Proceedings of the 9th International Conference
on Nonlinear Analysis and Convex Analysis
(Chiangrai, Thailand, 2015), 1–10

CONVEXITY AND RELIABILITY IN ENGINEERING OPTIMIZATION

R. TYRRELL ROCKAFELLAR

Abstract. An important idea in engineering is "probability of failure," which can be modeled as the probability that a particular random variable will exceed zero. Problem formulations in optimal engineering design often entail placing an upper bound on that probabilitity, or minimizing it subject to constraints involving other random variables that likewise depend on the design parameters. However, probability of failure can behave poorly in that context and even discontinuously.

A better alternative, now under development as a product of the convex analysis of random variables, is "buffered probability of failure." It is more conservative, has nicer properties, and is easier to compute and work with in optimization. Moreover it is attractive in concept because it takes into account not just a probability threshold but also the expected value of a random variable in the tail of its distribution beyond the threshold.

1. Introduction

Optimization is all about making decisions that are "best" from some perspective, within the limitations of resources. However, this can be a complicated matter in situations where the consequences of a decision will only play out in an uncertain future. That is where reliability of consequences must enter when setting up an optimization model.

Important examples in engineering are seen in the design of structures such as buildings, bridges, tunnels, reservoirs, vehical frames, ship hulls, airplane wings, offshore platforms, and so forth. Any design involves the specification of the values of many decision variables associated, for instance, with lengths, widths, thicknesses, and proportions of different materials. These affect the strength and durability of the structure, in particular, but how much strength

---

2010 *Mathematics Subject Classification.* 46N10, 91B30.

Key words and phrases. Optimization, reliability, probability of failure, buffered probability of failure, random variables, cumulative distribution functions, quantiles, superquantiles, VaR, CVaR, stochastic modeling, convex analysis .

and durability might really be needed? That question is essential to controlling costs, but it isn't easily answered. The stresses and impacts that a stucture might face in its lifespan are not known with certainty in advance.

A common approach in these circumstances is to base decisions on statistics of the past. A bridge can be designed to withstand a 100-year-flood, which is to say, the scale of flooding that seems to come every 100 years or so. A tower can be designed to withstand an earthquake of force 7 on the Richter scale, and geologists may be able to say how likely such an earthquake is in the next 50 years. Such prescriptions place constraints on the "probability of failure" of the structure, as will soon be explained. Alternatively, one might think of minimizing such probability subject to constraints on other aspects of the design like cost and performance.

Whether treating probability of failure as a constraint or as an objective in optimization, the critical thing to keep in mind is its dependence on the decision variables. Unfortunately, that dependence can be troublesome and lead to instability of results. Furthermore the concept itself has be criticized as taking only an abrupt threshold into account. There is no attention paid to the scale of cost or damage when "failure" occurs. This has led to the development of other ways of looking at reliability which can help to counteract such shortcomings.

The purpose of this article is to give an overview of these developments and the way they illustrate the role that convex analysis can have in serious problems of practical significance.

To get closer to the issues mathematically, we can think of the optimization framework as centered on the choice of "design variables" $x_1, \ldots, x_n$. The aspect of the design $x = (x_1, \ldots, x_n)$ that must be controlled for purposes of reliability is expressed by a function

$$c(x, v) = c(x_1, \ldots, x_n, v_1, \ldots, v_d) \tag{1.1}$$

involving additional "data variables" $v_1 \ldots, v_d$ which are not subject to choice and yet may be uncertain in their values, for instance by involving observations in the future which can only be made after the design has been finalized. Perhaps $c(x, v)$ calibrates a sort of hazard or a degree of damage beyond some acceptable level. Anyway we can imagine it has been set up so that

$$c(x, v) \leq 0 \text{ is deemed to be acceptable, but } c(x, v) > 0 \text{ signals "failure."} \tag{1.2}$$

The uncertainty in $v = (v_1, \ldots, v_d)$ can be modeled stochastically by saying that we really have a *vector-valued random variable* $V = (V_1, \ldots, V_d)$. Then $c(x, V)$ is an *x-dependent scalar-valued random variable* with its probability distribution induced by that of V. The usual focus for reliability, in line with (1.2), would be on the magnitude of

$$\text{probability of failure for } x = \text{prob}\big\{ c(x,V) > 0 \big\}. \tag{1.3}$$

In optimization one might to keep this below some specified upper bound or to choose x to minimize it subject to a collection of design constraints.

