How to Estimate Expected Shortfall When Probabilities Are Known with Interval or Fuzzy Uncertainty

Christian Servin
Information Technology Department
El Paso Community College
El Paso, TX 79915, USA
cservin@gmail.com

Hung T. Nguyen
Department of Mathematical Science
New Mexico State University
Las Cruces, NM 88003, USA
and Faculty of Economics
Chiang Mai University, Thailand
hunguyen@nmsu.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 7996, USA
vladik@utep.edu

***Abstract*—To gauge the risk corresponding to a possible disaster, it is important to know both the probability of this disaster and the expected damage caused by such potential disaster ("expected shortfall"). Both these measures of risk are easy to estimate in the ideal case, when we know the exact probabilities of different disaster strengths. In practice, however, we usually only have a partial information about these probabilities: we may have an interval (or, more generally, fuzzy) uncertainty about these probabilities. In this paper, we show how to efficiently estimate the expected shortfall under such interval and/or fuzzy uncertainty.**

I. Formulation of the Problem

How to gauge risk. In the ideal world, there should be no risk: all engineering designs should be 100% reliable. To achieve such reliability, civil engineers use the record of historic floods, tsunamis, hurricanes, earthquakes, and other natural disasters to estimate the largest possible strength of such a disaster, and design the buildings, bridges, and other structures so that they can withstand such disasters.

The historic experience shows, however, that there is always a possibility that the disaster strength S exceeds the estimated threshold s_0: this was the reason why the hurricane Katrina devastated New Orleans, why in 2011, Fukushima nuclear power station in Japan was destroyed by an unusually high tsunami, etc.

Since we cannot have a threshold s_0 that would *guarantee* that the disaster strength never exceeds s_0, the next best thing is to select a threshold s_0 for which the probability of exceeding s_0 does not exceed a given small number p_0, i.e., for which, we probability $p \stackrel{\text{def}}{=} 1 - p_0 \approx 1$, we have $S \le s_0$.

The choice of the threshold probability p_0 depends on the situation. For example:

- for manned space flights, NASA selected $p_0 = 10^{-3}$; smaller values were not feasible because of high uncertainty associated with space flights;
- on the other extreme, for reliability of a cell forming a computer memory, we need $p_0 \ll 10^{-9}$, because otherwise, if we allow $p_0 \gg 10^{-9}$, at least one of the billions of cells will always go wrong.

In addition to knowing the threshold s_0, it is also desirable to also know how much damage will come, on average, if this threshold is exceeded. For each possible value S of the corresponding disaster strength, we can estimate the corresponding damage X; the stronger the disaster, the larger the damage. Let x_p denote the damage corresponding to the threshold value s_0; then, the condition $S \ge s_0$ is equivalent to $X \ge x_p$.

In these terms, the probability that the disaster strength exceeds the threshold s_0 is equal to $\text{Prob}(X > x_p)$. In addition to this probability, it is desirable to also know the conditional expectation of the damage under the condition that the disaster strength exceeds the threshold x_p, i.e., the value

$$\text{ES}_p \stackrel{\text{def}}{=} E[X \mid X \ge x_p].$$

The corresponding conditional expectation is known as *expected shortfall*. These two values:

- the threshold x_p, and
- the expected shortfall ES_p,

is how we gauge the risk.

Similar two measures are used in finance to describe the risk that an investment would result in a big loss; see, e.g., [4].

How to estimate the expected shortfall in the ideal case, when we know the probability distribution describing damage. In the ideal case, we know the probability distribution that describes possible values of the damage X. A usual way to describe a probability distribution is by describing its *cumulative distribution function* (cdf) $F(x) \stackrel{\text{def}}{=} \text{Prob}(X \le x)$; see, e.g., [6].

In terms of cdf, the probability of exceeding the threshold value x_0 is simply equal to $1 - F(x_p)$. Thus, we have $1 - F(x_p) = p_0$ and hence, $F(x_p) = 1 - p_0 = p$. For each probability p, the value x_p for which $F(x_p) = p$ is known as the p-th *quantile*. For example:

- for $d = 0.5$, we get the median;
- for $d = 0.25$ and $d = 0.75$, we get *quartiles*,

etc. In mathematical terms, the function that maps d to x_d is an *inverse* function to the cdf $F(x)$.

The conditional expectation can then be computed as the ratio

$$\frac{\int_{x_p}^{\infty} x\,dF(x)}{1-p}. \tag{1}$$

In practice, we only have partial information about the probabilities. In practice, we rarely know the exact values of all the probabilities, we only have partial information about these probabilities. This may mean that, instead of the exact values $F(x)$ corresponding to different values x, we only know an *interval* $[\underline{F}(x), \overline{F}(x)]$ that contains the actual (unknown) value $F(x)$. Such situation when, for each d, we only know the corresponding intervals, is known as a *probability box* or, for short, a *p-box*; see, e.g., [1], [2].

