Continuous Decision Variables with Multiple Continuous Parents

Barry R. Cobb
Virginia Military Institute
cobbbr@vmi.edu

Abstract

This paper introduces an influence diagram (ID) model that permits continuous decision variables with multiple continuous parents. The marginalization operation for a continuous decision variable first develops a piecewise linear decision rule as a continuous function of the next continuous parent in the deletion sequence. Least squares regression is used to convert this rule to a piecewise linear function of all the decision variable's continuous parents. This procedure is incorporated into an iterative solution algorithm that allows more refined decision rules to be constructed once the non-optimal regions of the state spaces of decision variables are identified. Additional examples serve to compare relative advantages of this technique to other ID models proposed in the literature.

1 Introduction

The *influence diagram* (ID) is a graphical and numerical representation for a decision problem under uncertainty (Howard and Matheson, 1984). The ID model is composed of a directed acyclic graph that shows the relationships among chance and decision variables in the problem, as well as a set of conditional probability distributions for chance variables and a joint utility function. An example of a decision problem under uncertainty is given in the following section.

1.1 Example

A firm facing uncertain demand must choose production capacity and set product prices (Göx, 2002). Product demand is determined as $Q(p,z) = 12 - p + z$, where P is the product price and Z is a random demand "shock." Assume $Z \sim N(0,1)$ and that the firm's utility (profit) function is

$$u_0(k,p,z) = \begin{cases} (p-1)\cdot(12-p+z)-k & \text{if } Q(p,z) \le k \\ (p-1)\cdot k - k & \text{if } Q(p,z) > k \,. \end{cases} \quad (1)$$

Notice that the firm's sales are limited to the minimum of product demand and production capacity (K). Figure 1 shows an ID model for the example. The chance and decision variables in the ID are depicted as ovals and rectangles, respecitively. The joint utility function appears as a diamond. Since there is an arrow pointing from Z to K and P, Z is a *parent* of K and P. The set of all parents of P is $Pa(P) = \{K, Z\}$. An arrow pointing to a chance node indicates the distribution for this node is conditioned on the variable at the head of the arrow. An arrow pointing to a decision node means that the value of the variable will be known when the decision is made.

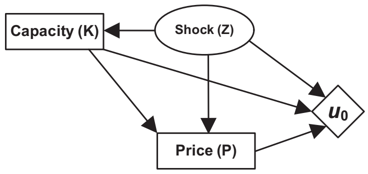

Figure 1: Influence Diagram Model.

1.2 Background

Although most ID models proposed in the literature assume that all decision variables take values in discrete (countable) state spaces, there are some exceptions. Shachter and Kenley (1989) introduce Gaussian IDs, where all continuous chance variables are normally distributed, all decision variables are continuous, and utility functions are quadratic.

The mixture-of-Gaussians ID (Poland and Shachter, 1993) requires continuous chance variables to be modeled as mixtures of normal distributions and allows continuous decision variables. Madsen and Jensen (2005) outline an improved solution procedure for IDs constrained under the same conditions as mixture-of-Gaussians IDs that is able to take advantage of an additive factorization of the joint utility function.

Cobb (2007) introduces an ID model which allows continuous decision variables with one continuous parent and continuous chance variables having any probability density function (pdf). Using this approach, pdfs and utility functions are approximated by mixtures of truncated exponentials (MTE) potentials (Moral et al., 2001), which allows the marginalization operation for continuous chance variables to be performed in closed form. This technique develops a piecewise linear decision rule for continuous decision variables and subsequently marginalizes them from the model as deterministic chance variables.

This paper builds upon the model in (Cobb, 2007) by allowing continuous decision variables to have multiple continuous parents. The marginalization operation for continuous decision variables and the iterative solution algorithm are introduced using examples. A longer working paper (Cobb, 2010) contains more formal definitions.

The remainder of this paper is organized as follows. In §2, notation and definitions are introduced. In §3, a procedure for marginalizing a continuous decision variable is presented using the example in §1.1. In §4, the results from the example problem are compared to an analytical solution. §5 describes solutions to additional examples before §6 concludes the paper.

