Dependent Decision Analysis

by

Robert T. Clemen
Fuqua School of Business
Duke University
Durham, NC 27708
clemen@mail.duke.edu

and

Terence Reilly
Lundquist College of Business
University of Oregon
Eugene, OR 97403
reilly@euclid.uoregon.edu

November, 1996

Dependent Decision Analysis

Abstract

Although any given decision analysis may unfold in a problem-specific way, a number of steps are commonly found in an analysis, including creating a deterministic model, performing a sensitivity analysis to identify critical variables, modeling uncertainty, evaluating the model by calculating expected value, expected utility, or risk profiles, and refining the model on the basis of expected-value-of-information or further sensitivity-analysis results. In fact, the steps in a conventional decision analysis have been formalized along precisely these lines.

In this paper, we describe how the conventional approach can be enhanced by introducing dependence assessments (Spearman's r) for the input variables in the initial sensitivity analysis, which in turn provide flexibility in the uncertainty-modeling stage. By leveraging the information in the dependence assessments, conditional relations and conditional distributions can be approximated with no further judgmental assessments. Alternatively, a complete joint distribution can be constructed. Such a model is based on the use of the dependence assessments to construct an appropriate copula representation of the joint distribution. We demonstrate how pairwise judgments of correlations, along with fractile assessments from the initial sensitivity analysis, can be used with the multivariate-normal copula to create a joint distribution for the critical random variables in a simple example.

(*Measures of Dependence, Spearman's* r, *Decision Analysis Process, Sensitivity Analysis, Copulas, Multivariate Normal Copula*)

1. Introduction

The term “decision analysis” (DA) connotes a general framework for modeling and analyzing decision situations. DA embraces a wide variety of tools for structuring decisions, assessing subjective probabilities and preferences, and analyzing decision models. In this paper, we focus on a specific decision-analysis paradigm developed and popularized by Ron Howard and his colleagues (e.g., see Howard and Matheson, 1983). This approach, which is sometimes called the “Stanford approach,” is especially designed to address decisions under uncertainty, and typically that uncertainty is modeled on the basis of subjective judgments by experts. Those subjective probability judgments typically include assessments of both marginal and conditional probability distributions.

An early step in the Stanford approach is using sensitivity analysis to determine which input variables affect the ranking of the alternatives. Based on the sensitivity-analysis results, an uncertainty model is created for those critical variables, and the analysis proceeds by using that model to calculate expected values, expected utility, or risk profiles (probability distributions over the consequence variables). The sensitivity analysis, however, often uses “one-way” sensitivity analyses in which the impact on the consequence variable is calculated as the inputs are varied one at a time, the others remaining fixed. Performed this way, the sensitivity analysis essentially ignores possible relationships, which we call dependence, among the input variables.

Reilly (1996) shows how to incorporate assessments of pairwise correlation (Spearman’s r) into the sensitivity analysis. The analysis is based on a singular-value decomposition, and Reilly shows that incorporating dependence into the sensitivity analysis can have unanticipated impacts on the model, the subsequent analysis, and ultimately on the alternative chosen. In this paper we show how Reilly’s dependent sensitivity analysis can serve as a platform for incorporating assessed dependence measures in different ways and at different stages in subsequent model construction and analysis. At the least, the incorporation of dependence measures can suggest a probabilistic structure and provide guidance for conditional probability assessments. At the other extreme, the analyst can construct a complete copula-based joint distribution using correlations and fractiles assessed in the dependent sensitivity analysis, thereby eliminating the need for further assessments.

Our proposed approach is neither a panacea nor a replacement for conventional DA. Although we demonstrate how correlations can enhance the DA process, the approach has costs and limitations, and we discuss those below. For now, we note that our approach is most useful when the uncertain variables are continuous rather than discrete. In addition, the use of correlation as a dependence measure limits the nature of the probabilistic relationships that can be modeled, although we will argue that such limitations may be reasonable in an initial, rough-cut model.

The discussion in this paper relies fundamentally on the presumption that experts can reliably assess measures of dependence. We will not discuss assessment techniques in this paper, but refer the reader to Clemen and Reilly (1997), who describe a variety of practical correlation-assessment techniques. As they note, Spearman's r can be assessed using familiar DA assessment techniques. Nevertheless, the assessments may be difficult for an expert to make.

In the next section, we present briefly the conventional DA process as described and implemented by Howard and his colleagues. This is followed by a description of our dependence-based decision-analysis process, for which we will use the acronym DDA. Section 3 provides a step-by-step demonstration of DDA, using the Eagle Airlines example from Clemen (1996) and demonstrating in detail how DDA can enhance the DA modeling process. Section 4 turns to a discussion of concerns associated with DDA, some of which have been mentioned above. Section 5 summarizes and concludes.

2. Conventional and Dependent Decision Analysis

2.1. Conventional DA. Figure 1, taken from Matheson and Howard (1968), shows the steps in the conventional DA process. The process begins by creating a deterministic model based on relationships among input variables. The model is constructed so that a given set of inputs determines a set of relevant outputs (e.g., profit, net present value). A convenient analogy for a deterministic model in present-day terms is a spreadsheet model; inputs would typically be constants or parameters in the spreadsheet, and by means of formulas referencing input cells, the spreadsheet model calculates outputs of interest to the decision maker.

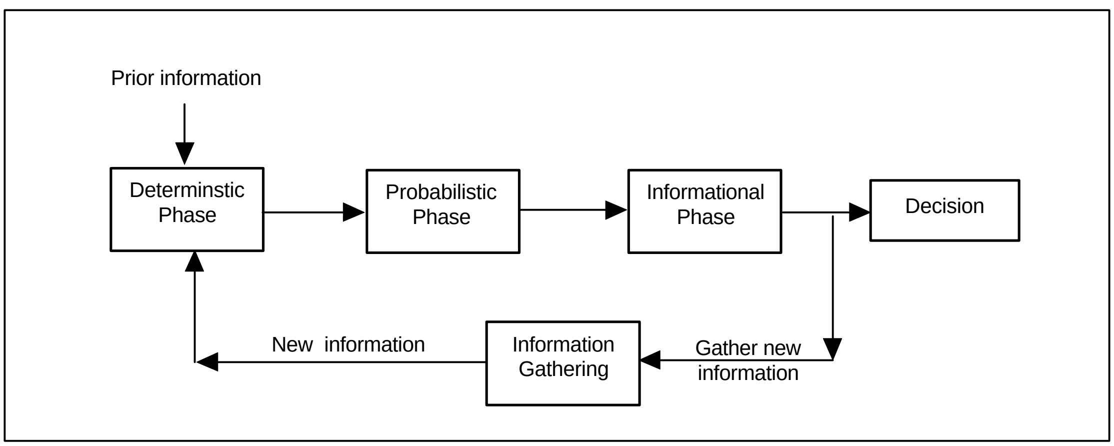

Figure 1. The Conventional DA Process (after Matheson and Howard, 1968).

The deterministic phase culminates in a deterministic sensitivity analysis. Although a sensitivity analysis might be conducted in different ways, a frequent and often recommended approach is one-way sensitivity analysis: For each continuous input variable, the expert assesses a "base value," which can be thought of as an estimate, and in practice is often assessed as the median of the expert's subjective probability distribution. In addition, high and low values for the variable are assessed; these might be the 10th and 90th percentiles of the expert's subjective distribution, for example. With a table of low, base, and high values for each input, a "Rainbow diagram" or a "Tornado diagram" (see, e.g., Clemen, 1996) graphically illustrates the degree to which the output variable changes as a single input is varied from high to low, with all other inputs fixed at their base values.

Based on the deterministic sensitivity analysis, the analyst chooses which variables to include in a probabilistic model and which to leave as constants, presumably at their base values. Although the choice of critical variables to include in the probabilistic model is more a matter of art than science, one would typically include those variables that generate the greatest variance in the output or those variables that, when varied from high to low, lead to different optimal alternatives.