For an optimization model to be effective computationally and produce results that can be trusted, the functions selected for objective and constraints need to have sensible properties such as differentiability, or at least continuity. Convexity, too, can be extremely helpful when available. Can the probability in (1.3) be expected to enjoy such properties as a function of x? Unfortunately, there is little assurance of that, for intrinsic reasons explained in the next section.

Anyway, the probability distribution of the random variable $c(x,V)$ induced by that of V can in general only be "shaped" to some degree by the choice of x. That influence may be inadequate for countering, as much as desired, the dangers coming from the uncertainty of V. This underscores the importance of working with a failure concept that captures reliability in a robust manner. That ought to include taking into account the scale of $c(x,V)$ outcomes in the failure zone. After all, in setting up $c(x,v)$ to have 0 as the critical value in (1.2) there may not be a really unmovable criterion. Optimization should be able to respond to more than just the probability of a violation but also whether the violations are likely to be just modest in size or potentially catastrophic.

That is a major motivation for the "buffered" probability concept that will be described below, which offers a remedy with many advantages. It first appeared in [3] as an outgrowth of an approach to risk in random variables that was developed in [6] and [7] through the replacement of quantiles of random variables by "superquantiles."

2. Failure addressed through quantiles and superquantiles

Insight into the potential difficulties in the behavior of the failure probability in (1.3) as a function of x can be gained through a look at this kind of probability as associated with a general (scalar) random variable X. with the goal of later specializing to $X = c(x,V)$.

Definition 2.1 (probability of failure of a random variable)**.** For a random variable X, the probability of failure (with respect to 0 being the threshold for failure) is

$$p_f(X) = \text{prob}\big\{ X > 0 \big\}. \tag{2.1}$$

Consider in connection with this the cumulative distribution function of X, which is the function F_X from $(-\infty, \infty)$ to $[0,1]$ given by

$$F_X(q) = \text{prob}\{\, X \leq q \,\}. \tag{2.2}$$

It nondecreasing and has limit 0 as q goes down to $-\infty$ and limit 1 as q goes up to ∞. Obviously in (2.1) we have

$$p_f(X) = 1 - F_X(0). \tag{2.3}$$

Shifts in $p_f(X)$ relative to shifts in X can thus be understood by looking at how $F_X(0)$ might change with X. As an elementary test, we can look at shifting X to $X - a$, which changes $F_X(0)$ to $F_{X-a}(0) = F_X(a)$. Already here, trouble emerges. The trouble comes from the fact that F_X is, in general only continuous from the right and can have jumps where the limit from the left is less than the limit from the right. Such a jump occurs for F_X at 0 exactly when there is a so-called *probability atom* in the distribution of X at 0, i.e., when $\text{prob}\{\, X = 0 \,\} > 0$. The associated discontinuity in F_X at 0 signals a discontinuity in (2.3) in shifting from X to $X - a$.

Of course, the probability distibutions encountered in applications often have an associated probability density function ρ so that

$$F_X(q) = \int_{-\infty}^{q} \rho(r)dr, \tag{2.4}$$

in which case F_X is not only continuous everywhere but also differentiable everywhere. But at the opposite extreme it is often necessary to consider random variables X that are only known through a finite number of observations, whether historical or generated empirically through sampling. One or more of the components V_k of the data random vector $V = (V_1, \ldots, V_d)$ might be of that type, or even be a 0-1 variable associated with a discrete event. Such discreteness would be passed on to $X = c(x, V)$ in the form of probability atoms. Continuity, not to speak also of differentiability, with respect to x could then fall apart.