2 Notation and Definitions

2.1 Notation

In this paper, we assume all decision and chance variables take values in finite-bounded, continuous (non-countable) state spaces. All variables are denoted by capital letters in plain text, e.g., A, B, C. Sets of variables are denoted by capital letters in boldface, with $\mathbf{Z}$ representing chance variables, $\mathbf{D}$ representing decision variables, and $\mathbf{X}$ indicating a set of variables whose components are a combination of chance and decision variables. If A and $\mathbf{X}$ are one- and multi-dimensional variables, respectively, then a and $\mathbf{x}$ represent specific values of those variables. The finite-bounded, continuous state space of $\mathbf{X}$ is denoted by $\Omega_{\mathbf{X}}$.

Example 1. *In the ID shown in Figure 1, the state spaces of the variables are* $\Omega_K = \{k : 0 \leq k \leq 14\}$, $\Omega_P = \{p : 1 \leq p \leq 9\}$, *and* $\Omega_Z = \{z : -3 \leq z \leq 3\}$. *This assumes the distribution for* Z *is normalized over the interval* $[-3, 3]$ *to solve the ID.*

MTE probability potentials are denoted by lower-case Greek letters, e.g., ϕ, ψ, φ, whereas MTE utility potentials are denoted by u_i, where the subscript i is normally zero for the joint utility function in the problem, and one for the initial MTE approximation to the joint utility function. The subscript can be increased to index additional MTE utility potentials in the initial representation or solution.

2.2 Mixtures of Truncated Exponentials (MTE) Potentials (Moral et al., 2001)

Let $\mathbf{X}$ be a mixed variable and let $\mathbf{Z} = (Z_1, \ldots, Z_c)$ and $\mathbf{D} = (D_1, \ldots, D_f)$ be the chance and decision variable parts of $\mathbf{X}$, respectively. Given a partition $\Omega_1, \ldots, \Omega_n$ that divides $\Omega_{\mathbf{X}}$ into hypercubes, an n-piece MTE potential $\phi : \Omega_{\mathbf{X}} \mapsto \mathcal{R}^+$ has components

$$\phi_h(\mathbf{z}, \mathbf{d}) = a_0 + \sum_{i=1}^{m} a_i \exp\left\{\sum_{j=1}^{c} b_i^{(j)} z_j + \sum_{\ell=1}^{f} b_i^{(c+\ell)} d_\ell\right\}$$

for $h = 1, \ldots, n$, where $a_i, i = 0, \ldots, m$ and $b_i^{(j)}$, $i = 1, \ldots, m$, $j = 1, \ldots, (c+f)$ are real numbers.

We assume all MTE potentials are equal to zero in unspecified regions. In this paper, all probability distributions and utility functions are approximated by MTE potentials.

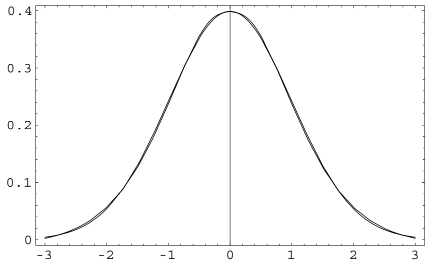

Figure 2: $N(0,1)$ pdf and MTE Potential ϕ_1.

Example 2. *The function $f_1(p) = p$ over the interval $[1, 9]$ can be approximated by the MTE potential*

$$u_P(p) = \begin{cases} -53.028072 + 54.05148 \exp\{0.017847(p-1)\} & \text{if } 1 \le p < 5 \\ -49.028072 + 54.05148 \exp\{0.017847(p-5)\} & \text{if } 5 \le p \le 9\,, \end{cases}$$

using the method described in (Cobb and Shenoy, 2008). Similar MTE potentials, $u_K(k)$ and $u_Z(z)$, are used to approximate the functions $f_2(k) = k$ on $[0, 14]$ and $f_3(z) = z$ on $[-3, 3]$. These approximations are are substituted into (1) to form the MTE utility function u_1 for the example problem.