With a set of critical variables in hand, the analyst is in a position to construct the joint probability distribution for the critical variables. The conventional approach relies on decomposing the joint distribution into a product of marginal and conditional distributions. If an influence diagram is used as a structuring tool, that diagram specifies the conditional relationships among the variables and suggests an appropriate ordering of the assessments. The high, base, and low values used in the sensitivity analysis provide starting points for marginal distribution assessments.

Following Figure 1, the next stage in the process is known as the information stage, in which value-of-information calculations are performed to determine which, if any, of the uncertain variables bear further study in an effort to reduce the current level of uncertainty. The results of this analysis can lead to further information-gathering activity, which in turn can lead the analyst to iterate through the process. Iteration continues until further modeling and analysis is infeasible or unwarranted, until further insights seem unlikely to emerge, or until the decision maker becomes satisfied that the optimal alternative, along with suitable intuition and explanation for its optimality, has been identified.

The preceding description of the conventional DA process is necessarily brief. Although Howard and his colleagues have written about and implemented many variations of the process, the stages discussed here and laid out in Figure 1 still comprise the core of conventional decision analysis as it is taught and practiced by many analysts. That analysts continue to find this paradigm useful is a tribute to its robustness. The collection by Howard and Matheson (1983) contains many articles describing the paradigm and its application in real-world settings.

2.2. Correlations and DDA. The premise of DDA is that an analyst can use additional information, in the form of correlations among variables, to enrich the overall decision-analysis process. Such enrichment begins early with a sensitivity analysis that includes an initial specification of the stochastic relationships among the variables and continues with the construction and analysis of a probabilistic model reflecting those relationships. In this section we describe in detail the ways in which this enrichment adds flexibility and insight to the conventional DA process. Our discussion below parallels the steps shown in Figure 1. To assist the reader, Figure 2 outlines the major elements of DDA and will help frame the discussion to follow.

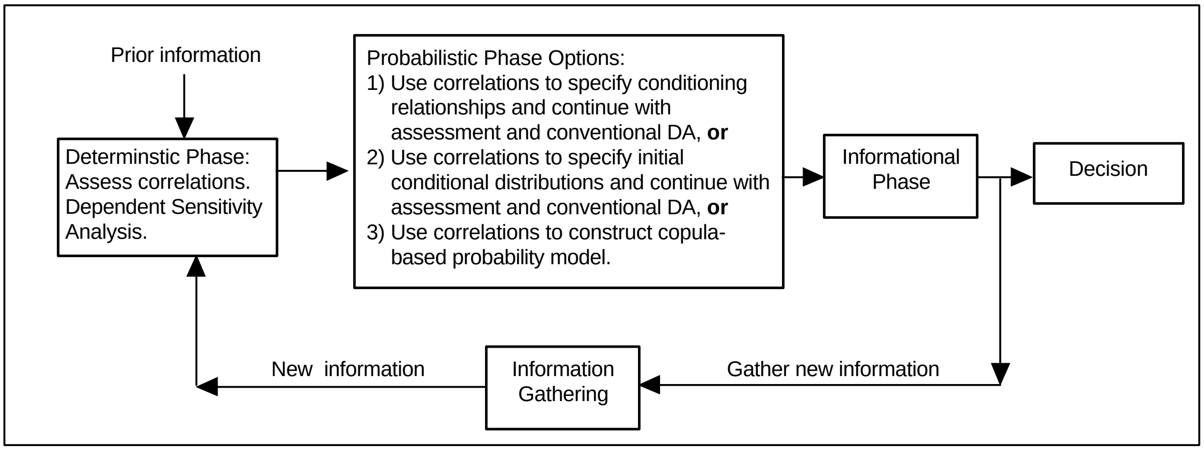

Figure 2. Steps in Dependent Decision Analysis. Beginning in the deterministic phase correlations are assessed and incorporated into the sensitivity analysis, the results of which set up several probability-modeling options in the probabilistic phase. The analyst may choose from among these options depending on analysis requirements and the expert's knowledge.

DDA begins in the same way as conventional DA; the analyst specifies a deterministic model that captures the algebraic relationships among the variables. At the stage of the deterministic sensitivity analysis, however, DDA includes assessments of correlation among the input variables. To understand the motivation for doing so, it suffices to think about a simple demand problem in which two inputs are quantity sold and price. One-way sensitivity analysis varies only one input at a time while keeping all others fixed, thus implicitly behaving as if the inputs can be treated independently. This may be entirely inappropriate for the demand problem, however, because basic economic thinking suggests that as price increases, demand will tend to decrease, and vice versa. Including this negative relationship between the two variables can give the analyst a clearer picture of the joint influence of these variables on total revenue.

One can perform two-way sensitivity analysis (e.g., Clemen, 1996), and the current version of TreeAge Software's DATA performs three-way sensitivity analysis by animating a series of two-way analyses. In principle, *n*-way sensitivity analysis is possible, although visualization of the results becomes problematic (to say the least). In such multi-way sensitivity analysis, contours of the output variable can be plotted with respect to the inputs. The analyst can then use these contours as the basis for a discussion with the expert to determine what combinations of the inputs -- what regions of the graphs -- are most likely to occur. In short, we are informally incorporating subjective information about the input variables' joint distribution.

Reilly (1996) describes in detail how to incorporate pairwise correlations into the sensitivity analysis as shown in Figure 2. The procedure is called dependent sensitivity analysis (DSA) and can be described briefly as follows: As before, the deterministic model along with low, base, and high values of the inputs are specified. At this point, the expert assesses Spearman's correlation ρ_{ij} for each (X_i, X_j) pair of input variables. Using the resulting correlation matrix $\mathbf{R}_S$, the analyst can construct a set of synthetic variables that are approximately uncorrelated. As long as $\mathbf{R}_S$ is positive-definite, there exists an $n \times n$ matrix $\mathbf{L}$ that is orthonormal (i.e., $\mathbf{L}^{-1} = \mathbf{L}^T$) that orthogonalizes $\mathbf{R}_S$:

$$\mathbf{R}_S = \mathbf{L}\,\mathbf{\Lambda}\,\mathbf{L}^T,$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{R}_S$. Now denote the column vector of n input variables by $\mathbf{X} = (X_1, ..., X_n)^T$, and the vector of base values by $\mathbf{B}$. We also require estimates of the standard deviations s_i of each X_i, which can be obtained by substituting the low, base, and high values into an appropriate moment-estimation formula (e.g., Moder and Rodgers, 1968; Perry and Grieg, 1975). Let $\mathbf{S}$ be an $n \times n$ diagonal matrix with diagonal elements s_i.

With these specifications, define a column vector of synthetic variables $\mathbf{Y} = (Y_1, ..., Y_n)^T = \mathbf{LS}^{-1}(\mathbf{X}-\mathbf{B})$. Reilly shows that the Y_i's are approximately uncorrelated, and their variances are equal to the eigenvalues of $\mathbf{R}_S$. Rewriting $\mathbf{X}$ in terms of $\mathbf{Y}$, we have $\mathbf{X} = \mathbf{SL}^T\mathbf{Y} + \mathbf{B}$. Substituting this expression for the X_i's into the deterministic model, we now have the ability to run a conventional sensitivity analysis on the Y_i's. As each Y_i is varied individually, the X_i's simultaneously covary according to correlation matrix $\mathbf{R}_S$. A tornado diagram identifies the influential Y_i's, which in turn determine (via the equations $\mathbf{Y} = \mathbf{LS}^{-1}(\mathbf{X}-\mathbf{B})$) the critical X_i's that warrant further probabilistic modeling. Reilly's analysis and examples show that performing the sensitivity analysis on the Y_i's can result in insights not revealed by the conventional approach and ultimately in different recommended alternatives.

Like sensitivity analysis in conventional DA, the output of DSA is a set of variables that are deemed to require probabilistic modeling. At this point, the flexibility of DDA becomes apparent in the multiple ways in which DDA can proceed as listed in Figure 2. Of course, it is possible to proceed as in conventional DA by taking the critical input variables and assessing appropriate marginal and conditional distributions. At the very least, however, the assessed correlation structure can provide an initial influence diagram and guidance in identifying conditional independence in the model so that the diagram is consistent with $\mathbf{R}_S$. We will demonstrate this in our example in Section 3.

