

# **Equivalence of methods for uncertainty propagation of real-valued random variables**

Helen M. Regan

Applied Biomathematics, 100 North Country Road, Setauket, NY 11733

Biology Department, San Diego State University, 5500 Campanile Drive, San Diego, CA 92184-4614

Scott Ferson

Applied Biomathematics, 100 North Country Road, Setauket, NY 11733

Daniel Berleant

Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011

Author to whom correspondence should be addressed:

Helen M. Regan

Biology Department

San Diego State University

5500 Campanile Drive

San Diego, CA 92107-4614

Phone: (619) 594 2738

Fax: (619) 594 5676

Email: [hregan@sciences.sdsu.edu](mailto:hregan@sciences.sdsu.edu)

## **Abstract**

In this paper we compare four probabilistic bounding methods for the reliable propagation of uncertainty through calculations involving the binary operations of addition, multiplication, subtraction and division. The methods we investigate are: *i*) dependency bounds convolution; *ii*) Distribution Envelope Determination; *iii*) imprecise probabilities; and *iv*) Dempster-Shafer belief functions. We show that while each of these methods was constructed for different types of applications, they converge to equivalent methods when they are restricted to cumulative distribution functions on the positive reals. For Dempster-Shafer belief functions we require the additional restriction of a partition of the positive reals. We also show that while some of the methods have been formally constructed to deal only with operations on random variables under an assumption of independence, all of the methods can be extended to deal with unknown dependencies and perfect positive and negative dependence among variables.

**Keywords:** dependency bounds, Dempster-Shafer belief functions, imprecise probabilities, uncertainty propagation

## Introduction

In this paper we address the general problem of performing convolutions of real-valued random variables under binary operations, when there are multiple sources of variation and uncertainty in the constituent random variables and in the dependency between them.

Suppose we wish to add, subtract, multiply or divide two real-valued random variables,  $A$  and  $B$ , to produce a new joint random variable. Denoting the binary operation by  $*$ , we wish to describe the distribution of  $A*B$  from the distributions of  $A$  and  $B$ . The general formulation of the joint distribution for  $A*B$ , given marginal distributions for  $A$  and  $B$ , can be described by the following convolution:

$$\text{insert convolution equation here} \tag{1}$$

where  $A$  and  $B$  are real-valued random variables,  $F_A(x)$  and  $F_B(y)$  are the marginal cumulative distribution functions (cdfs) for  $A$  and  $B$  respectively, and  $*$  is a binary operation, specifically  $*$   $\in \{+, -, \times, \div\}$ . It is well known that analytic solutions for equation (1) may only be derived for simple functions  $F_A(x)$  and  $F_B(y)$  (e.g. when  $F_A(x)$  and  $F_B(y)$  are uniform distribution functions, however, there other examples where analytic solutions may be derived). Numerical integration algorithms are usually used to solve the integral in equation (1), even for the simple cases when analytic solutions are available. Monte Carlo simulations are the best known, and most commonly applied, numerical methods for solving convolutions of this type.

The Monte Carlo method for performing convolutions of the type in equation (1) involves selecting uniformly distributed random deviates,  $U_x$  and  $U_y$ , between 0 and 1 for each random variable  $A$  and  $B$  respectively. The inverse functions,  $F_A^{(-1)}(U_x)$  and  $F_B^{(-1)}(U_y)$ , are then calculated and  $x*y = F_A^{(-1)}(U_x) * F_B^{(-1)}(U_y)$  is formed. The process is repeated many times such that the inverse cumulative distribution functions  $F_A^{(-1)}$  and  $F_B^{(-1)}$  are sampled sufficiently well

across their full range of values. Once a satisfactory number of samples are calculated for  $A*B$  the cumulative distribution function  $F_{A*B}$  can be formed via simple frequency-based cumulative probabilities.

The convolution in equation (1) and the Monte Carlo method described above to solve it, both rely on a couple of important assumptions. First, the random variables  $A$  and  $B$  are assumed to be independent. There are many practical situations where there is a specified dependency relationship between the random variables. More importantly, there are many more cases where the dependency relationship is only partially specified (e.g. there exists a restricted range of relevant dependency structures for  $A$  and  $B$  but which of these is the true structure is unknown) or completely unknown. The Monte Carlo method described above can be extended to cases where a specific dependency structure is known. For example, when  $A$  and  $B$  are perfectly positively correlated (i.e. when the value of  $A$  is large relative to its full range, so is the value of  $B$ , likewise when  $A$  is small relative to its full range, so is the value of  $B$ ) then  $U_x = U_y$  for each pair of samples in the simulation. Conversely, if  $A$  and  $B$  are perfectly negatively correlated,  $U_x = 1 - U_y$  for each pair of samples in the simulation. There are also methods for Monte Carlo simulation when  $A$  and  $B$  exhibit other types of specific dependency relationship (e.g. when the correlation coefficient for  $A$  and  $B$  is a specific value between 0 and 1; see REFS for details). However, for the more general and realistic case where the dependency relationship is only partially known or completely unknown, there is no strategy within the Monte Carlo methodology for performing convolutions under binary operations.

A second assumption behind equation (1) and the Monte Carlo method described above, is that the entire suite of sources of uncertainty in the random variables  $A$  and  $B$  is represented in the distribution functions  $F_A(x)$  and  $F_B(y)$ , and that these functions are known without error. In

many practical instances uncertainty will arise from natural variation and from more than one source of observation error (Regan et al. 2001, Morgan and Henrion 1990). A single cumulative distribution function  $F_A(x)$  is intended to model the uncertainty in  $A$  due to variation. However, the shape of  $F_A(x)$  may also be uncertain due to observation error (Regan et al. 2002; Regan et al. in press). When this is the case, the true distribution for the variable of interest,  $A$ , will be one of a range of likely candidates, bounded within upper and lower cdfs,  $\overline{F}_A(x)$  and  $\underline{F}_A(x)$  (Figure 1). 2<sup>nd</sup>-order Monte Carlo methods are sometimes employed to deal with convolutions of variables described by lower and upper bounds on the respective cdfs. 2<sup>nd</sup>-order Monte Carlo simulations are carried out in much the same way as described above, however,  $\overline{F}^{(-1)}(U)$  and  $\underline{F}^{(-1)}(U)$  are also calculated for  $A$  and  $B$ , where the uniformly distributed random deviates may or may not be correlated depending on what is known about the dependence between  $A$  and  $B$ . In addition to specifying the functions  $F_A(x)$  and  $F_B(y)$ , 2<sup>nd</sup>-order Monte Carlo simulations require assumptions regarding the realistic range of possible input distributions. Hence, the more uncertain a variable is, the more data is required to assign bounds on its cdf. This is counterintuitive. The more data available, the more certain we should be of the form of the cdf. Furthermore, the bounding cdfs,  $\overline{F}_A(x)$  and  $\underline{F}_A(x)$ , are usually expressed as upper and lower percentiles on the full range of cdfs. For instance,  $\underline{F}_A(x)$  and  $\overline{F}_A(x)$  might represent the 5<sup>th</sup> and 95<sup>th</sup> percentile cdfs of the full range of possible bounding cdfs on  $F_A(x)$ . Hence, they do not encompass the full extent of uncertainty due to variation and observation error. As a result, 2<sup>nd</sup>-order Monte Carlo simulation is often an unsatisfactory and misleading treatment of compounding uncertainty. This is due, in part, to the fact that Monte Carlo methods were not originally designed to deal with multiple sources of uncertainty in random variables and so their extension to 2<sup>nd</sup>-order simulation for the

treatment of both variation and observation error is somewhat *ad hoc*.

In this paper we present methods that were specifically designed for the treatment of unknown dependency between variables, or for multiple forms of uncertainty in the variables of interest, or for both. We analyze and compare four well-developed methods for propagating uncertainty through calculations when the uncertainty in the input variables can be represented as bounds on probability distributions. These methods are:

- i. dependency bounds convolution [8],
- ii. Distribution Envelope Determination, or DEnv [7],
- iii. imprecise probabilities [9],
- iv. Dempster-Shafer belief functions [10].

We show that while each of these methods was constructed for a different purpose, they all converge on the same method when the domain is restricted to the positive reals,  $\mathfrak{R}^+$ , and the operands are expressed in the form of cumulative distribution functions (cdfs) and, for Dempster-Shafer belief functions, the additional restriction that focal elements be a partition of the positive real line (this caveat will be explained in detail in latter sections). We also show that while some of the methods have been formally constructed to deal only with independent random variables, they all can be extended to deal with unknown dependencies as well as perfect positive and negative dependence amongst variables.

The structure of this paper is as follows: in the next section we present a background to the problem of convolutions of random variables with unknown dependency structure, and the formulation of its solutions. In the following sections we present methods for the construction of bounds on random variables and convolutions of bounded variables (where “bounded” here means the form of the cdf for a random variable is uncertain but lies within known bounds). In

the remainder of the paper we show how the methods presented here are equivalent under the set of restrictions outlined above and how they can be extended to include treatments of uncertainty in the cumulative distribution functions and uncertainty in the dependency between random variables. This equivalence relies on constructing bounding cdfs for each random variable, such as in Figure 1, then deconstructing these bounds into a set of intervals with associated probability mass and performing convolutions of the set of intervals.

### Unknown dependency between random variables

In the mid-1900s Kolmogorov posed the following well-known problem: given marginal cumulative probability distributions  $F_A(x)$  and  $F_B(y)$  for random variables  $A$  and  $B$ , what is the resultant joint distribution for  $A+B$  when the dependency structure between  $A$  and  $B$  is unknown? Fréchet had solved a similar problem for *events*  $A$  and  $B$ , with respective probabilities  $P(A)$  and  $P(B)$ , and the binary operations conjunction (AND,  $\wedge$ ) and disjunction (OR,  $\vee$ ):

$$\text{Insert Frechet bounds here} \tag{2}$$

However, these bounds do not immediately and obviously carry over to the case where  $F_A(x)$  and  $F_B(y)$  are marginal distributions. While there were many attempts at solving Kolmogorov's problem, it was not until 1987 that it was fully solved by Frank, Nelson, and Schweizer [16]. Williamson and Downs [8] generalized this solution to convolutions of two random variables under unknown dependency for binary operations  $* \in \{+, -, \times, \div\}$ . Both treatments of Kolmogorov's problem rely on the theory of copulas [17] (first introduced by Sklar in 1959 [REF]) and are summarized in the following theorem:

Theorem 1 (modified from [8, 16])

Let  $A$  and  $B$  be almost surely positive (i.e. positive almost everywhere) random variables with distributions  $F_A$  and  $F_B$ , and let  $Z = A * B$ , where  $*$   $\in \{+, -, \times, \div\}$ . The lower and upper dependency bounds ( $\underline{F}_Z$  and  $\overline{F}_Z$  and respectively) for  $F_Z$  are given by

$$\begin{aligned}\underline{F}_{A+B}(z) &= \max_{x+y=z} \{ \max[F_A(x) + F_B(y) - 1, 0] \} \\ \overline{F}_{A+B}(z) &= \min_{x+y=z} \{ \min[F_A(x) + F_B(y), 1] \} \\ \underline{F}_{A-B}(z) &= \max_{x+y=z} \{ \max[F_A(x) - F_B(-y), 0] \} \\ \overline{F}_{A-B}(z) &= 1 + \min_{x+y=z} \{ \min[F_A(x) - F_B(-y), 0] \} \\ \underline{F}_{A \times B}(z) &= \max_{x \times y=z} \{ \max[F_A(x) + F_B(y) - 1, 0] \} \\ \overline{F}_{A \times B}(z) &= \min_{x \times y=z} \{ \min[F_A(x) + F_B(y), 1] \} \\ \underline{F}_{A \div B}(z) &= \max_{x \times y=z} \{ \max[F_A(x) - F_B(1/y), 0] \} \\ \overline{F}_{A \div B}(z) &= 1 + \min_{x \times y=z} \{ \min[F_A(x) - F_B(1/y), 0] \}\end{aligned}$$

In fact, in the case where  $*$  is addition or subtraction (check this – I changed it from multiplication), the bounds hold for any pair of marginals,  $F_A$  and  $F_B$ , not just those for positive  $A$  and  $B$ . This is proved in [16] (for addition) and [8] (for all the arithmetic operators  $*$   $\in \{+, -, \times, \div\}$ ). Note that here, the upper and lower bounds on the convolution result from uncertainty in the dependency of  $A$  and  $B$ , and not from uncertainty in the marginal distributions. The result above (and its operationalizations outlined below) is referred to here as dependency bounds convolution. A noticeable and satisfying feature of the theorem above is that it relies on an analogy of Fréchet’s bounds. While this is not surprising in and of itself (the theory of copulas relies, to an extent, on Fréchet’s bounds), the fact that it is extremely nontrivial to extend Fréchet’s bounds to solve Kolmogorov’s problem is testament to the fact that what often appears obvious in mathematics, at face value, is very often confoundingly difficult to prove with the

theory at hand. In this case, the theory of copulas was necessary to make the extension from Fréchet bounds to dependency bounds convolution.