For working with probability of failure in a context of optimization, the message from these observations is that the assumptions needed for the justification of many solution algorithms may not be met. Moreover, if calculations are somehow carried out, the results might not be trustworthy. Tiny changes in decision variables, perhaps just due to “noise” in the numerics, could cause abrupt changes in the objective or constraint functions, thereby making solutions be unstable.

The complications we have been describing can also be examined from a different perspective through the connection between probability constraints and the quantiles of random variables.

Definition 2.2 (quantiles of a random variable)**.** For a random variable X and a probability level $p \in (0,1)$, the pth-quantile of X is

$$q_p(X) = \text{ lowest } q \text{ such that } F_X(q) \geq p. \tag{2.5}$$

If F_X is a continuous and always increasing on $(-\infty, \infty)$, its inverse F_X^{-1} exists and is well defined, in which case $q_p(X) = F_X^{-1}(p)$. However, F_X might not always be increasing and instead be constant on some intervals, so that for a particular p there would an interval of values of q with $F_X(q) = p$. Because F_X is continuous from the right, such an interval would at least always have a lowest q. On the other hand, there could be a discontinuity in F_X at q such that $F_X(r) \geq p$ for $r > q$ but $F_X(r) < p$ for $r < q$. Then $q_p(X) = q$ in (2.5).

These observations relate to our investigation of "failure" through the consequence of (2.5) that

$$p_f(X) \leq 1 - p \iff q_p(X) \leq 0, \tag{2.6}$$

Troubles in the behaviour of $p_f(X)$ can be thus be seen equally well as troubles in the behavior of the quantile $q_p(X)$ with respect to shifts in X or p.

A way of getting around this difficulty is to replace quantiles $q_p(X)$ by "superquantiles" $\bar{q}_p(X)$ that have superior properties derived through convex analysis. In the literature of finance, the quantile $q_p(x)$ is called the *value-at-risk* $\mathrm{VaR}_p(X)$, so the alternative, as developed for that subject in [6] and [7], was first therefore called, in parallel, the *conditional value-at-risk* and denoted by $\mathrm{CVaR}_p(X)$. Later, as it became apparent that the idea could be useful not just in finance but also in many branches of engineering, the application-indenpendent term "superquantile" with notation $\bar{q}_p(X)$ was offered in [3] as more convenient perhaps in that wider context, and that is also where the definition of "buffered" probability of failure was first published.

The superquantile concept requires extracting from the distribution of X its *upper p-tail distribution* for any $p \in (0,1)$. The goal behind that is to make sense of "worst-case portions" of the distribution of X, such as "the worst $(1-p)100\%$ of outcomes of X" (e.g. the worst 10% when $p = 0.9$). If F_X lacks discontinuities, that is easy. The upper p-tail distribution is simply the conditional probability distribution for X subject to $X \geq q_p(x)$, inasmuch as the interval $[q_p(X), \infty)$ then has probability $1-p$ exactly. However, if there is a probability atom for X at $q_p(X)$, the interval $[q_p(X), \infty)$ may have probability more than $1-p$ while the interval $(q_p(X), \infty)$ has probability less than $1-p$. Some adjustment is needed in that situation to capture with probability $1-p$ the worst outcomes of X having such likelihood. In essence, the adjustment is to "split the atom" to make the altered probability of the tail come out as $1-p$, as pointed out in [7]. Technically this can be carried out as follows.

Definition 2.3 (tail distributions)**.** The upper p-tail distribution of a random variable X at a probability level $p \in (0,1)$ is the probability distribution on $[q_k(X), \infty)$ for which the cumulative distribution function, derived from that of X, is

$$F_X^p(q) = \begin{cases} \frac{1}{1-p}[F_x(q) - p] \text{ for } q \geq q_p(X), \\ 0 \text{ for } q < q_p(X). \end{cases} \tag{2.7}$$

In the absence of a probability atom, this reverts to the conditional probability distribution with respect to X being in $[q_p(X), \infty)$. as indicated.