Even more generally, for each x, we may have several intervals $[\underline{F}(x), \overline{F}(x)]$ corresponding to different degrees of certainty $\alpha \in [0,1]$, i.e., in effect, a *fuzzy number*; see, e.g., [3], [5], [7].

How to gauge risk under such an uncertainty? For different distributions $F(x) \in [\underline{F}(x), \overline{F}(x)]$ within a given p-box, we get different values of quantiles x_p for which $F(x_p) = p$. One can easily check that:

- the smallest value x_p corresponds to the largest values $\overline{F}(x)$ of the cdf; while
- the largest value x_p corresponds to the smallest values $\underline{F}(x)$ of the cdf.

Thus, possible values of the quantile x_p form an interval $[\underline{x}_p, \overline{x}_p]$ in which $\overline{F}(\underline{x}_p) = \underline{F}(\overline{x}_p) = p$.

Such quantile intervals are often useful when we perform computations with p-boxes; see, e.g., [1], [2].

We can use this idea to handle the case when we have a fuzzy-valued function $\mathbf{F}(x)$, if we take into account the known fact that for all possible computation $y = f(x_1, \ldots, x_n)$ with fuzzy numbers, the alpha-cut

$$\mathbf{y}(\alpha) \stackrel{\text{def}}{=} \{y : \mu(y) \geq \alpha\}$$

of the result is equal to the range

$$f(\mathbf{x}_1(\alpha), \ldots, \mathbf{x}_n(\alpha)) =$$

$$\{f(x_1, \ldots, x_n) : x_1 \in \mathbf{x}_1(\alpha), \ldots, x_n \in \mathbf{x}_n(\alpha)\}$$

of the values $f(x_1, \ldots, x_n)$ when each x_i belongs to the corresponding α-cut $\mathbf{x}_i(\alpha) = \{x_i : \mu_i(x_i) \geq \alpha\}$; see, e.g., [3], [5].

Thus, to find the α-cut of the quantile x_p, it is sufficient to compute the interval $[\underline{x}_p, \overline{x}_p]$ in situation when each $F(x)$ belongs to the corresponding α-cut of the fuzzy number $\mathbf{F}(x)$.

Such straightforward computation is possible because the dependence of the probability $p = F(x_p)$ on the unknown function $F(x)$ is monotonic (namely, increasing), and that each cdf $F(x)$ is also an increasing function. So, if we increase the values of the function $F(x)$, then for newly increased function $\widetilde{F}(x)$, we will have $\widetilde{F}(x_p) > F(x_p) = p$. Thus, to find the value $\widetilde{x}_p$ for which $\widetilde{F}(\overline{x}_p) = p$, we need to decrease the value x_p: $\widetilde{x}_p < x_p$.

So, to find the range of possible values $[\underline{x}_p, \overline{x}_p]$ of the quantile x_p, we do not need to enumerate all possible functions $F(x)$ from the p-box $[\underline{F}(x), \overline{F}(x)]$ – because of monotonicity, we know that:

- the smallest value $\underline{x}_p$ of the quantile x_p is attained when the cdf $F(x)$ attains its largest possible value, i.e., when $F(x) = \overline{F}(x)$ for all x; and
- the largest value $\overline{x}_p$ of the quantile x_p is attained when the cdf $F(x)$ attains is smallest possible value, i.e., when $F(x) = \underline{F}(x)$ for all x.

For expected shortfall, different probability distributions from the p-box, in general, lead to different values. We would like to find the range of possible values of the expected shortfall. Estimating this range, however, is not a very straightforward task: when we increase $F(x)$, it is not a priori clear whether the corresponding integral increases or decreases. Thus, we cannot immediately come up with a simple expression for the range of the expected shortfall.

It is therefore necessary to come up with efficient algorithms for estimating the range of the expected shortfall under p-box (interval) uncertainty. Similarly, for the fuzzy case, we need to be able to transform the fuzzy-valued cdf $\mathbf{F}(x)$ into a fuzzy value for the resulting expected shortfall.

What we do in this paper. In this paper, we provide the efficient algorithms for computing the expected shortfall under interval (p-box) and fuzzy uncertainty.