The resulting function will contain values of variables in the limits of the domain. In other words, to create a true MTE potential where the limits are hypercubes, values for two of the variables must be substituted. MTE potentials defined in this way require replacement of linear terms when integration is used to marginalize variables in the ID solution. This is discussed in (Cobb and Shenoy, 2006).

Example 3. *The MTE approximation ϕ_1 to the $N(0,1)$ pdf (see Cobb et al., (2006) for numerical details) that approximates the distribution for the variable Z in the example from §1.1 is shown in Figure 2, overlaid on the actual $N(0,1)$ distribution. The MTE function is normalized on the interval $[-3, 3]$.*

2.3 Fusion Algorithm

IDs are solved in this paper by applying the fusion algorithm of Shenoy (1993), which is relevant for the case where the joint utility function factors multiplicatively. This algorithm involves deleting the variables in an elimination sequence that respects the information constraints in the problem. The sequence is chosen so that decision variables are eliminated before chance or decision variables that are immediate predecessors.

When a variable is to be deleted from the model, all probability and/or utility potentials containing this variable in their domains are combined via pointwise multiplication, then the variable is marginalized from the result. The appropriate marginalization operation depends on whether the variable being marginalized is a chance variable (in which case marginalization is accomplished by integrating over the domain of the chance variable being removed) or a decision variable. Formal definitions of combination and marginalization of chance variables can be found in (Cobb, 2010).

3 Marginalizing Decision Variables

Assume we want to eliminate a decision variable D with parents $\mathbf{X} = \{X_1, X_2 \ldots, X_n\}$ from the ID. The variables in $\mathbf{X}$ may be either chance or decision variables, and the subscripts on variables in $\mathbf{X}$ serve to number the variables as they appear in the deletion sequence for the problem. The set of parents excluding X_1 is denoted by $\mathbf{X}' = \mathbf{X} \setminus X_1$. Eliminating the decision variable is a four-step process:

(1) Combine all potentials containing D in their domain, create discrete approximations to Ω_D and $\Omega_{\mathbf{X}'}$, and find the discrete value of D that maximizes utility for each region of a hypercube of Ω_{X_1} for each (discrete) $\mathbf{x}' \in \Omega_{\mathbf{X}'}$.

(2) For each (discrete) $\mathbf{x}' \in \Omega_{\mathbf{X}'}$, create a decision rule for D as a piecewise linear function of X_1.

(3) Use least squares regression to create a piecewise linear decision rule for D as a function of $\mathbf{X}$.

(4) Convert D to a deterministic chance variable, and marginalize D using the procedure in §3.4.

This process has similarities to the procedure for marginalizing a continuous decision variable proposed by Cobb (2007); however, employing regression in Step 3 enables this new operation to permit continuous decision variables with multiple continuous parents. The steps are introduced by illustrating the removal of P from the ID of §1.1 using the deletion sequence P, K, Z.

In this solution, we utilize $v = 8$ discrete values and regions at each step in the process when we are required to discretize or sub-divide the state space of a continuous variable.

3.1 Step 1–Discrete Approximation

The purpose of this step in the marginalization process is to find a relationship—given a value of Z—between the optimal price (P) and production capacity (K) by examining the utility function for various values of P.

In this step, discrete values p_u, $u = 1, \ldots, 8$, for P are assigned as $\{1.5, 2.5, \ldots, 8.5\}$. Assign discrete values z_t, $t = 1, \ldots, 8$, to the chance variable Z, the most distant parent of P in the deletion sequence. Based on the state space Ω_Z, these discrete values are $\{-2.625, -1.875, -1.125, \ldots, 2.625\}$. For each discrete value z_t, create an MTE utility function $u_1(k, p, z_t)$ by substituting $Z = z_t$ in u_1.

