UKAEA-CCFE-PR(22)58

Ander Gray, Scott Ferson, Edoardo Patelli

Distribution-Free Risk Analysis

Distribution-Free Risk Analysis

Ander Gray, Scott Ferson, Edoardo Patelli

This is a preprint of a paper submitted for publication in
International Journal of Approximate Reasoning

DISTRIBUTION-FREE RISK ANALYSIS

A PREPRINT

Ander Gray
Institute for Risk and Uncertainty
University of Liverpool
Ander.Gray@liverpool.ac.uk

Scott Ferson
Institute for Risk and Uncertainty
University of Liverpool
Scott.Ferson@liverpool.ac.uk

Edoardo Patelli*
Civil and Environmental Engineering
University of Strathclyde
Edoardo.Patelli@strath.ac.uk

January 17, 2022

ABSTRACT

Elementary formulas for propagating information about means and variances through mathematical expressions have long been used by analysts. Yet the precise implications of such information are rarely articulated. This paper explores distribution-free techniques for risk analysis that do not require simulation, sampling or approximation of any kind. We describe best-possible bounds on risks that can be inferred given only information about the range, mean and variance of a random variable. These bounds generalise the classical Chebyshev inequality in an obvious way. We also collect in convenient tables several formulas for propagating range and moment information through calculations involving 7 binary convolutions (addition, subtraction, multiplication, division, powers, minimum, and maximum) and 9 unary transformations (scalar multiplication, scalar translation, exponentiation, natural and common logarithms, reciprocal, square, square root and absolute value) commonly encountered in risk expressions. These formulas are rigorous rather than approximate, and in most cases are either exact or mathematically best possible. The formulas can be used effectively even when only interval estimates of the moments are available. Although most discussions of moment propagation assume stochastic independence among variables, this paper shows the assumption to be unnecessary and generalises formulas for the case when no assumptions are made about dependence.

Keywords uncertainty propagation, moment propagation, distribution-free risk analysis, probability box, dependence, interval arithmetic

1 Introduction

Many authors have suggested propagating means and variances of variables through mathematical expressions as a crude form of risk analysis. This approach is sometimes called first-order error analysis, and it is a widely used approach for making risk estimates. In traditional probability theory, these calculations are called moment propagation and are considered a fundamental part of mathematical statistics (for example, see Wilks (1962)). Despite this wide use, there has always been a disconnect between moment propagation and what these calculations would imply about risks of extreme values of the variable. For instance, after reviewing some moment propagation formulas, Cullen and Frey (1999, page 184) gave a rather pessimistic conclusion:

> Although the results of [the formulas] are useful in some cases for propagating the mean and variance trough a simple linear model, they do not imply anything about the shape of the model output distribution. Thus, if we were interested in making predictions regarding the 95th percentile of the model output for a linear function of independent random variables, we would not have sufficient information based solely on the properties of the mean and variance to do this.

Their pessimistic view is based on the fact that the well-known formulas for moment propagation

- require stochastic independence,

*Corresponding author: Edoardo.Patelli@strath.ac.uk

- require moments to be perfectly known (point values), and
- give no information about output distributions without assumptions (e.g. normality).

In this paper, we suggest that one can combine the methods of moment propagation with elementary interval analysis to obtain results that are better than can be obtained from either analysis separately. Rowe (1988) considered the problem of computing moments of certain kinds of transformations such as exp, log, sqrt, etc. from sparse structural information such as first moments and ranges of the operands. We extend this approach to the context of convolutions between poorly characterised random variables, and provide formulae for moment propagation which require no assumptions about stochastic dependence. Rowe's methods, together with the present extension, creates what may be characterised as a *distribution-free* risk analysis that lets analysts compute bounds on uncertain expressions without making assumptions about the precise distributions of the underlying variables. We also show that information about moments actually does enable us to make rigorous conclusions about the shape and, indeed, the percentiles of the output distributions that will be useful in many real-world risk assessments (contra Cullen and Frey 1999, page 184).

2 Means and variances always 'exist'

Mathematically, the distribution of a random variable may fail to have a mean or variance. For instance, Student's t distribution with two degrees of freedom theoretically has no variance because its formula does not converge to a finite value. Similarly, the quotient of independent unit normals, which follows a Cauchy distribution, has neither a variance nor mean. Wiwatanadate and Claycamp (2000) suggested that a risk calculation based on simple formulas for means and variances can only be applied in situations where the moments all exist.

As a practical matter, however, we do not consider the nonexistence of moments to be of any real significance for risk analysts. Infinite means and variances are merely mathematical bêtes noires that need not concern the practically minded. All random variables relevant to real-world risk analyses come from bounded distributions. As an example, consider human body weight. There are no infinitely massive body weights (despite recent trends in western dietary health). The largest recorded human body weight was 635 kg. Although a person could probably exceed this weight, perhaps even substantially, there are clearly bounds that human body mass cannot exceed. Therefore, as a practical matter, even a very comprehensive risk analysis need never include a mathematically infinite distribution for body weight. Similar arguments apply to other variables. Analysts concerned with infinite tails of distributions are addressing mathematical problems, not risk analysis problems. All the moments of any bounded distribution are finite and therefore 'exist' in the mathematical sense.

On the other hand, just because the moments are finite, does not imply they are determinate. In fact, it may usually be the case that only an indeterminate estimate of a mean or variance is available. In such situations, we can use intervals to represent the value, whatever it is, in some range. We can then use interval arithmetic (Moore, 1966) to manipulate the estimate and propagate it through calculations even though we cannot specify its value precisely.

3 Propagating range and moment information

In this section, we review formulas for bounds on the range and first two moments (mean and variance) for imprecisely specified random variables. Bounds are considered "rigorous" or "true" bounds if they are certain to contain the value (given the assumptions). All of the formulas in the tables in this paper are rigorous, so the true moments are guaranteed to be inside the given bounds so long as the inputs are within their respective bounds. This means that none of the table entries is merely approximate. Bounds are considered "best possible" if they cannot be any tighter. If a formula in the table is exact or best possible, it is displayed in boldface. Most of the other formulas yield fairly narrow results and are still quite good for practical purposes even though they may not be mathematically best possible.

Table 1 summarises formulas that can be used to estimate the least and greatest possible value of a distribution arising from a transformation or convolution. In this and the following tables, X and Y are two random numbers and k is an arbitrary constant. $\underline{X}$ and $\overline{X}$ denote respectively the least and greatest possible value of X. EX denotes the expectation or mean of X, and VX denotes its variance. Following Rowe (1988), we define the variance with a denominator of n instead of $n-1$, and emphasise that the quantities under consideration are moments of finite data populations, which are not necessarily samples of anything. In other respects, the random variables are arbitrary except for restrictions implied by the mathematical operations. For instance, the entries in the square root rows assume X cannot take on negative values, and the rows for division assume that the random variable Y does not straddle zero.

The formulas in Table 1 are essentially a synopsis of standard interval arithmetic (Moore, 1966) and, apart from the row for subtraction perhaps, are probably not very surprising. Monotone increasing transformations are especially easy, because the endpoint of the transformation is just the transformation of the endpoint. For instance, the least possible

Table 1: Rigorous formulas for least and greatest possible values of 9 transformations and 7 convolutions of random variables (all the formulations in this table are mathematically best-possible).

	Least possible value	*Greatest possible value*
$k+X$ (shifting)	$\mathbf{k}+\underline{\mathbf{X}}$	$\mathbf{k}+\overline{\mathbf{X}}$
kX (rescaling)	$\begin{cases}\mathbf{k}\underline{\mathbf{X}}, & \text{if } 0 \le k \\ \mathbf{k}\overline{\mathbf{X}}, & \text{if } k < 0\end{cases}$	$\begin{cases}\mathbf{k}\overline{\mathbf{X}}, & \text{if } 0 \le k \\ \mathbf{k}\underline{\mathbf{X}}, & \text{if } k < 0\end{cases}$
e^X	$\mathbf{e}^{\underline{\mathbf{X}}}$	$\mathbf{e}^{\overline{\mathbf{X}}}$
$\ln(X)$ for $0 < X$	$\mathbf{ln}(\underline{\mathbf{X}})$	$\mathbf{ln}(\overline{\mathbf{X}})$
$\log_{10}(X)$ for $0 < X$	$\mathbf{log_{10}}(\underline{\mathbf{X}})$	$\mathbf{log_{10}}(\overline{\mathbf{X}})$
$\frac{1}{X}$ for $0 \notin X$	$\mathbf{1}/\overline{\mathbf{X}}$	$\mathbf{1}/\underline{\mathbf{X}}$
X^2	$\begin{cases}\mathbf{0}, & \text{if } 0 \in X \\ \mathbf{min}(\underline{\mathbf{X}}^{\mathbf{2}}, \overline{\mathbf{X}}^{\mathbf{2}}), & \text{otherwise}\end{cases}$	$\mathbf{max}(\underline{\mathbf{X}}^{\mathbf{2}}, \overline{\mathbf{X}}^{\mathbf{2}})$
$\lvert X\rvert$ (absolute value)	$\begin{cases}\mathbf{0}, & \text{if } 0 \in X \\ \mathbf{min}(\lvert\underline{\mathbf{X}}\rvert, \lvert\overline{\mathbf{X}}\rvert), & \text{otherwise}\end{cases}$	$\mathbf{max}(\lvert\underline{\mathbf{X}}\rvert, \lvert\overline{\mathbf{X}}\rvert)$
$\sqrt{X}$ for $0 \le X$	$\sqrt{\underline{\mathbf{X}}}$	$\sqrt{\overline{\mathbf{X}}}$
$X+Y$	$\underline{\mathbf{X}}+\underline{\mathbf{Y}}$	$\overline{\mathbf{X}}+\overline{\mathbf{Y}}$
$X-Y$	$\underline{\mathbf{X}}-\overline{\mathbf{Y}}$	$\overline{\mathbf{X}}-\underline{\mathbf{Y}}$
$X \times Y$	$\mathbf{min}(\underline{\mathbf{X}}\underline{\mathbf{Y}}, \underline{\mathbf{X}}\overline{\mathbf{Y}}, \overline{\mathbf{X}}\underline{\mathbf{Y}}, \overline{\mathbf{X}}\overline{\mathbf{Y}})$	$\mathbf{max}(\underline{\mathbf{X}}\underline{\mathbf{Y}}, \underline{\mathbf{X}}\overline{\mathbf{Y}}, \overline{\mathbf{X}}\underline{\mathbf{Y}}, \overline{\mathbf{X}}\overline{\mathbf{Y}})$
$\frac{X}{Y}$ for $0 \notin Y$	$\mathbf{min}(\underline{\mathbf{X}}/\underline{\mathbf{Y}}, \underline{\mathbf{X}}/\overline{\mathbf{Y}}, \overline{\mathbf{X}}/\underline{\mathbf{Y}}, \overline{\mathbf{X}}/\overline{\mathbf{Y}})$	$\mathbf{max}(\underline{\mathbf{X}}/\underline{\mathbf{Y}}, \underline{\mathbf{X}}/\overline{\mathbf{Y}}, \overline{\mathbf{X}}/\underline{\mathbf{Y}}, \overline{\mathbf{X}}/\overline{\mathbf{Y}})$
X^Y for $0 < X$ or $0 < Y$	$\mathbf{min}(\underline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \underline{\mathbf{X}}^{\overline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\overline{\mathbf{Y}}})$	$\mathbf{max}(\underline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \underline{\mathbf{X}}^{\overline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\overline{\mathbf{Y}}})$
$\min(X,Y)$	$\mathbf{min}(\underline{\mathbf{X}}, \underline{\mathbf{Y}})$	$\mathbf{min}(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$
$\max(X,Y)$	$\mathbf{max}(\underline{\mathbf{X}}, \underline{\mathbf{Y}})$	$\mathbf{max}(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$

value of the square root of some variable is simply the square root of the least possible value of the variable. The relevant endpoints are reversed for monotone decreasing transformations. For instance, the greatest possible value of the reciprocal of some variable is the reciprocal of its least possible value. Non-monotone functions, such as absolute value, are more troublesome to account for because values inside the range of the variable can play a role in determining the endpoints of the transformation of the variable. For instance, the least possible value of the absolute value of some variable that ranges between $+2$ and -2 is zero (which is neither endpoint).