A second way in which DDA can proceed is to use the assessed fractiles of the input variables, along with $\mathbf{R}_S$, to generate an initial set of marginal and conditional distributions. These distributions can then be reviewed by the expert and modified as necessary to match the expert's beliefs. From this point the decision analysis can proceed along conventional lines. Although specifying the required marginal distributions is straightforward based on the assessed fractiles, constructing a set of conditional distributions requires specific modeling assumptions regarding the joint distribution. In the Section 3 example, we will demonstrate how straightforward derivations of conditional distributions follow from a particular copula family, the one that arises from the multivariate normal distribution.

Related to both of the two approaches described so far is the possibility of using the assessed correlations as consistency checks. That is, correlations can be calculated from assessed marginal and conditional distributions, and those correlations should be consistent with correlations assessed during the sensitivity analysis. Inconsistency may require reassessment of either the correlations or the probability distributions, which may in turn require the expert to rethink the relationships among the variables.

A third way to proceed with DDA is to use the assessed correlations, along with the assessed fractiles, to construct a copula representation of the joint distribution of the random variables. See Dall'Aglio, Kotz, and Salinetti (1991) for background information on copulas. Jouini and Clemen (1996), MacKenzie (1994), and Yi and Bier (1995) contain examples of copulas in decision- and risk-analysis applications. The essential mathematical results that we require here are the following:

> Sklar's Theorem (1959): Given a joint distribution function $F(x_1, \ldots, x_n)$ for random variables $X_1, \ldots, X_n$ with marginals $F_1(x_1), \ldots, F_n(x_n)$, F can be written as a function of its marginals:
>
> $$F(x_1, \ldots, x_n) = C[F_1(x_1), \ldots, F_n(x_n)],$$
>
> where $C(u_1, \ldots, u_n)$ is a joint distribution with uniform marginals. Moreover, if each F_i is continuous, C is unique.

The function C is called a *copula*. Sklar's theorem is important because it tells us that any joint distribution can be written in copula form. The second result we need is that, under the assumption that each F_i and C are differentiable, the joint density $f(x_1, \ldots, x_n)$ can be written as

$$f(x_1, \ldots, x_n) = f_1(x_1)\times \cdots \times f_n(x_n)\; c[F_1(x_1), \ldots, F_n(x_n)], \tag{1}$$

where $f_i(x_i)$ is the density corresponding to $F_i(x_i)$ and $c = \partial^n C/(\partial F_1 \ldots \partial F_n)$ is called the *copula density*. Equation (1) states that the joint density can be written as a product of the marginal densities and the copula density. The conventional DA approach capitalizes on decomposing the joint density as a product of marginals and conditionals. In contrast, DDA generates insights into the joint density via the copula representation. For example, if the X_i's are independent, then $c = 1$ and $f(x_1, \ldots, x_n) = f_1(x_1)\times \cdots \times f_n(x_n)$, the familiar formula for n independent uniform random variables. From the representation in (1) it is clear that the copula density c encodes information about the dependence among the X_i's, and for this reason c is sometimes called a *dependence function*.

The fact that the copula is a function of the marginals is especially useful in DDA because it allows the "coupling" of the marginals into a joint distribution using the assessed fractiles and correlations from the sensitivity-analysis stage. Doing this requires two steps. First is modeling the marginal distributions in some way, which may require additional assessments or fitting a member of a distribution family (normal, exponential, beta, etc.) to those assessments. Standard techniques from decision and risk analysis are available to accomplish this (Morgan and Henrion, 1990; Clemen, 1996). The second step is to create a copula that captures the dependence among the random variables. Given $\mathbf{R}_S$, several copula families are available to accomplish this, one of which is the copula c_N that arises from the multivariate normal distribution. Like other copula families, the multivariate normal copula allows any marginal distribution for the X_i's (beta, gamma, lognormal, etc.). It is called the normal copula because it encodes dependence in precisely the same way that the multivariate normal distribution does using only pairwise correlations among the variables, but it does so for variables with arbitrary marginals. The flexibility and analytical tractability of the multivariate normal copula leads us to use it in our demonstration of DDA below.

We begin construction of our copula model by recalling that $\mathbf{R}_S$ is a matrix of Spearman rank-correlation coefficients but that the multivariate normal distribution typically is parameterized in terms of Pearson product-moment correlations. Thus, for each element ρ_{ij} of $\mathbf{R}_S$, calculate the corresponding product-moment correlation r_{ij} for the multivariate normal as $r_{ij} = 2\sin(\pi\rho_{ij}/6)$ (Kruskal, 1958), and construct matrix $\mathbf{R}$ with elements r_{ij}. Now rearrange equation (1), applying it to the n-dimensional multivariate normal density $\phi^{(n)}(y_1, \ldots, y_n \mid \mathbf{R})$:

$$c_N[\Phi(y_1), \ldots, \Phi(y_n) \mid \mathbf{R}_S] = \phi^{(n)}(y_1, \ldots, y_n \mid \mathbf{R}) / [\phi(y_1) \times \cdots \times \phi(y_n)], \qquad (2)$$

where Φ and ϕ denote the univariate standard normal distribution and density, respectively. Substituting the expressions for the normal densities and performing a bit of algebra leads to:

$$c_N[\Phi(y_1), \ldots, \Phi(y_n) | \mathbf{R}] = \exp\{-\mathbf{y}^{\mathrm{T}}(\mathbf{R}^{-1} - \mathbf{I})\mathbf{y}/2\} / |\mathbf{R}|^{1/2}, \qquad (3)$$

where $\mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$, and $\mathbf{I}$ is the $n \times n$ identity matrix.

Using c_N as a dependence function for random variables with arbitrary marginals $F_1(x_1), \ldots, F_n(x_n)$ requires using the normal inverse transformation (e.g., Kelly and Krzysztofowicz, 1996a, b). We define $Y_i = \Phi^{-1}[F_i(X_i)]$ and substitute these and (3) into (1) to obtain the joint density:

$$\begin{aligned} f(x_1, \ldots, x_n | \mathbf{R}) &= f_1(x_1) \times \cdots \times f_n(x_n) \times \exp\{-\mathbf{y}^{\mathrm{T}}(\mathbf{R}^{-1} - \mathbf{I})\mathbf{y}/2\} / |\mathbf{R}|^{1/2} \\ &= f_1(x_1) \times \cdots \times f_n(x_n) \times \qquad (4) \\ &\exp\{-(\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_n(x_n)])(\mathbf{R}^{-1} - \mathbf{I})(\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_n(x_n)])^{\mathrm{T}}/2\} / |\mathbf{R}|^{1/2} \end{aligned}$$

This joint density has the specified marginals and, because Spearman's rank-order correlation is invariant under monotone 1-1 transformations of the original variables, the X_i's have the specified rank-order correlations $\mathbf{R}_S$. Calculating the density for specific values $x_1,\ldots, x_n$ is relatively easy, requiring n inversions of the standard normal distribution.