For each value z_t, determine the (discrete) value in Ω_P that maximizes the utility function $u_1(k, p, z_t)$ for each region of a hypercube of Ω_K. For example, when $Z = z_3 = -1.125$, the utility functions $u_1(k, p_u, z_3)$ appear as shown in Figure 3. From the diagram, it is apparent that $u_1(k, 8.5, z_3) \approx u_1(k, 7.5, z_3)$ when $K = 2.75$, $u_1(k, 7.5, z_3) \approx u_1(k, 6.5, z_3)$ when $K = 4.05$, and $u_1(k, 6.5, z_3) \approx u_1(k, 5.5, z_3)$ when $K = 5.35$.

The results of this step of the operation are the sets of points, $\Phi_{1,t}$, and decision variable values, $\Psi_{1,t}$, for $t = 1, \ldots, 8$. For instance, $\Phi_{1,3} = \{0, 2.75, 4.05, 5.35, 14\}$ and $\Psi_{1,3} = \{8.5, 7.5, 6.5, 5.5\}$, where the three in the sub-

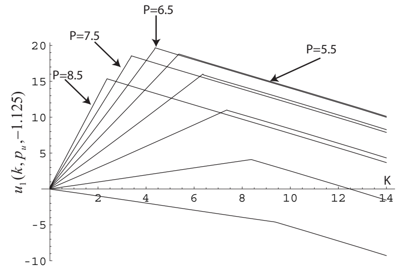

Figure 3: The Utility Functions $u_1(k, p_u, z_3)$.

scripts is an index on the related value $z_3 = -1.125$. The set $\Phi_{1,3}$ can be used to determine intervals where the optimal discrete value of price is invariant, and the set $\Psi_{1,3}$ contains the optimal values for P corresponding to these intervals.

This procedure is derived from the operation for marginalizing a discrete decision variable in a hybrid ID (Cobb and Shenoy, 2008).

3.2 Step 2–Piecewise Linear Decision Rule

The purpose of this step is to express the relationship between optimal price (P) and production capacity (K) by estimating a continuous function $P = f(K)$, given a value for Z.

Continuing from §3.1, when $-1.5 \leq z \leq -0.75$, $\Phi_{1,3}$ is used to determine $k = \left(\frac{0+2.75}{2}, \frac{2.75+4.05}{2}, \frac{4.05+5.35}{2}, \frac{5.35+14}{2}\right) = (1.375,\ 3.4,\ 4.7,\ 9.675)$, with a corresponding set of points, $p = (8.5,\ 7.5,\ 6.5,\ 5.5)$ defined as in $\Psi_{1,3}$. The equation for the line connecting the coordinates $\{(k = 1.375, p = 8.5), (k = 3.4, p = 7.5)\}$ is $p(k) = 9.17901 - 0.49383k$. Similar equations are determined using other sets of adjacent coordinates and these form a piecewise linear decision rule for P as

$$\hat{\Psi}_{1,3}(k) = \begin{cases} 9 & 0 \leq k < 0.3625 \\ 9.17901 - 0.49383k & 0.3625 \leq k < 3.4 \\ 10.11540 - 0.76923k & 3.4 \leq k < 4.7 \\ 7.44472 - 0.20101k & 4.7 \leq k \leq 14\ . \end{cases}$$

The equation for the first (last) line segment is extrapolated until the result of the function is greater (less) than the endpoint of Ω_P, in which case the function is defined as the maximum (minimum) value in Ω_P. A similar decision rule is developed for the regions with midpoints z_t, $t = 1, \ldots, 8$. A function $\hat{\Psi}_1$ is comprised of the resulting piecewise functions as

$$\hat{\Psi}_1(k,z) = \begin{cases} \hat{\Psi}_{1,1}(k) & -3 \leq z < -2.25 \\ \vdots & \vdots \\ \hat{\Psi}_{1,8}(k) & 2.25 \leq z \leq 3 \; . \end{cases}$$

3.3 Step 3–Least Squares Regression

This step further refines the decision rule to be a compact piecewise linear function for optimal price given values of K and Z.