The formulas in Table 2 review the basic arithmetic operations on moments without dependence assumptions. These formulas generally yield intervals rather than precise values. In part, the results are indeterminate because we are not specifying the stochastic dependence between the random variables X and Y (this reason is reflected in the occasional appearance of the $\pm$ operator in the table). This indeterminism would be present even if the estimates of means and variances used as inputs were precise. But, of course, these inputs may well start out as intervals, perhaps because they were previously computed using the tabulated formulas or because they were imprecisely estimated from statistical data or by subjective judgement.

Some of these formulas, such as those in the first two rows, are elementary and can be found in any textbook on mathematical statistics (e.g. Wilks (1962)). Rowe (1988) describes several bounds on transformations of random variables that have constant-sign derivatives, including exponentiation, logarithms, reciprocal, square and square root. Rowe showed how to make use of information about the minimum and maximum values to obtain surprisingly tight bounds on the mean and variance with simple closed-form expressions. These expressions do not require approximation and are extremely fast when implemented on a computer. In the table, we use rowe (Rowe's mean estimate) and rowevar (Rowe's variance estimate) to denote his functional templates:

$$\text{rowe}(t) = \text{env}(Pt(\underline{X}) + (1-P)t(EX + \frac{VX}{EX - \underline{X}}), Qt(\overline{X}) + (1-Q)t(EX + \frac{VX}{EX - \overline{X}})) \tag{1}$$

$$\text{rowevar}(t) = \text{env}(\frac{t(\underline{\nu}) - t(\underline{X})}{(\underline{\nu} - \underline{X})^2}(VX + (\underline{\nu} - EX)^2), \frac{t(\overline{\nu}) - t(\overline{X})}{(\overline{\nu} - \overline{X})^2}(VX + (\overline{\nu} - EX)^2)) \tag{2}$$

Table 2: Rigorous formulas for the mean and variance for 9 transformations and 7 convolutions of random variables (best-possible formulations in boldface).

	Mean	*Variance*
$k + X$ (shifting)	$\mathbf{k + EX}$	$\mathbf{VX}$
kX (rescaling)	$\mathbf{kEX}$	$\mathbf{k^2VX}$
e^X	rowe(exp)	rowevar(exp)
$\ln(X)$ for $0 < X$	rowe(ln)	rowevar(ln)
$\log_{10}(X)$ for $0 < X$	rowe($\log_{10}$)	rowevar($\log_{10}$)
$\frac{1}{X}$ for $0 \notin X$	rowe(reciprocal)	rowevar(reciprocal)
X^2	$(\mathbf{EX})^\mathbf{2} + \mathbf{VX}$	rowevar(square)
$\|X\|$ (absolute value)	$\begin{cases} \mathbf{EX}, & \text{if } 0 \leq \underline{X} \\ -\mathbf{EX}, & \text{if } \overline{X} \leq 0 \\ [\|EX\|, \|EX\| + \sqrt{VX}(\pi - \text{atan}(\frac{\|EX\|}{\sqrt{VX}}))], & \text{if } 0 \in X \end{cases}$	$\max(0, EX^2 + VX - E(\|X\|)^2)$
$\sqrt{X}$ for $0 \leq X$	rowe($\surd$)	rowevar($\surd$)
$X + Y$	$\mathbf{EX + EY}$	$(\sqrt{\mathbf{VX}} \pm \sqrt{\mathbf{VY}})^\mathbf{2}$
$X - Y$	$\mathbf{EX - EY}$	$(\sqrt{\mathbf{VX}} \pm \sqrt{\mathbf{VY}})^\mathbf{2}$
$X \times Y$	$\mathbf{EXEY} \pm \sqrt{\mathbf{VXVY}}$	"Homespun variance"
$\frac{X}{Y}$ for $0 \notin Y$	$E(X \times (1/Y))$	$V(X \times (1/Y))$
X^Y for $0 < X$ or $0 < Y$	$E(\exp(\ln(X) \times Y))$	$V(\exp(\ln(X) \times Y))$
$\max(X, Y)$	"Bertsimas max"	env($\max(VX, VY), 0)$
$\min(X, Y)$	"Bertsimas min"	env($\max(VX, VY), 0)$

where t denotes one of the transformations exp, ln, $\log_{10}$, square root or reciprocal ($1/X$), and where *env* denotes the interval envelope:

$$\text{env}(a, b) = [\min(a, b), \max(a, b)] \tag{3}$$

P and Q in equation 1 are

$$P = 1/(1 + (EX - \underline{X})^2/VX) \tag{4}$$

$$Q = 1/(1 + (EX - \overline{X})^2/VX) \tag{5}$$

and ν in equation 2 is the anti-transformation of the Rowe mean estimate (which generally gives an interval result):

$$\nu = t^{-1}(\text{rowe}(t)). \tag{6}$$

For example, the mean of $\ln(X)$ would be estimated by

$$\text{env}(P \ln(\underline{X}) + (1 - P) \ln(EX + \frac{VX}{EX - \underline{X}}), Q \ln(\overline{X}) + (1 - Q) \ln(EX + \frac{VX}{EX - \overline{X}})) \tag{7}$$

and the variance would be estimated by

$$\text{env}(\frac{\ln(\underline{\nu}) - \ln(\underline{X})}{(\underline{\nu} - \underline{X})^2}(VX + (\underline{\nu} - EX)^2), \frac{\ln(\overline{\nu}) - \ln(\overline{X})}{(\overline{\nu} - \overline{X})^2}(VX + (\overline{\nu} - EX)^2)) \tag{8}$$

where ν is the exp (antilog) of the mean estimate. Thus, if X ranges over $[10, 30]$ and has a mean of 15 and a variance of 3, then the mean of $\ln(X)$ is sure to be within the interval $[2.699, 2.704]$, and a variance sure to be in $[0.006437, 0.02002]$, and has a range of $[2.3025, 3.4012]$. Although these templates are a bit complicated for manual calculation, they are very amenable to implementation on a computer and require only two dozen elementary floating-point operations and four evaluations of the transformation function. Rowe's approach works for all transformations that have constant-sign first derivatives.

The derivation of $V[X+Y]$ is straightforward and illustrates several important points. It starts with the familiar general formulation for the variance of a sum

$$V[X+Y] = VX + VY + 2\,\mathrm{Cov}[X,Y]. \tag{9}$$

Even if we do not know what the covariance between the two variables is, we can still bound it quantitatively. We know that, for any pair of random numbers X and Y, their correlation coefficient

$$\rho = \mathrm{Cov}(X,Y)/\sqrt{VXVY}, \tag{10}$$

surely lies within the interval $[-1,+1]$. This implies that their covariance is somewhere within the interval $[-\sqrt{VXVY}, +\sqrt{VXVY}]$. (For compactness, we can write this in interval expressions as $\pm\sqrt{VXVY}$ so long as we keep in mind that it refers to an entire interval, rather than merely a pair of values.) These bounds can can also be found from Cauchy–Schwarz inequality $|\mathrm{Cov}[X,Y]| \leq \sqrt{VXVY}$. This means, then, that

$$V[X+Y] = VX + VY \pm 2\,\sqrt{VXVY}, \tag{11}$$

which is a perfect square that can be simplified to

$$(\sqrt{VX} \pm \sqrt{VY})^2. \tag{12}$$

This is the formula that appears in the table above. As an example, suppose that X has its mean at 10 and a variance of 1, and that Y has a mean of 25 and a variance of 5. The variance of the sum $X+Y$ is sure to lie within the interval $[1.52786, 10.4722]$, no matter what statistical dependency there might be between X and Y. These bounds are best-possible in the sense that they cannot be any tighter given only the stated information about the two moments for each variable.

The rearrangement that changed the perfect square polynomial into the simpler form turns out to have been important. The methods of elementary interval analysis can be sensitive to repeated variables that might introduce the (same) uncertainty into an expression multiple times. Rearranging a formulation so that no uncertain variable appears more than once assures that the calculation will always yield optimal results when the inputs are intervals. For example, if X has mean of $[9,11]$ and variance $[0.8,1.2]$, and Y has mean $[24,26]$ and variance $[4,6]$, then our best possible estimate for the variance of the sum $X+Y$ is the interval $[0,12.5666]$. Despite the imprecision of this result, we will see in section 6 that it actually puts a very strict limit on the exceedance risks and other probabilistic statements associated with the imprecisely known random variable.

The derivation for $V[X-Y]$ is the same as for $V[X+Y]$ beginning from the usual formula

$$V[X-Y] = VX + VY - 2\,\mathrm{Cov}[X,Y]. \tag{13}$$

The derivation for $E[XY]$ is also straightforward. Beginning from the covariance identity

$$\mathrm{Cov}[X,Y] = E[XY] - EXEY \tag{14}$$

and rearranging for

$$E[XY] = EXEY + \mathrm{Cov}[XY], \tag{15}$$

and bounding the covariance with $\pm\sqrt{VXVY}$ gives

$$E[XY] = EXEY \pm \sqrt{VXVY}. \tag{16}$$

Some of the formulae for moment propagation under any dependence are too lengthy to be placed in Table 2. We therefore expand them here. Goodman (1960) provides a formula for the variance of product:

$$V[XY] = (EX)^2VY + (EY)^2VX + 2EXEYE_{11} + 2EXE_{12} + 2EYE_{21} + E_{22} - E_{11}^2 \tag{17}$$

where E_{ij} are the higher bivariate moments: $E_{ij} = E[(X - EX)^i (Y - EY)^j]$ (e.g. E_{11} is covariance). These are generally not tracked by our method, however they may be expressed in terms of the marginal moments and the other formulae described here. We show this in appendix A. However a simpler, and tighter, formula for the variance of the product under no assumption about the dependence may be derived, which we show now. Beginning with $V[X] = E[X^2] - E[X]^2$, $\text{Cov}[X, Y] = E[XY] - EXEY$ and

$$\text{Cov}[X^2, Y^2] = E[X^2Y^2] - E[X^2]E[Y^2], \tag{18}$$

the variance of $V[XY]$ is

$$\begin{aligned}
V[XY] &= E[X^2Y^2] - E[XY]^2 && \text{(from variance identity)}\\
&= \text{Cov}[X^2, Y^2] + E[X^2]E[Y^2] - E[XY]^2 && \text{(using (18))}\\
&= \text{Cov}[X^2, Y^2] + E[X^2]E[Y^2] - (\text{Cov}[X, Y] + EXEY)^2 && \text{(covariance identity)}\\
&= \text{Cov}[X^2, Y^2] + (VX + EX^2)(VY + EY^2) - (\text{Cov}[X, Y] + EXEY)^2.
\end{aligned}$$

The above is expression now written in terms of means, variances and covariances. The covariance may be further expanded in terms of the Pearson correlation coefficient:

$$\text{Cov}[X, Y] = \rho_{XY}\sqrt{VXVY},$$

$$\text{Cov}[X^2, Y^2] = \rho_{X^2Y^2}\sqrt{V[X^2]V[Y^2]},$$

and since it is required that $\rho \in [-1, 1]$:

$$V[XY] = (VX + EX^2)(VY + EY^2) + [-1, 1]\sqrt{V[X^2]V[Y^2]} - ([-1, 1]\sqrt{VXVY} + EXEY)^2, \tag{19}$$

and where $V[X^2]$ may be evaluated with rowevar. Note that in the above expression [-1,1] are intervals, and must be evaluated using interval arithmetic, the using formulas in Table 1. The above formula for the variance of the product without any dependence assumptions is tight, and we believe has not been described elsewhere. This is what we call the "Homespun variance" in Table 2. Note that (19) is not be best possible, since we are ignoring the potential information on $\rho_{X^2Y^2}$ from knowledge that $\rho_{XY} \in [-1, 1]$, i.e. $\rho_{X^2Y^2}$ may not span the entire interval $[-1, 1]$ in all situations. However setting $\rho_{X^2Y^2} = [-1, 1]$ is rigorous and gives a tight and easy to evaluate bound.

The formulas for division can be seen as an application of those for reciprocal and multiplication. That is, we define X/Y as $X * (1/Y)$, which is perfectly valid and yields rigorous results. Similarly, we evaluate X^Y as $e^{\ln(x)*Y}$.