Conditional densities are also easily calculated using the multivariate normal copula model. Let $\mathbf{R}$ and $\mathbf{y}$ be partitioned as follows:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{n-1} & \mathbf{r} \\ \mathbf{r}^{\mathrm{T}} & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = (\mathbf{y}_{n-1}, y_n),$$

where $\mathbf{y}_{n-1} = (y_1, \ldots, y_{n-1})^{\mathrm{T}}$, $\mathbf{R}_{n-1}$ is the $(n-1) \times (n-1)$ correlation matrix for $(Y_1, \ldots, Y_{n-1})$, and $\mathbf{r} = (r_{1n}, \ldots, r_{(n-1)n})^{\mathrm{T}}$. From (3), the definition of conditional probability, and the formula for the standard multivariate normal density, we have

$$f(x_n | x_1, \ldots x_{n-1}, \mathbf{R}_S)$$

$$= f_n(x_n) \times \frac{\phi^{(n)}(\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_n(x_n)] | \mathbf{R})}{\phi(\Phi^{-1}[F_n(x_n)]) \times \phi^{(n-1)}(\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_{n-1}(x_{n-1})] | \mathbf{R}_{n-1})},$$

which, upon substituting in the expressions for the normal densities and reducing, becomes

$$f(x_n | x_1, \ldots, x_n, \mathbf{R})$$

$$= f_n(x_n) \exp\left\{ -0.5 \left[\frac{(\Phi^{-1}[F_n(x_n)] - \mathbf{r}^{\mathrm{T}} \mathbf{R}_{n-1}^{-1} \mathbf{y}_{n-1})^2}{(1 - \mathbf{r}^{\mathrm{T}} \mathbf{R}_{n-1}^{-1} \mathbf{r})} + (\Phi^{-1}[F_n(x_n)])^2 \right] \right\} (1 - \mathbf{r}^{\mathrm{T}} \mathbf{R}_{n-1}^{-1} \mathbf{r})^{-1/2}. \quad (5)$$

Although the development of the formulas above appears complex, the resulting expressions in (4) and (5) are relatively straightforward and easy to use in computations. For example, the conditional density in (5) is the product of a marginal density, an exponential term, and a scaling constant.

With the joint distribution specified, expected values, expected utilities, and risk profiles can be calculated directly from the copula model. As in conventional DA, DDA can proceed with value-of-information analysis; doing so will typically require the calculation of conditional distributions from the copula. As before, analysis and modeling may iterate until clarity of action is obtained.

The copula approach can be very efficient. For example, the analyst can use the assessed fractiles and correlations from DSA to create an initial model for a quick, rough-cut probabilistic analysis, followed by further refinement of the assessments and model as deemed necessary. If the expert has done a satisfactory job in the initial assessment of the fractiles and correlations, no further assessment may be needed at all. If more careful assessment is required, however, it may be helpful to check the consistency of the copula model's conditional distributions against assessed conditionals or conditional expected values (via regression equations, for example). More fundamentally, the expert may at this point invoke causal as well as correlational reasoning in thinking about the relationships among the variables; causal reasoning naturally leads to conditional distributions rather than a copula representation. Finally, it may be the case that some of the variables can be modeled with a copula and some with conditional distributions.

3. An Example: Eagle Airlines

Clemen (1996) describes the hypothetical decision faced by Dick Carothers, owner of the fledgling Eagle Airlines. Carothers is considering purchasing a used aircraft. His decision criterion is whether the airplane will generate more profit than a money-market alternative investment. Reilly (1996) modifies the model slightly in his sensitivity-analysis example, and we will use Reilly's model here. The influence diagram in Figure 3 portrays the deterministic model, and Table 1 lists Carother's low, base, and high values for the input variables. For illustrative purposes, we take these to be the assessed 10th, 50th, and 90th percentiles of the respective probability distributions.

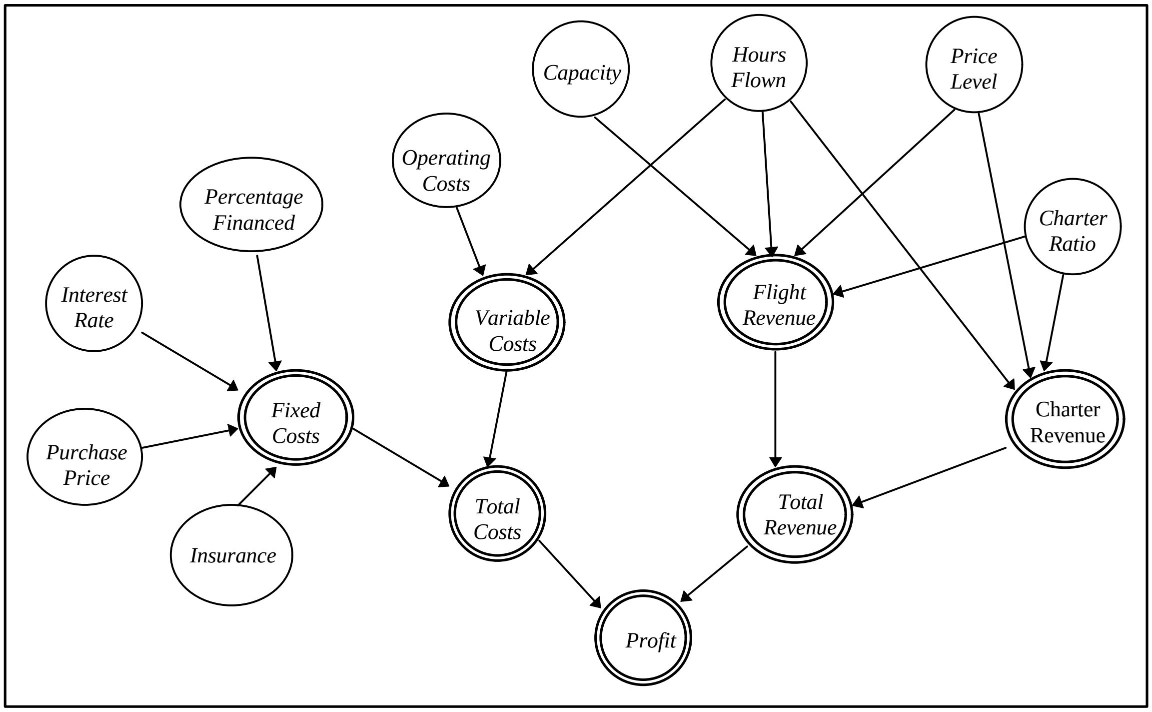

Figure 3. The initial influence diagram for Eagle Airlines.

Variable (*X*)	Fractile:	**Low** 0.10	**Base** 0.50	**High** 0.90
Charter Ratio		45%	50%	70%
Capacity (*C*)		40%	50%	60%
Price Level (*P*)		$95	$100	$108
Hours Flown (*H*)		500	800	1000
Operating Cost per Hour (*O*)		$230	$245	$260
Percentage Financed		30%	40%	50%
Interest Rate		10.50%	11.50%	13.00%
Insurance		$18,000	$20,000	$25,000
Purchase Price		$85,000	$87,500	$90,000

Table 1. Input variables and assessed fractiles for use in Eagle Airlines sensitivity analysis.

3.1. Dependent Sensitivity Analysis. Reilly assesses correlations for the all pairs of variables listed in Table 1 and uses the resulting correlation matrix $\mathbf{R}_S$ to demonstrate DSA. We summarize the results and interpretation here and refer the reader to his paper for the correlations and detailed analysis. Let *P*, *H*, *C*, and *O* represent *Price Level*, *Hours Flown*, *Capacity*, and *Operating Cost*, respectively. Likewise, we will let *P**, *H**, *C**, and *O** represent standardized versions of these variables. Because there are nine original variables (X_i), there are likewise nine synthetic Y_i's, and Reilly finds two of these to be critical. One (Y_2) is a price-and-demand factor and is approximately

$$Y_2 = -0.60\, H^* + 0.52\, P^* - 0.53\, C^*.$$

(Following Reilly, we have left out the other six variables because their coefficients or *loadings* are relatively small.) Y_2 is shown to cross the critical threshold in the model (when $Y_2 = 1.11$), and moreover this happens when *P* is relatively high ($102.78) and *H* and *C* are relatively low (676 hours and 46%, respectively). This result is in contrast to conventional one-way sensitivity analysis in which only a relatively low value of *P*, considered in isolation from *H* and *C*, crosses the threshold. The interpretation here is that the high price, coupled with low demand as one might expect, leads to a low level of profit.

The second critical factor (Y_7) is approximately

$$Y_7 = 0.66\, P^* + 0.55\, C^* - 0.39\, O^*.$$

This factor is a revenue-versus-cost factor, and the threshold is crossed for relatively low values of *P* and *C* ($99.12 and 48.96%, respectively) and a high value for *O* ($246.20).