The bounds expressed in Theorem (1) are point-wise best possible. This means that the bounds are wide enough to permit inclusion of all possible dependencies of  $A$  and  $B$  and no wider. That is, every point  $\overline{F}_{A*B}(z)$  and  $\underline{F}_{A*B}(z)$  refer to the cumulative probability of the realization of  $A*B$  under some dependency between the two random variables and there is no possible dependency structure that lies outside these bounds.

We will not dwell on the theory of copulas employed to solve Kolmogorov's problem, because, as in the case of equation (1), analytic expressions are difficult to extract for many realizations of  $F_A(x)$  and  $F_B(y)$ . Rather, we are interested in the operationalization of this theorem. Williamson and Downs convert the assignments in Theorem (1) into a numerical algorithm to accommodate discretized cumulative distributions functions (see [8] and [15] for implementations of this). The approach to performing convolutions of  $A$  and  $B$  under binary operations described by Williamson and Downs [8] involves discretizing each precise cdf,  $F_A(x)$  and  $F_B(y)$ , into upper and lower cdfs with  $n$  steps of equal height throughout (in fact, for some cases the number of steps may be different for each random variable). The way this is achieved is by first partitioning the cumulative probability into  $n$  equal parts from 0 to 1 (see Figure 2 for an example of this). This will result in a partition of the probability scale into  $n$  equal intervals, each of length  $p_i = 1/n$  where  $i = 1, \dots, n$ . The partition of the probability scale  $[0, 1]$ , will consist of the intervals  $[0, 1/n], (1/n, 2/n], (2/n, 3/n], \dots, ((n-1)/n, 1]$ . This can be summarized as

$(\sum_{i=0}^j p_i, \sum_{i=0}^{j+1} p_i]$  where  $j = 0, \dots, n-1$ , and  $p_0 = 0$ . For each interval along the cumulative probability axis the inverse cumulative distribution function is calculated as

$$\left[ F^{(-1)}\left(\sum_{i=0}^j p_i\right), F^{(-1)}\left(\sum_{i=0}^{j+1} p_i\right) \right] = [a_{L_{j+1}}, a_{U_{j+1}}] \quad (3)$$

where  $j = 0, \dots, n-1$ , and  $p_0 = 0$  (see Figure 2 for a diagram of this). (For distributions with infinite tails we first truncate the tails at extreme finite ends). Note that in equation (3)

$a_{L,j} = a_{U_{j-1}}, \forall j = 1, \dots, n$  because the intervals along the cumulative probability form a

partition. For a random variable  $A$  this results in the following step functions for upper and lower cdfs:

$$\bar{F}_A(x) = \begin{cases} 0 & \text{if } x < a_{L1} \\ \sum_{i=0}^j p_i = \frac{j}{n} & \text{if } a_{Lj} \leq x < a_{L,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } x \geq a_{Ln} \end{cases} \quad (4a)$$

$$\underline{F}_A(x) = \begin{cases} 0 & \text{if } x \leq a_{U1} \\ \sum_{i=0}^j p_i = \frac{j}{n} & \text{if } a_{Uj} < x \leq a_{U,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } x > a_{Un} \end{cases} \quad (4b)$$

where  $a_{L,j} = a_{U_{j-1}}, \forall j = 1, \dots, n$  and  $p_i = p_{i+1}, \forall i = 1, \dots, n-1$  (see also Figure 2). The original function  $F_A(x)$  is enclosed by  $\bar{F}_A(x)$  and  $\underline{F}_A(x)$ . Note that these upper and lower bounds do not refer to bounds due to any dependency considerations in a convolution. They are merely the result of discretizing the cdf in such a way that all values  $F_A(x)$  are surely contained within the bounds  $\bar{F}_A(x)$  and  $\underline{F}_A(x)$ . For example, in Figure 2,  $n = 5$ , and the precise cdf for random variable  $A$  is discretized into upper and lower cdfs, each with 5 steps at equal increments in the cumulative probability. In this case  $p_i = 1/5, \forall i = 1, \dots, 5$ .

Now suppose we wish to perform the convolution of random variables  $A$  and  $B$  under a binary operation given the discretization of their respective cumulative distribution functions as specified in equation (4). We will first demonstrate how this works under an assumption of independence between the random variables (because this is the simplest case) and then extend this to perfect positive and negative dependence and then finally we will use a modification of Theorem 1 to address the more difficult case of unknown dependency. The numerical method for calculating the bounds on the cumulative distribution of  $A*B$  using the bounds specified in equation (4) (and similar bounds calculated on  $F_B(y)$ ) involves calculating the inverse of the discretized upper and lower cdfs for each random variable. In fact, in equation (3) we have already specified the inverse intervals of the discretized upper and lower cdf for random variable  $A$ . For the random variable  $B$  the inverse intervals are specified as

$$[F_B^{(-1)}\left(\sum_{k=0}^l q_k\right), F_B^{(-1)}\left(\sum_{k=0}^{l+1} q_k\right)] = [b_{L_{k+1}}, b_{U_{k+1}}] \quad (5)$$

where  $l = 0, \dots, m - 1$ , and  $q_0 = 0$ . Note that here the cumulative probability is partitioned into  $m$  intervals for  $F_B(y)$ , whereas for  $F_A(x)$  the cumulative probability is partitioned into  $n$  intervals.

Also note that the intervals in equations (3) and (5) may be specified as

$$[\bar{F}_A^{(-1)}\left(\sum_{i=1}^j p_i\right), \underline{F}_A^{(-1)}\left(\sum_{i=1}^j p_i\right)] = [a_{L_j}, a_{U_j}] \quad (6a)$$

$$[\bar{F}_B^{(-1)}\left(\sum_{k=1}^l q_k\right), \underline{F}_B^{(-1)}\left(\sum_{k=1}^l q_k\right)] = [b_{L_l}, b_{U_l}] \quad (6b)$$

by taking the inverse of equations (4) at each upper interval terminal  $\sum_{i=1}^j p_i, \forall j = 1, \dots, n$  and

$\sum_{k=1}^l q_k, \forall l = 1, \dots, m$  (also see Figure 2). Similar relationships would be obtained by considering

each lower interval terminal, however, to perform the convolution it is convenient here to consider only the upper terminals of the probability intervals. Throughout this paper, the following inequalities hold for the intervals in equation (6):  $a_{L,j} \leq a_{U,j}$ ,  $a_{L,j} \leq a_{L,j+1}$  and  $a_{U,j} \leq a_{U,j+1}$ . We refer to equations (6) as a *deconstruction* into intervals of the lower and upper bounds on the cdfs for random variables  $A$  and  $B$ . Each interval  $[a_{L,j}, a_{U,j}]$  has probability mass  $p_i = 1/n$  associated with it. Likewise, each interval  $[b_{L,i}, b_{U,i}]$  has an associated probability mass  $q_k = 1/m$ .

In order to perform a convolution of  $A$  and  $B$  under a binary operation  $* \in \{+, -, \times, \div\}$ , regardless of the dependency structure between them, we appeal to interval arithmetic. Interval arithmetic, as suggested by its name, refers to arithmetic performed on intervals and is carried out in the following way for general intervals  $A' = [a_1, a_2]$  and  $B' = [b_1, b_2]$  where  $a_1 \leq a_2$  and  $b_1 \leq b_2$  (in keeping with Theorem 1 we will assume that the intervals are surely positive, but in fact the calculations hold for addition and subtraction on any pair of intervals not just those that are positive):

$$\begin{aligned}
[a_1, a_2] + [b_1, b_2] &= [a_1 + b_1, a_2 + b_2] \\
[a_1, a_2] - [b_1, b_2] &= [a_1 - b_2, a_2 - b_1] \\
[a_1, a_2] \times [b_1, b_2] &= [a_1 \times b_1, a_2 \times b_2] \\
[a_1, a_2] \div [b_1, b_2] &= [a_1 \div b_2, a_2 \div b_1]
\end{aligned} \tag{7}$$

The bounds in equations (7) are calculated to give the maximum interval length for the result on the RHS given the interval lengths in the inputs on the LHS. In this way, if two uncertain quantities  $a'$  and  $b'$  surely lie somewhere within the bounds  $A'$  and  $B'$ , the bounds  $A' * B'$  will surely contain the uncertainty quantity  $a' * b'$ . For a binary operation under the assumption of independence, all pair-wise combinations of  $[a_{L,j}, a_{U,j}]$  and  $[b_{L,i}, b_{U,i}]$  are selected and the

calculations in equations (7) are applied according to which binary operation is of interest. For example, if we wish to calculate the convolution  $A + B$ , we would calculate

$$[a_{Lj}, a_{Uj}] + [b_{Ll}, b_{Ul}] = [a_{Lj} + b_{Ll}, a_{Uj} + b_{Ul}] \quad \forall j = 1, \dots, n \text{ and } l = 1, \dots, m.$$

If  $F_A$  is discretized into  $n$  steps and  $F_B$  is discretized into  $m$  steps, there will be  $n \times m$  pair-wise combinations. Recall that  $[a_{Lj}, a_{Uj}]$  and  $[b_{Ll}, b_{Ul}]$  have associated probability masses of  $p_i = 1/n$  and  $q_k = 1/m$  respectively. The result of a binary operation applied to each pair-wise combination of  $[a_{Lj}, a_{Uj}]$  and  $[b_{Ll}, b_{Ul}]$  must also have an associated probability mass. Since

we are initially operating under an assumption of independence, this will simply be  $\frac{1}{n \times m}$  (recall that the probability of two independent events is the product of the probabilities for each event).

The lower bound for the sum  $A+B$  is then constructed by sorting the values  $a_{Lj} + b_{Ll}$  into ascending order. This will give the set of values  $\{(a+b)_{Lh} = a_{Lj} + b_{Ll} : h = 1, \dots, n \times m\}$  from which the upper bound can be constructed as

$$\bar{F}_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x < (a+b)_{L1} \\ \frac{h}{n \times m} & \text{if } (a+b)_{Lh} \leq x < (a+b)_{L,h+1}, \quad h = 2, \dots, nm-1 \\ 1 & \text{if } x \geq (a+b)_{L,nm} \end{cases} \quad (8a)$$

where  $+_{ind}$  refers to addition under independence. A similar construction for the lower bound for  $A + B$  yields

$$F_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x \leq (a+b)_{U1} \\ \frac{h}{n \times m} & \text{if } (a+b)_{Uh} < x \leq (a+b)_{U,h+1}, \quad h = 2, \dots, nm-1 \\ 1 & \text{if } x > (a+b)_{U,nm} \end{cases} \quad (8b)$$

This construction is identical to equations (4), where  $a$  is replaced by  $a + b$ .