Continuing from §3.2, the state space of K—the next parent of P in the deletion sequence—is divided into 8 regions, $[0, 1.75], \ldots, [12.25, 14]$, with the m-th region denoted by $[k^d_{m-1}, k^d_m]$ where $k^d_m = k_{min} + m \cdot (k_{max} - k_{min})/v$ for $m = 0, \ldots, v$. Define $\varepsilon = 0.1$, $n_K = \lfloor (k_{max} - k_{min})/\varepsilon \rfloor + 1$, and $n_Z = \lfloor (z_{max} - z_{min})/\varepsilon \rfloor + 1$. The function $\hat{\Psi}_1$ is used to output a series of ordered data points $\{\hat{\Psi}_1(k_i, z_j), k_i, z_j\}$ for each $k_i = k_{min} + (i-1) \cdot \varepsilon$, $i = 1, \ldots, n_K$ and $z_j = z_{min} + (j-1) \cdot \varepsilon$, $j = 1, \ldots, n_Z$. These ordered data points are sorted into ascending order according to the values k_i and grouped into $v = 8$ tables, where the m-th table contains points such that all $k_i \in [k^d_{m-1}, k^d_m]$ for each $m = 1, \ldots, v$. In other words, for each value k_i that appears in the m-th table, each pair (k_i, z_j) appears exactly once, along with the corresponding values $\hat{\Psi}_1(k_i, z_j)$. Each table is used to create the matrices required to estimate a linear equation $\hat{p}(k,z) = b_{2m} + b_{3m} \cdot k + b_{4m} \cdot z$ via least squares regression.

For example, with $\varepsilon = 0.1$, $n_K = 141$, and $n_Z = 61$, so 61 values for Z are matched with each of the 18 values of K in the second interval, $[1.75, 3.5]$, defined using Ω_K. Thus, $18 \times 61 = 1098$ data points are used to define the (1098×1) matrix Υ_2 and the (1098×3) matrix Λ_2 as follows:

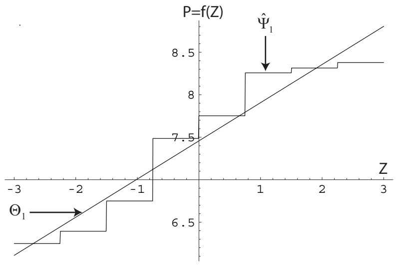

Figure 4: The Decision Rule $\Theta_1(4.375, z)$ for P.

$$\Upsilon_2 = \begin{bmatrix} \hat{\Psi}_1(k_{19}, z_1) \\ \vdots \\ \hat{\Psi}_1(k_{19}, z_{61}) \\ \vdots \\ \hat{\Psi}_1(k_{36}, z_1) \\ \vdots \\ \hat{\Psi}_1(k_{36}, z_{61}) \end{bmatrix} \quad \Lambda_2 = \begin{bmatrix} 1 & k_{19} & z_1 \\ \vdots & \vdots & \vdots \\ 1 & k_{19} & z_{61} \\ \vdots & \vdots & \vdots \\ 1 & k_{36} & z_1 \\ \vdots & \vdots & \vdots \\ 1 & k_{36} & z_{61} \end{bmatrix}$$

The least squares regression estimators are determined as $\mathbf{b}_2 = [\; b_{22} \;\; b_{32} \;\; b_{42} \;]^\top = \left(\Lambda_2^\top \Lambda_2\right)^{-1} \Lambda_2^\top \Upsilon_2$. Following this process in each region of the state space of K creates the following piecewise linear decision rule:

$$\Theta_1'(k,z) = \begin{cases} 9.0194 - 0.3631k + 0.0986z & 0 \leq k < 1.75 \\ 9.0265 - 0.3808k + 0.3246z & 1.75 \leq k < 3.5 \\ \vdots & \vdots \\ 8.1670 - 0.1731k + 0.5340z & 12.25 \leq k \leq 14 \; . \end{cases}$$

A revised piecewise linear decision rule is then determined as

$$\Theta_1(k,z) = \begin{cases} p_{min} & \Theta_1'(k,z) < p_{min} \\ \Theta_1'(k,z) & p_{min} \leq \Theta_1'(k,z) \leq p_{max} \\ p_{max} & \Theta_1'(k,z) > p_{max} \; . \end{cases}$$

Using this revised formulation of the decision rule ensures that the assigned values are contained in Ω_P. A graphical view of the decision rule for P as a function of Z given that $K = 4.375$ is shown in Figure 4. The decision rule Θ_1 is a refinement of the decision rule $\hat{\Psi}_1$.