The Bertsimas et al. (2006) formulae for the expectation of max is

$$E[\max(X, Y)] = \text{env}(\max(EX, EY), \max(\overline{X}, \overline{Y})) \cap (EX + EY - \text{env}(\min(EX, EY), \min(\underline{X}, \underline{Y}))), \tag{20}$$

and min is

$$E[\min(X, Y)] = \text{env}(\min(EX, EY), \min(\underline{X}, \underline{Y})) \cap (EX + EY - \text{env}(\max(EX, EY), \max(\overline{X}, \overline{Y}))). \tag{21}$$

3.1 Independence need not be assumed (but can be)

Unlike the formulations usually given for moments of the sums, products, quotients, etc. of random variables (e.g., Wiwatanadate and Claycamp (2000)), the formulas in Table 2 do *not* assume that X and Y are stochastically independent. Our formulas are guaranteed to give correct results whenever their inputs enclose the respective extremes, means and variances. However, if an analyst is willing to assume independence, then the formulas in Table 2 can be improved substantially. Table 3 gives the preferred formulas for such cases. We hasten to point out that an independence assumption is extremely strong, and it is very widely abused in risk analysis. Some uses of the assumption border on the ridiculous, such as the assumption that body weight and skin surface area are independent, or the assumption, echoed even in the paper of Wiwatanadate and Claycamp (2000), that body mass and height are independent.

Table 3: Improved formulas for the mean and variance for convolutions of random variables under an assumption of stochastic independence (best-possible formulations in boldface).

	Mean	*Variance*
$X+Y$	$\mathbf{EX+EY}$	$\mathbf{VX+VY}$
$X-Y$	$\mathbf{EX-EY}$	$\mathbf{VX+VY}$
$X\times Y$	$\mathbf{EXEY}$	$\mathbf{(EX)^2VY+(EY)^2VX+VXVY}$
$\frac{X}{Y}$ for $0\notin Y$	$E(X\times(1/Y))$	$V(X\times(1/Y))$
X^Y for $0<X$ or $0<Y$	$E(\exp(\ln(X)\times Y))$	$V(\exp(\ln(X)\times Y))$
$\max(X,Y)$	$\begin{cases}\mathbf{EX}, & \text{if } Y<X\\ \mathbf{EY}, & \text{if } X<Y\\ \text{“Bertsimas max"}, & \text{otherwise}\end{cases}$	$\begin{cases}\mathbf{VX}, & \text{if } Y<X\\ \mathbf{VY}, & \text{if } X<Y\\ \text{env}(\max(VX,VY),0), & \text{otherwise}\end{cases}$
$\min(X,Y)$	$\begin{cases}\mathbf{EX}, & \text{if } X<Y\\ \mathbf{EY}, & \text{if } Y<X\\ \text{“Bertsimas min"}, & \text{otherwise}\end{cases}$	$\begin{cases}\mathbf{VX}, & \text{if } X<Y\\ \mathbf{VY}, & \text{if } Y<X\\ \text{env}(\max(VX,VY),0), & \text{otherwise}\end{cases}$

Analysts should take care to use assumptions of independence and the formulas of Table 3 only when justified by theoretical argument or comprehensive empirical information. In contrast, the formulas of Tables 1 and 2 are appropriate for all situations and need not be justified by special argument or evidence.

3.2 Using the formulas with interval inputs

Even if one starts out with point estimates for means and variances, applying the formulas in the tables generally yields interval results. Thus, if uncertainty is to be propagated through multiple arithmetic operations, interval estimates for the moments must be handled. The above formula can be readily evaluated with intervals for EX and VX and will surely bound the transformed mean and variances; however the tightness of the result depends on the number times a variable appears in the expression. If the variables appears just once, then the result is tightest possible. But if variables appear multiple times in an expression (for example in the rowe and rowevar templates), then the same uncertainty will be introduced multiple times, and the result will be artificially inflated. This is the well known *repeated variables problem*, and has several numerical solutions such as significance arithmetic (Hyman, 1982), affine arithmetic (Rump and Kashiwagi, 2015), Taylor models (Makino, 1998) and more recently zone arithmetic (Gray et al., 2021a). Where possible, expressions can be rearranged in such a way that the variables appear only once, for example realising that $a^2+a=(a+\frac{1}{2})^2-\frac{1}{4}$. It has been suggested that this process can be automated by an uncertainty compiler (Gray et al., 2019).

A simple-to-implement solution (although more computationally expensive than the above suggestions) is *sub-intervalisation*, where the interval is split into n (usually linearly spaced) sub-intervals, and the expression is evaluated n times with each sub-interval. The resulting range is then the union of the propagated sub-intervals. Usually the main drawback from this method is that it suffers from the curse of dimensionality, that is if a function has m inputs, then n^m interval calculations are required. However, since the expressions proposed in this paper usually require 2 variables (EX and VX), and at most 4, to be sub-intervalised, this is an appropriate technique here. Around 15 sub-intervals is sufficient to substantially reduce the effect of repeated variables, and does not dramatically impact the performance of the method.

3.3 When some moments are missing

The above formulas may still be applied when some, or both, of the moments are unknown. In such cases, it is possible to bound the mean and variance from the range of the random variable. The range $\underline{X}$, $\overline{X}$ provides simple bounds on the mean

$$EX\in[\underline{X},\overline{X}].$$

The variance may also be bounded from the range (Ferson, 2002):

$$VX\in[0,(\overline{X}-\underline{X})^{2/4}].$$

That is, it is not possible to find a random variable with range $[\underline{X}, \overline{X}]$ and with variance greater than $(\overline{X} - \underline{X})^{2/4}$. The lower bound of the variance is zero, since scalars are also included. The upper bound on the variance may be further tightened when bounds $[\underline{EX}, \overline{EX}]$ on the mean are known. Say that m is the centre point of the range $m = (\overline{X} + \underline{X})/2$ and

$$
\begin{aligned}
v_1 &= (\overline{X} - \underline{X})(\overline{X} - \underline{EX}) - (\overline{X} - \underline{EX})^2 \\
v_2 &= (\overline{X} - \underline{X})(\overline{X} - \overline{EX}) - (\overline{X} - \overline{EX})^2 \\
v_3 &= \begin{cases} (\overline{X} - \underline{X})(\overline{X} - m) - (\overline{X} - m)^2, & \text{if } m \in [\underline{EX}, \overline{EX}] \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

then

$$
VX \in [0, \max(v_1, v_2, v_3)]. \tag{22}
$$

If EX is known precisely, v_1 and v_2 are the same, and v_3 only plays a role when the interval mean intersects the mid point.

The above formulas are hard theoretical constraints on means and variances, and are useful in other ways apart from filling in missing information. They can be used as constraints on any initial moment information. For example, if it is initially known that $EX \in [2, 5]$, $VX \in [2, 3]$, and $X \in [3, 6]$, the above constrains tighten this information to $EX \in [3, 5]$, $VX \in [2, 2.25]$, and $X \in [3, 6]$. They also tell you if any initial information is inconsistent (outside theoretical bounds). For example the mean cannot be outside the range $[\underline{EX}, \overline{EX}] \cap [\underline{X}, \overline{X}] = \emptyset$, and initial variance must be inside the bounds provided by (22).

These constraints can also be applied after any of the propagation formulas. Those formulas which are best possible will not be tighten further, however those which are not best possible and the Rowe formulas may sometimes be tightened. They also act as a self-verification for the method. It is possible to check that the propagation formulas in this paper will not produce answers outside of theoretically possible bounds. Although it is meticulous to apply these constrains by hand after every operation, they can easily be done by computer. In an object orientated or typed programming language, they are simple to implement in a variable's constructor, so that when any new moment variable is created either by user input or after an operation, these constrains automatically check for inconsistencies and tighten answers.

4 Homespun bounds

The following are simple approaches to bound the variance of certain transformations which are different from that of Rowe (1988):

4.1 Square

Consider the variance of the square of a variable:

$$
\begin{aligned}
V[X^2] &= E[(X^2 - E[X^2])^2] && \text{(by definition)} \\
&= E[X^4 - 2X^2E[X^2] + E[X^2]^2] && \text{(expand the square)} \\
&= E[X^4] - 2E[X^2]^2 + E[X^2]^2 && \text{(distribute expectation operator)} \\
&= E[X^4] - E[X^2]^2 && \text{(simplify)} \\
&= E[X^4] - (EX^2 + VX)^2 && \text{(use bounds on mean)}
\end{aligned}
$$

Thus, we can bound the variance if we can specify the forth moment. At first, this might seem like it would not provide any improvement. It turns out however, that even the crudest bound on the forth moment

$$
E[X^4] \in [0, \max(|\underline{X}, \overline{X}|)^4],
$$

can give force to this simple bound on the variance, so that

$$V[X^2] \in \max(0, [0, \max(|\underline{X}, \overline{X}|)^4] - (EX^2 + VX)^2). \tag{23}$$

These naive bounds on the variance can sometimes be better than those given by Rowe (1988), although they are not so always. Because both sets of bounds are rigorous, they can be combined straightforwardly via intersection. As Rowe emphasises, this simplicity of combination is one of the important advantages of working with rigorous bounds. The composite bounds would therefore be

$$V[X^2] = \text{rowevar}(X^2) \ \cap \ \max(0, [0, \max(|\underline{X}, \overline{X}|)^4] - (EX^2 + VX)^2). \tag{24}$$

This bound can however be further improved by using the Rowe bound on $E[X^4]$ (see section 4.4)

$$V[X^2] = \text{rowevar}(X^2) \ \cap \ (\text{rowe}(X^4) - (EX^2 + VX)^2). \tag{25}$$

For instance, if the range of X is $[0, 8]$ and the mean and variance are both 3, the Rowe bound on the variance is $[38.5846, 422.585]$. The homespun bound on $V[X^2]$ is $[48, 324.48]$, which is a significant improvement.

4.2 Square root

Consider the variance of the square root of a variable:

$$\begin{aligned}
V[\sqrt{X}] &= E[(\sqrt{X} - E[\sqrt{X}])^2] && \text{(by definition)}\\
&= E[X - 2\sqrt{X}E[\sqrt{X}] + E[\sqrt{X}]^2] && \text{(expand the square)}\\
&= EX - 2E[\sqrt{X}]E[\sqrt{X}] + E[\sqrt{X}]^2 && \text{(distribute expectation)}\\
&= EX - 2E[\sqrt{X}]^2 + E[\sqrt{X}]^2 && \text{(simplify)}\\
&= EX - E[\sqrt{X}]^2 && \text{(simplify)}\\
&= EX - \text{rowe}(\sqrt{X})^2 && \text{(use rowe formula)}
\end{aligned}$$

Therefore, if the range of X is $[0, 8]$ and the mean and variance are both 3, then the rowevar bounds are $[0.1483, 1.584]$. The homespun bound on $V[\sqrt{X}]$ are $[0.1565, 0.75]$, which is an improvement on both sides.

4.3 Exponential

Even the bounds on the variance of an exponential can apparently be improved. Using the same approach as above, we can show that $V[e^X] = E[e^{2X}] - E[e^X]^2$.

4.4 Other integer powers

The formulas from Rowe (1988) may be readily applied to any unary transformation which have a constant-sign derivative in the range of the random variable. Thus, the rowe and rowevar formulas apply to any even powered transformation (e.g. X^4, X^6, X^8), since their first derivatives (X^3, X^5, X^7) are monotonic in $\mathbb{R}$. However, the formulas may also be applied to odd powers if the variable is completely positive or negative ($X \in (-\infty, 0]$ or $X \in [0, \infty)$), since the the derivative of any odd powered transformation (which is an even powered transformation) has a single minima at 0, and is monotonic on either side. However, in the case where the range of X does extend into both the positive and negative portion of the reals, a loose bound on the odd powers may still be found by evaluating it as

$$X^a = X^{a-1} \times X, \tag{26}$$

for example X^3 may be evaluated as $X^2 \times X$. The multiplication must however be performed using the Fréchet formulas from Table 2 and not the independence formulas of Table 3. This is because the variables X^2 and X are dependent on one another through the squaring transformation, which must be accounted for. Although it must be noted that this dependence is not precisely accounted for, since Table 2 accounts for all possible dependencies. As a result the bounds may be quite loose, but importantly are still rigorous. As an example, the following range, expectation and variance:

$$X \in [1,5], \qquad EX = 3, \qquad VX \in [1,3],$$

give

$$E[X^4] = [117.89, 271.2033], \qquad V[X^4] = [2820.02, 95597.0304],$$

$$E[X^3] = [33.8274, 55.7302], \qquad V[X^3] = [227.738, 3791.15].$$

However if the range of X is extended to the negative portion of the reals, $X \in [-1,5]$, and the moments are kept the same, the following is yielded:

$$E[X^4] = [104.448, 271.2033], \qquad V[X^4] = [631.118, 95950.8271],$$

$$E[X^3] = [7.483, 57.97], \qquad V[X^3] = [0, 5625].$$

Although both bounds did widen as the range was extended, the bound on the cube widened substantially due to the Fréchet multiplication. However, it is not too wide as to be uninformative, especially considering we began with an interval for the variance.