The approach for calculating the joint distribution of  $A$  and  $B$  under an assumption of perfect positive or negative dependence is the same as that for independence except for a few minor differences. Suppose that  $F_A$  and  $F_B$  are discretized to give upper and lower cdfs as described above, but this time with the same number of equal steps in the cumulative probability. That is, both  $F_A$  and  $F_B$  are discretized into  $n$  steps each. Since the cumulative probability is discretized uniformly for both cdfs, the addition of two random variables under the assumption of perfect positive dependence is simply the addition of values of each random variable at equal level along the probability axis. Hence, instead of adding all pairwise combinations of  $[a_{Lj}, a_{Uj}]$  and  $[b_{Ll}, b_{Ul}]$ , we calculate

$$[a_{Lj}, a_{Uj}] + [b_{Lj}, b_{Uj}] = [a_{Lj} + b_{Lj}, a_{Uj} + b_{Uj}] \quad \forall j = 1, \dots, n .$$

The associated probability mass for each resultant interval is simply  $1/n$ . The construction of upper and lower bounds is then identical to the case for independence.

For perfect negative dependence, intervals of one random variable at each fractile,  $r$ , are added to intervals of the other random variable at the opposite fractile,  $1-r$  (here a fractile is a probability along the cumulative probability scale). Hence, for addition, we have the following:

$$[a_{Lj}, a_{Uj}] + [b_{L_{n-j+1}}, b_{U_{n-j+1}}] = [a_{Lj} + b_{L_{n-j+1}}, a_{Uj} + b_{U_{n-j+1}}], \quad \forall j = 1, \dots, n$$

The associated probability mass and the method for construction of upper and lower bounds are identical to the case for perfect positive dependence.

Our description so far of the Williamson and Downs operationalization of the dependency bounds convolution has not yet dealt with the case for which it was originally formulated—convolutions with unknown dependency between random variables. However, we have established the main components to which we will apply Williamson and Downs'

operationalization. These main components are: the upper and lower bounds on the marginal cdfs; the intervals resulting from a deconstruction of the bounds on the marginal cdfs; and their associated probabilities. For dependency bounds convolution of two random variables  $A$  and  $B$ , each with cdfs that have been discretized into  $n$  equal steps along the cumulative probability, to give lower and upper bounds on the respective cdfs, the following modification of Theorem 1 applies:

Theorem 2 [8]

Let  $A$  and  $B$  be almost surely positive (i.e. positive almost everywhere) random variables with discretized distributions to give upper and lower bounds  $\overline{F}_A$  and  $\underline{F}_A$  for random variable  $A$  and  $\overline{F}_B$  and  $\underline{F}_B$  for random variable  $B$ . Let  $Z = A * B$ , where  $* \in \{+, -, \times, \div\}$ . The upper and lower dependency bounds ( $\overline{F}_Z$  and  $\underline{F}_Z$  respectively) for  $F_Z$ , the distribution of  $Z$ , are given by

$$\begin{aligned}\underline{F}_{A+B}(z) &= \max_{x+y=z} \{ \max[\underline{F}_A(x) + \underline{F}_B(y) - 1, 0] \} \\ \overline{F}_{A+B}(z) &= \min_{x+y=z} \{ \min[\overline{F}_A(x) + \overline{F}_B(y), 1] \} \\ \underline{F}_{A-B}(z) &= \max_{x+y=z} \{ \max[\underline{F}_A(x) - \overline{F}_B(-y), 0] \} \\ \overline{F}_{A-B}(z) &= 1 + \min_{x+y=z} \{ \min[\overline{F}_A(x) - \underline{F}_B(-y), 0] \} \\ \underline{F}_{A \times B}(z) &= \max_{x \times y=z} \{ \max[\underline{F}_A(x) + \underline{F}_B(y) - 1, 0] \} \\ \overline{F}_{A \times B}(z) &= \min_{x \times y=z} \{ \min[\overline{F}_A(x) + \overline{F}_B(y), 1] \} \\ \underline{F}_{A \div B}(z) &= \max_{x \times y=z} \{ \max[\underline{F}_A(x) - \overline{F}_B(1/y), 0] \} \\ \overline{F}_{A \div B}(z) &= 1 + \min_{x \times y=z} \{ \min[\overline{F}_A(x) - \underline{F}_B(1/y), 0] \}\end{aligned}$$

We see that the equations in Theorem 2 are analogous to the interval arithmetic specified above, where the upper and lower bounds on the cdfs form the terminals of the intervals. If the bounds on  $F_A$  and  $F_B$  are deconstructed into  $n$  intervals, as described above, then the following holds for addition, for each  $j = 1, \dots, n$ :

$$\begin{aligned}
\overline{F}_{A+B}^{(-1)}\left(\sum_{i=1}^j p_i\right) &= \max_{1 \leq k \leq j-1} \left[ \overline{F}_A^{(-1)}\left(\sum_{i=1}^k p_i\right) + \overline{F}_B^{(-1)}\left(\sum_{i=1}^j p_i - \sum_{i=1}^k p_i\right) \right] \\
\underline{F}_{A+B}^{(-1)}\left(\sum_{i=1}^j p_i\right) &= \min_{j \leq k \leq n} \left[ \underline{F}_A^{(-1)}\left(\sum_{i=1}^k p_i\right) + \underline{F}_B^{(-1)}\left(\sum_{i=1}^k p_i - \sum_{i=1}^j p_i + 1\right) \right]
\end{aligned} \tag{9}$$

(Refer to [8, 16] for formulation of the inverses of the functions in Theorem 2). Since

$\sum_{i=1}^j p_i = \frac{j}{n}$ ,  $\forall j = 1, \dots, n$  and using equations (6), we can write the following:

$$\begin{aligned}
\overline{F}_{A+B}^{(-1)}\left(\frac{j}{n}\right) &= \max_{k=1, \dots, j-1} \left[ \overline{F}_A^{(-1)}\left(\frac{k}{n}\right) + \overline{F}_B^{(-1)}\left(\frac{j}{n} - \frac{k}{n}\right) \right] \\
&= \max_{k=1, \dots, j-1} \left[ a_{L,k} + b_{L, \frac{(j-k)}{n}} \right]
\end{aligned} \tag{10a}$$

$$\begin{aligned}
\underline{F}_{A+B}^{(-1)}\left(\frac{j}{n}\right) &= \min_{k=j, \dots, n} \left[ \underline{F}_A^{(-1)}\left(\frac{k}{n}\right) + \underline{F}_B^{(-1)}\left(\frac{k}{n} - \frac{j}{n} + 1\right) \right] \\
&= \min_{k=j, \dots, n} \left[ a_{U,k} + b_{U, \frac{(k-j+n)}{n}} \right]
\end{aligned} \tag{10b}$$

Similar relations can be derived for subtraction, multiplication and division using Theorem 2 (see

[8] for details). The bounds  $\overline{F}_{A+B}(z)$  and  $\underline{F}_{A+B}(z)$  are then formed by taking the inverse of

equations (10). This is equivalent to sorting the values  $\overline{F}_{A+B}^{(-1)}\left(\frac{j}{n}\right)$  from least to greatest (for the

upper bound) and then assigning cumulative probabilities  $\frac{j}{n}$ , (for  $j = 1, \dots, n$ ) to these values in

ascending order. The same procedure applies to the lower bound  $\underline{F}_{A+B}(z)$  using the values

$\underline{F}_{A+B}^{(-1)}\left(\frac{j}{n}\right)$ . The resulting bounds are point-wise best possible for the discretizations of  $F_A$  and  $F_B$ .

Hence, we have presented convolutions of  $A$  and  $B$  under the binary operations

$* \in \{+, -, \times, \div\}$  for discretizations of cdfs  $F_A$  and  $F_B$  under assumptions of independence, perfect

positive and negative dependence and the general case of unknown dependency. In the following section we present an alternative method for calculating dependency bounds, known as Distribution Envelope Determination, when the cdfs  $F_A$  and  $F_B$  are uncertain. It is our aim to show that the two approaches to convolution are equivalent for independence, perfect positive and negative dependence and unknown dependence.

### **Distribution Envelope Determination**

Distribution Envelope Determination is a method for determining dependency bounds on cdfs for convolutions of random variables  $A$  and  $B$  under binary arithmetic operations when the input cdfs may themselves be uncertain. It was devised by Berleant and Goodman-Strauss in 1998 [7] independently of the dependency bounds convolution method described above and differs in the procedure for determining the bounds and in the fact that multiple sources of uncertainty are incorporated in the input cdfs.

Berleant and Goodman-Strauss start by describing  $A$  and  $B$  with probability density functions (pdfs),  $f_A(x)$  and  $f_B(y)$  respectively. These pdfs may be discretized using histograms (Figure 3a). This is achieved by forming a partition of the range of values of  $A$  and  $B$  and calculating the probability under the curves  $f_A(x)$  and  $f_B(y)$  for each interval in the partition. To handle the problem of discretization error, each histogram interval may be interpreted as stating a range (from its left side to its right side) and a probability mass (its area) distributed over its range, with no assumption made about how the mass is distributed within the range.

Generalizing the histogram concept, we can allow the intervals of  $A$  and  $B$  to overlap. Each interval is interpreted as expressing a probability mass distributed within a range, just as described above (Figure 3b). Let us call this range of intervals with accompanying probabilities a

*thicket* to suggest a collection of perhaps overlapping bars. Binary arithmetic on thickets is described by Berleant and Goodman-Strauss [7]. The type of algorithm used is called Distribution Envelope Determination, or DEnv. In DEnv, each random variable is first discretized into a thicket. This allows operations on distributions to become operations on sets of intervals and their associated probabilities. For example, suppose the discretization of  $f_A(x)$  yields the 2-element set

$$\left\{ P(x \in [1,2]) = \frac{1}{2}, P(x \in [2,4]) = \frac{1}{2} \right\}$$

and the discretization of  $f_B(y)$  yields the 3-element set

$$\left\{ P(y \in [2,3]) = \frac{1}{4}, P(y \in [3,4]) = \frac{1}{2}, P(y \in [4,5]) = \frac{1}{4} \right\}.$$

More generally, we can write a discretization for a pdf  $f_A(x)$  as  $\{P(x \in A_i) = p_i : i = 1, \dots, n\}$ , where  $P$  represents a probability measure on intervals of the domain of  $A$ , and where the various  $A_i$  might or might not overlap.

Now suppose we wish to calculate the probability distribution for a new random variable  $Z$ , where each sample of  $Z$  is the sum of a sample of  $A$  and a sample of  $B$ , i.e.  $Z = A * B$ . For ease of exposition we start with the assumption of independence, writing it as  $Z = A +_{ind} B$  where “ $+_{ind}$ ” indicates addition of samples of random variables under the assumption of independence. A thicket that defines the probability distribution of  $Z$  is obtained by performing interval addition on each pair of intervals  $A_i$  and  $B_j$  in the thicket discretizations of  $A$  and  $B$  to get  $Z_{ij} = A_i + B_j$  where “ $+$ ” here indicates interval addition, as defined in the previous section. (Note that when adding the values of the random variable, specification of the dependency structure is irrelevant. It is only in the assignment of probability mass to the resultant intervals that dependency

between the random variables becomes an issue). Because  $A$  and  $B$  are independent,

$P(Z_{ij}) = P(A_i) \times P(B_j)$ . For the two simple thickets described earlier,

$$\left\{ P(x \in [1, 2]) = \frac{1}{2}, P(x \in [2, 4]) = \frac{1}{2} \right\} \text{ and}$$

$$\left\{ P(y \in [2, 3]) = \frac{1}{4}, P(y \in [3, 4]) = \frac{1}{2}, P(y \in [4, 5]) = \frac{1}{4} \right\},$$

this results in the *joint distribution matrix* shown in Table 1.

A joint distribution matrix implies bounds on a corresponding cumulative distribution function (cdf), which are constructed as follows. To obtain the left envelope (Figure 4 contains a left and a right envelope), the distribution of the probability mass of each interval in the joint distribution matrix is assumed to be concentrated at the lower bound of the interval, because this will cause the cumulation to rise as fast as possible while still being consistent with the joint distribution matrix. The envelope is then constructed by numerical integration (the low bounds are sorted from smallest to greatest, and the cumulative distribution curve is stepped increased by the amount of the probability concentrated at that low bound). This envelope is referred to here as the left envelope (which corresponds to the upper bounding cdf). To construct the right envelope, the probability mass for each interval is assumed to be concentrated at the *upper* bound of the interval and then the probability mass is integrated across all the upper bounds. Figure 4 displays the bounds for the cumulative distribution of  $A +_{ind} B$ , constructed from Table 1.