3.4 Step 4–Removing the Decision Variable

Continuing from §3.3, since a value for P will be completely determined by observed values of K and Z, P can be replaced in the joint utility function as $u_2(k,z) = u_1(\Theta_1(k,z),k,z)$. The substitution of Θ_1 for P in u_1 is accomplished on a piecewise basis. For instance, when $5.25 \leq k \leq 7$, Θ_1 is defined as $f_1(k,z) = 8.16701 - 0.17307k + 0.47450z$ for all $z \in \Omega_Z$. When $k \geq 12 - p + z$, $0 \leq z \leq 3$, and $1 \leq p \leq 5$, u_1 is defined as

$$f_2(p,k,z) = -1067.4336 - 94.5901\exp\{0.0102k\} + \cdots + 2311.6851\exp\{0.0179p + 0.02380z\} + \cdots$$

The calculation of $u_2(k,z) = u_1(\Theta_1(k,z),k,z)$ includes the result of the substitution $f_2(f_1(k,z),k,z)$, with the ensuant expression included in u_2 where the domains of the two functions overlap, or

$$f_2(f_1(k,z),k,z) = -1067.4336 - 94.5901\exp\{0.0102k\} + \cdots + 2311.6851\exp\{0.0179 \cdot f_1(k,z) + 0.02380z\} + \cdots$$

for $5.25 \leq k \leq 7$ and $0.8269k - 0.5255z \geq 3.8330$. A similar substitution of each piece of Θ_1 is made into each piece of u_1 to create the MTE utility function u_2.

3.5 Results

To complete the example problem, the decision variable K is marginalized using the process in §3.1 through 3.4, except that since K has only one parent (Z), Step 3 (least squares regression) is not performed. The decision rule for K as a function of Z is determined as $\Theta_2(z) = 5.2646 + 0.5833z$ for all $z \in \Omega_Z$. To marginalize K, a new utility function u_3 is determined as $u_3(z) = u_2(\Theta_2(z), z)$. The firm's expected utility is then calculated as $\int_{\Omega_Z} \phi_1(z) \cdot u_3(z)\, dz = 24.6394$.

The ID method presented in this paper is sensitive to the state spaces assigned to the decision variables. In other words, if the continuous interval of possible optimal values for the decision variables can be narrowed, the accuracy of the decision rules can be improved.

In this example, the ID decision rule $\Theta_2(z)$ only selects values for K in the interval $[3.5146, 7.0146]$. Similarly, the decision rule $\Theta_1(k,z)$ only allows for values of P in the interval $[4.142, 9]$. In a second iteration of the ID solution procedure, we can replace the original state spaces of the decision variables P and K with these intervals and obtain a better approximation to the true optimal decision rules and profit function.

To complete the second iteration for the example, the same marginalization procedure is used to develop decision rules for P as a function of $\{K, Z\}$ and K as a function of Z.

Cobb (2009) explains additional details of the iterative algorithm.

4 Comparison

In the example problem, the firm knows the true value, $Z = z$, at the time it chooses capacity. Göx (2002) uses this fact to find an analytical solution for the optimal capacity of $k^*(z) = \frac{10+z}{2}$. This result hinges on several restrictive assumptions, including the linearity of the demand function and the symmetric form of the distribution for Z. By choosing an example with an analytical solution, we can compare the results from the ID solution as a means of determining its accuracy. The ID method can then be extended to cases where an analytical solution is not available (see §5).