To summarize, DSA identifies only Y_2 and Y_7 to be critical, and from the expressions for these two synthetic variables, four of the original variables are found to contribute the most to the sensitivity of profit. These are the four variables that are to be modeled probabilistically. Their assessed fractiles and correlations (Table 2) serve as inputs to the probabilistic phase in DDA.

		Low	Base	High	Correlations		
					Price	*Hours*	
Variable (*X*)	Fractile:	0.10	0.50	0.90	*Level*	*Flown*	*Capacity*
Price Level		$95	$100	$108			
Hours Flown		500	800	1000	-0.50		
Capacity		40%	50%	60%	-0.25	0.50	
Operating Cost per Hour		$230	$245	$260	0	0	0.25

Table 2. Fractiles and Spearman correlations for four critical variables in Eagle Airlines

3.2. Specifying Qualitative Relationships. The first option in the probabilistic phase of DDA is to use the critical variables and assessed correlations as a basis for designating conditional relationships among the critical variables, e.g., arcs among these variables in the influence diagram. Approximate conditional expected values – regressions – can be easily calculated directly from $\mathbf{R}_S$. If a regression coefficient is zero (or approximately so), no arc is included in the influence diagram. It is an approximation because the Spearman correlations are used to approximate the product-moment correlations required in the regression calculus. Although this is far from a perfect approach and may overlook important nonlinear or more complex relationships among the variables, this step is intended only as an initial trial structure to be refined in further modeling and assessment.

As in conventional DA, the choice of a conditioning order is at the discretion of the analyst, who may specify an order based on the structure of the problem or, if appropriate, the expert's causal reasoning. (See Burns and Clemen (1993) for a related discussion.) For Eagle Airlines, we take P as the first variable in the assessment sequence, implying that we will eventually assess its marginal distribution. We then take H given P; the assessed pairwise correlation value of -0.50 means that $E(H^* \mid P^*) = -0.5\, P^*$. These two variables tend to move in opposite directions, and so an arc is required from *Price Level* to *Hours Flown*.

The third conditional relationship is C given H and P. In this case, we use the regression calculus to determine what arcs must be included. To do this, define vector $\rho_{C|H,P}$ and matrix $\mathbf{R}_{S:H,P}$:

$$\rho_{C|H,P} = (\rho_{C,H}, \rho_{C,P}) = (0.50, -0.25); \quad \mathbf{R}_{S:H,P} = \begin{bmatrix} 1 & \rho_{H,P} \\ \rho_{H,P} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -0.50 \\ -0.50 & 1 \end{bmatrix},$$

where the subscripts C, H, and P denote *Capacity*, *Hours Flown*, and *Price Level*, respectively. Using the standardized variables, we calculate

$$E(C^* \mid H^*, P^*) = \rho_{C|H,P}\ \mathbf{R}_{S:H,P}^{-1}(H^*, P^*)^{\mathrm{T}} = 1/2\ H^* + 0\ P^*.$$

This result suggests that *Capacity* may be conditionally independent of *Price Level* given *Hours Flown*. Likewise, the regression for O given C, H, and P can be calculated:

$$E(O^* \mid C^*, H^*, P^*) = 1/3\ C^* - 1/6\ H^* + 0\ P^*,$$

which suggests that *Operating Cost* may be conditionally independent of *Price Level*, given *Capacity* and *Hours Flown*. Taken together, these results lead to an influence diagram (Figure 4) that shows the conditional relationships among *Price Level*, *Hours Flown*, *Capacity*, and *Operating Cost*. This influence diagram would serve as a point of departure for further discussion and assessment with the expert.

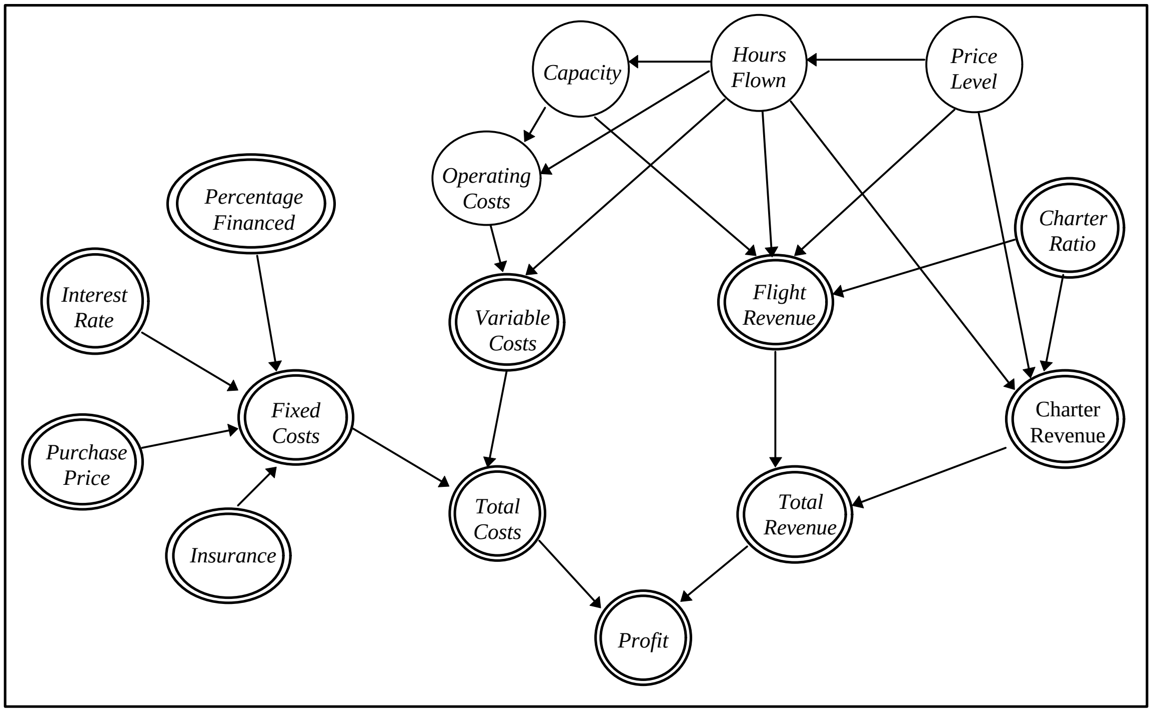

Figure 4. Influence diagram with conditional relationships specified for critical chance variables. The arcs connecting *Price Level*, *Hours Flown*, *Capacity*, and *Operating Cost* provide initial guidance to the analyst and expert regarding the probabilistic relationships and assessment order for these variables.

3.3. Initial Marginal and Conditional Distributions. The analyst may continue with DDA by using the assessed fractiles for the critical variables, along with $\mathbf{R}_S$, to create initial marginal and conditional distributions. To demonstrate, we fit marginal distributions to the four critical variables as indicated in Table 3 so that the modeled distributions have approximately the same 0.10, 0.50, and 0.90 fractiles as shown in Table 2.

