Distribution-free uncertainty propagation

Scott Ferson[1] and Ander Gray[2]
Institute for Risk and Uncertainty, University of Liverpool, UK
[1]*Scott.Ferson@liverpool.ac.uk*, [2]*Ander.Gray@liverpool.ac.uk*

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 of uncertainty propagation that do not require simulation, sampling or approximation of any kind. We describe best-possible bounds on exceedance risks (probabilities of extreme events) that can be inferred given only information about the range, mean and variance of a random variable. These bounds generalize the classical Chebyshev inequality in an obvious way, yet apparently have not been described elsewhere. 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 uncertain 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 generalizes formulas for the case when no assumptions are made about dependence. These formulas can be viewed as a distribution-free risk analysis.

Keywords: uncertainty propagation, moment propagation, distribution-free risk analysis, imprecise probabilities

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 (see, for example, 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 current moment propagation methodologies:

- Require stochastic independence
- Require moments to be perfectly known (point values)
- 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 characterized 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 characterized 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. Wiwatandate 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 I summarizes 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 emphasize 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 I 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 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 II review the basic arithmetic operations on moments without dependence assumptions. These formulas generally yield intervals rather than precise values. In part, the results

Table I. 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)	$\begin{cases} \mathbf{k}\underline{\mathbf{X}}, & \text{if } 0 \leq k \\ \mathbf{k}\overline{\mathbf{X}}, & \text{if } k < 0 \end{cases}$	$\begin{cases} \mathbf{k}\overline{\mathbf{X}}, & \text{if } 0 \leq k \\ \mathbf{k}\underline{\mathbf{X}}, & \text{if } k < 0 \end{cases}$
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 $0 \leq X$	$\sqrt{\underline{\mathbf{X}}}$	$\sqrt{\overline{\mathbf{X}}}$
$X + Y$	$\underline{\mathbf{X}} + \underline{\mathbf{Y}}$	$\overline{\mathbf{X}} + \overline{\mathbf{Y}}$
$X - Y$	$\underline{\mathbf{X}} - \overline{\mathbf{Y}}$	$\overline{\mathbf{X}} - \underline{\mathbf{Y}}$
$X \times Y$	$\mathbf{min}(\underline{\mathbf{X}}\underline{\mathbf{Y}}, \underline{\mathbf{X}}\overline{\mathbf{Y}}, \overline{\mathbf{X}}\underline{\mathbf{Y}}, \overline{\mathbf{X}}\overline{\mathbf{Y}})$	$\mathbf{max}(\underline{\mathbf{X}}\underline{\mathbf{Y}}, \underline{\mathbf{X}}\overline{\mathbf{Y}}, \overline{\mathbf{X}}\underline{\mathbf{Y}}, \overline{\mathbf{X}}\overline{\mathbf{Y}})$
$\frac{X}{Y}$ 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 $0 < Y$	$\mathbf{min}(\underline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \underline{\mathbf{X}}^{\overline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\overline{\mathbf{Y}}})$	$\mathbf{max}(\underline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \underline{\mathbf{X}}^{\overline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\underline{\mathbf{Y}}}, \overline{\mathbf{X}}^{\overline{\mathbf{Y}}})$
$\min(X, Y)$	$\mathbf{min}(\underline{\mathbf{X}}, \underline{\mathbf{Y}})$	$\mathbf{min}(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$
$\max(X, Y)$	$\mathbf{max}(\underline{\mathbf{X}}, \underline{\mathbf{Y}})$	$\mathbf{max}(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$

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 tabled formulas or because they were imprecisely estimated from statistical data or by subjective judgment.

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:

Table II. 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	$\mathbf{(EX)^2 + VX}$	rowevar(square)
$\lvert X\rvert$ (absolute value)	$\begin{cases} \mathbf{EX}, & \text{if } 0 \leq \underline{X} \\ -\mathbf{EX}, & \text{if } \overline{X} \leq 0 \\ [\lvert EX\rvert, \lvert EX\rvert + \sqrt{VX}(\pi - \text{atan}(\frac{\lvert EX\rvert}{\sqrt{VX}}))], & \text{if } 0 \in X \end{cases}$	$\max(0, EX^2 + VX - E(\lvert X\rvert)^2)$
$\sqrt{X}$ for $0 \leq X$	rowe($\surd$)	rowevar($\surd$)
$X + Y$	$\mathbf{EX + EY}$	$\mathbf{(\sqrt{VX} \pm \sqrt{VY})^2}$
$X - Y$	$\mathbf{EX - EY}$	$\mathbf{(\sqrt{VX} \pm \sqrt{VY})^2}$
$X \times Y$	$\mathbf{EXEY \pm \sqrt{VXVY}}$	"Goodman"
$\frac{X}{Y}$ 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"	$\text{env}(\max(VX, VY), 0)$
$\min(X, Y)$	"Bertsimas min"	$\text{env}(\max(VX, VY), 0)$

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

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

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 second derivatives.