A negative integer power X^{-a} may be evaluated using the reciprocal $X^{-a} = \frac{1}{X^a}$. Care must be taken that X is either completely positive or negative ($0 \notin X$) since this would give a division by 0 for which the moments are not defined.

4.5 Other integer roots, real powers and interval powers

All integer roots (e.g. $\sqrt[3]{X}, \sqrt[4]{X}, \sqrt[5]{X}$) also have constant sign derivatives when $X \geq 0$, and therefore the Rowe formulas apply here also. For example when $X \in [2,6]$, $EX \in [3, 3.5]$, and $VX \in [2,3]$ gives $\sqrt[3]{X} \in [1.258, 1.818]$, $E[\sqrt[3]{X}] \in [1.38, 1.51]$, and $V[\sqrt[3]{X}] \in [0.0288, 0.0768]$.

Interestingly, the same also applies to any real (including negative) powers (e.g. $X^{1.5}, X^{-3.2}, X^{6.7}$) when X is positive. Here we again evaluate a negative power using reciprocation: $X^{-a} = \frac{1}{X^a}$. Note however that any negative number to the power of a real number (which is not an integer), for example $-2^{1.2}$, will generally give an imaginary number. We therefore only define these operations for positive X. Using the same X as the previous example: $X^{-4.7} \in [5 \times 10^{-4}, 1.0]$, $E[X^{-4.7}] \in [0.01, 0.4291]$ and $V[X^{-4.7}] \in [2 \times 10^{-5}, 0.245]$.

Another useful observation is that for any two real (including negative) numbers $a_1 <= a_2$ gives $X^{a_1} <= X^{a_2}$ in the range $X \in [1, \infty)$ and gives $X^{a_1} >= X^{a_2}$ in $[0,1]$. That is, $f(x) = x^a$ monotonically increases in a for the range $[1, \infty)$ and is monotonically decreases in $[0,1]$, and is equal to 1 when $x = 1$ for any a. We can use this construct some very simple formulae for evaluating interval powers:

$$X^{[a_1,a_2]} = \begin{cases} \text{env}(X^{a_1}, X^{a_2}, 1), & \text{if } 0 \in [a_1,a_2] \text{ or } 1 \in X \\ \text{env}(X^{a_1}, X^{a_2}), & \text{otherwise} \end{cases} \tag{27}$$

$$E[X^{[a_1,a_2]}] = \begin{cases} \text{env}(E[X^{a_1}], E[X^{a_2}], 1), & \text{if } 0 \in [a_1,a_2] \text{ or } 1 \in X \\ \text{env}(E[X^{a_1}], E[X^{a_2}]), & \text{otherwise} \end{cases} \tag{28}$$

$$V[X^{[a_1,a_2]}] = \begin{cases} \text{env}(V[X^{a_1}], V[X^{a_2}], 0), & \text{if } 0 \in [a_1,a_2] \text{ or } 1 \in X \\ \text{env}(V[X^{a_1}], V[X^{a_2}]), & \text{otherwise} \end{cases} \tag{29}$$

where essentially the extremums of the transformed variable can be found through a simple evaluation of the endpoints of the inputs. Note that $X^0 = 1$ for any X.

Table 4: Improved formulas for the mean and variance for convolutions of random variables using correlation coefficient ρ_{XY} (best-possible formulations in boldface).

	Mean	*Variance*
$X+Y$	$\mathbf{EX+EY}$	$\mathbf{VX+VY+2}\ \rho_{\mathbf{XY}}\sqrt{\mathbf{VXVY}}$
$X-Y$	$\mathbf{EX-EY}$	$\mathbf{VX+VY-2}\ \rho_{\mathbf{XY}}\sqrt{\mathbf{VXVY}}$
$X\times Y$	$\mathbf{EXEY}+\rho_{\mathbf{XY}}\sqrt{\mathbf{VXVY}}$	$(VX+EX^2)(VY+EY^2)-$ $(\rho_{XY}\sqrt{VXVY}+EXEY)^2\pm\sqrt{V[X^2]V[Y^2]}$
$\frac{X}{Y}$ for $0\notin Y$	$E(X\times(1/Y))$	$V(X\times(1/Y))$

5 Using bounds on correlations

In this section we summarise formulas for moment arithmetic which use any general bound on the correlation coefficient between variables. The previous sections we have seen how to perform arithmetic between moment variables under the assumption of independence and without any dependence assumptions (Fréchet). However, the formulas can also be improved if some covariance information between variables is known. For example, if it is known that X and Y have are positively correlated (that their person correlation coefficient is positive $\rho_{XY}>=0$) the answers provided in Table 2 can be tightened. Stochastic dependence plays an important role in arithmetic operations between random variables. That is, the output of a function depends not only on what the random variables X and Y are (on their distribution or their moments), but also on how they are correlated; sometimes dramatically so. If any correlation information is known about the variables, it should be incorporated and the answer tightened.

One thing to note is that the covariance can always be related to the Pearson correlation coefficient and the variances:

$$\text{Cov}[X,Y]=\rho_{XY}\sqrt{VXVY}.$$

The above formula also works for interval variances and correlation. For example, if $VX\in[2,3]$, $VY\in[4,5]$, and $\rho_{XY}\in[-1,-0.5]$ gives an exact bound on the covariance: $\text{Cov}[X,Y]\in[-3.87299,-1.41421]$. Setting the correlation coefficient to its maximum bound of $\rho_{XY}\in[-1,1]$ gives a maximum bound on the covariance $\text{Cov}[X,Y]\in[-3.87299,3.87299]$.

For sum and subtraction, the usual formulas for the variance (9) and (13) yield

$$V[X+Y]=VX+VY+2\ \rho_{XY}\sqrt{VXVY},$$

and

$$V[X-Y]=VX+VY-2\ \rho_{XY}\sqrt{VXVY}.$$

Equation (15) for the mean of the product gives

$$E[XY]=EXEY+\rho_{XY}\sqrt{VXVY},$$

and equation (19) for variance of the product gives

$$V[XY]=(VX+EX^2)(VY+EY^2)-(\rho_{XY}\sqrt{VXVY}+EXEY)^2\pm\sqrt{V[X^2]V[Y^2]}.$$

Table 4 summarises formulas for moment arithmetic using a correlation coefficient ρ_{XY}. Any interval value which is a subset or equal to $[-1,1]$ may be used for the ρ_{XY} in these formulas. For example, if $EX=EY=2$, $VX=VY=1$, and $\rho_{XY}\in[0,1]$ (positively correlated), then $V[X+Y]=[2,4]$ and $V[X-Y]=[0,2]$. Indeed the previous independence and Fréchet formulas may be derived from Table 2, by setting different interval values for the correlation coefficient. For example setting $\rho_{XY}=[-1,1]$ give Fréchet, and $\rho_{XY}=0$ gives independence. Note that $\rho_{XY}=0$ gives $\rho_{X^2,Y^2}=0$, so the formula for $V[XY]$ is simplified further for independence. Further, formulas for perfectly dependent ($\rho=1$) and oppositely dependent ($\rho=-1$) random variables may also be derived, and are summarised in Tables 5 and 6 respectively.

Table 5: Improved formulas for the mean and variance for convolutions of random variables under an assumption of perfect dependence (best-possible formulations in boldface).

	Mean	*Variance*
$X+Y$	$\mathbf{EX+EY}$	$(\sqrt{\mathbf{VX}}+\sqrt{\mathbf{VY}})^{\mathbf{2}}$
$X-Y$	$\mathbf{EX-EY}$	$(\sqrt{\mathbf{VX}}-\sqrt{\mathbf{VY}})^{\mathbf{2}}$
$X\times Y$	$\mathbf{EXEY}+\sqrt{\mathbf{VXVY}}$	$(VX+EX^2)(VY+EY^2)-$ $(EXEY+\sqrt{VXVY})^2\pm\sqrt{V[X^2]V[Y^2]}$
$\frac{X}{Y}$ for $0\notin Y$	$E(X\times(1/Y))$	$V(X\times(1/Y))$

Table 6: Improved formulas for the mean and variance for convolutions of random variables under an assumption of opposite dependence (best-possible formulations in boldface).

	Mean	*Variance*
$X+Y$	$\mathbf{EX+EY}$	$(\sqrt{\mathbf{VX}}-\sqrt{\mathbf{VY}})^{\mathbf{2}}$
$X-Y$	$\mathbf{EX-EY}$	$(\sqrt{\mathbf{VX}}+\sqrt{\mathbf{VY}})^{\mathbf{2}}$
$X\times Y$	$\mathbf{EXEY}-\sqrt{\mathbf{VXVY}}$	$(VX+EX^2)(VY+EY^2)-$ $(EXEY-\sqrt{VXVY})^2\pm\sqrt{V[X^2]V[Y^2]}$
$\frac{X}{Y}$ for $0\notin Y$	$E(X\times(1/Y))$	$V(X\times(1/Y))$

6 What do the range and moments say about risks?

What does knowing something about the mean and variance of a random number tell us about the probability distribution of that variable? Generally, people expect that it is unlikely for a random value to be many standard deviations away from the mean. But what exactly is the chance of being, say, 5 standard deviations (or more) larger than the mean? If we assume the underlying distribution is standard normal, the risk is roughly 1 in 3.5 million. Such a value seems very small and might be considered an acceptable risk by planners and decision makers.

But what can one say about such risks *without assuming normality*? What inferences can be drawn about the risks of exceedance that are free of assumptions about the particular shape of the distribution? This question was posed by Chebyshev (1874) and answered by Markoff (1900) for the case when only the mean and variance are known. The answer we need for risk analysis is embodied in a version of the classical Chebyshev inequality (Feller, 1968; Allen, 1990). The upper bound on the probability that the variable X will exceed a value as large as x is

$$\mathbb{P}(x\le X)\le\begin{cases}1/(1+(x-EX)^2/VX), & \text{if } EX<x\\ 1, & \text{if } x\le EX\end{cases}\tag{30}$$

where EX and VX are the mean and variance of X. The lower bound on the same probability is

$$\mathbb{P}(x\le X)\ge\begin{cases}1/(1+VX/(x-EX)^2), & \text{if } x<EX\\ 1, & \text{if } EX\le x\end{cases}\tag{31}$$

Although the derivation of the Chebyshev inequality is an elementary exercise in mathematical statistics (Ash (1993) page 271f), the one-sided version used here (without the absolute value sign) is uncommon and maybe unfamiliar to many practitioners.

If we use the Chebyshev inequality to ask how large the chance might be without any assumption about the shape of the underlying distribution (with mean 0 and variance 1 at 5 standard deviations), we find it is somewhere between zero and $1/(1+(5-0)^2/1)=0.03846$, or 1 in 26. Omitting the normality assumption causes the risk to go from 0.000000286 to almost $[0,0.04]$, which represents a potential risk increase of over five orders of magnitude. What engineer designing a safety system for a nuclear power plant, or for that matter, the razor burn guard on an electric shaver, would be happy with a potential risk of 1 in 26?

The Chebyshev inequality may be used to compute bounds on the cumulative distribution function (cdf) of an imprecisely known random variable. The graph below depicts these bounds for a random variable whose mean is zero and whose variance is unity. Such a characterisation of an imperfectly known random variable is called a probability box, or p-box (Ferson et al., 2015). Although the area inside the bounds is integrable, the tails extend to infinity in both directions. Nevertheless, the p-box can be used in practical risk analyses by truncating the tails at some appropriate percentile.