To be more precise, we produce an algorithm for computing the expected shortfall under interval (p-box) uncertainty. Based on what we mentioned earlier about fuzzy computations, to find the α-cut of the expected shortfall, it is sufficient to compute the range of possible values of ES_p of the expected shortfall in situation when each $F(x)$ belongs to the corresponding α-cut of the fuzzy number $\mathbf{F}(x)$. In other words, from the algorithmic viewpoint, the problem of computing the expected shortfall under fuzzy uncertainty can be indeed reduced to the case of interval (p-box) uncertainty.

II. Analysis of the Problem

Let us find an equivalent expression for ES_p. To find the range of possible values of expected shortfall ES_p, let us find an equivalent expression for ES_p that would simplify the computation of this range.

According to formula (1), we have

$$\mathrm{ES}_p = \frac{1}{1-p} \cdot I, \tag{3}$$

where

$$I \stackrel{\text{def}}{=} \int_{x_p}^{\infty} x\,dF(x). \tag{4}$$

Thus:

- the expected shortfall ES_p attains its smallest possible value $\underline{\mathrm{ES}}_p$ when the integral I attains its smallest possible value $\underline{I}$, and
- the expected shortfall ES_p attains its largest possible value $\overline{\mathrm{ES}}_p$ when the integral I attains its largest possible value $\overline{I}$.

The integral I has an infinite upper bound. This integral can be thus represented as a limit of integrals I_T with a finite upper bound T when $T \to \infty$:

$$I = \lim_{T\to\infty} I_T, \tag{5}$$

where

$$I_T \stackrel{\text{def}}{=} \int_{x_p}^{\infty} x\, dF(x). \tag{6}$$

Thus, for very large T, we have $I \approx I_T$.

The integral I_T can be integrated by part:

$$I_T = x \cdot F(x)|_{x_p}^{T} - \int_{x_p}^{T} F(x)\, dx =$$

$$T \cdot F(T) - x_p \cdot F(x_p) - \int_{x_p}^{T} F(x)\, dx. \tag{7}$$

For large T, we have $F(T)$ practically equal to

$$\lim_{T\to\infty} F(T) = 1,$$

so $T \cdot F(T) = T$ and

$$I_T = T - x_p \cdot F(x_p) - \int_{x_p}^{T} F(x)\, dx. \tag{8}$$

By definition of a quantile x_p, we have $F(x_p) = p$, so

$$I_T = T - x_p \cdot p - \int_{x_p}^{T} F(x)\, dx. \tag{9}$$

When does the integral I_T attain its largest and smallest possible values? According to our analysis:

- the expected shortfall ES_p attains its largest possible value when the value of the integral I_T (corresponding to a very large value T) is the largest possible, and
- the expected shortfall ES_p attains its smallest possible value when the value of the integral I_T (corresponding to a very large value T) is the smallest possible.

Let us thus analyze when the integral I_T attains its largest and smallest possible values.

When does the integral I_T attain its largest possible value? Let us start with the largest possible value. Different cdfs $F(x)$ from the given p-box result, in general, in different values of the integral I_T. Let $x_p^{\max}$ be the value corresponding to the cdf $F^{\max}(x)$ for which this integral is the largest possible. This means, in particular, that among all cdfs $F(x)$ with the same value of the p-th quantile $x_p^{\max}$ (i.e., for which $F(x_p^{\max}) = p$), this particular cdf $F^{\max}(x)$ leads to the largest possible value of the integral I_T.

One can easily see, from the expression (9), that the integral I_T is a decreasing function of the values $F(x)$. Thus, this integral is the largest when all the values $F(x)$ are the smallest. What limitations on the values $F(x)$ do we have?

- We have a limitation $\underline{F}(x) \le F(x) \le \overline{F}(x)$ coming from the fact that we only consider cdfs from a given p-box $[\underline{F}(x), \overline{F}(x)]$.
- We also have a limitation $F(x) \ge F(x_p) = p$, which, for values $x \ge x_p$, comes from the requirement $F(x_p) = p$ and from the fact that each cdf is an increasing function of x.

These constraints $\underline{F}(x) \le F(x) \le \overline{F}(x)$ and $F(x) \ge p$ can be equivalently described by a single constraint

$$\max(\underline{F}(x), p) \le F(x) \le \overline{F}(x). \tag{10}$$

Thus, the smallest possible values of $F(x)$ correspond to

$$F(x) = \max(\underline{F}(x), p). \tag{11}$$

When $\underline{F}(x) \ge p$, we have

$$\max(\underline{F}(x), p) = \underline{F}(x)$$

and hence $F(x) = \underline{F}(x)$. As we described earlier, the equality $\underline{F}(x) = p$ is equivalent to $x = \overline{x}_p$, thus the condition $\underline{F}(x) \ge p$ is equivalent to $x \ge \overline{x}_p$.