In the case of perfect positive dependence, whenever one random variable is sampled, the value of the second random variable is sampled at the same fractile. For perfect negative dependence, a value of one random variable at fractile  $r$  of its distribution is convolved with a value of the other random variable at the opposite fractile,  $1-r$ , of its distribution. All the

intervals in the joint distribution matrix remain the same, as do the probabilities of the marginal intervals; it is just the calculation of the probabilities associated with the resultant joint intervals that change according to the dependency structure of the two variables. The joint distribution matrix for  $A+B$  under assumptions of perfect positive and perfect negative dependence for the example above, appear in Table 2. Bounds on the cumulative distribution function for  $A+B$  are then constructed from the  $p_+$ 's or  $p_-$ 's of Table 2 in the same way as described above for the case where  $A$  and  $B$  are assumed independent.

Berleant and Goodman-Strauss [4] also provide a method for calculating  $A*B$ ,  $*$   $\in$   $\{+, -, \times, \div\}$ , when the dependency relationship between  $A$  and  $B$  is unknown. As noted above, different dependency relationships result in different probability assignments to the intervals, which in turn result in different envelopes on the cumulative probability distributions. For the case of an unknown dependency relationship between  $A$  and  $B$ , envelopes around the cumulative distribution of a random variable  $Z = A*B$  must enclose all the envelopes that result from all possible specific dependency relationships. To find a point on these dependency envelopes, the probability masses associated with the intervals in the joint distribution matrices (such as appear in Tables 1 and 2) must be considered variable, but this variability is limited by the constraints imposed by the discretized operand distributions ( $A$  and  $B$  in Tables 1 and 2) that form the margins of the joint distribution matrix [14]. In particular, the sum of the resultant probabilities in a row or column of a joint distribution matrix is constrained to equal the probability of the marginal cell associated with that row or column. For example, the constraints imposed by the joint distribution matrix of Table 1 are:

<b>Column constraints</b>	<b>Row constraints</b>
---------------------------	------------------------

$\frac{1}{2} = p_{11} + p_{12} + p_{13}$	$\frac{1}{4} = p_{11} + p_{21}$
$\frac{1}{2} = p_{21} + p_{22} + p_{23}$	$\frac{1}{2} = p_{12} + p_{22}$
	$\frac{1}{4} = p_{13} + p_{23}$

The constraints imposed by the joint distribution matrix shown in Table 2 are exactly the same as those imposed by Table 1, because the two tables differ only in the alternative sets of values for the various  $p_{ij}$  (named  $p_{ij+}$  and  $p_{ij-}$  in Table 2), yet those sets of values are consistent with exactly the same constraints imposed by the marginal cells of the tables.

The variability of the  $p_{ij}$ 's, though limited, enables the goal of finding values for them that lead to left or right extremes in the cumulation, i.e., points on the overall dependency assumption-free envelopes. Since these constraints, exemplified in the previous paragraph, are linear, the desired extremes may be found using linear programming (or for very small problems, careful inspection) to maximize (for the left envelope) or minimize (for the right envelope) the probability mass cumulated up to a given value of  $Z=A*B$ ,  $* \in \{+, -, \times, \div\}$ . For example, the maximum cumulation possible for  $x + y$ ,  $x \in A$  and  $y \in B$ , for values of  $x + y \leq 4.5$ , given the marginals of Table 1, occurs for the dependency relationship expressed by the various  $p_{ij}$  shown in Table 3. Each  $p_{ij}$ , if assumed to be concentrated at the low bound of its corresponding interval, would cause the cumulation to rise faster than any other distribution of the  $p_{ij}$ 's over their intervals. Then the entire amount of  $p_{11}$ ,  $p_{21}$ , and  $p_{12}$  would appear in the cumulation at  $x + y = 4.5$ , implying that the left envelope has a height of  $p_{11} + p_{21} + p_{12} = 0.75$  at  $x + y = 4.5$ . Note that the envelopes themselves are not necessarily cdfs resulting from any single dependency relationship. Rather, different portions of them may be implied by different dependency

relationships. The bounds derived in this fashion are point-wise best possible in the same way as the bounds resulting from dependency bounds convolution.

### Deconstruction of bounded cdfs into a thicket

Thickets derived directly from discretization of a continuous pdf will typically be histograms. (If the pdf has infinite tail(s) then either the tails are truncated or the extreme low terminal of the first interval would have a low bound of  $-\infty$ , or the extreme high terminal of the final interval would have a high bound of  $\infty$ , or both.) However thickets might also be derived from the interior cells of a joint distribution matrix. For example, the thicket for the sum  $A+B$  in Table 1 would consist of one interval for each  $p_{ij}$ . This interval has mass  $p_{ij}$  and has a width and location on the horizontal axis defined by the interval in the same cell as  $p_{ij}$ . This thicket will have overlapping intervals.

A thicket may also be derived directly from envelopes via deconstruction, in the same way that a discretized cdf was deconstructed into intervals in the previous section. Figure 5 illustrates this process. More formally, deconstruction of a pair of envelopes, such as appear in Figure 5, into a thicket occurs as follows: we are given two discretized bounds on the cumulative distribution function for some random variable  $A$ ,

$$\bar{F}_A(x) = \begin{cases} 0 & \text{for } x < a_{L1} \\ q_h & \text{for } a_{Lh} \leq x < a_{L,h+1}, h = 1, \dots, k-1 \\ 1 & \text{for } x \geq a_{Lk} \end{cases} \quad (11a)$$

$$\underline{F}_A(x) = \begin{cases} 0 & \text{for } x < a_{U1} \\ r_j & \text{for } a_{Uj} \leq x < a_{U,j+1}, j = 1, \dots, m-1 \\ 1 & \text{for } x \geq a_{Um} \end{cases} \quad (11b)$$

such that  $\forall x, \overline{F}_A(x) \geq \underline{F}_A(x)$ , i.e.  $\overline{F}_A(x)$  is the left envelope (or upper bound) and  $\underline{F}_A(x)$  is the right envelope (or lower bound), and such that  $q_{h+1} \geq q_h$  and  $r_{j+1} \geq r_j$ . Note that equations (11) are general in that they could have resulted from the DEnv procedure applied to two pdfs or they could have been derived as discretizations of bounds assigned to a cdf due to multiple sources of uncertainty in the random variable  $A$ , as in Figure 1. We wish to construct a thicket such that the lower bounds of the intervals in the thicket are elements of the set  $\{a_{Lh} : h = 1, \dots, k\}$  and the upper bounds are elements of the set  $\{a_{Uj} : j = 1, \dots, m\}$ .

Define the inverse of  $\overline{F}_A(x)$  and  $\underline{F}_A(x)$  as follows:

$$\overline{F}_A^{(-1)}(v) = \begin{cases} a_{L1} & \text{if } 0 \leq v \leq q_1 \\ a_{Lh} & \text{if } q_{h-1} < v \leq q_h, \quad h = 2, \dots, k-1 \\ a_{Lk} & \text{if } q_{k-1} < v \leq 1 \end{cases} \quad (12a)$$

$$\underline{F}_A^{(-1)}(v) = \begin{cases} a_{U1} & \text{if } 0 \leq v \leq r_1 \\ a_{Uj} & \text{if } r_{j-1} < v \leq r_j, \quad j = 2, \dots, m-1 \\ a_{Um} & \text{if } r_{m-1} < v \leq 1 \end{cases} \quad (12b)$$

The thicket,  $\{P(a \in A_i) = p_i : i = 1, \dots, k + m - 1\}$ , is then constructed as follows (see also Figure

5). Let  $S = \{q_1, \dots, q_{k-1}, r_1, \dots, r_{m-1}, 1\}$ . Sort the elements of  $S$  from smallest to largest and rename

them  $s_1 \dots s_{k+m-1}$ . Let  $p_i$  represent the value  $P\left(v \in \left[\overline{F}_A^{(-1)}(s_i), \underline{F}_A^{(-1)}(s_i)\right]\right)$ ,  $i = 1, \dots, k + m - 1$ . Then

$$p_i = \begin{cases} s_i & \text{for } i = 1 \\ s_i - s_{i-1} & \text{for } i = 2, \dots, k + m - 1. \end{cases} \quad (13)$$

The corresponding intervals for each  $p_i$  will be  $\left[\overline{F}_A^{(-1)}(s_i), \underline{F}_A^{(-1)}(s_i)\right]$ ,  $i = 1, \dots, k + m - 1$ . One

outcome of such a decomposition into intervals is that while the intervals may overlap, and in some cases may even be nested with one terminal coinciding (i.e. two intervals  $[a, b]$  and  $[a, c]$

or  $[a, b]$  and  $[c, b]$  may result) one interval will never be completely subsumed within another (i.e. it is impossible to decompose equations (12) into intervals  $[a, b]$  and  $[c, d]$  where  $a < c$  and  $d < b$ ). Once the pair of envelopes has been decomposed into a set of intervals with associated probabilities in this way, the random variable can be convolved with another, using a binary operation and the procedure for Distribution Envelope Determination described above. Thus, random variables can be convolved when they are represented as:

- a) discretizations of continuous pdfs;
- b) envelopes resulting from binary operations applied to two discretized pdfs;
- c) discretized left and right envelopes bounding a cdf;
- d) any combination of the above.

Theorem 3.

*Let  $A$  and  $B$  be positive real-valued random variables described by lower and upper bounds on the cdf  $\bar{F}_A, \underline{F}_A, \bar{F}_B$  and  $\underline{F}_B$ . For uniform discretizations of the input bounds on cdfs, dependency bounds convolution and Distribution Envelope Determination result in identical bounds on the cdf resulting from the convolution  $A*B$  under assumptions of independence, perfect positive and perfect negative dependence and under unknown dependence between the random variables.*

Proof

We first discretize the bounds  $\bar{F}_A, \underline{F}_A, \bar{F}_B$  and  $\underline{F}_B$  uniformly into  $n$  steps along the cumulative probability axis (according to equations (4)) to give:

$$\bar{F}_A(x) = \begin{cases} 0 & \text{if } x < a_{L1} \\ \sum_{i=0}^j p_i & \text{if } a_{Lj} \leq x < a_{L,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } x \geq a_{Ln} \end{cases} \quad (14a)$$

$$\underline{F}_A(x) = \begin{cases} 0 & \text{if } x \leq a_{U1} \\ \sum_{i=0}^j p_i & \text{if } a_{Uj} < x \leq a_{U,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } x > a_{Un} \end{cases} \quad (14b)$$

$$\bar{F}_B(y) = \begin{cases} 0 & \text{if } y < b_{L1} \\ \sum_{i=0}^j p_i & \text{if } b_{Lj} \leq y < b_{L,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } y \geq b_{Ln} \end{cases} \quad (14c)$$

$$\underline{F}_B(y) = \begin{cases} 0 & \text{if } y \leq b_{U1} \\ \sum_{i=0}^j p_i & \text{if } b_{Uj} < y \leq b_{U,j+1}, j = 1, \dots, n-1 \\ 1 & \text{if } y > b_{Un} \end{cases} \quad (14d)$$

where here  $p_i = p_{i+1} = \frac{1}{n}$ ,  $\forall i = 1, \dots, n-1$  but  $a_{L,j} \leq a_{U,j-1}$ ,  $\forall j = 1, \dots, n$  and likewise for the terminals  $b_{L,j} \leq b_{U,j-1}$ , i.e. in this case the intervals  $[a_{L,j}, a_{U,j}]$  may overlap, likewise for the intervals  $[b_{L,j}, b_{U,j}]$ . This last qualification is important because for *dependency bounds convolution* this is an extension of the formulation to accommodate bounds that treat multiple sources of uncertainty in the random variables. If  $a_{L,j} = a_{U,j-1}$ ,  $\forall j = 1, \dots, n$  then the bounds in (14) are a discretization of a single precise cdf, however if  $a_{L,j} < a_{U,j-1}$  for some  $j = 1, \dots, n$ , then equations (14) specify a discretization of bounding cdfs on the uncertain function. Equations (14) are a special case of the bounds formulated for *Distribution Envelope Determination*, in that instead of a non-uniform partition of the cumulative probability scale, here we have identical uniform partitions of the cumulative probability for both random variables,  $A$  and  $B$ . Hence, the

cumulative probabilities expressed as  $q_h$  and  $r_j$  in equations (11) are expressed here as  $\sum_{i=1}^j p_i$ .