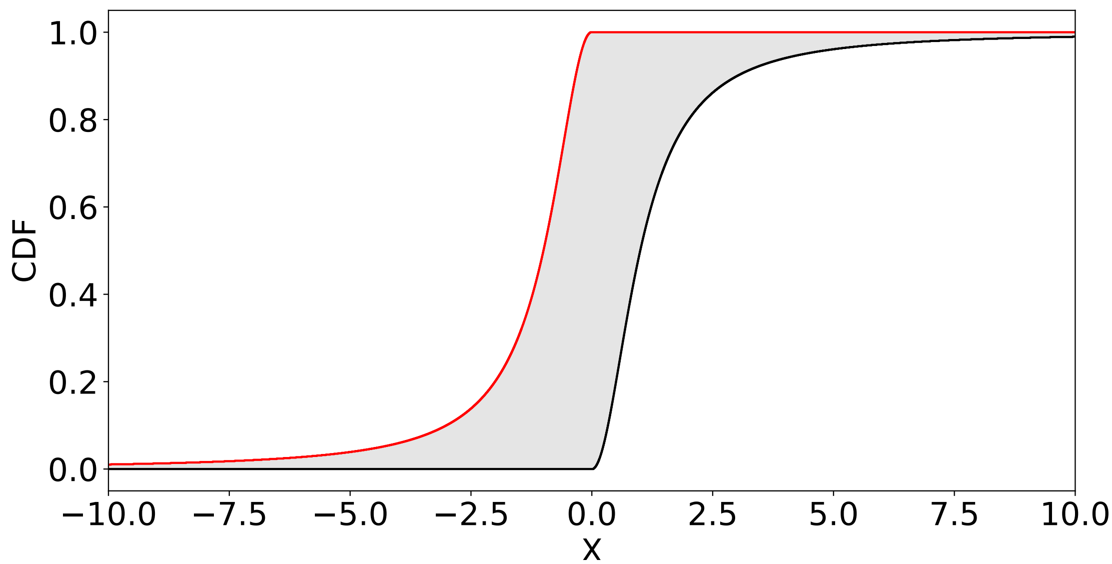

Figure 1: Bounds on the cdf of all random variables with mean 0 and variance 1

These bounds are rigorous in the sense that they enclose all distributions, no matter what shape they have, that have the prescribed mean and variance. One should be careful to note that neither the upper nor the lower bound corresponds to a distribution that has the specified mean and variance. For instance, it is obvious that the upper bound $\overline{F}$ (left red curve) describes a distribution whose mean is smaller than 0, and the lower bound $\underline{F}$ (right black curve) describes a bound with a mean larger than 0. In fact, it is discrete distributions that force the Chebyshev bounds out so far (Karlin and Studden, 1966; Smith, 1993) . The bounds are actually the envelope of many distributions elbowing out at the edges. Although the Chebyshev bounds are not attainable by any single distribution, they are best possible bounds on the risks given the stated constraints. This means that the bounds could not be any tighter and still contain all distributions that do have the given mean and variance. The breadth of these bounds (grey shaded region) might be surprising to someone who has not considered just how strong assumptions about distribution shape really are.

Most of the risk analysis-relevant questions can be asked from the p-box. For example, the probability that the random variable falls in some set can be found from the bounding cdfs $\underline{F}$ and $\overline{F}$. However since the p-box characterises an imperfectly known random variable, the p-box will return an interval probability instead of a single precise probability. This interval probability bounds the contribution to the risk from all of the possible distributions the random variable could have. Using the p-boxes two bounding cdfs $\underline{F}$ and $\overline{F}$, the interval probability that the random variable X is less than or equal to some value x is

$$\underline{\mathbb{P}}(X \leq x) = \underline{F}(x)$$
$$\overline{\mathbb{P}}(X \leq x) = \overline{F}(x).$$

The exceedance probability $\mathbb{P}(X > x)$ is

$$\underline{\mathbb{P}}(X > x) = 1 - \overline{F}(x)$$
$$\overline{\mathbb{P}}(X > x) = 1 - \underline{F}(x).$$

The interval probability that the random variable is in some interval $U = [a, b]$ may also be computed as

$$\underline{\mathbb{P}}(U) = \max(0, \underline{F}(b) - F(a))$$
$$\overline{\mathbb{P}}(U) = \overline{F}(b) - \underline{F}(a).$$

As an example for the above p-box, $\mathbb{P}(X \leq -7.5) = [0, 0.018]$ (at most 1 in 55), $\mathbb{P}(X > 2.5) = [0, 0.138]$ (at most 2 in 15), and $\mathbb{P}([-1, 1]) = [0.5, 1]$ (at least 1 in 2).

The Chebyshev bounds can be tightened substantially in some cases by the addition of knowledge about one endpoint of the range, i.e., either the minimum or the maximum of the underlying distribution. This improvement is expressed in the classical Cantelli inequalities, which give rigorous and best possible bounds on the distribution function for a nonnegative random variable X having mean EX and variance VX. The Cantelli inequalities are a combination of the Markov and Chebyshev inequalities. The upper bound on the probability that the variable X will be no larger than a value x is

$$\mathbb{P}(x \leq X) \leq \begin{cases} 0, & \text{if } x \leq 0 \\ 1/(1 + (x - EX)^2/VX), & \text{if } 0 \leq x \leq EX \\ 1, & \text{if } EX < x \end{cases} \tag{32}$$

This function forms the left side of a p-box for X. The right side is the lower bound on the same probability, which is

$$\mathbb{P}(x \leq X) \geq \begin{cases} 0, & \text{if } x \leq EX \\ 1 - EX/x, & \text{if } EX \leq x \leq EX + VX/EX \\ 1/(1 + VX/(x - EX)^2), & \text{if } EX + VX/EX < x \end{cases} \tag{33}$$

If the minimum value of X is not zero, we can encode the information in a new variable Y whose minimum value is zero with the transformations

$$\begin{aligned} Y &= X - \underline{X}, \\ EY &= EX - \underline{X}, \\ VY &= VX, \end{aligned}$$

then apply the inequalities to obtain the p-box for Y, and finally back-transform this p-box to get the bounds in terms of the original variable X by adding $\underline{X}$ to it. If it is the maximum, rather than the minimum that is known, we can use the encoding

$$\begin{aligned} Z &= -X \\ EZ &= -EX, \\ VZ &= VX, \end{aligned}$$

and then apply the inequalities (possibly also encoding to make the new minimum zero), and finally negate the resulting Z p-box to reexpress it in terms of the original variable. Figure 2 show bounds on random variables using the mean, variance and either a maximum or a minimum, and are an improvement over the Chebyshev p-box.

6.1 min-max-mean-variance bounds

An obvious generalisation of the Cantelli inequalities discussed in the previous section would be rigorous and best possible bounds on a distribution function given the minimum, maximum, mean and variance of the underlying random variable. We derived the bounds for this case ourselves and have not seen them published before. However, they are obvious enough to have been discovered in prior work somewhere, and we would not be surprised to find that they are well known to someone.

Because there are four specifications known about the variable, the distributions defining the maximal and minimal cumulative probabilities will be discrete distributions on three points (Karlin and Studden, 1966), with at least one of the points at an extremum of the range. Suppose, to start, that this point is at the smallest possible value $\underline{X}$ of the random variable X. Call the mass at this point p_0. Let the other two masses, p_1 and p_2, be at points x_1 and x_2. To find the left side of the p-box, i.e., the upper bound on the probability, at point x_1, we seek to maximise the quantity $p_1 + p_2$. Equivalently, we could minimise p_2. In fact, x_2 can be chosen to minimise p_2. There are then three constraints over four variables (p_0, p_1, p_2, x_2) and we look for $\min(p_2)$ as a function of EX, VX and x_1. The constraints are that on the total probability, the mean, and the variance:

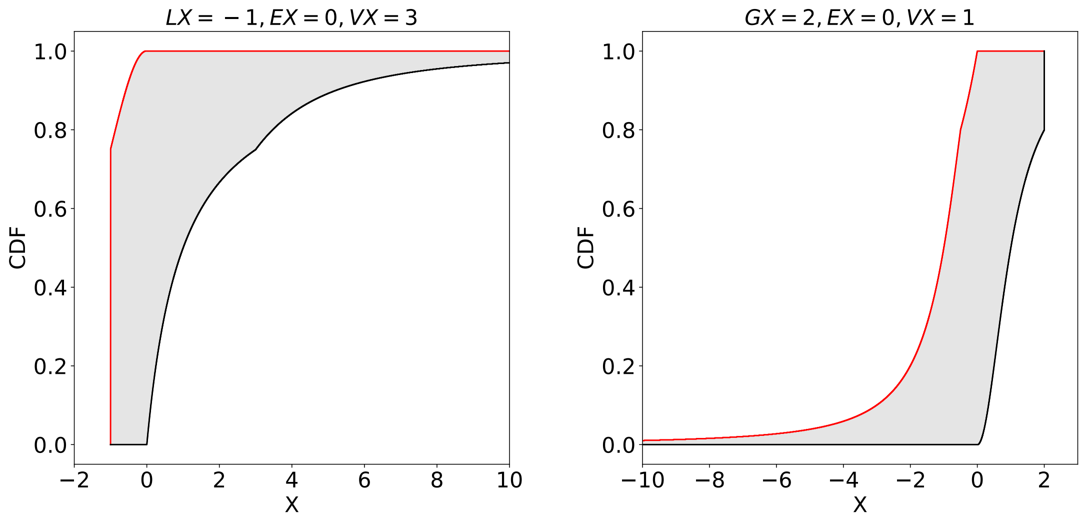

Figure 2: Bounds on random variables with (left) $\underline{X} = -1, EX = 0, VX = 3$, and (right) $\overline{X} = 2, EX = 0, VX = 1$

$$\begin{gathered} p_0 + p_1 + p_2 = 1, \\ p_0 \underline{X} + p_1 x_1 + p_2 x_2 = EX, \\ p_0(\underline{X} - EX)^2 + p_1(x_1 - EX)^2 + p_2(x_2 - EX)^2 = VX. \end{gathered}$$

Without an unrecoverable loss of generality, we can assume $\underline{X} = 0$ and $\overline{X} = 1$. (We can always use rescaling to account for more general situations.) Solving simultaneous equations yields:

$$p_2 = (VX + (EX)^2 - x_1 EX)/(x_2(x_2 - x_1)).$$

Minimising p_2 with respect to x_2 mean (by inspection) that we should make x_2 as large as possible. Thus, let $x_2 = \overline{X} = 1$. Then

$$\min(p_2) = (VX + (EX)^2 - x_1 EX)/(1 - x_1),$$

so

$$\max(p_0 + p_1) = 1 - (VX + (EX)^2 - x_1 EX)/(1 - x_1).$$

Varying x_1 gives the limit for every value in $[\underline{X}, \overline{X}] = [0, 1]$. This bound is simultaneous with the Chebyshev and Cantelli bounds and should therefore be combined with them.

Here are the resulting general expressions for the min-max-mean-variance bounds. The left side of the p-box, which is the upper bound on the cumulative probability, is

$$\mathbb{P}(x \leq X) \leq \begin{cases} 0, & \text{if } x \leq \underline{X} \\ 1/(1 + (EX - x)^2/VX), & \text{if } \underline{X} \leq x \leq EX + VX/(EX - \overline{X}) \\ 1 - (\mu^2 - \mu y + \sigma^2)/(1 - y), & \text{if } EX + VX/(EX - \overline{X}) < x < EX + VX/(EX - \underline{X}) \\ 1, & \text{if } EX + VX/(EX - \underline{X}) \leq x, \end{cases} \tag{34}$$

where $y = (x - \underline{X})/(\overline{X} - \underline{X})$, $\mu = (EX - \underline{X})/(\overline{X} - \underline{X})$, and $\sigma^2 = VX/(\overline{X} - \underline{X})^2$. The right side of the p-box, which is the lower bound on the same probability, is

$$\mathbb{P}(x \leq X) \geq \begin{cases} 0, & \text{if } x \leq EX + VX/(EX - \overline{X}) \\ 1 - (\mu(1+y) - \sigma^2 - \mu^2)/y, & \text{if } EX + VX/(EX - \overline{X}) < x < EX + VX/(EX - \underline{X}) \\ 1/(1 + VX/(x - EX)^2), & \text{if } EX + VX/(EX - \underline{X}) \leq x < \overline{X} \\ 1, & \text{if } \overline{X} \leq x. \end{cases} \tag{35}$$

When specifying p-boxes with the min-max-mean-variance inequalities, analysts must take care to respect feasibility constraints, i.e, those discussed in section 3.3. Figures 3 and 4 shows examples min-max-mean-variance bounds.