Variable	**Distribution**	**Parameters**	**Range**
Price Level (*P*)	Beta	a = 9, b = 15	[$81.94, $133.96]
Hours Flown (*H*)	Beta	a = 4, b = 2	[66.91, 1135.26]
Capacity (*C*)	Beta	a = 20, b = 20	[0, 1]
Operating Cost (*O*)	Normal	m= 245, s = 11.72	(-∞, +∞)

Table 3. Marginal distributions for Eagle Airlines probability model

Along with the marginal distributions, we use $\mathbf{R}_S$ to specify a normal copula-based joint density and from that joint density derive the required conditional densities. Denote the marginal density and cumulative distribution for P as $f_b(p)$ and $F_b(p)$, respectively, and similarly for H and C. Likewise, let $f_N(o)$ and $F_N(o)$ denote the marginal density and cumulative distribution for the normally distributed O. It is straightforward to derive the conditional densities $f(h \mid p)$, $f(c \mid h)$, and $f(o \mid c, h)$:

$$f(h|p) = f_{\mathrm{b}}(h)\exp\left\{-0.5\left[\left(\Phi^{-1}[F_{\mathrm{b}}(h)] + 0.518\Phi^{-1}[F_{\mathrm{b}}(p)]\right)^2 / 0.732 - \left(\Phi^{-1}[F_{\mathrm{b}}(h)]\right)^2\right]\right\} / 0.732^{1/2}$$

$$f(c|h) = f_{\mathrm{b}}(c)\exp\left\{-0.5\left[\left(\Phi^{-1}[F_{\mathrm{b}}(c)] - 0.518\Phi^{-1}[F_{\mathrm{b}}(h)]\right)^2 / 0.732 - \left(\Phi^{-1}[F_{\mathrm{b}}(c)]\right)^2\right]\right\} / 0.732^{1/2}$$

$$f(o|c,h) = f_N(o)\exp\left\{-0.5\left[\left(\Phi^{-1}[F_N(o)] - 0.357\Phi^{-1}[F_{\mathrm{b}}(c)] + 0.185\Phi^{-1}[F_{\mathrm{b}}(h)]\right)^2 / 0.907 + \left(\Phi^{-1}[F_N(o)]\right)^2\right]\right\} / 0.907^{1/2}$$

These expressions are no more than realizations of Equation (5) in which we have substituted in the correlation values from $\mathbf{R}$ and the marginal densities from Table 3.

Figure 5 shows a portion of the Eagle Airlines event tree based on a discrete approximation of the continuous conditional densities above. The branch values displayed are the 0.05, 0.50, and 0.95 fractiles of each marginal or conditional density, and the probabilities on the top, middle, and bottom in each set of three branches are 0.185, 0.63, and 0.185, respectively, *per* the extended Pearson-Tukey formula (Keefer and Bodily, 1983). The complete event tree is available on request from the authors.

The bivariate relationships among the variables reveal themselves readily in Figure 5. For example, as *Price* increases, *Hours Flown* tends to decrease, reflecting the negative correlation between these two variables. As *Hours Flown* increases, its positive correlation with *Capacity* leads to an increase in the latter. The same is true when we examine changes in the distribution of *Operating Cost* given *Capacity*. In the context of DDA, Figure 5 represents initial conditional distributions for the probability model in the Eagle Airlines example. These distributions can be examined in detail by the expert and refined as needed. When the expert is satisfied that the model adequately captures his or her knowledge, analysis can proceed.

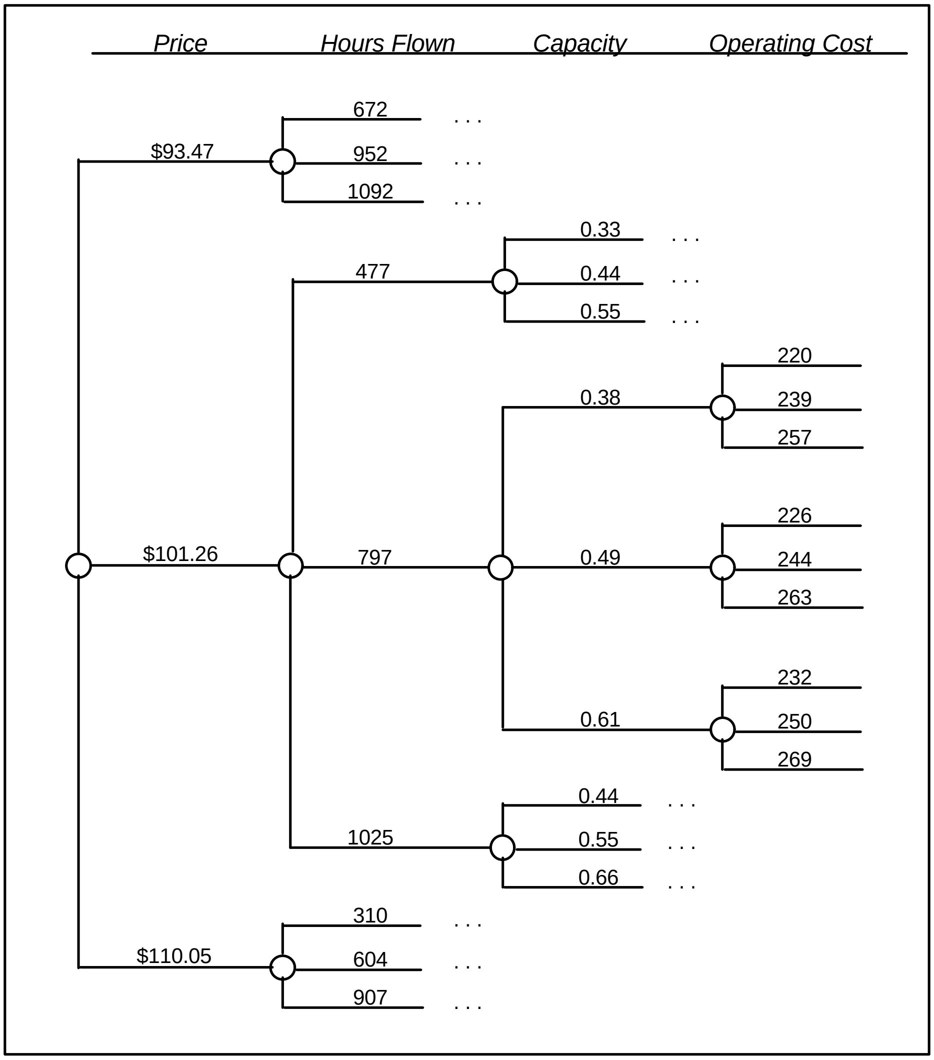

Figure 5. Event tree from multivariate normal copula.

3.4. A copula-based joint density. Using copulas to construct the conditional densities and discrete approximations in the previous section suggests using the copula-based joint density directly in the decision analysis. If the expert and analyst are convinced that the correlation assessments adequately capture the relationships among the variables, this approach is a very efficient way to generate the probabilistic model.

No convention currently exists for representing a copula-based joint distribution in an influence diagram. Any convention adopted must accommodate the presence of predecessors and successors that are external to the copula model. At the same time, directed arcs among the variables that are related by the copula are inappropriate. We suggest placing the copula-related variables in proximity to each other and within a shaded region as exemplified in Figure 6. This convention suggests that each variable in the joint distribution will require assessment of its marginal distribution, and that measures of dependence among the variables will be necessary to specify the copula.

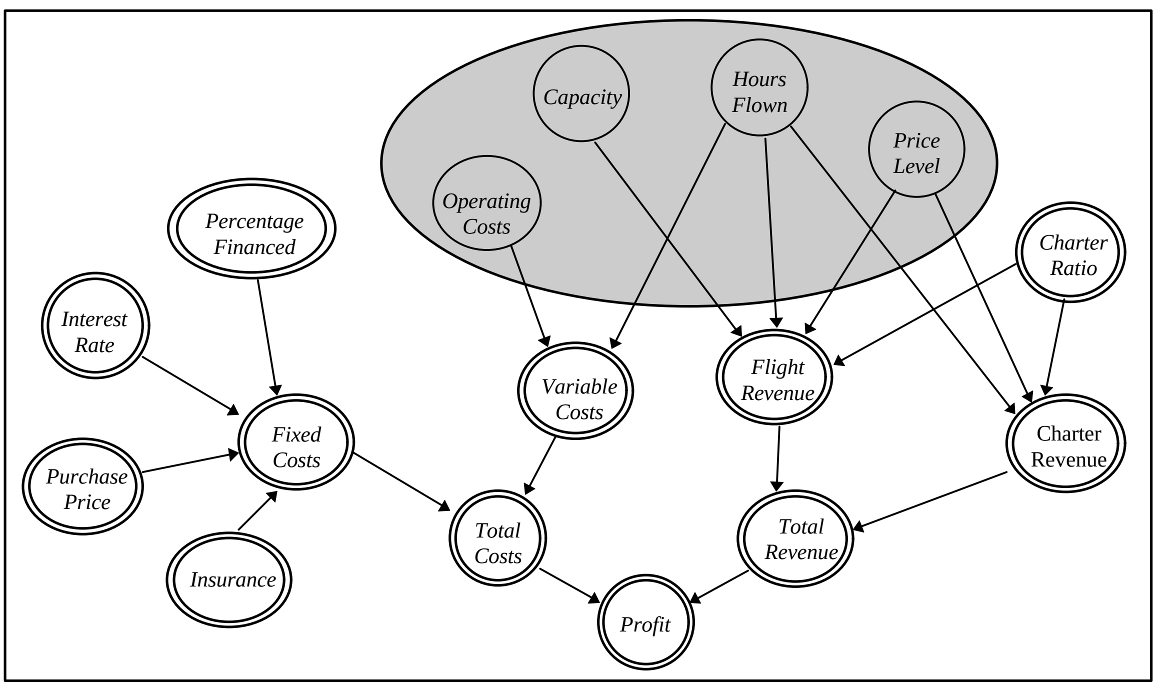

Figure 6. Representing a copula-based probability model in an influence diagram.