The decision rule Θ_2 for K determined using two iterations of the ID solution procedure is shown in Figure 5 with the analytical capacity decision rule. This decision rule has seven linear pieces. The mean squared error (MSE) (Winkler and Hays, 1970) can be used as a measure of the difference between the analytical and estimated decision rules. The MSE is calculated as

$$\int_{\Omega_Z} \phi_1(z) \cdot \left(\Theta_2(z) - k^*(z)\right)^2 \, dz = 0.01469\,.$$

The MSE after the first iteration is 0.07682, so revising the state space and performing the second iteration improves the accuracy of the decision rule. The decision rule Θ_1 for P is used in the determination of Θ_2, so this MSE measurement is a measure of the accuracy of the decision rules developed the ID solution. The

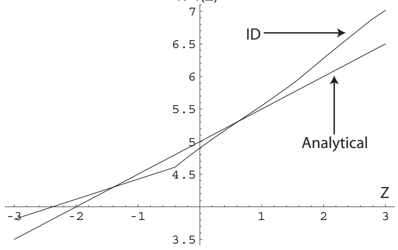

Figure 5: Decision Rules for $K = f(Z)$.

expected profit is 25.0792, as compared to the first iteration and analytical values of 24.6394 and 25.25, respectively.

5 Additional Examples

This section briefly describes two additional examples derived from the problem in §1.1 (for additional details, see (Cobb, 2010)).

5.1 Non-Gaussian Chance Variable

One advantage of using the ID model described in this paper is that it can accommodate non-Gaussian chance variables directly without using a mixture-of-Gaussians representation.

Suppose that the firm has established production capacity at a minimum of 3.5 units and a maximum of 6.5 units. The random variable K represents the percentage of additional capacity (above minimum) available (which fluctuates with changes in labor and machine utilization) and is modeled with a $Beta(3,3)$ distribution. The distribution for K is approximated by the MTE potential ϕ_2 determined using the method discussed by Cobb et al. (2006). The MTE approximation ϕ_1 to the distribution for Z remains the same.

Although P now has two parents (K and Z) that are chance variables (one of which is non-Gaussian), the procedure for marginalizing P from the ID proceeds in exactly the same way as in the previous example.

5.2 Nonmonotonic Decision Rule

Suppose P and K are decision variables as in §1.1, but that the unit variable cost of \$1 is replaced in the joint utility function by z^2, i.e. unit variable costs are now higher for values of the demand shock Z farther from zero. This utility function is approximated with an MTE utility function as in Example 2.

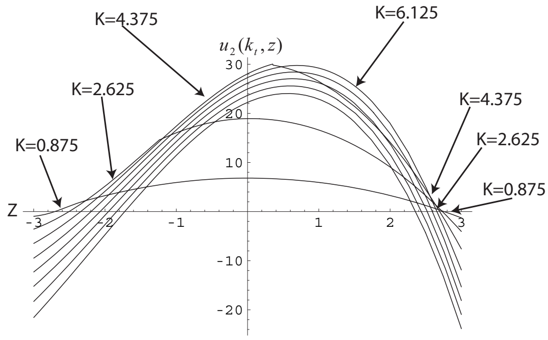

Figure 6: The Utility Functions $u_2(k_t, z)$ in the Example with Revised Unit Variable Cost.

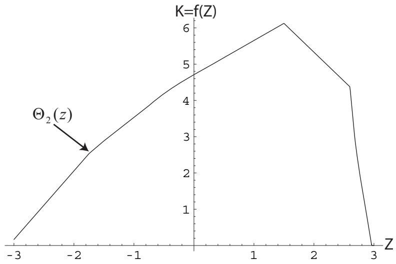

Figure 7: The Decision Rule Θ_2 for K.

Figure 6 shows the utility function (for eight discrete values of K) after marginalizaton of P and illustrates that the optimal value for K must be determined as a nonmonotonic function of Z. For instance, when $-2.25 \leq z \leq -1.15$ or $2.65 \leq z \leq 2.75$, a capacity of $K = 2.625$ is optimal, whereas if $-1.15 \leq z \leq 0.45$ or $2.55 \leq z \leq 2.65$, the best value of K is 4.375. Ultimately, these points are used to create the decision rule Θ_2 for K as a function of Z (see Figure 7).