Equations (14) can be deconstructed into intervals using equations (6) for dependency bounds convolution or equations (12) for DEnv. Note that both deconstructions will give the same set of intervals, namely  $[a_{L_j}, a_{U_j}]$  and  $[b_{L_j}, b_{U_j}]$  for  $j = 1, \dots, n$ , with accompanying probability mass  $p_i$ .

This is straightforward for dependency bounds convolution. To see this for the DEnv

deconstruction, note that here  $S = \{s_j = \sum_{i=1}^j p_i, \forall j = 1, \dots, n\}$ , and hence each  $p_i$  will have a

corresponding interval  $\left[ \overline{F}_A^{(-1)}\left(\sum_{i=1}^j p_i\right), \underline{F}_A^{(-1)}\left(\sum_{i=1}^j p_i\right) \right] = [a_{L_j}, a_{U_j}]$ . Likewise for the deconstruction

of the bounding cdfs for  $B$ . It is now a simple matter to show that dependency bounds

convolution and DEnv give the same results under assumptions of independence and perfect

positive and perfect negative dependence. For an assumption of *independence* both methods

apply interval arithmetic to all pairwise combinations of intervals and assign a probability mass

of  $\frac{1}{n^2}$  to each interval. For both dependency bounds convolution and DEnv, the resultant

bounding cdfs are then constructed as (for addition but analogous equations hold for the other

binary operations):

$$\overline{F}_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x < (a+b)_{L1} \\ \frac{h}{n^2} & \text{if } (a+b)_{Lh} \leq x < (a+b)_{L,h+1}, \quad h = 2, \dots, n^2 - 1, \\ 1 & \text{if } x \geq (a+b)_{Ln^2} \end{cases}$$

$$\underline{F}_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x < (a+b)_{U1} \\ \frac{h}{n^2} & \text{if } (a+b)_{Uh} \leq x < (a+b)_{U,h+1}, \quad h = 2, \dots, n^2 - 1. \\ 1 & \text{if } x \geq (a+b)_{Un^2} \end{cases}$$

where the terminals  $(a + b)_{Lh}$  and  $(a + b)_{Uh}$  have been sorted in ascending order with respect to  $h$ .

For perfect positive dependence, only intervals that correspond to identical fractiles of the cumulative probability are convolved via interval arithmetic. For both methods this results in intervals of the form  $[a_{Lj} + b_{Lj}, a_{Uj} + b_{Uj}]$  (for addition, but analogous equations hold for the other binary operations). Since identical probability masses accompany each interval of  $A$  and  $B$ , the resultant probability mass associated with  $[a_{Lj} + b_{Lj}, a_{Uj} + b_{Uj}]$  is  $\frac{1}{n}$ ,  $\forall j = 1, \dots, n$  for both dependency bounds convolution and DEnv. The bounding cdfs are then constructed as above, however, here

$$\bar{F}_{A+ind B}(x) = \begin{cases} 0 & \text{if } x < (a + b)_{L1} \\ \frac{h}{n} & \text{if } (a + b)_{Lh} \leq x < (a + b)_{L,h+1}, \quad h = 2, \dots, n-1 \\ 1 & \text{if } x \geq (a + b)_{Ln} \end{cases}$$

$$\underline{F}_{A+ind B}(x) = \begin{cases} 0 & \text{if } x < (a + b)_{U1} \\ \frac{h}{n} & \text{if } (a + b)_{Uh} \leq x < (a + b)_{U,h+1}, \quad h = 2, \dots, n-1 \\ 1 & \text{if } x \geq (a + b)_{Un} \end{cases}$$

where the terminals  $(a + b)_{Lh}$  and  $(a + b)_{Uh}$  have been sorted in ascending order with respect to  $h$ .

Since all intervals have identical probability mass, for perfect negative dependence interval arithmetic is applied to  $[a_{Lj}, a_{Uj}]$  and  $[b_{L_{n-j-1}}, b_{U_{n-j-1}}]$   $\forall j = 1, \dots, n$  and the associated probability mass for each resultant interval is  $\frac{1}{n}$ . The bounding cdfs are then constructed as (for addition but analogous construction holds for the other binary operations):

$$\overline{F}_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x < (a+b)_{L1} \\ \frac{h}{n} & \text{if } (a+b)_{Lh} \leq x < (a+b)_{L,h+1}, \quad h = 2, \dots, n-1 \\ 1 & \text{if } x \geq (a+b)_{Ln} \end{cases}$$

$$\underline{F}_{A+_{ind}B}(x) = \begin{cases} 0 & \text{if } x < (a+b)_{U1} \\ \frac{h}{n} & \text{if } (a+b)_{Uh} \leq x < (a+b)_{U,h+1}, \quad h = 2, \dots, n-1 \\ 1 & \text{if } x \geq (a+b)_{Un} \end{cases}$$

where the terminals  $(a+b)_{Lh}$  and  $(a+b)_{Uh}$  have been sorted in ascending order with respect to  $h$ .

For a proof of equivalence of the two methods under no assumptions of density dependence we appeal to the fact that both dependency bounds convolution and DEnv give point-wise best possible bounds on the resultant convolution. Point-wise best possible bounds, by definition, are unique, hence it follows that that both methods must give identical bounds when the deconstruction into intervals with associated probabilities for both methods is identical.  $\square$

So we see that Williamson and Downs' convolution of two random variables is equivalent to the thicket approach of convolving random variables in the sense that the discretized cdfs can be reduced to a set of intervals with associated probabilities. Indeed, the only differences between dependency bounds convolution and thickets are that the method proposed by Williamson and Downs requires cdfs to be discretized uniformly along the cumulative probability, whereas the discretization may be irregular in the case of thickets. Furthermore, for unknown dependencies, analytic methods are invoked to determine the resultant bounds in dependency bounds convolution, whereas linear programming methods are adopted in the case of thickets.

## Imprecise probabilities

In the two methods described above, uncertainty about the values of a random variable is represented as bounds on cdfs. In this section we will focus on uncertainty in the probability associated with the values of random variables, i.e. imprecise probabilities. For imprecise probabilities, bounds are assigned to the probability of an event in the form of upper and lower probabilities, i.e. intervals on probability values. The imprecision can reflect the quality of the information used to assign the probability value, indeterminate beliefs about events, indecision, group beliefs and decisions, and others [9].

The construction of an imprecise probability involves assigning upper and lower probabilities to distinct events. Upper or lower probabilities are real-valued functions ( $\bar{P}$  and  $\underline{P}$ , respectively) defined on a domain  $S$  (usually referred to as the sample space) which is an arbitrary subset of a universal set (or universe of discourse)  $\Omega$ . A probability measure  $P$  must satisfy the following conditions, sometimes known as the probability calculus, but more commonly referred to as Kolmogorov's axioms:

- i.  $0 \leq P(X_i) \leq 1, \quad \forall X_i \in S;$
- ii.  $P(S) = 1, \quad P(\emptyset) = 0$  where  $\emptyset$  denotes the empty set;
- iii.  $P(X_i \cup X_j) = P(X_i) + P(X_j), \quad \forall X_i, X_j \in S$  such that  $X_i \cap X_j = \emptyset$ .

Upper and lower probability measures  $\bar{P}$  and  $\underline{P}$ , on the other hand, have the basic properties (Walley 19??):

- a.  $0 \leq \underline{P}(X_i) \leq \bar{P}(X_i) \leq 1, \quad \forall X_i \in S;$
- b.  $\underline{P}(S) = \bar{P}(S) = 1, \quad \underline{P}(\emptyset) = \bar{P}(\emptyset) = 0;$

c.  $\overline{P}(X_i) + \underline{P}(X_i^c) = 1$  where  $X_i^c$  is the complement of  $X_i$ ;

d.  $\underline{P}(X_i) + \underline{P}(X_j) \leq \underline{P}(X_i \cup X_j) \leq \underline{P}(X_i) + \overline{P}(X_j) \leq \overline{P}(X_i \cup X_j) \leq \overline{P}(X_i) + \overline{P}(X_j)$ ,

$\forall X_i, X_j \in S$  such that  $X_i \cap X_j = \emptyset$

An imprecise probability is characterized as  $[\underline{P}(A), \overline{P}(A)]$ , i.e. it is an interval bounded below by the lower probability assignment  $\underline{P}$  and above by the upper probability assignment  $\overline{P}$ . It differs from the intervals constructed in the previous sections in that here, the intervals consist of probabilities, whereas the intervals in the previous sections consisted of values of the random variable coupled with an associated probability. Our task here is to convert a set of imprecise probabilities to a set of intervals of the random variable  $A$ , with associated probabilities on those intervals.

Here we restrict the universal set  $\Omega$  to the positive reals (in keeping with Theorems 1 to 3). Once upper and lower probabilities have been assigned for each event in the sample space we can then define upper and lower distribution functions of a random variable  $A$  with respect to the upper and lower probability functions. The upper and lower cumulative distribution functions  $\overline{F}_A(x)$  and  $\underline{F}_A(x)$  given  $\overline{P}(A)$  and  $\underline{P}(A)$  are defined as

$$\begin{aligned}\overline{F}_A(x) &= \overline{P}(A \leq x) = 1 - \underline{P}(A > x) \\ \underline{F}_A(x) &= \underline{P}(A \leq x) = 1 - \overline{P}(A > x)\end{aligned}\tag{15}$$

respectively, for all positive real values  $x$ . Hence, for any pair of upper and lower probability functions defined on  $\mathfrak{R}$ , there exist corresponding upper and lower cumulative distribution functions. Our analysis of imprecise probabilities now becomes an analysis of upper and lower cdfs. It is important to point out that converting a set of upper and lower probability assignments

into upper and lower cumulative distribution functions is information losing by virtue of property d) above. That is, the resultant bounding cdfs may not necessarily be deconstructed to give back the original imprecise probabilities from which the bounding cdfs were constructed. Imprecise probabilities resulting from a deconstruction of equations (15) will generally result in wider bounds than the original  $\bar{P}(A)$  and  $\underline{P}(A)$  that were used to construct equations (15). However, for our purposes it suffices to construct the bounding cdfs and then deconstruct these into a set of *intervals of the random variable* with associated probabilities, rather than the original *probability intervals* of connected subsets of the reals. For cases where  $\bar{F}_A(x)$ ,  $\underline{F}_A(x)$ ,  $\bar{F}_B(y)$  and  $\underline{F}_B(y)$  are step functions, as in Figures 4 and 5, deconstruction must be performed in such a way as to ensure equal probability mass assignments for each resultant interval. This may involve selecting a partition of the cumulative probability with a resolution fine enough that multiple copies of an interval results. For instance, if the deconstruction of probability bounds for random variable  $B$  resulted in a set of intervals with associated probability mass of 0.1 each, then for the probability bounds on  $A$  appearing in Figure 2, there would be two occurrences of each interval, each with associated probabilities 0.1.

Convolving two random variables  $A$  and  $B$  is now a matter of discretizing each cdf or each pair of bounds on a cdf and deconstructing into thickets via equations (6) (or equivalently, equations (12)) with equal steps along the cumulative probability axis. The methods described in the previous sections are then applied to calculate the resultant upper and lower cdfs for  $A*B$  under assumptions of independence, perfect positive or negative dependence or under no assumptions of dependency structure. So we see that for the special case of an  $\mathfrak{R}^+$ -valued domain, a reconstruction of imprecise probabilities into bounds on cdfs coincides with both

DEnv and dependency bounds convolution. Once a set of intervals with associated probabilities are obtained from the bounding cdfs for each random variable, it readily follows from Theorem 3 that applying either of the dependency bounds methods described above to a convolution of random variables will lead to the same results.