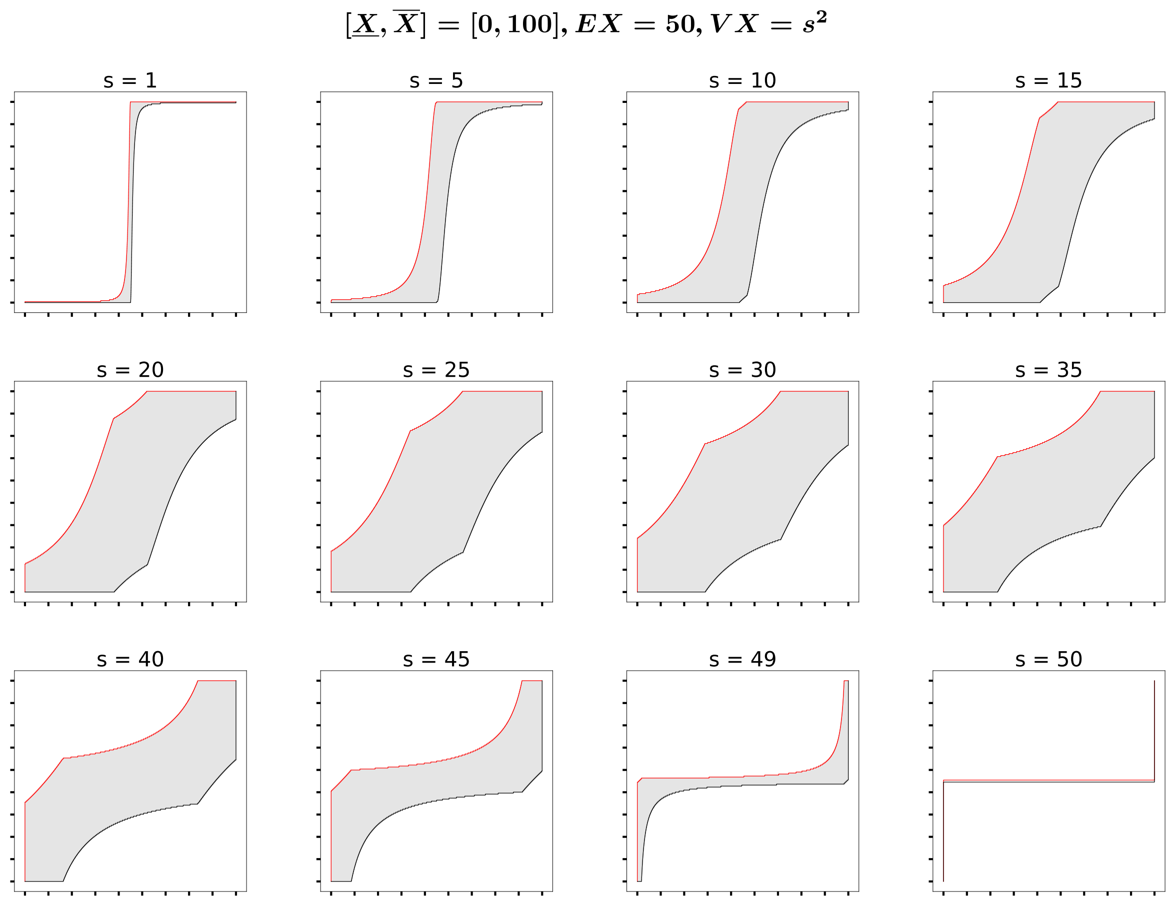

Figure 3: Bounds on random variables with $[\underline{X}, \overline{X}] = [0, 100]$, $EX = 50$, $VX = s^2$

Because the Cantelli inequalities are essentially a superimposition of the Markov and Chebyshev inequalities, one might expect these min-max-mean-variance inequalities to be a simple extension that superimposes Chebyshev and two Markov inequalities (to account for the minimum and for the maximum from different directions). In fact, the Markov inequality does not even play a role in the present functions, except at the two cusps where the new functions coincide with both Chebyshev and Markov.

These bounds are somewhat tighter that the Cantelli inequalities, with improvements to the upper part of the left bound and the lower part of the right bound (i.e., the least important parts for risk analysis). The main significance of this result is its comprehensiveness. Indeed, these bounds generalise many of the inequalities we have discussed in this paper. If $\overline{X} = \infty$ (or if $\underline{X} = -\infty$), then these bounds become the Cantelli inequalities. If both endpoints of the range are infinite, then the Chebyshev inequality is retrieved. If VX is unknown, that is, if its estimate is $[0, \infty]$, then this inequality degenerates to Rowe's inequality based on the minimum, maximum and mean (Rowe, 1988). If both VX is

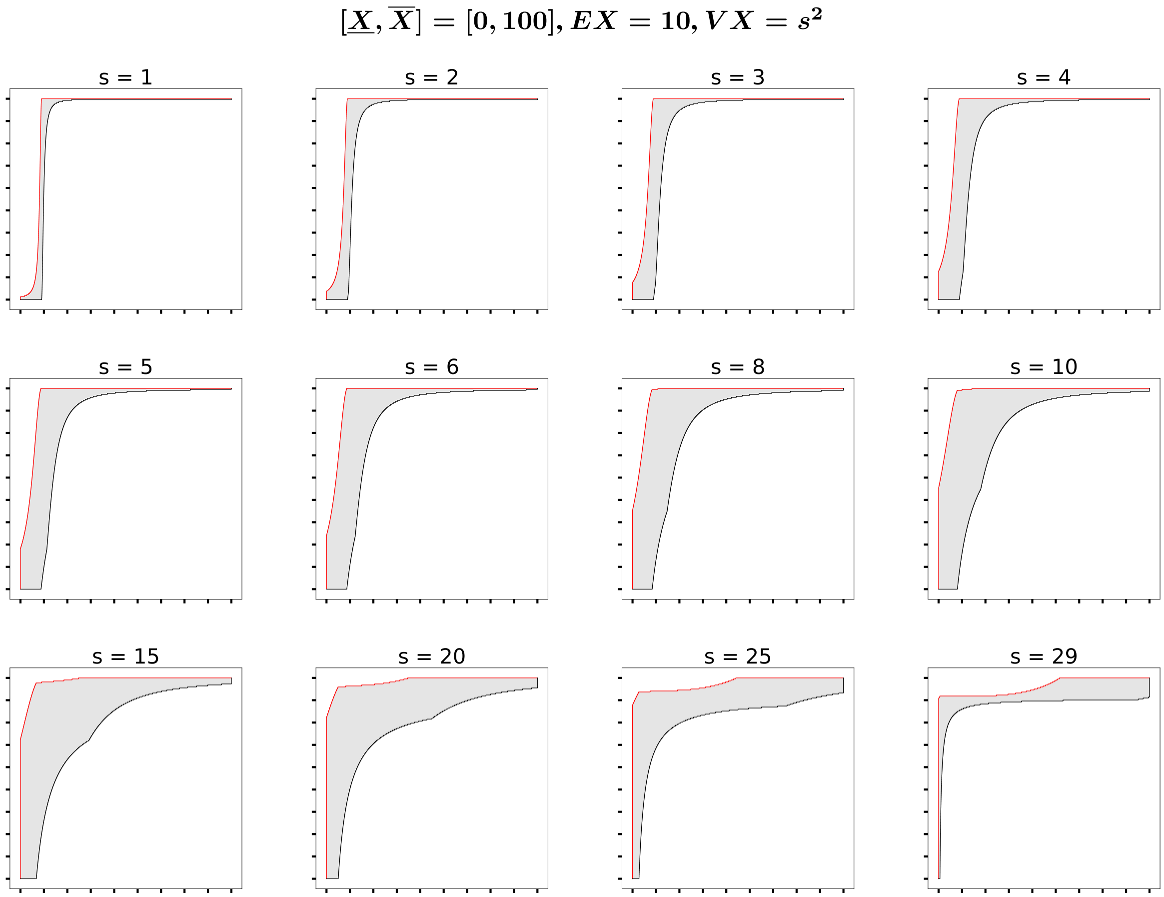

Figure 4: Bounds on random variables with $[\underline{X}, \overline{X}] = [0, 100]$, $EX = 10$, $VX = s^2$

unknown and $\overline{X} = \infty$, then it degenerates to the Markov inequality. And if mean and variance are entirely unknown, it reduces to the interval determined by the range.

7 Covariance tracking

Like in interval arithmetic expressions where multiple repeated occurrences of the same interval will lead to artificially inflated answers, a dependency problem also exits in moment arithmetic. This extra puffiness in interval and moment arithmetic happens because dependency information is only known between variables at the beginning of expressions, and is lost as variables are used in operations. The formulas in Table 2 are still valid and will give a rigorous propagation when stochastic dependencies are unknown. However, incorporating and tracking covariance information could be used as a form of crude form of dependence tracking in moment arithmetic. This covariance information can be used with the correlated formulas in Table 4 to tighten answers. As an example, consider the following simple expression:

$$Z = (X + Y)X,$$

where initially the mean, variance, and ranges of X and Y are known, as well as $\text{Cov}[X, Y]$. The initial sum $K_1 = X + Y$ can be evaluated exactly using the correlated formulas in Table 4, however the product $K_1 \times X$ must be evaluated with the Fréchet formulas in Table 2, since $\text{Cov}[X + Y, X]$ is unknown, leading to an artificial inflation of uncertainty. In this section, we argue that simple expressions for covariance algebra can be used to calculate and track covariance between variables to tighten results from repeated variables. Although covariance tracking would be an intensive task when done by hand, it may be easily automated in software, which we discuss in section 7.2.

The covariance $\text{Cov}[X + Y, X]$ can be calculated as

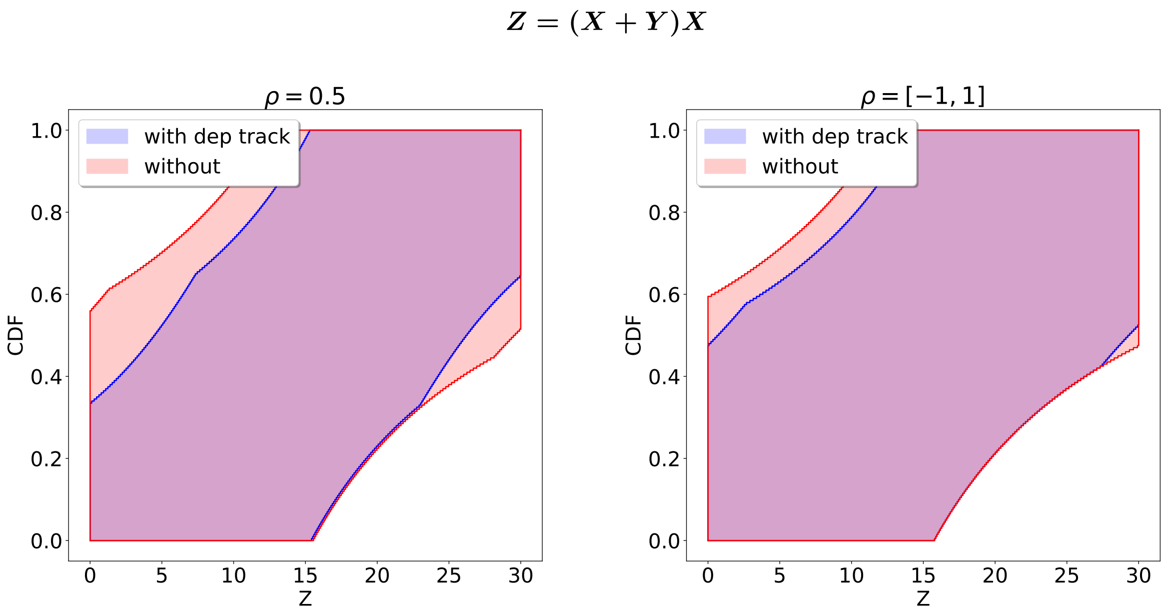

Figure 5: Comparison between moment arithmetic with covariance tracking (blue) and without (red), for the expression $Z = (X + Y)X$. Left shows where initial correlation is $\rho_{XY} = 0.5$, and right shows where the initial dependence is unknown $\rho_{XY} = [-1, 1]$

$$\begin{aligned}
\text{Cov}[X+Y, X] &= E[((X+Y) - E[X+Y])(X - EX)] \\
&= E[(X+Y)X - (X+Y)EX - E[X+Y]X + EXE[X+Y]] \\
&= E[X^2] + E[XY] - E[X+Y]EX - E[X+Y]EX + EXE[X+Y] \\
&= E[X^2] - (EX)^2 + E[XY] - EXEY.
\end{aligned}$$

And since $\text{Cov}[X, Y] = E[XY] - EXEY$ and $VX = E[X^2] - (EX)^2$:

$$\text{Cov}[X+Y, X] = VX + \text{Cov}[X, Y]. \tag{36}$$

By symmetry we can show that

$$\text{Cov}[X+Y, Y] = VY + \text{Cov}[X, Y]. \tag{37}$$

Beginning with $EX = 2$, $VX = 1$, $X \in [0, 3]$, $EY = 5$, $VY = 0.5$, $Y \in [2, 7]$, and $\rho_{XY} = 0.5$, gives $\text{Cov}[K_1, X] = 1.35355$, which may then be used in the product $K_1 \times X$. Without covaraince tracking, the following moments are yielded:

$$EZ = [12.514, 15.486], \qquad VZ = [0, 196.078],$$

and covariance tracking leads to the following contraction:

$$EZ = [15.3535, 15.3536], \qquad VZ = [0, 116.955].$$

This improvement is best seen in the contraction in the p-boxes, which is shown in the left of Figure 5. Interestingly, even if one begins with a unknown dependence between the initial variables, covariance tracking can still give a contraction. Beginning with $\rho_{XY} = [-1, 1]$, and the same X and Y as before, gives $\text{Cov}[K_1, X] = [0.2928, 1.70711]$,

Table 7: Formulas for covariances between operands and outputs of operations (best-possible formulations in boldface).