Definition 2.4 (superquantiles of a random variable)**.** For a random variable X and a probability level $p \in (0,1)$, the pth-superquantile is

$$\bar{q}_p(X) = \text{ average value in the upper } p\text{-tail distribution,} \tag{2.8}$$

or in other words, in figurative language, "the average outcome of X in the worst $(1-p)\%$ of instances." In the absence of a probability atom at the quantile $q_p(X)$, this is the conditional expectation of X with respect to having $X \geq q_p(X)$.[1]

This definition of the superquantile $\bar{q}_p(X)$ may seem fraught with more to worry about the definition of the quantile, which it moreover depends on. Luckily, though, a sort of miracle formula comes to the rescue in this situation,[2] as will be explained later.

With superquantiles in hand, we can proceed to define buffered probability of failure. A relation with quantiles that will play into it hinges on the "essential supremum" of a random variable X, namely

$$\sup X = \text{ lowest } q \in (-\infty, \infty] \text{ such that } X \leq q \text{ almost surely,} \tag{2.9}$$

which may be ∞. From the definitions of $q_p(X)$ and $\bar{q}_p(X)$ it is clear that $q_p(X) \leq \bar{q}_p(X)$ always, and that $\bar{q}_p(X)$ is continuous and nonincreasing with respect to p with

$$\bar{q}_p(X) \to \sup X \text{ as } p \to 1, \text{ but } \bar{q}_p(X) \to EX \text{ as } p \to 0. \tag{2.10}$$

In further detail, one has

$$\bar{q}_p(X) \geq EX \text{ and } q_p(X) < \bar{q}_p(X) < \sup X \qquad \text{as long as } 0 < p < 1 - \text{prob}\{\, X = \sup X \,\}, \tag{2.11}$$

[1]From this point on, we restrict attention to random variables X having $E|X| < \infty$, so as to be able to work securely with expected values.

[2]The splitting of the probability atom $\bar{q}_p(X)$ in the general case was proposed in [7]. Keeping instead, in all cases, to the conditional expectation of X that $X \geq q_p(X)$, produces what is called by some in finance the tail-VaR of X at level p. It does not lead to the highly supportive properties of superquantiles.

moreover with $\bar{q}_p(X)$ increasing (strictly) with respect to p in that domain. This covers all of $p \in (0,1)$ unless $\text{prob}\{\, X = \sup X \,\} > 0$, i.e., $\sup X$ is finite and there is a probability atom for X at that value, but in the exceptional case, however, (2.11) has to be complemented by

$$q_p(X) = \bar{q}_p(X) = \sup X \;\text{ for }\; p \geq 1 - \text{prob}\{\, X = \sup X \,\}. \tag{2.12}$$

However, our main interest lies in the fact that, for any q in $(EX, \sup X)$ there is a unique p such that $\bar{q}_p(X) = q$. That follows from the noted continuity and strict montonicity of the superquantile on the indicated domain.

Definition 2.5 (buffered probability of failure)**.** For a random variable X, the buffered probability of failure (with respect to 0 being the threshold for failure) is

$$\bar{p}_f(X) = \begin{cases} \text{the unique } 1-p \text{ such that } \bar{q}_p(X) = 0, \text{ as long as } EX \leq 0 < \sup X, \\ 0 \;\text{ if }\; \sup X \leq 0, \;\text{ but }\; 1 \;\text{ if }\; EX > 0. \end{cases} \tag{2.13}$$

Thus, $\bar{p}_f(X)$ gives, in the main case, the probability $1-p$ at which the worst $(1-p)\%$ of outcomes of X average out to 0 in the sense of the p-tail distribution.