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

$$V(XY) = (EX)^2VY + (EY)^2VX + 2EXEYE_{11} + 2EXE_{12} + 2EYE_{21} + E_{22} - E_{11}^2 \tag{9}$$

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 Bertsimas (Bertsimas et al, 2006) formulae for the expectation of max is:

$$E[\max(X,Y)] = \text{env}(\max(EX,EY), \max(\overline{X},\overline{Y})) \cap (EX+EY-\text{env}(\min(EX,EY),\min(\underline{X},\underline{Y}))) \tag{10}$$

and min being:

$$E[\min(X,Y)] = \text{env}(\min(EX,EY), \min(\underline{X},\underline{Y})) \cap (EX+EY-\text{env}(\max(EX,EY),\max(\overline{X},\overline{Y}))) \tag{11}$$

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., Wiwatandate and Claycamp 2000), the formulas in Table II 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 II can be improved substantially. Table III 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 Wiwatandate and Claycamp (2000), that body mass and height are independent.

Analysts should take care to use assumptions of independence and the formulas of Table III only when justified by theoretical argument or comprehensive empirical information. In contrast,

Table III. 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 $0<Y$	$E(\exp(\ln(X)\times Y))$	$V(\exp(\ln(X)\times Y))$
$\max(X,Y)$	$\begin{cases}\mathbf{EX}, & \text{if } Y<X\\ \mathbf{EY}, & \text{if } X<Y\\ \text{“Bertsimas max”}, & \text{otherwise}\end{cases}$	$\begin{cases}\mathbf{VX}, & \text{if } Y<X\\ \mathbf{VY}, & \text{if } X<Y\\ \text{env}(\max(VX,VY),0), & \text{otherwise}\end{cases}$
$\min(X,Y)$	$\begin{cases}\mathbf{EX}, & \text{if } X<Y\\ \mathbf{EY}, & \text{if } Y<X\\ \text{“Bertsimas min”}, & \text{otherwise}\end{cases}$	$\begin{cases}\mathbf{VX}, & \text{if } X<Y\\ \mathbf{VY}, & \text{if } Y<X\\ \text{env}(\max(VX,VY),0), & \text{otherwise}\end{cases}$

the formulas of Tables I and II 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 one must propagate uncertainty through multiple arithmetic operations, one needs to be able to handle interval estimates for the moments. 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 (such as in the variance of the product in Table III), then the interval 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 2021). Where possible, expressions can be rearranged in such a way that the variables appear only once, for example realising that $a^2+a=(a+\frac{1}{2})^2-\frac{1}{4}$. This process may be automated by an uncertainty compiler, as suggested by 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 only require 2 variables (EX and VX) to be sub-intervalised, this is an appropriate technique here.

4. 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 Markov (1887) 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, page 152; Allen 1990, page 79). The upper bound on the probability that the variable X will exceed a value as large as x is:

$$Prob(x \leq X) \leq \begin{cases} 1/(1+(x-EX)^2/VX), & \text{if } EX < x \\ 1, & \text{if } x \leq EX \end{cases} \tag{12}$$

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

$$Prob(x \leq X) \geq \begin{cases} 1/(1+VX/(x-EX)^2), & \text{if } x < EX \\ 1, & \text{if } EX \leq x \end{cases} \tag{13}$$

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 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:

$$Prob(x \le X) \le \begin{cases} 0, & \text{if } x \le 0 \\ 1/(1+(x-EX)^2/VX), & \text{if } 0 \le x \le EX \\ 1, & \text{if } EX < x \end{cases} \quad (14)$$

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:

$$Prob(x \le X) \ge \begin{cases} 0, & \text{if } x \le EX \\ 1 - EX/x, & \text{if } EX \le x \le EX + VX/EX \\ 1/(1+VX/(x-EX)^2), & \text{if } EX + VX/EX < x \end{cases} \quad (15)$$

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 LX to it. If it is the maximum, rather than the minimum that is known, we can use the encoding:

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

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.

Using the above formulation, it is possible to construct a p-box using a minimum, mean and variance and one using the maximum, mean and variance. A p-box using both endpoints, mean and variance can thus be found by intersecting these two p-boxes. Figure 1. shows this for range $= [0, 6]$, $EX = 3$ and $VX = 5$.