Operation	*Covariance*	*Formula*
$k+X$	Cov$[k+X, X]$	$\boldsymbol{VX}$
kX	Cov$[kX, X]$	$\boldsymbol{kVX}$
$1/X$	Cov$[1/X, X]$	$1-$ rowe(reciprocal)EX
X^n	Cov$[X^n, X]$	$E[X^{n+1}] - E[X^n]EX$
$X+Y$	Cov$[X+Y, X]$	$\boldsymbol{VX} + \textbf{Cov}\boldsymbol{[X,Y]}$
$X-Y$	Cov$[X-Y, X]$	$\boldsymbol{VX} - \textbf{Cov}\boldsymbol{[X,Y]}$
$X-Y$	Cov$[X-Y, Y]$	$\textbf{Cov}\boldsymbol{[X,Y]} - \boldsymbol{VY}$
$X\times Y$	Cov$[XY, X]$	$E[X^2Y] -$ Cov$[X,Y]EX - EX^2EY$
X/Y	Cov$[X/Y, X]$	$E[X^2/Y] - E[X/Y]EX$
X/Y	Cov$[X/Y, Y]$	$E[X] - E[X/Y]EY$

which is substantially tighter than the widest possible covariance of $[-1.70711, 1.70711]$. The improvement using covariance tracking in the p-boxes for the Fréchet case is shown on the right of Figure 5.

Table 7 summarises formulas for calculating the covariances between inputs and outputs of operations. The derivation of these formulas is quite simple, and follows the same reasoning as the derivation of (36). Some of the formulas require a Fréchet evaluation, for example in the formula for Cov$[XY, X]$ there exists a $E[X^2Y]$, which can be calculated as $X\times Y$ using a correlated multiplication, followed by a Fréchet multiplication with X. Because the formulas in Table 7 sometimes yield covariances outside the widest possible bound of $[-\sqrt{VZVX}, +\sqrt{VZVX}]$, the software should always use intersection to tighten the results appropriately.

7.1 Covariances of a third variable

The covariance formulas in Table 7 work for inputs and outputs of operations. However, if there exits a third variable in the expression, covariances between newly generated variables and this third variable may also be found. Consider for example the following expression:

$$W = (X+Y)(Z-Y),$$

where initially the covariance matrix of X, Y and Z is known, in addition to their moments and ranges. The expression can be evaluated with the the correlated summation $K_1 = X+Y$ and subtraction $K_2 = Z-Y$, however to tightly evaluate $K_1 \times K_2$, we need to calculate Cov$[K_1, K_2]$.

The covariance between a sum $X+Y$ and a *different* variable Z can be calculated as follows:

$$\begin{aligned}
\text{Cov}[X+Y, Z] &= E[((X+Y) - E[X+Y])(Z-EZ)] \\
&= E[(X+Y)Z - (X+Y)EZ - E[X+Y]Z + EZE[X+Y]] \\
&= E[XZ] + E[YZ] - E[X+Y]EZ - E[X+Y]EZ + EZE[X+Y] \\
&= E[XZ] + E[YZ] - EXEZ - EYEZ.
\end{aligned}$$

And since Cov$[X, Z] = E[XZ] - EXEZ$:

$$\text{Cov}[X+Y, Z] = \text{Cov}[X, Z] + \text{Cov}[Y, Z]. \tag{38}$$

Following the same reasoning, we can show that

$$\text{Cov}[X-Y, Z] = \text{Cov}[X, Z] - \text{Cov}[Y, Z]. \tag{39}$$

Table 8: Formulas for covariances between results of operations and another variable (best-possible formulations in boldface).

Operation	*Covariance*	*Formula*
$k+X$	$\text{Cov}[k+X,Z]$	$\mathbf{Cov}[\boldsymbol{X},\boldsymbol{Z}]$
kX	$\text{Cov}[kX,Z]$	$\boldsymbol{k}\mathbf{Cov}[\boldsymbol{X},\boldsymbol{Z}]$
$1/X$	$\text{Cov}[1/X,Z]$	$E[X/Z]-E[1/X]EZ$
X^n	$\text{Cov}[X^n,Z]$	$E[X^nZ]-E[X^n]EZ$
$X+Y$	$\text{Cov}[X+Y,Z]$	$\mathbf{Cov}[\boldsymbol{X},\boldsymbol{Z}]+\mathbf{Cov}[\boldsymbol{Y},\boldsymbol{Z}]$
$X-Y$	$\text{Cov}[X-Y,Z]$	$\mathbf{Cov}[\boldsymbol{X},\boldsymbol{Z}]-\mathbf{Cov}[\boldsymbol{Y},\boldsymbol{Z}]$
$X\times Y$	$\text{Cov}[XY,Z]$	$E[XYZ]-\text{Cov}[X,Y]EZ-EXEYEZ$
X/Y	$\text{Cov}[X/Y,Z]$	$E[XZ/Y]-E[X/Y]EZ$

Beginning with $EX=2$, $VX=1$, $X\in[0,3]$, $EY=5$, $VY=0.5$, $Y\in[2,7]$, $EZ=12$, $VZ=3$, $Z\in[10,20]$, and $\rho_{XY}=\rho_{XZ}=\rho_{YZ}=0.5$, the following covarainces can be calculated:

$$\begin{aligned}
\text{Cov}[K_1,Y] &= \text{Cov}[X+Y,Y]=VY+\text{Cov}[X,Y]=0.8536,\\
\text{Cov}[K_1,Z] &= \text{Cov}[X+Y,Z]=\text{Cov}[X,Z]+\text{Cov}[Y,Z]=1.4784,\\
\text{Cov}[K_2,K_1] &= \text{Cov}[Z-Y,K_1]=\text{Cov}[Z,K_1]-\text{Cov}[Y,K_1]=0.62484.
\end{aligned}$$

The product $W=K_1\times K_2$ may then be evaluated with $\text{Cov}[K_1,K_2]$, which gives the following moments:

$$EW=[49.62484,49.62485],\qquad VW=[0,1066.67],$$

which is a significant improvement of the results without covaraince tracking:

$$EW=[44.836,53.164],\qquad VW=[0,2897.8763].$$

Figure 6 shows the improvement in the corresponding p-boxes. The p-box bounds may also be further reduced by using sub-intervalisation for the range estimate. This additional contraction is shown on the right of Figure 6.

Table 8 summaries covariance algebra expressions of outputs of operations and another variable. These formulas may be derived along the same lines as (38). Note that $\text{Cov}[XY,Z]$ may also be derived from the results of Bohrnstedt and Goldberger (1969), who provide an exact covariance of the product of three variables.

7.2 Automating covariance tracking

The covariance formulas of Tables 7 and 8 are quick to evaluate, and may be called by software whenever a new variable is created either from a unary transformation, or from a binary operation between two moment variables. The formulas in Table 7 may be used to calculate the dependence between the operands and outputs of operations, and Table 8 find the dependence between any other variable not used in the operation. Whenever the formulas produce bounds outside of $[-\sqrt{VZVX},+\sqrt{VZVX}]$, software may automatically contract the range to at least this bound. This dependency information may be stored in a partially defined covariance matrix, the missing elements of which may be filled using Tables 7 and 8. Whenever a new variable is generated after an operation, new entries for this variable can be added to this covariance matrix. To illustrate this, consider the previous example $W=(X+Y)(Z-Y)$. In a computer program, this would be evaluated as $K_1=X+Y$, $K_2=Z-Y$, followed by $W=K_1\times K_2$. Say that initially the covariance matrix of X,Y and Z is known:

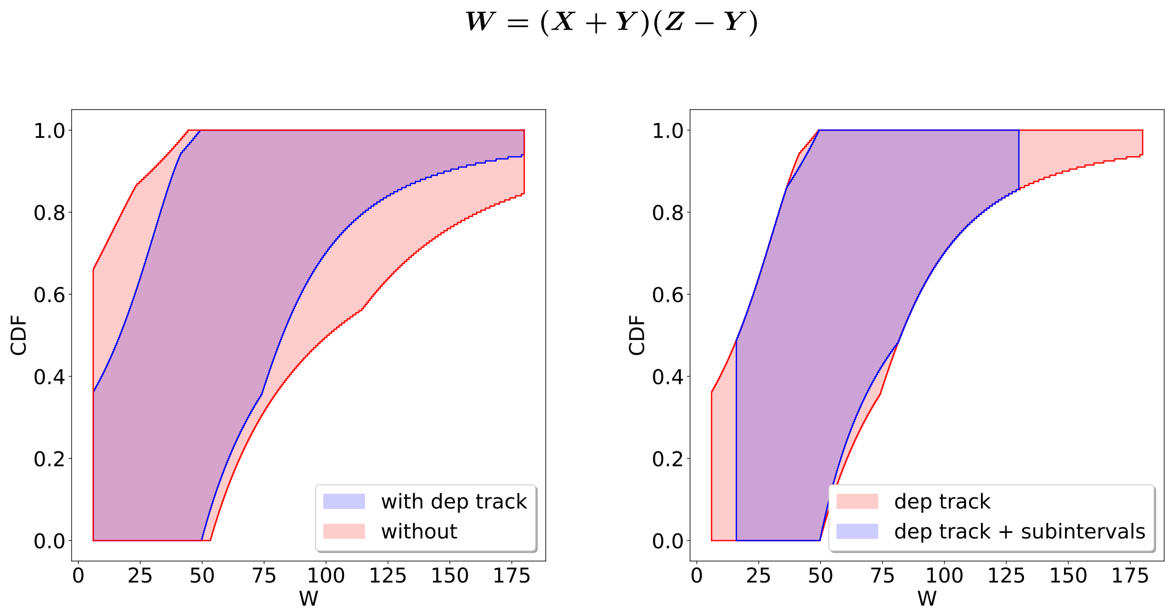

Figure 6: Left shows a comparison arithmetic with covariance tracking (blue) and without (red), for $W = (X+Y)(Z-Y)$. Right shows a further contraction when sub-intervalisation is used for the range.

$$\begin{pmatrix} VX & \cdots & \cdots \\ \text{Cov}[X,Y] & VY & \cdots \\ \text{Cov}[X,Z] & \text{Cov}[Y,Z] & VZ \end{pmatrix}.$$

When the first operation is called $K_1 = X + Y$, the covariance $\text{Cov}[X,Y]$ may be looked-up in the matrix. After the operation new rows and columns may be added for K_1, with the entries found using the covariance Tables 7 and 8:

$$\begin{pmatrix} VX & \cdots & \cdots & \cdots \\ \text{Cov}[X,Y] & VY & \cdots & \cdots \\ \text{Cov}[X,Z] & \text{Cov}[Y,Z] & VZ & \cdots \\ VX + \text{Cov[X, Y]} & VY + \text{Cov[X, Y]} & \text{Cov}[X,Z] + \text{Cov}[Y,Z] & VK_1 \end{pmatrix}.$$

Similarly after the second operation, entries may be added for K_2:

$$\begin{pmatrix} VX & \cdots & \cdots & \cdots & \cdots \\ \text{Cov}[X,Y] & VY & \cdots & \cdots & \cdots \\ \text{Cov}[X,Z] & \text{Cov}[Y,Z] & VZ & \cdots & \cdots \\ \text{Cov}[X,K_1] & \text{Cov}[Y,K_1] & \text{Cov}[Z,K_1] & VK_1 & \cdots \\ \text{Cov}[X,Z] - \text{Cov}[X,Y] & \text{Cov}[Y,Z] - VY & VZ - \text{Cov}[Y,Z] & \text{Cov}[Z,K_1] - \text{Cov}[Y,K_1] & VK_2 \end{pmatrix},$$

and so on. This way covariances may always be known between variables in the computer program. One obvious issue with this is that as more variables are added, the number of covariances to be calculated rises quickly. However, this can be reduced by only calculating the necessary entries, and no more. For example, in the calculation of $(X+Y)(Z-Y)$, the covariances for $\text{Cov}[K_1, X]$, $\text{Cov}[K_2, X]$, $\text{Cov}[K_2, Y]$ and $\text{Cov}[K_1, Z]$ are unused, and may omitted from the calculation. In terms of automation, these entries may be left blank in the covariance matrix, and only calculated when required, for example if X and W are used in operation later. Whenever a binary operation between two variable is called, their covariance may be looked up in the matrix, and if missing calculated then. This way only the required number of entries is calculated.

