Buffered Probability of Exceedance and Buffered Service Level: Definitions and Properties

Stan Uryasev

October 14, 2014

RESEARCH REPORT #2014-3

Department of Industrial and Systems Engineering
University of Florida, Gainesville, FL 32611

Abstract

This note is motivated by the Rockafellar and Royset (2010) paper, which defined so called Buffered Probability of Failure (bPOF). Compared to standard Probability of Failure, bPOF has exceptional mathematical properties. In optimization problems, a constraint on bPOF can be reduced to a convex constraint on the superquantile (Conditional Value-at-Risk). bPOF is an inverse function of superquantile at zero point. Rockafellar and Royset (2014) showed that the inverse function of superquantile is a a cumulative distribution function for an auxilary random variable which they call a superdistribution function (SDF). To help understand SDF, we discuss an interpretation of this function called Buffered Service Level (bSL), arising in the context of 'servers' that must meet an uncertain demand. Similar to bPOF, we also define Buffered Probability of Exceedance (bPOE) by equation bPOE = 1 – SDF. Mafusalov and Uryasev (2014) demonstrated that bPOE is a quasi-convex function of a random variable, and consequently, SDF is a quasi-concave function of a random variable. Therefore, bPOE and SDF can be very efficiently optimized. Norton and Uryasev (2014) showed that when a random function is positively homogeneous w.r.t. control variables, bPOE (and consequently SDF) optimization can be reduced to convex programming, and even linear programming. Several engineering areas, such as supply chain management, environmental and financial engineering, can benefit from these recent mathematical results. The objective of this note is to explain bPOE and SDF concepts and explore new mathematical properties of these notions.

Superdistribution Function (SDF) and Buffered Probability of Exceedance (bPOE)

Let W be some parameter (threshold) and L be a random value having continuous distribution. For instance, we have a server with capacity W and we want to satisfy some random demand L.

We define *Probability of Exceedance* (POE) for parameter W as

$$p_W(L) = Pr[L > W].$$

Accordingly, *Cumulative Distribution Function* (CDF) for parameter W equals

$$\alpha_W(L) = 1 - p_W(L) = 1 - Pr[L > W] = Pr[L \le W].$$

Suppose that W is the size of a server and L is random demand, then CDF, in this case, is called *Service Level*. For instance, when service level equals $\alpha_W(L) = 0.95$, demand is satisfied in 95% cases.

The defined two characteristics, W and $\alpha_W(L)$ do not provide any information about the magnitude of values of L exceeding W. Frequently, in engineering applications, the information on the possible values of L exceeding the threshold W may be quite important. In the context of a server, we can ask the question: How large should a server capacity be to cover demands exceeding W? By this, we mean an additional (emergency, redundant, auxiliary, "umbrella") server, working only when the main server can not satisfy the demand. Suppose that the demand L is normally distributed; in this case (theoretically), there is NOT a fixed capacity which can satisfy the demand for all possible values of L. You may say that the normal distribution is a pure mathematical concept which does not really exist in practice. However, quite frequently, in practical applications, losses have so-called "heavy tails", i.e., losses can be very high with very small probabilities (as is the typical case in financial or environmental applications; recall losses from the hurricane Katrina). We can pose a less demanding question: what is the capacity of the server to cover on average the demand exceeding W? We say that the server with capacity $V = E[L \mid L > W]$ covers *on average* the demand exceeding W, where $E[L \mid L > W]$ is the conditional expectation that the demand L exceeds level W. By definition, a server with capacity V has *Buffered Service Level* (bSL), $\bar{\alpha}_V(L)$, and *Buffered Probability of Exceedance* (bPOE), $\bar{p}_V(L) = 1 - \bar{\alpha}_V(L)$. Actually, bSL is a new distribution function, defined by Rockafellar and Royset (2014) which they call *Superdistribtion Function* (SDF). Here is the formal definition of SDF and bPOE.

Definition 1.

Assume L is a continuously distributed random value and V a threshold in the range, $EL \le V \le \sup L$. Suppose that for a different threshold $W(V)$, which is a function of V, the following equation is satisfied:

$$V = E[L \mid L > W(V)].$$

Then, by definition, SDF*, which is a function of* L *and* V, *equals*

$$\bar{\alpha}_V(L) = Pr[L \le W(V)],$$

and, by definition*,* bPOE *equals*

$$\bar{p}_V(L) = 1 - \bar{\alpha}_V(L) = Pr[L > W(V)].$$

bPOE answers the question: What is the fraction of largest outcomes which, on average, equals V ? The random excess over the threshold $W(V)$, by definition, equals:

$$[L - W(V)]^+ = \max[0, L - W(V)] \ .$$

For continuous distributions, the conditional expectation $V = E[L \mid L > W(V)]$ is related to the expected excess $E[L - W(V)]^+$ as follows:

$$V = E[L \mid L > W(V)] = W(V) + \frac{E[L - W(V)]^+}{Pr[L > W(V)]} = W(V) + \frac{E[L - W(V)]^+}{\overline{p}_V(L)} .$$

We will call V the *Buffered Parameter* or *Buffered Threshold.* Evidently the buffered threshold V exceeds $W(V)$, i.e. $V \geq W(V)$.