### Dempster-Shafer belief functions

Dempster-Shafer (D-S) belief functions are probabilities that are constructed from evidence [10]. The difference between traditional probability and Dempster-Shafer theory is that in the traditional probabilistic framework evidence is associated with a single event, whereas in Dempster-Shafer theory evidence can be associated with multiple possible events, or sets of events (Sentz and Ferson 2002). Dempster-Shafer theory relies on three main functions: the basic probability assignment (bpa), the belief function (denoted  $Bel$ ) and the plausibility function (denoted  $Pl$ ).

Suppose we have a universal set  $\Omega$ , the power set of which is denoted by  $\wp(\Omega)$ . A lower probability  $\underline{P}$ , defined on  $\wp(\Omega)$ , is a belief function (denoted  $Bel$ ) if and only if it can be written in the form

$$\underline{P}(A) = Bel(A) = \sum_{B \subset A} m(B) \quad \forall B \subset A \subseteq \Omega, \quad (16)$$

where  $m$  is a measure on  $\wp(\Omega)$  such that

- i)  $0 \leq m(B) \leq 1 \quad \forall B \subset \Omega$ ,
- ii)  $m(\emptyset) = 0$ ,
- iii)  $\sum_{B \subseteq \Omega} m(B) = 1$ .

The function  $m$  is called the basic probability assignment. A focal element is any subset  $B$  of  $\Omega$  such that  $m(B) \geq 0$ . The basic probability assignment is not a probability in the classical sense, in general (although the Dempster-Shafer theoretical framework can support this interpretation, it cannot be equated with classical probability in general). It represents the proportion of all relevant and available evidence that supports the claim that a particular element of  $\Omega$  belongs to the set  $B$ . For example, suppose there has been a theft and the suspects are  $\Omega = \{\text{Mary, John, Peter}\}$ . Further suppose that the probability the thief wears glasses is 0.6. If Mary is the only suspect that wears glasses then, in the Dempster-Shafer theory of evidence,  $m(\{\text{Mary}\}) = 0.6$ ,  $m(\Omega) = 0.4$ , and  $m(A) = 0$  for all other  $A \subset \Omega$ . The value  $m(B)$  relates to  $B$  only and does not imply any additional probability assignments for any other subsets of  $\Omega$  including subsets or complements of the set  $B$ . Hence the belief function  $Bel$  is the sum of the basic probability assignments for all proper subsets of the set of interest,  $A$ .

The upper probability can be defined in terms of its conjugate lower probability as

$$\bar{P}(A) = 1 - \underline{P}(A^c) = Pl(A) = \sum_{\substack{B \text{ such that} \\ A \cap B \neq \emptyset}} m(B) \quad \forall A, B \subseteq \Omega \quad (17)$$

where  $A^c$  is the complement of  $A$ . In Dempster-Shafer theory, the upper probability is referred to as the plausibility function (denoted  $Pl$ ). It is the sum of all the basic probability assignments for all sets that intersect the subset of interest,  $A$ .  $Bel(A)$  and  $Pl(A)$  may be viewed as lower and upper bounds on the actual probability  $P(A)$ . Table 4 displays the upper and lower probability assignments (i.e. degrees of plausibility and belief, respectively) for all  $A \subseteq \Omega$  in the example above.

Yager [18] shows how to construct upper and lower bounds on the cumulative bpa via the belief and plausibility functions when the domain is restricted to  $\mathfrak{R}$ . Let's suppose that the

universal set is the reals,  $\mathfrak{R}$ , and the bpa defined over all subsets of  $\mathfrak{R}$  is denoted by  $m$ . The focal elements of the set of all subsets of  $\mathfrak{R}$ , with respect to  $m$ , are denoted  $A_i, i = 1, \dots, n$ .

Furthermore, let  $G(x)$  denote the set of numbers less than or equal to  $x$ , i.e.

$$G(x) = \{ z : z \leq x, x \in \mathfrak{R} \}.$$

The lower and upper bounds on the cdf can be constructed as

$$\begin{aligned} \bar{F}(x) &= \bar{P}(G(x)) = Pl(G(x)) \\ &= \sum_i m(A_i) \cdot \max_{z \leq x} (I_{A_i}(z)) \end{aligned} \quad (18a)$$

$$\begin{aligned} \underline{F}(x) &= \underline{P}(G(x)) = Bel(G^c(x)) \\ &= \sum_i m(A_i) \cdot \left( 1 - \max_{z > x} (I_{A_i}(z)) \right) \end{aligned} \quad (18b)$$

where  $G^c(x)$  is the complement of  $G(x)$ , and the indicator function  $I_{A_i}$  takes the value 1 if  $z$  is in  $A_i$  and 0 otherwise. Hence, an unknown real variable  $A$ , the values of which are described by a Dempster-Shafer framework with basic probability assignment  $m_A$ , has a cdf,  $F_A$ , that lies between upper and lower bounds as constructed above. That is,

$$\underline{F}_A(x) \leq F_A(x) \leq \bar{F}_A(x).$$

Such a construction is generally information losing in the same way that a conversion of imprecise probabilities to a pair of bounding cdfs is information losing—the cumulative bounds cannot be deconstructed, in general, to give the original belief and plausibility functions. Usually the bounds resulting from a deconstruction of equations (18) will be wider than those used to construct (18) in the first place. However if the events  $A_i$  are intervals that form a partition of the real line then the cumulative bounds will be of the form in Figure 2 and a deconstruction of the

bounds to intervals with associated probability mass will give the original intervals  $A_i$  with respective basic probability assignments  $m_A(A_i)$ . This is straightforward from equations (18).

Suppose now that we have two Dempster-Shafer frameworks for two unknown real variables,  $A$  and  $B$ , with basic probability assignments  $m_A$  and  $m_B$  and focal elements denoted by  $A_i$  and  $B_j$ , respectively, where  $A_i$  and  $B_j$  are partitions of the positive real line. Further suppose we wish to convolve the variables  $A$  and  $B$  via a binary operation to form  $A*B$ . Yager's method for calculating the resultant probability assignment under an assumption of independence is

$$m(D) = \sum_{\substack{\forall i,j \\ A_i * B_j = D}} m_A(A_i) \times m_B(B_j)$$

where  $D = \{A_i * B_j, \forall i, j\}$ . The resultant probability assignment is also a Dempster-Shafer framework for  $A*B$  since  $m: D \rightarrow [0, 1]$  satisfies i) to iii) above [18]. Hence cumulative bounds can be constructed according to equations (18). Note that if  $A_i$  and  $B_j$  are intervals then  $D$  is the result of interval arithmetic on the two intervals. The associated basic probability assignment is then simply calculated as the product of the two respective bpas. This is identical to the DEnv determination on thicket and Williamson and Downs algorithm for convolving independent random variables.

In order to see that Yager's method for convolving two random variables is equivalent to the method described by Berleant and Goodman-Strauss [7] and Williamson and Downs [8] for independent random variables, we simply need to decompose the upper and lower bounds on the cumulative distribution function (in equations 18) into a collection of intervals and probabilities and show that, for the variable  $A$ , say, the thicket consists precisely of the intervals  $A_i$  with associated probabilities  $m_A(A_i)$ . As explained above, this is indeed the case when  $A_i$  forms a partition of the real line.

Let's suppose that  $A_i = [a_{L_i}, a_{U_i}]$  and that  $a_{L_i} \leq a_{L_{i+1}}$  and  $a_{U_i} \leq a_{U_{i+1}}$  and  $a_{L_{i+1}} = a_{U_i}$ ,

$\forall i = 1, \dots, n$ , that is, that  $A_i$  forms a partition of the positive real line. The upper and lower

bounds on the cdf can be rewritten as

$$\bar{F}_A(x) = \begin{cases} 0 & \text{for } x < a_{L_1} \\ \sum_{j=1}^i m_A(A_j) \cdot \max_{z \leq x} (I_{A_i}(z)) & \text{for } a_{L_i} \leq x < a_{L_{i+1}}, \quad i = 1, \dots, n-1, \\ 1 & \text{for } x \geq a_{L_n} \end{cases} \quad (19a)$$

$$\underline{F}_A(x) = \begin{cases} 0 & \text{for } x \leq a_{U_1} \\ \sum_{j=1}^i m_A(A_j) \cdot \left(1 - \max_{z > x} (I_{A_i}(z))\right) & \text{for } a_{U_i} < x \leq a_{U_{i+1}}, \quad i = 1, \dots, n-1, \\ 1 & \text{for } x > a_{U_n} \end{cases} \quad (19b)$$

Applying the decomposition formula in equation (6) to the inverses of the upper and lower

bounds on the cdf in equations (19a) and (19b) we obtain  $\{P(A \in A_i) = m_A(A_i) : i = 1, \dots, n\}$ , that

is, the set of intervals  $A_i$  with associated probabilities  $m_A(A_i)$ . A similar deconstruction works

for the upper and lower bounds on the cdf describing the random variable  $B$ . It is now a simple

matter to convolve the two variables as  $A*B$ . Hence, the convolution of independent random

variables described by a Dempster-Shafer framework is identical to that of thickets and

dependency bounds convolution when  $A_i$  and  $B_j$  form a partition of the positive real line.

Furthermore, the formalism of DEnv on thickets allows us to extend Yager's method [18] of

convolution to the case where the dependency structure is perfectly positive or negative, or

unknown in the manner described in the sections above. The fact that we can construct

cumulative bounds on a Dempster-Shafer framework and then decompose these into a set of

intervals with associated bpas that coincide with the original D-S framework allows us to do this.

It is not clear, however, that a D-S framework is obtained for the resulting set of intervals and bpas under assumptions of perfect positive or negative dependence or unknown dependence when DEnv determination is used to convolve the two random variables. The conditions under which a D-S framework is obtained for convolutions under dependency assumptions other than independence are yet to be proved. The cumulative bounds resulting from convolutions under assumptions of unknown dependence may not be best possible for a D-S structure, they may actually be wider. It is also not clear that our conclusions for the D-S framework hold for intervals  $A_i$  that are overlapping or disconnected. Our conclusions also do not necessarily allow for the combination of evidence once a convolution under unknown dependency has been performed. While our conclusions only hold for a very restricted set of D-S frameworks and may lose much of the richness of D-S structures, we do show that dependency bounds convolution can be applied to D-S structures by virtue of the fact that cumulative bounds can be constructed on a precise cdf. Conditions under which a D-S structure can be obtained upon dependency bounds convolution for overlapping intervals of the real line is a worthwhile pursuit and currently under investigation. An important problem that remains to be solved is the determination of the full suite of conditions under which D-S structures can be convolved under dependency bounds convolution or DEnv to give resulting D-S structures.

## **Conclusion**

In this paper we have reviewed four probabilistic methods for the reliable propagation of uncertainty. We showed that each method can be reformulated and/or extended so that all methods are essentially the same for the restricted domain of the positive reals  $\mathfrak{R}^+$ , and with additional restrictions for Dempster-Shafer frameworks.

The convergence of the methods to equivalent formulations for the domain  $\mathfrak{R}^+$  now makes it possible to do calculations with Dempster-Shafer belief structures and imprecise probabilities under a variety of dependency assumptions (but see caveats above for D-S structures). These include assumptions of independence, perfect positive and negative dependence, and unknown dependency. The software RAMAS Risk Calc [15], originally designed for dependency bounds convolution, can thus be used for any of the approaches described here.

Our results enable reliable propagation of uncertainty through calculations when very little is known about the dependency structure, and require minimal knowledge (or assumptions) about the underlying true probability distribution. The treatment here extends the flexibility and utility of these methods to problems beyond those with simple and often unrealistic assumptions of independence, resulting in the reliable propagation of uncertainty through calculations.

### **Acknowledgements**

The authors wish to thank Lev Ginzburg and David Myers for valuable comments and discussion. This work has been much improved by the comments of Jim Hall, Arnold Neumaier and one anonymous reviewer. This work was supported by a National Cancer Institute grant to Applied Biomathematics (9R44CA81741) and by a Power Systems Engineering Research Center grant to Iowa State University. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Cancer Institute.