We fit the multivariate normal copula density by substituting the marginal densities (f_i) from Table 3, the corresponding distributions (F_i), and **R** into (4). Let $Y_P = \Phi^{-1}[F_\beta(P)]$, and similarly for *H*, *C*, and *O*, and let $\mathbf{y} = (y_P, y_H, y_C, y_O)$. Substituting into (3),

$$f(p,h,c,o) = f_\beta(p) f_\beta(h) f_\beta(c) f_N(o) \exp\{-\mathbf{y}'(\mathbf{R}^{-1} - \mathbf{I})\mathbf{y}/2\}/0.486,$$

where

$$\mathbf{R}^{-1} - \mathbf{I} = \begin{bmatrix} 0.366 & 0.715 & -0.014 & 0.004 \\ 0.715 & 0.777 & -0.787 & 0.205 \\ -0.014 & -0.787 & 0.506 & -0.393 \\ 0.004 & 0.205 & -0.393 & 0.103 \end{bmatrix}.$$

We calculated expected values, standard deviations, and risk profiles for the profit resulting from the aircraft purchase using three different approaches. The first uses the marginal-and-conditional discrete approximation from Figure 5. Second is a simulation using the multivariate normal copula model directly by generating multivariate-normal y_i's, evaluating the standard normal distribution function $\Phi(y_i)$ for each variate, and transforming to $x_i = F_i^{-1}[\Phi(y_i)]$. The third model is the same as the second except that the variables are taken to be independent. Both simulations were run for 10,000 trials.

The results are displayed in Table 4 and Figure 7. If we take the copula-based simulation as the benchmark, several observations can be made:

- Incorporating dependence can be important. The simulation that treated the variables as independent substantially overestimated the variability in profit.
- The copula-based discrete approximation underestimated the expected value in this case but very accurately estimated the standard deviation.
- The risk profile for the discrete approximation does a reasonably good job of replicating the copula simulation risk profile, especially in the tails.

	Expected Value	Standard Deviation
Conditional discrete approximation	$11,684	$20,221
Copula simulation (10,000 trials)	$12,417	$20,206
Independent simulation (10,000 trials)	$12,426	$23,628

Table 4. Expected profit and standard deviation for Eagle Airlines calculated from three models.

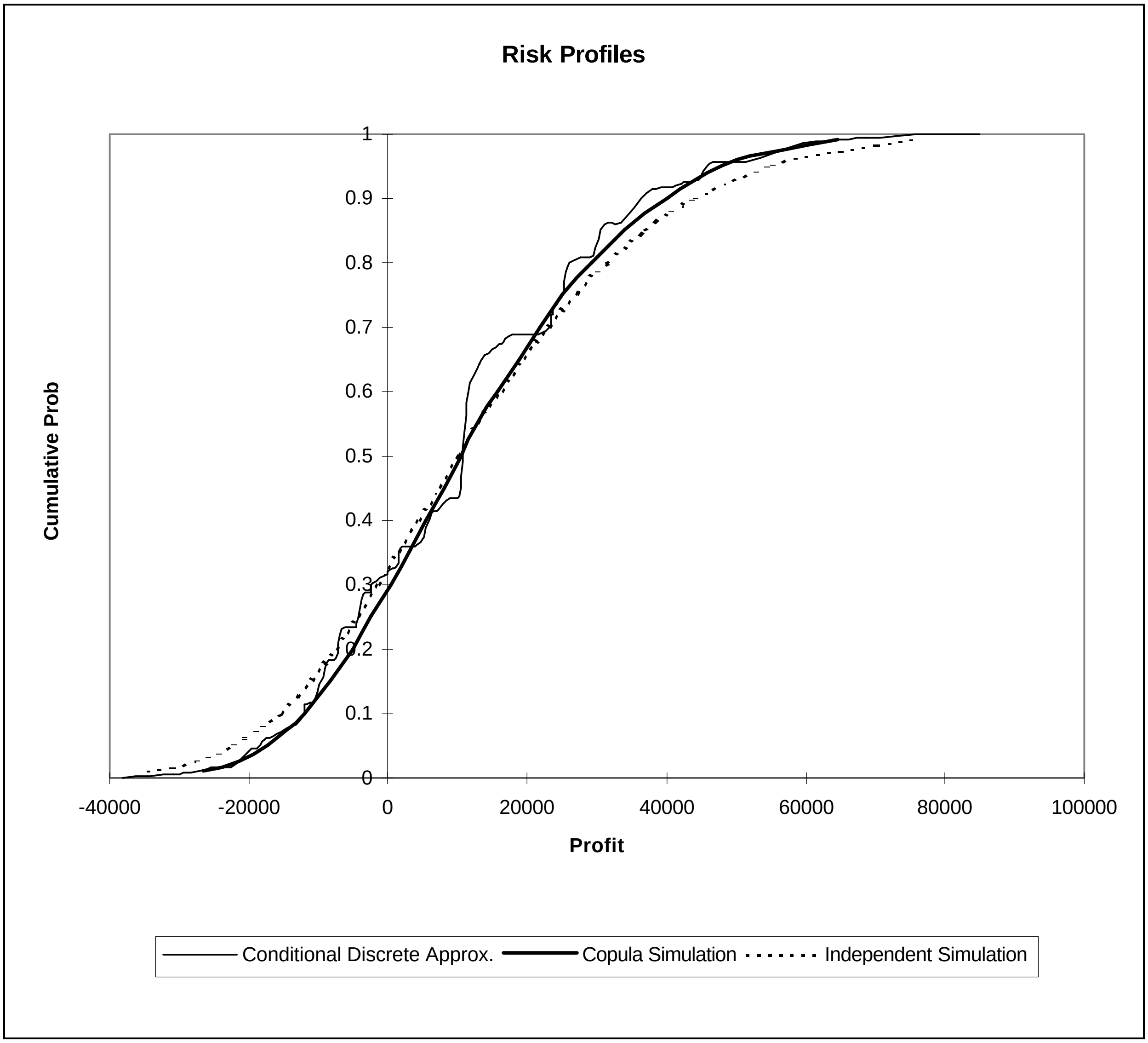

Figure 7. Risk profiles for Eagle Airlines from three models.

4. Using Copulas in DDA: Assessment, Modeling, and Analysis Issues

The description of DDA in Section 2 and the subsequent demonstration in the context of Eagle Airlines show the potential for DDA as an important tool for decision analysts. By obtaining the correlation assessments required for the initial DSA, the analyst has the opportunity to use this information in many different ways throughout the analysis, always so that DDA dovetails with conventional DA modeling and analysis. DDA can indeed be very efficient, requiring in some cases no more assessments than those made for the sensitivity analysis.

We discussed ways in which the analyst can use the correlations, for example, to create an initial influence diagram, to specify initial conditional distributions, as a way to check consistency of assessments, or to construct a full copula model. It is interesting to note that if the correlations are used only to create an initial influence-diagram structure and no more, then valuable information is being ignored. The notion of using the correlations to create initial conditional distributions might appear to be a reasonable compromise between not using the correlations and going all the way to the copula-based joint distribution. Although there may be situations in which this is appropriate, constructing the conditional distributions does require specifying the entire joint distribution in the first place.