On the other hand, when $\underline{F}(x) < p$, i.e., equivalently, when $x < \overline{x}_p$, we have $F(x) = p$. Thus,

$$\int_{x_p^{\max}}^{T} F(x)\, dx = \int_{x_p^{\max}}^{\overline{x}_p} p\, dx + \int_{\overline{x}_p}^{T} \underline{F}(x)\, dx =$$

$$(\overline{x}_p - x_p^{\max}) \cdot p + \int_{\overline{x}_p}^{T} \underline{F}(x)\, dx. \tag{12}$$

Thus, the expression (9) takes the form

$$I_T = T - x_p^{\max} \cdot p - (\overline{x}_p - x_p^{\max}) \cdot p - \int_{\overline{x}_p}^{T} \underline{F}(x)\, dx. \tag{13}$$

The two terms $x_p^{\max} \cdot p$ and $(\overline{x}_p - x_p^{\max}) \cdot p$ can be easily combined into a single term $\overline{x}_p \cdot p$, so

$$I_T = T - \overline{x}_p \cdot p - \int_{\overline{x}_p}^{T} \underline{F}(x)\, dx. \tag{14}$$

Since $\overline{x}_p$ is the quantile corresponding to the lower endpoint $\underline{F}(x)$ of the p-box, we can therefore conclude that the expression (14) is the value of the integral I_T corresponding to $F(x) = \underline{F}(x)$.

Thus, the largest value of the integral I_T – and hence, of the expected shortfall – is attained when $F(x) = \underline{F}(x)$.

When does the integral I_T attain its smallest possible value? Let us now consider the smallest possible value. Let $x_p^{\min}$ be the value corresponding to the cdf $F^{\min}(x)$ for which this integral is the largest possible. This means, in particular, that among all cdfs $F(x)$ with the same value of the p-th quantile $x_p^{\min}$ (i.e., for which $F(x_p^{\min}) = p$), this particular cdf $F^{\min}(x)$ leads to the smallest possible value of the integral I_T.

Since the integral I_T is a decreasing function of the values $F(x)$, this integral is the smallest when all the values $F(x)$ are the largest. Under the limitations (10), the largest possible values are $F(x) = \overline{F}(x)$.

Thus, the smallest value of the integral I_T – and hence, of the expected shortfall ES_p – is attained when $F(x) = \overline{F}(x)$.

III. Resulting Algorithm

Problem: reminder. We want to find the range $[\underline{\mathrm{ES}}_p, \overline{\mathrm{ES}}_p]$ of possible values of the expected shortfall ES_p when cdf $F(x)$ is in the given p-box $[\underline{F}(x), \overline{F}(x)]$.

Conclusion. The above analysis leads to the following conclusion:

- The largest possible value $\overline{\mathrm{ES}}_p$ of the expected shortfall ES_p is attained when $F(x) = \underline{F}(x)$ for all x.
- The smallest possible value $\overline{\underline{\mathrm{ES}}_p}$ of the expected shortfall ES_p is attained when $F(x) = \overline{F}(x)$ for all x.

Thus, we arrive at the following algorithm:

Resulting algorithm.

- First, we compute the expected shortfall ES_p corresponding to $F(x) = \underline{F}(x)$. This shortfall we be the desired upper endpoint $\overline{\mathrm{ES}}_p$ of the desired interval $[\underline{\mathrm{ES}}_p, \overline{\mathrm{ES}}_p]$.
- Then, we compute the expected shortfall ES_p corresponding to $F(x) = \overline{F}(x)$. This shortfall we be the desired upper endpoint $\underline{\mathrm{ES}}_p$ of the desired interval $[\underline{\mathrm{ES}}_p, \overline{\mathrm{ES}}_p]$.

Comment. Of course, these two computations, of $\overline{\mathrm{ES}}_p$ and $\underline{\mathrm{ES}}_p$, do not have to be performed sequentially: if parallel computers are available, we can perform the computation of these two endpoints in parallel.

Acknowledgment

This work was supported in part by the Faculty of Economics of Chiang Mai University and by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

The authors are thankful to Scott Ferson and Paul Embrechts for valuable discussions.

References

[1] S. Ferson, *Risk Assessment with Uncertainty Numbers: RiskCalc*, CRC Press, Boca Raton, Florida, 2002.

[2] S. Ferson, V. Kreinovich, W. Oberkampf, and L. Ginzburg, *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*, Sandia National Laboratories, Report SAND2007-0939, 2007.

[3] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[4] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, Tools*, Princeton University Press, Princeton, New Jersey, 2010.

[5] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[6] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2011.

[7] L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.