In comparing buffered probability of failure with ordinary probability of falure, we therefore have

$$\bar{p}_f(X) \geq p_f(X) \;\text{ always.} \tag{2.14}$$

along with the characterization in parallel to (2.6) that

$$\bar{p}_f(X) \leq 1-p \quad\Longleftrightarrow\quad \bar{q}_p(X) \leq 0. \tag{2.15}$$

In engineering applications as described in Section 1, the usual focus for reliability is on the probability of failure associated with a design x as in (1.3), which in the notation now available can be expressed as the function

$$\varphi(x) = p_f(c(x,V)). \tag{2.16}$$

The alternative offered by the newer developments is to focus instead on the buffered probability of failure associated with x as given by

$$\bar{\varphi}(x) = \bar{p}_f(c(x,V)). \tag{2.17}$$

What would be the potential benefits of passing to this alternative in formulating objectives or constraints in a problem of optimization? One benefit, already suggested for motivation, is that the buffered version is able to bring into consideration the magnitude of the failures $c(x,V) > 0$ that could occur. In this respect it is more cautious and provides better safeguards in a design. Anyway, upper bounds on the buffered probability of failure automatically induce upper bounds on the ordinary probability of failure through (2.14).

Another benefit, with major impact, is that buffered probability of failure is vastly easier to work with in optimization because of various properties it enjoys by way of convex analysis. That leads to superior behavior of $\bar{\varphi}(x)$ in (2.17) as opposed to $\varphi(x)$ in (2.16) and is the subject of the next section.

3. Advantageous properties derived through convex analysis

Buffered probability of failure gains its power through the advantages that superquantiles $\bar{q}_p(X)$ have over quantiles $q_p(X)$. A central ingredient is a surprising minimization formula for computing $\bar{q}_p(X)$ which at the same time yields the corresponding $q_p(X)$.

In order to present the formula compactly, we need to look more closely first at quantiles. In parallel to $q_p(X)$ being defined to be the lowest q such that $F_X(q) \geq p$, we now also need for

$$q_p^+(X) = \text{ highest } q \text{ such that } F_X(q) \leq p \tag{3.1}$$

for $p \in (0,1)$. It is easy to see that $q_p^+(X) = q_p(X)$ except in the case of there being more than one q for which $F_X(q) = p$. Then the set of all such q is the closed, bounded interval $[q_p(X), \bar{q}_p^+(X)]$.

The following result was obtained in [6] for random variables with continuous distributions and extended in [7] to general distributions.[3]

Theorem 3.1 (minimization formula for superquantiles). *For a random variable X and any probability level $p \in (0,1)$, one has*

$$\min_{-\infty<C<\infty} \left\{ C + \frac{1}{1-p} E[\max\{0, X - C\}] \right\} = \bar{q}_p(X). \tag{3.2}$$

The minimum is indeed attained, and in fact

$$\operatorname*{argmin}_{-\infty<C<\infty} \left\{ C + \frac{1}{1-p} E[\max\{0, X - C\}] \right\} = [q_p(X), q_p^+(X)], \tag{3.3}$$

with this minimizing set reducing to just $q_p(X)$ unless $F_X(q) = p$ for more than one q.

The obvious significance of these expressions is that *superquantiles can be calculated without any need to deal with conditional probability distributions*, and without having to forgo knowing the corresponding quantiles. There is no need, in computational practice, to cope with the "split atom" in the definition of the upper p-tail distribution for X.

[3]Here E denotes expectation, and we persist in restricting attention to random variables X such that $E|X|$ is finite.

Another strong feature is that the expression in C and X that is minimized is convex as a function of (C, X). This leads to the next statement, where part (a) comes from [6] and [7] and part (b) is then evident from (2.15).