As we noted in the introduction, DDA is not a panacea. In the remainder of this section we discuss a number of issues related to DDA.

4.1. Measures of dependence. DSA and DDA rely crucially on the ability of the expert to assess correlations. We grant that the assessment of correlations is not a trivial task, but practical techniques for correlation assessment do exist (Clemen and Reilly, 1997). As in probability assessment, training is important for valid assessments; the questions the expert must answer to assess correlations are complex. Clemen and Reilly do not argue that eliciting knowledge via marginal distributions and correlations is simpler than the conventional approach using marginals and conditionals, only that it is different. In some cases correlational reasoning may be more appropriate than causal or conditional reasoning, or an expert's background may make assessment of correlations easier than conditional distributions. Correlation and the idea of "percentage of variance explained" are fundamental statistical notions that many scientists learn and use.

On a technical issue, it is not just the assessment of the individual correlations that matters, but the assessment of $\mathbf{R}_S$ overall; the analyst must ensure that $\mathbf{R}_S$ is positive definite. If $\mathbf{R}_S$ is poorly conditioned (analogous to multicollinearity in regression), this may lead to some very counterintuitive results; for example, conditional distributions and risk profiles may appear deceptively narrow, reflecting the mathematics of the specific model, even though such a result may be inconsistent with the expert's intuition.

Finally, Reilly (1996) suggests using only rough assessments of the correlations (-0.75, -0.50, -0.25, 0, 0.25, 0.5, and 0.75) reflecting strong, moderate, or weak correlation. We agree with this approach as an initial step in assessing correlations; for purposes of DSA, it seems reasonable to ease the expert's assessment burden in this way. For the probabilistic phase, it may be worthwhile to refine the initial assessments using techniques from Clemen and Reilly (1997).

4.2. Modeling and analysis with copulas. The perceptive reader will have noticed that there are many ways to construct and use copula-based joint densities that were not mentioned in Sections 2 and 3. For example, we focused on the multivariate normal copula, but others may also be useful, such as the block-uniform family described by MacKenzie (1994). Also, we used the extended Pearson-Tukey discrete approximation method, but other methods may give better results. More to the point, however, modeling the entire joint distribution with a copula raises the question of how best to discretize a full joint distribution. How does one select appropriate representative points, and what probabilities should be applied? This question amounts to choosing representative scenarios in a complex multivariate space. Some efforts have been made in this direction (e.g., DeVuyst, Preckel, and Liu, 1996).

Finally, the calculation of conditional distributions is straightforward. One implication of this is that the search for conditional independence becomes less critical; it yields no real savings either in assessment or computational complexity. The ease of calculating conditional distributions will be especially useful for more complex models or analyses. For example, in a value-of-information analysis, it would be necessary to separate out one or more variables from the copula, conditioning decisions and the remaining copula variables on the information variable. For inference applications, the ease of calculating the conditional distributions should make it straightforward to propagate information through a network.

As mentioned in the introduction, DDA is most useful for situations involving continuous variables, but many important applications that use expert judgment, especially in artificial intelligence, involve the use of categorical variables, for which causal reasoning may be more appropriate. Also, the conventional approach has the advantage that any plausible relationship can be modeled. The use of pairwise correlations, on the other hand, limits the nature of the relationships that can be modeled to "monotone" dependence: The qualitative nature of a relationship between two variables does not depend on other variables (e.g., the sign of a regression coefficient does not change depending on the values of other conditioning variables). For a more complete discussion, see Kelly and Krzysztofowicz (1996a, b). In many cases, however, it appears that expert knowledge of variable relationships may be adequately modeled in terms of pairwise dependence by using correlations.

5. Conclusion. We have argued that decision analysts can benefit by explicitly incorporating measures of dependence, especially assessments of correlations, throughout the modeling-and-analysis process. We have shown how correlations can yield insights at the sensitivity-analysis stage and how the correlations can be used in the subsequent probability-modeling stage via the use of copulas. Throughout, our argument has been that this approach complements the set of tools available to the analyst, providing important insights and potentially streamlining the construction and analysis of the probabilistic model.

It is important to realize that good decision analysis does incorporate thinking about relationships among variables. Conventionally, this is done through careful thinking about conditional relationships, often aided by causal reasoning on the part of the expert. This approach has proven itself as a useful way to cope with complex knowledge elicitation and modeling situations. Its primary drawback is that the assessment burden grows exponentially with the number of variables in the model. As we have demonstrated, judicious use of correlations can potentially reduce the assessment burden, guide the knowledge-elicitation process, and provide important consistency checks for refining the assessments. For many experts and many situations, thinking in terms of correlations among variables may be more natural than the conventional conditional approach. Finally, although more work is required early to incorporate correlations into DSA and DDA, the additional insight to be gained throughout the decision-analysis process makes the effort well worthwhile.

Acknowledgments

We thank Bob Winkler and George MacKenzie for many discussions on both general and technical issues relating to the use of copulas in decision analysis. This work was supported by the National Science Foundation under Grant SBR 95-96176.

References

Burns, W. J., and R. T. Clemen, “Covariance Structure Models and Influence Diagrams,” *Management Science*, 39 (1993), 816-834.

Clemen, R. T., and T. Reilly, “Correlations and Copulas for Decision and Risk Analysis,” Duke University (1997), draft.

Clemen, R. T., *Making Hard Decisions: An Introduction to Decision Analysis, 2nd Ed*., Duxbury, Belmont, CA, 1996.

Dall’Aglio, G., S. Kotz, and G. Salinetti, *Advances in Probability Distributions with Given Marginals*, Kluwer, Dordrecht, Netherlands, 1991.

DeVuyst, E. A., P. V. Preckel, and S. Liu, “Discrete Approximations of Joint Probability Distributions,” University of Illinois (1996), draft..

Howard, R. A., and J. Matheson, *The Principles and Applications of Decision Analysis,* SDG, Inc., Menlo Park, CA, 1983.

Jouini, M., and R. T. Clemen, “Copula Models for Aggregating Expert Opinions,” *Operations Research*, 44 (1996), 444-457.

Keefer, D., and S. E. Bodily, “Three-Point Approximations for Continuous Random Variables,” *Management Science*, 29 (1983), 595-609.

Kelly, K. S., and R. Krzysztofowicz, “A Bayesian Model and Choice of Expert,” University of Virginia (1996a), draft.

Kelly, K. S., and R. Krzysztofowicz, “A Bayesian Model of Multiple Experts,” University of Virginia (1996b), draft.

Kruskal, W., “Ordinal Measures of Association,” *Journal of the American Statistical Association*, 53 (1958), 814-861.

MacKenzie, G. R., *Approximately Maximum-Entropy Multivariate Distributions with Specified Marginals and Pairwise Correlations*, PhD Dissertation, University of Oregon (1994).

Matheson, J. E., and R. A. Howard, “An Introduction to Decision Analysis,” Stanford Research Institute, Palo Alto, CA, 1968. Reprinted in Howard and Matheson (Eds.), *The Principles and Applications of Decision Analysis*, SDG, Inc., Menlo Park, CA (1983), 17-55.

Moder, J. J., and E. G. Rodgers, “Judgment Estimates of the Moments of PERT Type Distributions,” *Management Science*, 15 (1968), B76-B83.

Morgan, M. G., and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge, 1990.

Perry, C., and I. D. Grieg, “Estimating the Mean and Variance of Subjective Distributions in PERT and Decision Analysis,” *Management Science*, 21 (1975), 1477-1480.

Reilly. T., “Sensitivity Analysis for Dependent Variables,” University of Oregon (1996), draft.

Sklar, A., “Fonctions de Répartition à *n* Dimensions et Leurs Marges,” Publications de l’Institut Statistique de l’Université de Paris, 8 (1959), 229-231.

Yi, W., and V. M. Bier, “An Application of Copulas to Accident Precursor Analysis,” University of Wisconsin, Madison (1996), draft.