Therefore, with bPOE, $\overline{p}_V(L)$, we can express the expected excess, as follows:

$$E[L - W(V)]^+ = \overline{p}_V(L)[V - W(V)] \ . \qquad (1)$$

The *Additional Buffered Capacity*, $V - W(V) \geq 0$, is intended to cover the excess demand happening with the buffered probability $\overline{p}_V(L)$.

So, the expected excess $E[L - W(V)]^+$ equals the product of Additional Buffered Capacity, $V - W(V)$, and bPOE $\overline{p}_V(L)$. Or, equivalently, bPOE equals the normalized expected excess, which is stated in the following Theorem.

Theorem 1. bPOE can be expressed as follows:

$$\overline{p}_V(L) = \frac{E[L - W(V)]^+}{V - W(V)} \ , \quad \text{where } V = E[L \mid L > W(V)]. \qquad (2)$$

Proof. Proof of this statement for general distributions is included in Mafusalov and Uryasev (2014). An idea of the proof is sketched before the theorem statement.

Example. This example is discussed in detail in Davis and Uryasev (2014). The dataset includes damage loss data for 242 from land-falling hurricanes in USA. Average loss for 242 observation equals $6.96 billions. Table 1 contains hurricane damage loss ($ billions) and corresponding POE (%), and bPOE(%). We observe that bPOE is significantly higher than POE. Mafusalov and Uryasev (2014) proved that with reasonable assumptions, bPOE is at least two times larger than POE.

Damage ($ billions)	POE (%)	bPOE (%)
100.0	1.2	3.0
50.0	**4.2**	**9.8**
10.0	15.3	69.4
1.0	48.2	100.0
0.1	79.3	100.0

Table 1. USA land-falling hurricane damage loss (billions $), POE (%), and bPOE(%). bPOE is significantly higher than POE. Line with $50 billion loss (bold face): average loss can be calculated by assuming that there is $50 billion "constant" loss with probability 9.8%.

Let us discuss the $50 billion loss row in Table 1 (marked by bold face). POE for $50 billion loss threshold equals 4.2%. We see that losses exceeding $50 billion are quite rare and it looks like these excess losses may not impact significantly the overall average loss. bPOE equals 9.8%; therefore, the expected loss in the 9.8% upper tail equals 50*0.098= $4.9 billion. The lower tail has probability 100-9.8 = 90.2%. The expected loss in the lower tail equals the difference of total expected loss and expected loss in the upper tail: 6.96-4.9 = $2.06 billion. So, the conditional expected loss in the lower tail equals the expected loss in this tail divided by probability of this tail: 2.06/0.902 = $2.28 billion. The expected loss for the original loss distribution containing 242 observations can be equivalently "packed" to a discrete distribution with two atoms: lower tail expected loss $2.06 billion is the product of $2.28 billion (lower tail conditional expected loss) and probability of this tail 0.902 and the upper tail expected loss $4.9 billion is the product of $50 billion (upper tail conditional expected loss) and probability 0.098. Such representation shows importance of losses in the upper tail of the distribution. So, the expected loss equals:

$$2.28*0.902+50*0.098 = 2.06+4.9 = \$6.96 \text{ billion.}$$

Lower tail contains $242*0.902 \approx 218$ observations and the upper tail, 242-218 =24 observations. By ordering losses in the dataset we found that $18.06 billion loss splits upper and lower tails. Suppose we setup an insurance fund with $18.06 billion capital. Then, according to Theorem 1, the expected uncovered excess in the upper tail equals the product of bPOE=0.098 and difference of the conditional expected loss in the upper tail, $50 billion, and the insurance fund value $18.06 billion:

$$\text{expected excess} = 0.098*(50-18.06) = \$3.13 \text{ billion.}$$

The expected excess $E[L-W]^+$ is quite popular in various engineering applications; in financial applications it is called Partial Moment One, in stochastic programming it is called Integrated Chance. In statistics, Quantile Regression is based on the Koenker and Basset error function, which is actually a weighted sum of two expected excesses of a residual. Therefore, various engineering and mathematical areas can benefit from the concept of buffered probability of exceedance, which interprets the expected excess as a constant loss with some probability.