## References

- [1] Ferson, S. Naïve Monte Carlo methods yield dangerous underestimates of tail probabilities, *Proceedings of the High Consequence Operations Safety Symposium*, Sandia National Laboratories, SAND94-2364, J.A. Cooper (Ed.), 507-514, 1994.
- [2] Ferson, S. and Burgman, M. Correlations, dependency bounds and extinction risk, *Biological Conservation* 73:101-105, 1995.
- [3] Ferson, S. What Monte Carlo methods cannot do, *Human and Ecological Risk Assessment* 2:990-1007, 1996.
- [4] Bukowski, J., L. Korn and D. Wartenberg. Correlated inputs in quantitative risk assessment: the effects of distributional shape. *Risk Analysis* 15:215-219, 1995.
- [5] Burmaster, D.E., K.J. Lloyd, and K.M. Thompson. The need for new methods to backcalculate soil cleanup targets in interval and probabilistic cancer risk assessments. *Human and Ecological Risk Assessment* 1:89-100, 1995.
- [6] Smith, A.N. and I.A. Watson. Common cause failure—a dilemma in perspective. *Reliability Engineering* 1:127-142, 1980.
- [7] Berleant, D. and Goodman-Strauss, C. Bounding the results of arithmetic operations on random variables of unknown dependency using intervals, *Reliable Computing* 4(2):147-165, 1998.
- [8] Williamson, R. and Downs, T. Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds, *International Journal of Approximate Reasoning*, 4:89-158, 1990.
- [9] Walley, P. *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London, 1991.

[10] Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, 1976.

[11] Ferson, S. and Ginzburg, L. Hybrid arithmetic, *Proceedings of the 1995 Joint ISUMA/NAFIPS Symposium on Uncertainty Modeling and Analysis*, IEEE Computer Society Press, Los Alamitos, California, 619–623, 1995.

#####I've deleted references 12 and 13 – DB

[14] Karloff, H. *Linear Programming*, Birkhauser, Boston, 1991.

[15] Ferson, S., Root, W. and Kuhn, R. *RAMAS Risk Calc: Risk Assessment with Uncertain Numbers*, Applied Biomathematics, Setauket, New York, 1999.

[16] Frank, M.J., Nelson, R.B., and Schweizer, B. Best-possible bounds for the distribution of a sum—a problem of Kolmogorov, *Probability Theory and Related Fields* 74:199–211, 1987.

[17] Nelson, R.B. *An Introduction to Copulas*, Lecture Notes in Statistics Vol. 139, Springer-Verlag, New York, 1999.

[18] Yager, R.R. Arithmetic and other operations on Dempster-Shafer structures, *International Journal of Man-Machine Studies*, 25:357–366, 1986.

[19] Springer, M. D., *The Algebra of Random Variables*, John Wiley & Sons, 1979.

[20] Alefeld, G. and J. Herzberger, *Introduction to Interval Computations*, Academic Press, 1983.

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

Regan HM, Colyvan M, Burgman MA. 2002a. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol Applications* 12(2):618–28

Sklar, A. Functions de  $\square$ epartition à  $n$  dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris, 8:229-231.

Regan, H.M., B.K. Hope, S. Ferson. Portrayal and analysis of uncertainty in a food-web exposure model. *Human and Ecological Risk Assessment*. Citation, 2002.

Regan, H.M., H.R. Akçakaya, S. Ferson, K.V. Root, S. Carroll and L.R. Ginzburg. Treatments of uncertainty and variability in ecological risk assessment of single-species populations. *Human and Ecological Risk Assessment*, (in press).

Table 1. Table of intervals with their associated probabilities (a “joint distribution matrix”) describing  $A +_{ind} B$ , the sum of distributions  $A$  and  $B$  under the assumption of independence. The interior intervals (**bolded**) are the result of interval addition on the pairs  $A_i$  and  $B_j$  and the probability associated with each interior interval is calculated as  $p_i \times q_j$  due to the assumption of independence of  $A$  and  $B$ .

$A +_{ind} B$ $x : x \in A \Rightarrow$ $y : y \in B \Downarrow$	$x \in A_1, A_1 = [1, 2]$ $p_1 = 1/2$	$x \in A_2, A_2 = [2, 4]$ $p_2 = 1/2$
$y \in B_1, B_1 = [2, 3]$ $q_1 = 1/4$	$x + y \in [3, 5]$ $p_{11} = 1/8$	$x + y \in [4, 7]$ $p_{21} = 1/8$
$y \in B_2, B_2 = [3, 4]$ $q_2 = 1/2$	$x + y \in [4, 6]$ $p_{12} = 1/4$	$x + y \in [5, 8]$ $p_{22} = 1/4$
$y \in B_3, B_3 = [4, 5]$ $q_3 = 1/4$	$x + y \in [5, 7]$ $p_{13} = 1/8$	$x + y \in [6, 9]$ $p_{23} = 1/8$

Table 2. Table of the intervals and associated probabilities that define the distribution function of  $A + B$  under assumptions of perfect positive and perfect negative dependence. The intervals are the result of interval addition on the pairs  $A_i$  and  $B_j$ . The probabilities  $p_{ij+}$  and  $p_{ij-}$  refer to the joint distribution of  $A$  and  $B$  under the assumptions of perfect positive and perfect negative dependence, respectively.

$A+B$ $b \downarrow$	$x \rightarrow$	$x \in [1, 2]$ $p_1 = \frac{1}{2}$	$x \in [2, 4]$ $p_2 = \frac{1}{2}$
$y \in [2, 3]$ $q_1 = \frac{1}{4}$		$x + y \in [3, 5]$ $p_{11+} = \frac{1}{4}$ $p_{11-} = 0$	$x + y \in [4, 7]$ $p_{21+} = 0$ $p_{21-} = \frac{1}{4}$
$y \in [3, 4]$ $q_2 = \frac{1}{2}$		$x + y \in [4, 6]$ $p_{12+} = \frac{1}{4}$ $p_{12-} = \frac{1}{4}$	$x + y \in [5, 8]$ $p_{22+} = \frac{1}{4}$ $p_{22-} = \frac{1}{4}$
$y \in [4, 5]$ $q_3 = \frac{1}{4}$		$x + y \in [5, 7]$ $p_{13+} = 0$ $p_{13-} = \frac{1}{4}$	$x + y \in [6, 9]$ $p_{23+} = \frac{1}{4}$ $p_{23-} = 0$

Table 3.

$A+B$	$x \rightarrow$	$x \in [1, 2]$	$x \in [2, 4]$
$y \downarrow$		$p_1 = \frac{1}{2}$	$p_2 = \frac{1}{2}$
$y \in [2, 3]$ $q_1 = \frac{1}{4}$		$x+y \in [3, 5]$ $p_{11} = 0$	$x+y \in [4, 7]$ $p_{21} = \frac{1}{4}$
$y \in [3, 4]$ $q_2 = \frac{1}{2}$		$x+y \in [4, 6]$ $p_{12} = \frac{1}{2}$	$x+y \in [5, 8]$ $p_{22} = 0$
$y \in [4, 5]$ $q_3 = \frac{1}{4}$		$x+y \in [5, 7]$ $p_{13} = 0$	$x+y \in [6, 9]$ $p_{23} = \frac{1}{4}$

Table 4. Basic probability assignments (b.p.a.s), and degrees of belief and plausibility (i.e. lower and upper probabilities) as assigned in the Dempster-Shafer theory of evidence framework for  $\Omega = \{M, J, P\}$ ,  $m(\{M\}) = 0.6$ ,  $m(\Omega) = 0.4$  and  $m(A) = 0$  for all other  $A \subset \Omega$ .

Subsets of $\Omega = \{M, J, P\}$	$\emptyset$	$\{M\}$	$\{J\}$	$\{P\}$	$\{M, J\}$	$\{M, P\}$	$\{J, P\}$	$\{M, J, P\}$
$m(A)$	0	0.6	0	0	0	0	0	0.4
$\text{Bel}(B)$	0	0.6	0	0	0.6	0.6	0	1
$\text{Plaus}(B)$	0	1	0.4	0.4	1	1	0.4	1

Figure 1. Upper ( $\bar{F}_A(x)$ ) and lower ( $\underline{F}_A(x)$ ) bounds on a cumulative distribution function  $F_A(x)$  for real-valued random variable  $A$ . While  $\bar{F}_A(x)$  and  $\underline{F}_A(x)$  form the upper and lower bounds on the cumulative distribution function, respectively, their inverses form lower and upper bounds, respectively, on subsets of the random variable.

Figure 2. Discretization of cumulative probability distribution. The function  $F_A(x)$  is discretized into two functions, a left bound  $\bar{F}_A(x)$ , and a right bound  $\underline{F}_A(x)$ , both with  $n$  steps of equal height.

Figure 2. Upper and lower bounds on the cumulative distribution functions for  $A+B$ . The upper bound on the distribution (the left-most curve) is obtained by concentrating the probability mass associated with each interval  $A_i + B_j$  at the lower bound on the interval and then integrating over the lower bounds on all intervals. The lower bound on the distribution (i.e. the right-most curve) is obtained analogously.

Figure 3. A probability distribution function discretized with a histogram. The height of the histogram bars refers to the probability mass under the pdf for each respective interval. 3a) shows a histogram for a partition of the real line. 3b) shows a histogram for overlapping intervals.

Figure 4. Bounds for the cumulative distribution of  $A +_{ind} B$ , constructed from Table 1. The bounds are formed by “stacking” each interval from lowest to highest and cumulating the associated probabilities.

Figure 5. Diagram of decomposition of probability bounds,  $\overline{F}_A(x)$  and  $\underline{F}_A(x)$ , into a thicket. The values along the vertical axis correspond to the probability mass associated with each interval resulting from the decomposition. The values along the horizontal axis provide the upper and lower bounds on each interval. For example, the interval  $A_1 = [\overline{F}_A^{(-1)}(s_1), \underline{F}_A^{(-1)}(s_1)] = [\overline{F}_A^{(-1)}(r_1), \underline{F}_A^{(-1)}(r_1)]$  has associated probability mass  $r_1$ , the interval  $A_2 = [\overline{F}_A^{(-1)}(s_2), \underline{F}_A^{(-1)}(s_2)] = [\overline{F}_A^{(-1)}(q_1), \underline{F}_A^{(-1)}(q_1)]$  has associated probability mass  $q_1 - r_1$  etc.

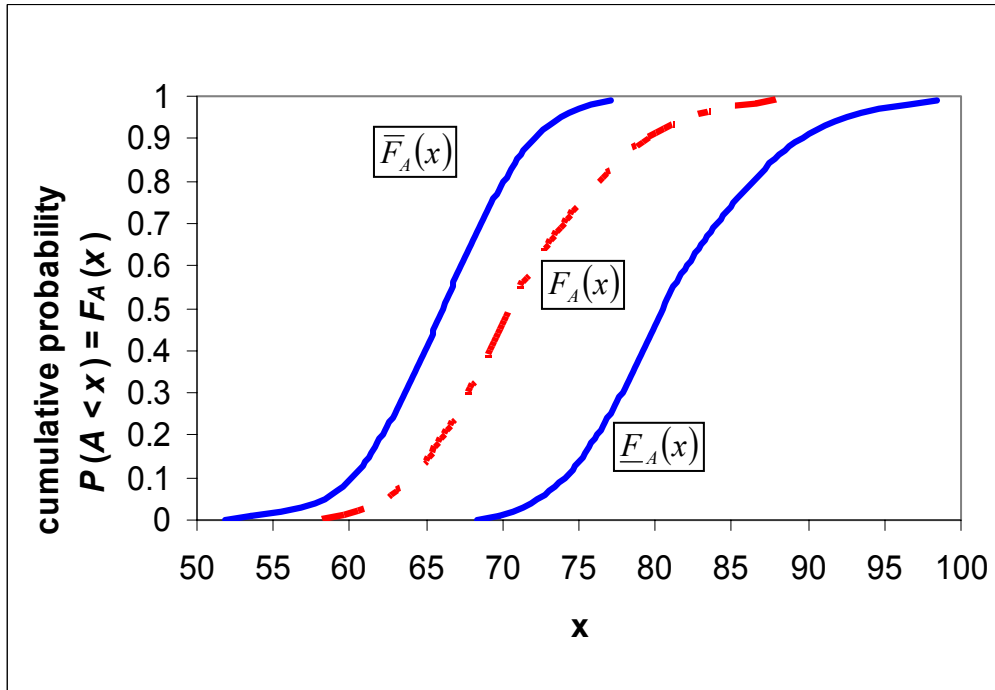


Figure 1.