6 Conclusions

This paper has introduced improvements to the model of Cobb (2007) that allow continuous decision variables in IDs to have multiple continuous parents. The framework proposed in this paper has some potential advantages over other ID models. To use the model in (Cobb, 2007) to solve the example in §1.1, we would have to impose one of the following restrictions: (1) model

K as a discrete decision variable or Z as a discrete chance variable, as the continuous decision variable P would be allowed to have only one continuous parent; or (2) discretize any pair of chance and/or decision variables. A comparison of the model in this paper to related models is provided in (Cobb, 2010).

The method in this paper permits non-Gaussian pdfs to be modeled without using mixtures of Gaussian distributions. This is in constrast to Gaussian IDs (Shachter and Kenley 1989) and mixtures-of-Gaussians IDs (Poland and Shachter, 1993; Madsen and Jensen, 2005). Additionally, those models determine only linear—as opposed to piecewise linear—decision rules, and thus cannot accomodate a case where the optimal decision rule is a nonmonotonic function of a decision rule's continuous parent(s), as in the example of §5.2.

Additional research is needed to demonstrate potential applications of the ID model and explain the compromise between computational cost and decision rule accuracy when parameters in the solution technique are altered. The model presented here is that the methodology has been designed to extend the ID model from (Cobb, 2007). There are other methods that could be employed to determine decision rules for continuous variables with multiple continuous parents, such as a straightforward grid search or a sampling technique. Future research will be aimed at exploring these methods and comparing them with those in this paper.

Acknowledgments

The author is grateful for the insightful comments of three anonymous reviewers. Support from the Spanish Ministry of Science and Innovation through project TIN2007-67418-C03-02 and by EFDR funds is gratefully acknowledged.

References

Cobb, B.R., 2007. Influence diagrams with continuous decision variables and non-Gaussian uncertainties. *Decision Anal.* **4**(3) 136–155.

Cobb, B.R., 2009. Efficiency of influence diagram models with continuous decision variables. *Decis. Support Syst.* **48** 257–266.

Cobb, B.R. 2010. Observing Multiple Continuous Sources of Information in Decision Problems. Working Paper, Virginia Military Institute, Lexington, VA. Available for download at: `www.vmi.edu/fswebs.aspx?tid=24697&id=24791`

Cobb, B.R., P.P. Shenoy. 2008. Decision making with hybrid influence diagrams using mixtures of truncated exponentials. *Eur. J. Oper. Res.* **186**(1) 261–275.

Cobb, B.R., P.P. Shenoy, R. Rumí. 2006. Approximating probability density functions in hybrid Bayesian networks with mixtures of truncated exponentials. *Statist. Comput.* **16**(3) 293–308.

Göx, R.F. 2002. Capacity planning and pricing under uncertainty. *J. Management Accounting Res.* **14**(1) 59–78.

Howard, R.A., J.E. Matheson. 1984. Influence diagrams. R.A. Howard, J.E. Matheson, eds. *Readings on the Principles and Applications of Decision Analysis II.* Strategic Decisions Group, Menlo Park, CA, 719–762.

Madsen, A.L., F. Jensen. 2005. Solving linear-quadratic conditional Gaussian influence diagrams. *Internat. J. Approx. Reason.* **38**(3) 263–282.

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

Poland, W.B., R.D. Shachter. 1993. Mixtures of Gaussians and minimum relative entropy techniques for modeling continuous uncertainties. D. Heckerman, E.H. Mamdani, eds. *Uncertainty in Artificial Intelligence: Proc. Ninth Conf.*, Morgan Kaufmann, San Francisco, CA, 183–190.

Shachter, R.D., C.R Kenley. 1989. Gaussian influence diagrams. *Management Sci.* **35**(5) 527–550.

Shenoy, P.P. 1993. A new method for representing and solving Bayesian decision problems. D.J. Hand, ed. *Artificial Intelligence Frontiers in Statistics: AI and Statistics III.* Chapman and Hall, London, 119–138.

Winkler, R.L., W.L. Hays. 1970. *Statistics: Probability, Inference, and Decisions.* Holt, Rinehart, and Winston, New York.