You may ask, what are the bSL and bPOE for the original server with capacity W ? Let us denote by $q_{\alpha}(L)$ the α -quantile of the random value L (also called Value-at-Risk or VaR in finance applications), which, by definition satisfies the equation

$$p_{q_{\alpha}(L)}(L) = Pr[L > q_{\alpha}(L)] = 1-\alpha \,.$$

Let us also denote by $\overline{q}_{\alpha}(L)$ the corresponding superquantile of the random value L (also called Conditional Value-at-Risk or CVaR in finance applications), which, equals

$$\overline{q}_{\alpha}(L) = E[L \,|\, L > q_{\alpha}(L)] \,.$$

So, by definition, $\overline{q}_{\alpha}(L)$, is the conditional average of the largest outcomes having probability $1-\alpha$. Now, with these notations, the bSL for capacity W*,* denoted by $\overline{\alpha}_{W}(L)$, satisfies the equation

$$\overline{q}_{\overline{\alpha}_{W}(L)}(L) = W \,, \tag{3}$$

and bPOE, $\overline{p}_{W}(L)$, for capacity W equals $\overline{p}_{W}(L) = 1-\overline{\alpha}_{W}(L)$. So, <u>by definition, bPOE</u>, $\overline{p}_{W}(L)$, <u>is a quantity satisfying the equation</u>

$$\overline{q}_{1-\overline{p}_{W}(L)}(L) = W \,. \tag{4}$$

The following inequality is always valid,

$$\overline{p}_{W}(L) \geq p_{W}(L) \,.$$

Therefore, bPOE, $\overline{p}_{W}(L)$, equals the POE, $p_{W}(L)$, plus the nonnegative *Buffer* = $\overline{p}_{W}(L) - p_{W}(L)$.

This observation motivates the term Buffered Probability of Excedance. Pay attention that the bSL, $\overline{\alpha}_{W}(L)$, is an <u>inverse function of the superquantile</u> $\overline{q}_{\alpha}(L)$, as defined by equation (3). Using as an input V, which is the conditional expected demand in the tail, we can rewrite the equation (1) in a general format as follows:

$$E[L - q_{1-\overline{p}_{V}(L)}(L)]^{+} = \overline{p}_{V}(L)[V - q_{1-\overline{p}_{V}(L)}(L)] \,. \tag{5}$$

Also, using as an input the confidence level α , the equation (1) can be presented in a general format as follows:

$$E[L - q_{\alpha}(L)]^{+} = (1-\alpha)[\overline{q}_{\alpha}(L) - q_{\alpha}(L)] \,, \tag{6}$$

where $\overline{p}_{\overline{q}_{\alpha}(L)}(L) = 1-\alpha$. In words, the last equation (6) reads as follows: the expected excess $E[L - q_{\alpha}(L)]^{+}$ over the quantile $q_{\alpha}(L)$ equals the product of bPOE, $\overline{p}_{\overline{q}_{\alpha}(L)}(L) = (1-\alpha)$, multiplied by the difference of superquantile $\overline{q}_{\alpha}(L)$ and quantile $q_{\alpha}(L)$. This sounds somewhat cumbersome, but actually the equation (6) provides a simple representation of the expected excess over the quantile.

Calculation and Optimization of Buffered Probability of Exceedance

Formula (2) can be used for the calculation of bPOE. This formula can be used, also, as an alternative definition of bPOE. A significant advantage of this formula comes from the following minimization representation of the buffered probability (which is stated in other format in Mafusalov and Uryasev (2014) and Norton and Uryasev (2014).

Theorem 2.

bPOE equals:

$$\bar{p}_V(L) = 0\,, \text{ if } V \geq \sup L\ ;$$

$$\bar{p}_V(L) = \min_{W<V} \frac{E[L-W]^+}{V-W}\,, \quad EL < V < \sup L \tag{6}$$

$$\bar{p}_V(L) = 1\,, \text{ otherwise.}$$

and an minimal $W(V)$ *in (6) equals* $W(V) = q_{1-\bar{p}_V(L)}(L)$ *and satisfies equation*

$$V = E[L \mid L > W(V)]\ .$$

Sketch of Proof.

The proof for general distributions is included in Mafusalov and Uryasev (2014). Here we provide only an idea of the proof for continuous distributions. Let us find and equation for and optimal W in (6) by taking derivative with respect to W and equating it to zero.

$$\nabla_W \frac{E[L-W]^+}{V-W} = \frac{\nabla_W E[L-W]^+}{V-W} + E[L-W]^+ \nabla_W \frac{1}{V-W} = -\frac{P[L>W]}{V-W} + \frac{E[L-W]^+}{(V-W)^2} = 0.$$

The last equation can be rearranged as follows:

$$V = W + \frac{E[L-W]^+}{P[L>W]} = W + E[L-W \mid L>W] = E[L \mid L>W]\ .$$

Formula (6) is proved.

