Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

European Journal of Operational Research 161 (2005) 240–255

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Stochastics and Statistics

Statistical dependence through common risk factors: With applications in uncertainty analysis

J. René van Dorp *

Department of Engineering Management and Systems Engineering, School of Engineering and Applied Science,
The George Washington University, 1776 G Street, N.W., Suite 110, Washington, DC 20052, USA

Received 7 April 2003; accepted 6 June 2003

Abstract

A model for building statistical dependence between marginal distribution with bounded support is discussed. The model is geared towards elicitation of dependence parameters through expert judgment. The resulting joint distribution may be useful in uncertainty analyses where dependence between random variables with a bounded support is present due to common risk factors, such as, e.g., in the classical Project Evaluation and Review Technique.

Keywords: Uncertainty modeling; Mixture of uniform distributions; Expert judgment

---

1. Introduction

> *"The concepts of dependence permeates the Earth and its inhabitants in a most profound manner. Examples of interdependent meteorological phenomena in nature and interdependence in the medical social, and political aspects of our existence, not to mention the economic structures are too numerous to be cited individually"*—Drouet and Kotz (2001).

The quote above expresses the need for modeling of dependence between uncertain phenomena. Dependent uncertainty analysis is usually performed with a generic software platform (@Risk, Crystall Ball) or with specialized programs such as UNICORN (see Cooke, 1995; Bedford and Cooke, 2002; Kurowicka and Cooke, 2002) or the Probability Bounds Analysis Software by Ferson (1997).

The long-standing issue of dependence between random variables has recently been discussed in application areas such as project risk analysis (see, e.g., Duffey and Van Dorp, 1998), accident probability analysis (see, e.g., Yi and Bier, 1998), Finance (see, e.g., Härdle et al., 2002) and decision analysis (see, e.g., Clemen and Reilly, 1999). Frees and Valdez (1998) introduced dependence in actuarial modeling. These authors unanimously suggested the copula approach (see, e.g., Sklar, 1959; Genest and Mackay, 1986; Nelsen, 1999) for dependence modeling. An advantage of the copula

---

* Tel./fax: +1-202-994-6638.
E-mail address: dorpjr@gwu.edu (J.R. van Dorp).

0377-2217/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2003.06.028

approach is that it utilizes the decomposition principle by separately describing the uncertainty aspect via the marginal distributions and dependence features between components via copula's.

Although, by now high dimensional sampling routines between a large number of random variables, say a 100 or more, is computationally not too difficult, the representation or modeling of dependence in models of that size in a meaningful manner is still quite cumbersome. With n specified random variables with known marginal distributions, building dependence usually requires specification of $\binom{n}{2}$ correlations (see, e.g., Law and Kelton, 1991). Applications with 100 random variables or more are feasible (see, e.g., Palisade Corporation, 1997), but specification of some $\binom{100}{2} = 4950$ correlations or more becomes a formidable task. Making this task even more daunting is that databases typically collect information at the individual random variable level, thereby not allowing for the assessment of correlations by means of classical statistical techniques. Hence, one is often compelled in models of this size to utilize the relaxed assumption of independence between the random variables or resort to a probability bounds analysis as suggested by Ferson (2001).

Instead, one may develop an approach to model statistical dependence between the random variables by identifying common risk factors as the source of dependence. The idea of common risk factors or common causes is not new and has already found wide appreciation in fault tree analysis for chemical and nuclear power plants (see, e.g., Haasl et al., 1981 or Zhang, 1989). Alternatively, common risk factors may be viewed as latent variables. Latent variable models have found wide application in the behavioral sciences (see, e.g., Bartholomew, 1987). Duffey and Van Dorp (1998) proposed eliciting dependence via expert judgment by using such common risk factors, however, only a single risk factor was allowed to influence the uncertainty distribution of a random variable which seems too restrictive for practical purposes. The dependence model herein extends the work in Duffey and Van Dorp (1998) by allowing multiple common risk factors to affect a single random variable. The extension utilizes a mixture of uniform random variables and its cumulative distribution function to allow for the above mentioned copula approach. A significant reduction is achieved in the required number of dependence parameters compared to the correlation matrix approach (600 in a dependence model with five common risk factors and 100 random variables) while allowing separate specification of marginal distributions.

In Section 2, a model for building multivariate dependence between random variables utilizing common risk factors will be discussed. The multivariate dependence of Section 2 utilizes a bivariate dependence model which is discussed in Section 3. In addition, Section 3 introduces a new dependence measure that in its interpretation resembles the well known R^2 measure in regression analysis. The models discussed in Sections 2 and 3 allow for elicitation of dependence parameters through the use of expert judgment in a meaningful manner. Section 4 discusses a theoretical result related to the dependence model in Section 2. In Section 5, the model is applied to a Project Evaluation and Review Technique (PERT) example (see, e.g., Winston, 1993). In the example, the effect of neglecting dependence will be benchmarked against a long-standing controversy regarding the use of beta distributions and triangular distributions in PERT analyses (see e.g., Clark, 1962; Grubbs, 1962; Kamburowski, 1997) with an assumption of independence between the random variables. Table 1 below summarizes the analysis results. A small project network consisting of 18 activities (see Fig. 1) and its accompanying minimal completion time was used to compare the effect of a mild dependence assumption amongst the durations of these activities against an existing controversy regarding

Table 1
Mean and standard deviation of the project completion time distribution using triangular, beta and TSP ($n = 5$) under independence and triangular distributions under dependence

	Mean	SD
Triangular-independence	155.15	5.04
Beta-independence	150.01	4.06
TSP ($n = 5$)–independence	149.85	2.96
Triangular-mild dependence	154.92	8.74