5. Conclusions

The limitations of linearity and independence mentioned by Cullen and Frey are real and serious, but they can be relaxed. In this paper we bring the following extensions to moment propagation:

- Independence between variables need not be assumed.

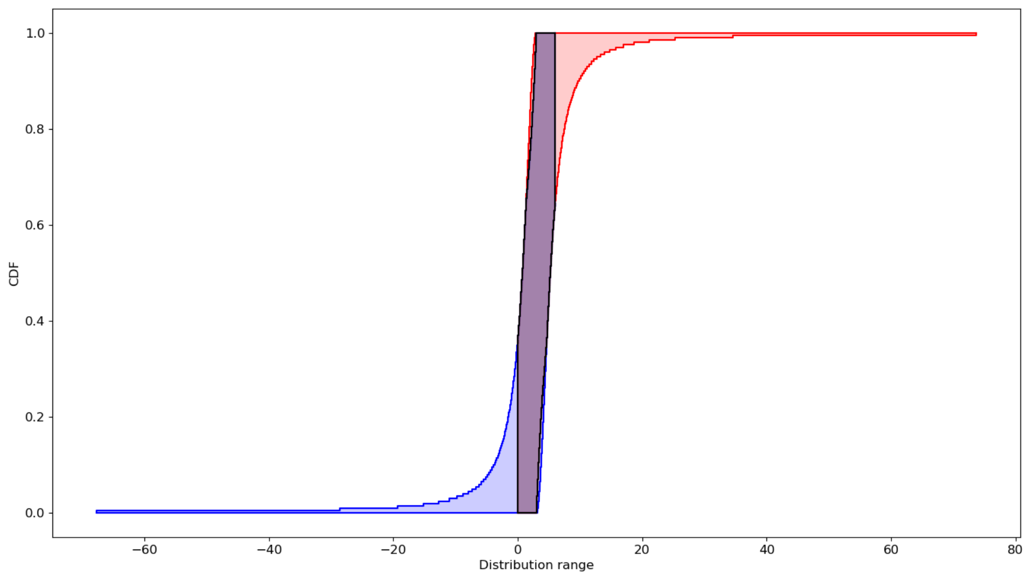

Figure 1. Shows a p-box a constructed using minimum, maximum, mean and variance information (black). Constructed by intersecting the p-box using Cantelli's inequalities with minimum, mean and variance (red) and maximum, mean and variance(blue)

– Moment propagation formulae may be evaluated with intervals.

– Assumptions about input distributions is no longer necessary.

– Distributional information may be obtained from moment and range information in models other than linear.

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, etc. In this paper, we provide convenient tables for moment propagation formulae for the independence case as well as 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, allowing for risks of extreme values to be calculated in models other than linear and without assumptions about input distributions, as is required by standard moment propagation practices.

Acknowledgements

The authors would like to thank the gracious 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.

References

Allen, A.O. Probability, Statistics and Queueing Theory with Computer Science Applications, second edition., Academic Press, Boston (1990)

Bertsimas, D., Natarajan, K., Teo, C.P., Tight bounds on expected order statistics, Probability in the Engineering and Informational Sciences 20 (4) (2006) 667.

Bogomolov, S., Forets, M., Frehse, G., Potomkin, K., Schilling, C., Juliareach: a toolbox for set-based reachability, in: Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control, 2019, pp. 39–44.

Chebyshev, P. L., Sur les valeurs limites des intégrales, Imprimerie de Gauthier-Villars, 1874.

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

Feller, W. An Introduction to Probability Theory and Its Applications. Volume 2, Second edition., John Wiley Sons, New York (1968)

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

Gray, A. and De Angelis, M., Ferson, S., and Patelli, E., What's Z X, when $Z = X + Y$? Dependency tracking in interval arithmetic with bivariate sets, Reliable Engineering Computing (2021).

Gray, N. and De Angelis, M. and Ferson, S., *Computing with Uncertainty: introducing Puffin the Automatic Uncertainty Compiler*. UNCECOMP, 2019.

Hyman, J. M., Forsig: an extension of fortran with significance arithmetic, Tech. rep., Los Alamos National Lab., NM (USA) (1982).

Makino, K., Rigorous analysis of nonlinear motion in particle accelerators. PhD thesis, Michigan State University, 1998.

Markoff, A., et al., Sur une question de maximum et de minimum: Proposée par m. tchebycheff, Acta Mathematica 9 (1887) 57–70.

Moore, R. E., Interval analysis, Vol. 4, Prentice-Hall Englewood Cliffs, 1966.

Rowe, N. C., 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) (1988) 1098–1113.

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

Wilks, S. S., Mathematical statistics. new york: Johnwiley & sons', Inc. Wilks Mathematical Statistics (1962).

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