Minimum formula (6) for bPOE is very similar to the minimum formula superquantile (CVaR) which was suggested in Rockafellar and Uryasev (2010). Minimum in this formula is achieved on quantile (VaR) with the 1-bPOE confidence level.

Optimization Example.

Let us demonstrate the power of formula (6) with the following *convex* optimization problem formulation obtained in Norton and Uryasev (2014). Let us consider that the loss $L(w,\theta) = w^T\theta$ is a linear function depending on control vector $w \in R^n$, $w \neq 0$ with a random vector of coefficients θ in R^n such that $E\theta \neq 0$. We are interested in minimizing the bPOE with $V = 0$, which corresponds to bPOF in Rockafellar and Royset (2010). If for some $w_* \neq 0$ we have $\sup w_*^T\theta \leq 0$, then $\min_{w\neq 0} \bar{p}_0(w^T\theta) = 0$ with $w_* \in \arg\min_{w\neq 0} \bar{p}_0(w^T\theta)$. Further on, suppose that $0 < \sup w^T\theta$ for $w \neq 0$. Then, there exists $\tilde{w} : E\tilde{w}^T\theta = \tilde{w}^T E\theta < 0$. According to Corollary 4 in Mafusalov and Uryasev (2014), for any random variable L, bPOE $\bar{p}_V(L)$ is a continuous strictly decreasing function of V at the interval $V \in [EL, \sup L)$. Therefore, $\bar{p}_0(\tilde{w}^T\theta) < 1$, and consequently, $\min_{w\neq 0} \bar{p}_0(w^T\theta) \leq \bar{p}_0(\tilde{w}^T\theta) < 1$.Note also that if $0 \leq Ew^T\theta = w^T E\theta$ for some $w \in R^n$, then, for all $W < 0$ we have

$$\frac{E[w^T\theta - W]^+}{0 - W} \geq \frac{E[w^T\theta - W]^+}{Ew^T\theta - W} = \frac{E[w^T\theta - W]^+}{E(w^T\theta - W)} \geq 1 .$$

This allows us to use formula (6) for bPOE minimization for all $w \neq 0$:

$$\min_{w\neq 0} \bar{p}_0(w^T\theta) = \min_{w\neq 0} \min_{W<0} \frac{E[w^T\theta - W]^+}{-W} = \min_{w\neq 0} \min_{W<0} E[\frac{1}{-W} w^T\theta + 1]^+ = \min_{w\neq 0} E[w^T\theta + 1]^+ .$$

Note also that since $E[0\theta + 1]^+ = 1$, then

$$\min_{w\neq 0} E[w^T\theta + 1]^+ = \min_{w\in R^n} E[w^T\theta + 1]^+ .$$

Finally,

$$\min_{w\neq 0} \bar{p}_0(w^T\theta) = \min_{w\in R^n} E[w^T\theta + 1]^+ .$$

The last optimization problem is a convex minimization problem in w, which can be even reduced to *linear programming* !!!

Rockafellar and Royset (2010) showed that for $V = 0$ a constraint on bPOF and a constraint on superquantile are equivalent. Similar statement for bPOE follows from formula (6) for all values of V, including $V = 0$ (it is formally proved in Mafusalov and Uryasev (2014)). Indeed, formula (6) implies that constraint

$$\bar{p}_V(L) \leq 1 - \alpha = const$$

can be replaced by the constraint

$$\frac{E[L-W]^+}{V-W} \le 1-\alpha \quad \Rightarrow \quad W + \frac{1}{1-\alpha} E[L-W]^+ \le V\,.$$

The last constraint is the constraint on the superquantile, see Rockafellar and Uryasev (2000).

References

[1] R.T. Rockafellar and J.O. Royset (2010), "On Buffered Failure Probability in Design and Optimization of Structures," Reliability Engineering & System Safety, Vol. 95, pp. 499-510.

[2] R.T. Rockafellar and J.O. Royset (2014), "Random Variables, Monotone Relations, and Convex Analysis," Mathematical Programming B, to apear.

[3] A. Mafusalov and S. Uryasev (2014), “Buffered Probability: Mathematical Properties and Optimization Algorithms," Working Paper (in preparation).

[4] M. Norton and S. Uryasev (2014), “Receiver Operating Characteristic (AUC) and Buffered AUC in Binary Classification," Working Paper (in preparation).

[5] R.T. Rockafellar and S. Uryasev (2000), “Optimization of Conditional Value-At-Risk.” The Journal of Risk, Vol. 2, No. 3, , 21-41.

[6] J.R. Davis and S. Uryasev (2014), “Analysis of Hurricane Damage using Buffered Probability of Exceedance.” Working Paper (in preparation).