3-dimensional conditional for X given Z and Y, which is represented by the 4-piece, 2-degree MOP function $h_4(z, y, x) = g_2(x - (z + y))$. Also, to find the marginal PDF of V requires the use of two convolutions as follows

$$h_5(v) = \int_{-\infty}^{\infty} g_2(z) \left(\int_{-\infty}^{\infty} h_1(z, y) \, h_4(z, y, v - (z + y)) \, \mathrm{d}y \right) \mathrm{d}z \qquad (5.1)$$
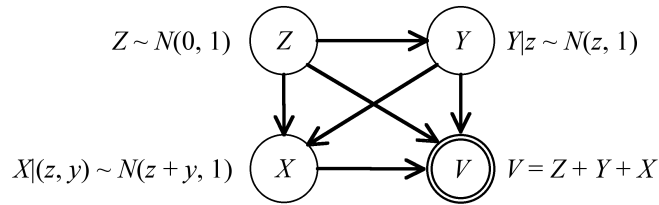
Figure 2: A Bayesian network with four variables

Notice that in equation 5.1, we can use limits $(-6, 6)$ for the inner integral (with respect to y) and $(-3, 3)$ for the outer integral (with respect to z). We computed $h_5(v)$ using the Piecewise method in $Mathematica^{©}$ v. 8.0.1.0 with exact limits, and using the DH method in $Mathematica^{©}$ v. 7.0.1.0. with $(-\infty, \infty)$ limits. Results are as follows. Using DH with $(-\infty, \infty)$ limits in $Mathematica^{©}$ v. 7.0.1.0 required a mean of 123.80 seconds (with a SE = 0.43 seconds). Using Piecewise with exact limits, $Mathematica^{©}$ v. 8.0.1.0 ran out of available memory (it was the only foreground application running besides the Finder). For obvious reasons, we did not replicate the test with $Mathematica^{©}$ v. 8.0.1.0 10 times.

If one uses the Piecewise method, it is possible to count the number of pieces in the result. Also, we can see the regions of each piece, and the region where the piecewise function is non-zero. If one uses the DH or PH methods, it is not quite as easy to count the number of pieces in the resulting marginal, or discern the region where the function is non-zero. Of course, we can plot, compute means and variances, etc., regardless of the method used.

Acknowledgements

References

[1] R. L. Burden and J. D. Faires. *Numerical Analysis.* Brooks Cole, 9th edition, 2010.

[2] P. P. Shenoy. Some issues in using mixtures of polynomials for inference in hybrid Bayesian networks. Working Paper 323, University of Kansas School of Business, Lawrence, KS, November 2010.

[3] E. Waring. Problems concerning interpolations. *Philosophical Transactions of the Royal Society of London*, 69:59–67, 1779.