8 Conclusions

The limitations of linearity and independence mentioned by Cullen and Frey (1999) are real and serious, but they can be relaxed. In this paper we extend moment propagation in several ways. We provide convenient tables for moment propagation formulae for the independence case and the case with no knowledge about dependence, and we suggest that one can combine the methods of moment propagation with elementary interval analysis to obtain results that are better than can be obtained from either analysis separately. We provide a method for bounding distributional information solely from moments and ranges, without assumptions about input distributions. Any non-linearity in the model will change distributional shape. In standard moment propagation, distributional information is usually only preserved through linear models, allowing for risks to be calculated. The methods of this paper relax such restrictions. Finally, we describe a crude form of dependency tracking based on calculating covariances between newly created and already existing variables. This covariance tracking, along with formulas for correlated moment propagation, may be used to reduce the effect of repeated variables. We show that using covariance tracking, the artificial uncertainty from repeated variables can be reduced even when no dependence information is initially known between inputs.

One important application of the methods to be developed in this paper is to the area of risk analysis. In this discipline, predictions are made about the magnitudes or probabilities of structural failures or other adverse extreme events such as patients receiving toxic doses of therapeutic drugs or endangered species going extinct. These forecasts are often computed from limited empirical information. In traditional “worst case" analyses, the elementary methods of interval analysis are applied to risk formulations estimating, for instance, the difference between a structure’s strength and some stress acting on it, or the delivered dose of a drug, or the population size of the endangered species, and so on. The methods described here provide a richer characterisation by the inclusion of the moments information, which may be used to inform risks of various kinds, particularly tail risks, which are most relevant in a risk analysis worried about worst-case outcomes.

Another useful application of moment arithmetic is its combination with p-box arithmetic (Gray et al., 2021b). In p-box arithmetic, cdf bounds of random variables are projected through expressions as a form of robust uncertainty propagation. Means and variances are calculated from p-box bounds, but the estimates are sometimes rather loose. When combined with the methods of this paper, means and variances may be tightly projected, which also inform cdf bounds, subsequently tightening p-boxes. Where p-box arithmetic provides tighter moments, the moment estimates from this method may also be tightened. Each method may inform and tighten the other, providing a more accurate analysis combined than either method alone. A similar argument can be made for rigorous possibilistic arithmetic (Hose and Hanss, 2019; Gray et al., 2021c), or any other rigorous bounding characterisation of an imprecisely known random variable which can be informed from moment and range information.

Software: `MomentArithmetic.jl`

We provide a Julia implementation of distribution-free risk analysis, `https://github.com/AnderGray/MomentArithmetic.jl`. Included is a custom Julia type which automatically checks the consistency of user provided moment information, and tightens it where possible. The package implements all of the formulas of this paper and uses Julia’s multiple dispatch to give an easy and automatic propagation of moments through general Julia functions with minimal user effort. After an operation, the moments are checked against theoretical bounds for self-verification, and also tightened where possible. Sub-intervalisation is used for formulas where variables are repeated and which use intervals. `IntervalArithmetic.jl` (Sanders et al., 2021) is used for interval arithmetic, and `ProbabilityBoundsAnalysis.jl` (Gray et al., 2021b) is used for bounding risks and p-box calculations.

Acknowledgements

Ander Gray would like to thank the support from the EPSRC iCase studentship award 15220067. We also gratefully acknowledge funding from UKRI via the EPSRC and ESRC Centre for Doctoral Training in Risk and Uncertainty Quantification and Management in Complex Systems. This research was supported by the EPSRC through grant EP/R006768/1, which is acknowledged for its funding and support. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

A Expansion of Goodman variance

The Goodman formula (Goodman, 1960) for the variance of product is:

$$V(XY) = (EX)^2VY + (EY)^2VX + 2EXEYE_{11} + 2EXE_{12} + 2EYE_{21} + E_{22} - E_{11}^2 \tag{40}$$

where E_{ij} are the higher bivariate moments: $E_{ij} = E[(X - EX)^i(Y - EY)^j]$ (e.g. E_{11} is covariance). These are generally not tracked by the method, however they may be expressed in terms of the marginal moments and the other formulae described here:

$$\begin{aligned} E_{11} &= E[(X - EX)(Y - EY)] \\ &= E[XY - XEY - YEX + EXEY] \\ &= E[XY] - EXEY \end{aligned}$$

$$\begin{aligned} E_{21} &= E[(X - EX)^2(Y - EY)] \\ &= E[X^2Y + E[X]^2Y - X^2E[Y] + 2E[X]E[Y]X - 2E[X]XY - E[X]^2E[Y]] \\ &= E[X^2Y] + E[X]^2E[Y] - E[X^2]E[Y] + 2E[X]^2E[Y] - 2E[X]E[XY] - E[X]^2E[Y] \\ &= E[X^2Y] - E[X^2]E[Y] + 2E[X]^2E[Y] - 2E[X]E[XY] \end{aligned}$$

$$\begin{aligned} E_{12} &= E[(X - EX)(Y - EY)^2] \\ &= E[XY^2 + XE[Y]^2 - E[X]Y^2 + 2E[X]E[Y]Y - 2E[Y]XY - E[X]E[Y]^2] \\ &= E[XY^2] + E[X]E[Y]^2 - E[X]E[Y^2] + 2E[X]E[Y]^2 - 2E[Y]E[XY] - E[X]E[Y]^2 \\ &= E[XY^2] - E[X]E[Y^2] + 2E[X]E[Y]^2 - 2E[Y]E[XY] \end{aligned}$$

$$\begin{aligned} E_{22} &= E[(X - EX)^2(Y - EY)^2] \\ &= E[E[X]^2E[Y]^2 - 2E[X]E[Y]^2X + E[Y]^2X^2 - 2E[X]^2E[Y]Y + 4E[X]E[Y]XY \\ &\quad -2E[Y]X^2Y + E[X]^2Y^2 - 2E[X]XY^2 + X^2Y^2] \\ &= E[X]^2E[Y]^2 - 2E[X]^2E[Y]^2 + E[Y]^2E[X^2] - 2E[X]^2E[Y]^2 + 4E[X]E[Y]E[XY] \\ &\quad -2E[Y]E[X^2Y] + E[X]^2E[Y^2] - 2E[X]E[XY^2] + E[X^2Y^2] \\ &= -3E[X]^2E[Y]^2 + E[X^2]E[Y]^2 + E[X]^2E[Y^2] + 4E[X]E[Y]E[XY] \\ &\quad -2E[Y]E[X^2Y] - 2E[X]E[XY^2] + E[X^2Y^2] \end{aligned}$$

The right hand sides of the above expressions may be evaluated with the formulas described in this paper.

References

Probability Allen, A. statistics and queuing theory with computer science applications, vol. 2, 1990.

Carol Ash. *The probability tutoring book: an intuitive course for engineers and scientists (and everyone else!)*. IEEE, 1993.

Dimitris Bertsimas, Karthik Natarajan, and Chung-Piaw Teo. Tight bounds on expected order statistics. *Probability in the Engineering and Informational Sciences*, 20(4):667, 2006.

George W Bohrnstedt and Arthur S Goldberger. On the exact covariance of products of random variables. *Journal of the American Statistical Association*, 64(328):1439–1442, 1969.

Pafnutii Lvovich Chebyshev. *Sur les valeurs limites des intégrales*. Imprimerie de Gauthier-Villars, 1874.

Alison C Cullen, H Christopher Frey, and Christopher H Frey. *Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs*. Springer Science & Business Media, 1999.

W Feller. An introduction t 0 probability theory and its application. 12.1, 1968.

Scott Ferson. *RAMAS Risk Calc 4.0 software: risk assessment with uncertain numbers*. CRC press, 2002.

Scott Ferson, Vladik Kreinovich, Lev Grinzburg, Davis Myers, and Kari Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2015.

Leo A Goodman. On the exact variance of products. *Journal of the American statistical association*, 55(292):708–713, 1960.

Ander Gray, Marco De Angelis, Scott Ferson, and Edoardo Patelli. What's $z - x$, when $z = x + y$? dependency tracking in interval arithmetic with bivariate sets. In *Proceedings of the 9th International Workshop on Reliable Engineering Computing (REC2021)*, 2021a.

Ander Gray, Scott Ferson, and Edoardo Patelli. Probabilityboundsanalysis.jl: Arithmetic with sets of distributions. *Submitted to the Proceedings of JuliaCon*, 2021b.

Ander Gray, Dominik Hose, Marco De Angelis, Michael Hanss, and Scott Ferson. Dependent possibilistic arithmetic using copulas. In *International Symposium on Imprecise Probability: Theories and Applications*, pages 169–179. PMLR, 2021c.

Nick Gray, Marco De Angelis, and Scott Ferson. Computing with uncertainty: introducing puffin the automatic uncertainty compiler. In *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP 2019)*. Institute of Structural Analysis and Antiseismic Research School of Civil ..., 2019.

Dominik Hose and Michael Hanss. Possibilistic calculus as a conservative counterpart to probabilistic calculus. *Mechanical Systems and Signal Processing*, 133:106290, 2019.

James M Hyman. Forsig: an extension of fortran with significance arithmetic. Technical report, Los Alamos National Lab., NM (USA), 1982.

Samuel Karlin and William J Studden. *Tchebycheff systems: With applications in analysis and statistics*, volume 15. Wiley, 1966.

Kyoko Makino. *Rigorous analysis of nonlinear motion in particle accelerators*. Michigan State University, 1998.

A Markoff. Sur une question de maximum et de minimum: Proposée par m. tchebycheff. *Acta Mathematica*, 9:57–70, 1900.

Ramon E Moore. *Interval analysis*, volume 4. Prentice-Hall Englewood Cliffs, 1966.

Neil C Rowe. Absolute bounds on the mean and standard deviation of transformed data for constant-sign-derivative transformations. *SIAM journal on scientific and statistical computing*, 9(6):1098–1113, 1988.

Siegfried M Rump and Masahide Kashiwagi. Implementation and improvements of affine arithmetic. *Nonlinear Theory and Its Applications, IEICE*, 6(3):341–359, 2015.

David P. Sanders, Luis Benet, and et. al. Juliaintervals/intervalarithmetic.jl: v0.19.2, September 2021. URL `https://doi.org/10.5281/zenodo.5519761`.

James E Smith. Moment methods for decision analysis. *Management science*, 39(3):340–358, 1993.

Samuel S Wilks. Mathematical statistics. new york: Johnwiley & sons'. *Inc. WilksMathematical Statistics1962*, 1962.

Phongtape Wiwatanadate and H Gregg Claycamp. Exact propagation of uncertainties in multiplicative models. *Human and Ecological Risk Assessment*, 6(2):355–368, 2000.