Theorem 3.2 (convexity consequences)**.**

(a) $\bar{q}(X)$ *is convex with respect to* X*:*

$$\bar{q}_p((1-\tau)X_0 + \tau X_1) \leq (1-\tau)\bar{q}_p(X_0) + \tau\bar{q}_p(X_1) \;\; \textit{for} \;\; \tau \in (0,1). \tag{3.4}$$

(b) $\bar{p}_f(X)$ *is quasi-convex with respect to* X*;*

$$\bar{p}_f((1-\tau)X_0+\tau X_1) \leq 1-p \;\; \textit{for} \;\; \tau \in (0,1) \;\; \textit{when} \;\; \bar{p}_f(X_0), \bar{p}_f(X_1) \leq 1-p. \tag{3.5}$$

Theorem 3.2 provides a stark contrast with the case of $q_p(X)$ and $p_f(X)$, for which such properties are far out of sight. Some properties do hold in common, for instance positive homogeneity

$$q_p(\lambda X) = \lambda q_p(X) \;\; \text{and} \;\; \bar{q}_p(\lambda X) = \lambda \bar{q}_p(X) \;\; \text{for} \;\; \lambda \geq 0, \tag{3.6}$$

and monotonicity with respect to the ordering $X_0 \leq X_1$ for random variables (meaning that the random variable $X_0 - X_1$ is ≤ 0 almost surely),

$$q_p(X_0) \leq q_p(X_1) \;\; \text{and} \;\; \bar{q}_p(X_0) \leq \bar{q}_p(X_1) \;\; \text{when} \;\; X_0 \leq X_1. \tag{3.7}$$

For superquantiles, these properties say that $\bar{q}_p(X)$ is "coherent" as a measure of risk in the sense of introduced in finance in [1].

Additional insights into the usefulness of superquantiles can be gleaned from [8]. An explanation of how the formula in Theorem 3.1 was originally deduced through consideration of conjugate functions in convex analysis is available in [4]

For applications to engineering design in the mode of Section 1, the combination of (3.7) with (3.4) yields the valuable fact, brought out in [7], that

$$c(x,v) \;\; \text{convex in} \;\; x \quad \Longrightarrow \quad \bar{q}_p(c(x,V)) \;\; \text{convex in} \;\; x. \tag{3.8}$$

Also in [7] is a prescription for how to employ this in a dramatic simplification of objectives and constraints when superquantiles are adopted in place of quantiles.

A result analogous to (3.8) holds for buffered probability of failure as well, namely

$$c(x,v) \;\; \text{convex in} \;\; x \quad \Longrightarrow \quad \bar{p}_f(c(x,V)) \;\; \text{quasi-convex in} \;\; x. \tag{3.9}$$

This is one of the significant ways that the function $\bar{\varphi}(x)$ in (2.17) is preferable to the function $\varphi(x)$ in (2.16) in an optimization framework.

More about risk-averse approaches to engineering design can be found in [5]. For some of the latest advances on buffered probability of failure and its generalization to "buffered probability of exceedance," see [2].

References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, *Coherent measures of risk*, Mathematical Finance **9** (1999), 203–227.

[2] A. Mafusalov and S. Uryasev, *Buffered probability of exceedance: mathematical properties and optimization algorithms*, Research Report 2014-1, ISE Dept., University of Florida.

[3] R. T. Rockafellar and J. O. Royset, *On buffered failure probability in design and optimization of structures*, Journal of Reliability Engineering and System Safety **99** (2010), 499–510.

[4] R. T. Rockafellar and J. O. Royset, *Random variables, monotone relations and convex analysis*, Mathematical Programming B **148** (2014), 297–331

[5] R. T. Rockafellar and J. O. Royset, *Engineering decisions under risk-averseness*, Journal of Risk and Uncertainty in Engineering Systems (accepted for publication 2014).

[6] R. T. Rockafellar and S. Uryasev, *Optimization of conditional value-at-risk*, Journal of Risk **2** (2000), 21–42.

[7] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, Journal of Banking and Finance **26** (2002), 1443–1471.

[8] R. T. Rockafellar and S. Uryasev, *The fundamental risk quadrangle in risk management, optimization and statistical estimation*, Surveys in Operations Research and Management Science **18** (2013), 33–53.

R. Tyrrell Rockafellar
University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350
E-mail address: rtr@uw.edu