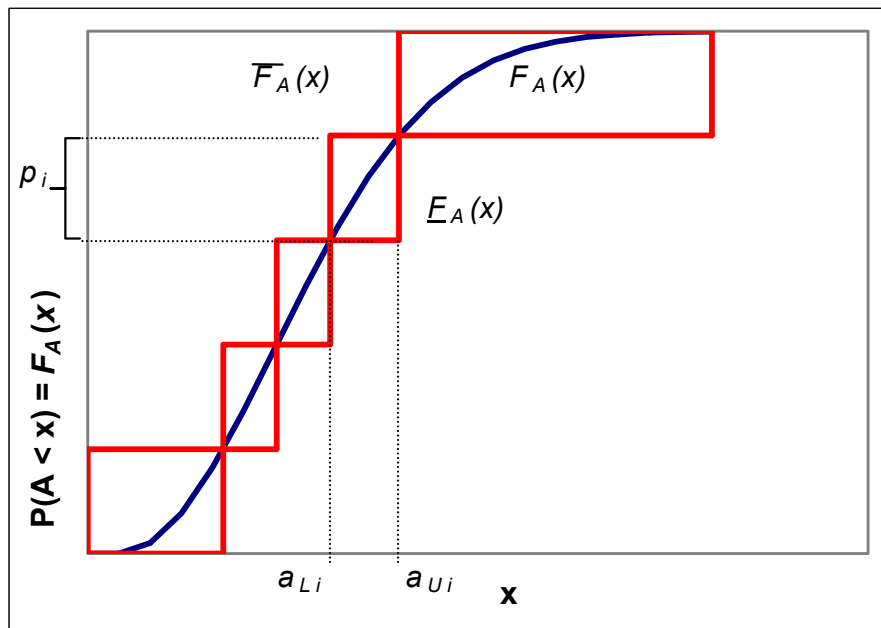
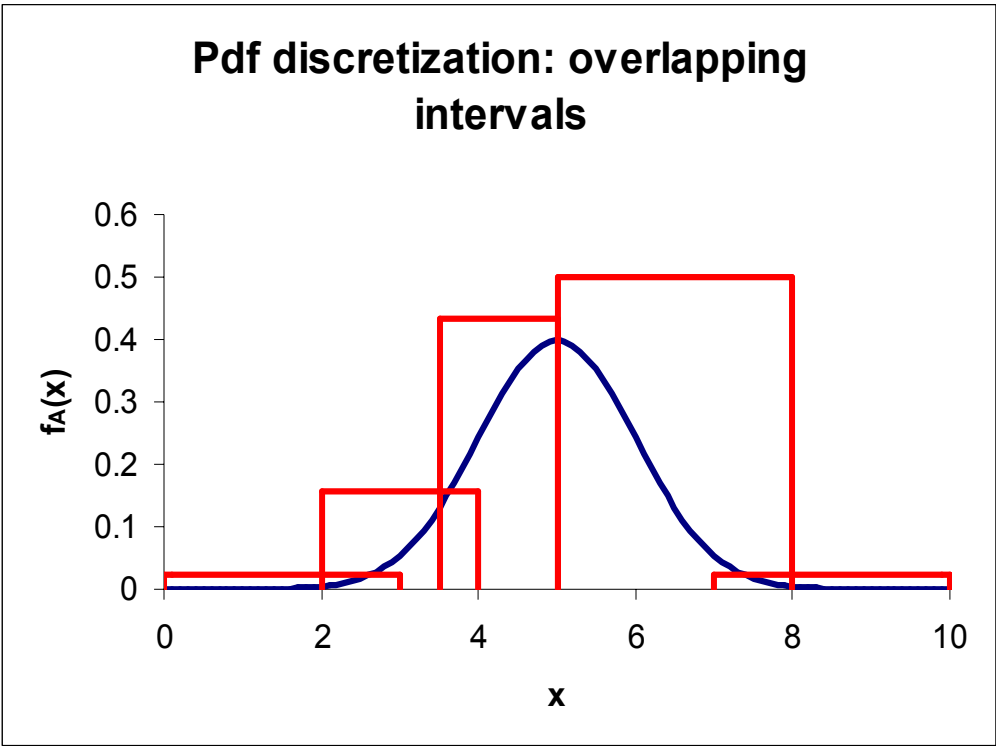
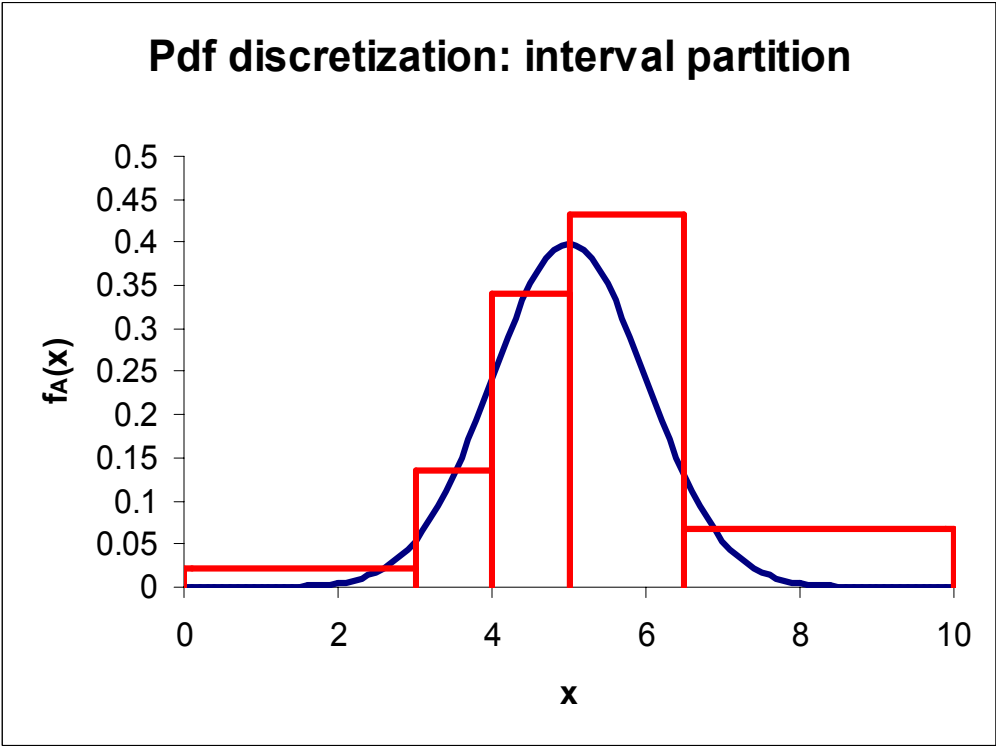


Figure 2



Figures 3a and 3b

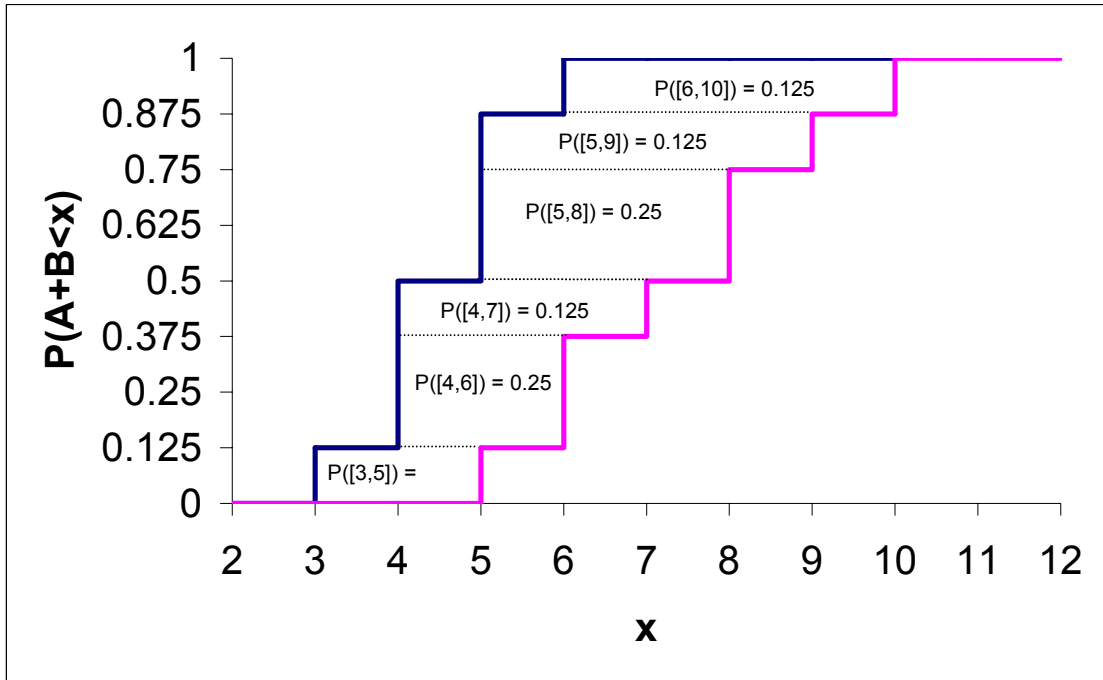


Figure 4

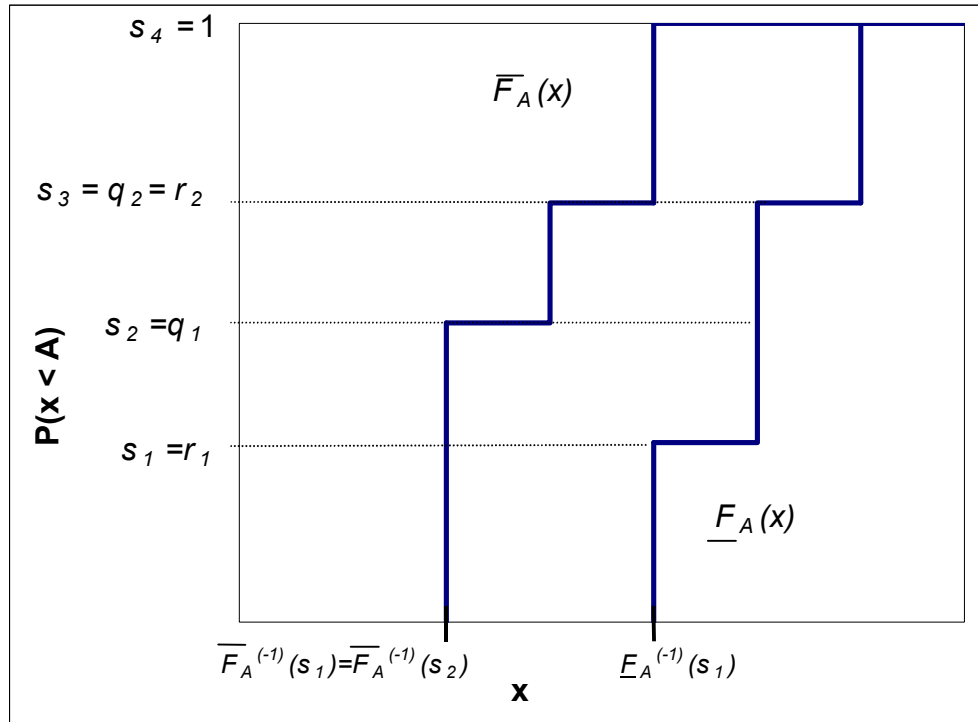


Figure 5