Uncertainty modeling of LOCA frequencies and break size distributions for the STP GSI-191 resolution

Alexander Galenko, Elmira Popova
agalenko@gmail.com, elmira@mail.utexas.edu
The University of Texas at Austin

1 Introduction

In the initial quantification (Crenshaw, 2012), Fleming et al. (2011) performed a substantial study designed to build upon the established EPRI risk-informed in service inspection program (EPRI, 1999). This methodology was used as primary basis to develop the size and location-specific rupture frequencies (events per year) for the initial quantification. Although the overall methodology appears to be sound based on peer review (Mosleh, 2011) and reasonableness of the values obtained, NRC feedback in the Pilot Project reviews has resulted in further review of the approach. In this paper we propose a new approach to assign location specific LOCA frequencies that are derived from the overall frequencies, as defined in Tregoning et al. (2008). We refer to this source reference as NUREG-1829.

The NUREG-1829 annual frequencies are not plant specific or plant-location specific. Yet they are used throughout the nuclear industry as an important input to PRA analyses, and therefore, they need to be preserved. Conservation of the NUREG-1829 break frequencies is our guiding principle.

In this document we will work with the six break size categories, defined in Table 1, NUREG-1829 Volume 1, page xxi as “effective break size” for the

PWR plants. Table 1 shows the mapping between NUREG-1829 notation and ours. In addition, the use of the term distribution will be equivalent to a distribution function (either cumulative distribution or probability density function) of a random variable used to model a specified uncertainty.

Table 1: LOCA categories notation map

Effective break size (inch) for PWR	Notation
$\frac{1}{2}$	cat_1
$1\frac{5}{8}$	cat_2
3	cat_3
7	cat_4
14	cat_5
31	cat_6

We should point out that the South Texas Project PRA analysis uses only 3 LOCA categories, small, medium, and large. Our proposed methodology can be applied to any finite number of categories.

In this document we will use the term location to represent weld locations. Overall there are two ways to come up with location-specific LOCA frequencies: bottom-up and top-down. The bottom-up approach requires location specific failure data to estimate the corresponding probability of a weld failure. If we assume there are M_j different locations in the plant where breaks of size cat_j can occur, $location_1, \ldots, location_{M_j}$, then using the law of total probability we can write:

$$P[cat_j] = \sum_{i=1}^{M_j} P[cat_j|location_i]P[location_i], j = 1, 2, \ldots, 6$$

where $P[cat_j]$ is the total probability of a cat_j LOCA, $P[cat_j|location_i]$ is the

conditional probability of a cat_j LOCA given location i has been chosen, and $P[location_i]$ is the probability that location i is chosen. In our application we will assume that all locations are equally likely to be selected, i.e.

$$P[location_i] = \frac{1}{M_j}, i = 1, \ldots, M_j$$

In the bottom-up approach we first must determine $P[cat_j|location_i]$ (using estimation or expert elicitation). If that happens, then we can multiply these probabilities by $1/M_j$, sum them up, and we will obtain the total probability of a cat_j LOCA. If the bottom-up approach is followed we claim that the resulting total cat_j LOCA probability will NOT equal the number provided in NUREG-1829 (or at least it is very unlikely to get that number). The approach taken by Fleming et al. (2011) is an inherently bottom-up approach. To preserve the NUREG-1829 frequencies Fleming et al. (2011) developed an approximation scheme. In their review, the NRC technical team raised several questions about the bottom-up approach, which has lead us to propose a different methodology.

The method that we propose is rooted in the top-down approach: start with the NUREG-1829 frequencies and develop an intuitive way to distribute the cumulative frequencies among different weld locations.

We will use again the cat_j LOCA as an illustrative example. Assume that we have computed the total probability of a cat_j LOCA from the NUREG-1829 frequency table using the formula

$$P[cat_j] = \frac{Frequency[cat_j]}{\sum_{i=1}^{6} Frequency[cat_i]}$$

Again we assume there are M_j different locations in the plant where breaks of size cat_j can occur, $(location_1, \ldots, location_{M_j})$, and they are equally likely,

i.e.

$$P[location_i] = \frac{1}{M_j}, i = 1, \ldots, M_j.$$

Note that the denominator is the total frequency of all LOCA sizes. We will assume that all $P[cat_j|location_i] = P[cat_j], i = 1, \ldots, M_j$. Then applying the law of total probability we see that the resulting probability of a cat_j LOCA equals exactly the NUREG-1829 probability and $P[cat_j \;\; at \;\; location_i] = P[cat_j]/M_j$.

The above methodology distributes equally the LOCA frequencies as defined in NUREG-1829 Table 1 between all locations that can experience breaks from one or more of the six size categories. The six break size categories (columns in NUREG-1829 Table 1) are ranges bounded by six discrete points. For a particular weld we need to be able to sample from the continuous range of break size values. In addition, we would like to be able to sample from the distribution of the frequencies. The rows in Table 1 from NUREG-1829 represents the distribution of the frequencies by reporting the mean, median, 5the and 95the percenttiles. We will use this information to fit six continuous distributions for each break size category.

2 Proposed methodology

2.1 Fitting distributions to the LOCA frequencies

We will first describe the distribution fit to the frequencies for each break size category. In theory, there are infinite number of distributions that one can fit to the LOCA frequencies represented in NUREG-1829: two split Lognormal distributions are used in NUREG-1829 and Gamma distributions are used in NUREG/CR 6928.

We chose to fit the bounded Johnson distribution, Johnson (1949) for the following reasons:

- It has four parameters that will allow us to match closely the four distributional characteristics provided by NUREG-1829. In order to get the parameters of the Johnson distribution we solve an optimization problem with constraints defined by the four distributional characteristics: the 5th percentile, the mean, the median, and the 95th percentile.
- It can have variety of shapes. In particular, skewed, symmetric, bimodal, or unimodal shapes can be obtained.

The cumulative distribution function (CDF) of the bounded Johnson is:

$$F[x] = \Phi\left\{\gamma + \delta f[(x-\xi)/\lambda]\right\},$$

where $\Phi[x]$ - CDF of a standard Normal (0,1) random variable, γ and δ are shape parameters, ξ is a location parameter, λ is a scale parameter, and $f(z) = log[z/(1-z)]$. The fitted parameters of the Johnson distribution for each of the six categories are given in Table 2. The comparison between the NUREG-1829 distributional characteristics of the LOCA frequencies and the fitted ones are presented in Table 3. The largest error in our estimation is 3.78% which we consider to be small enough.

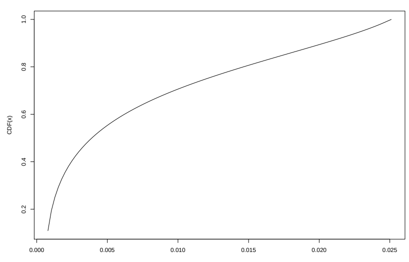
Figures 1 and 2 show the fitted CDFs and probability density functions (PDFs) of the Johnson distribution for each category.

Table 2: Fitted Johnson Parameters

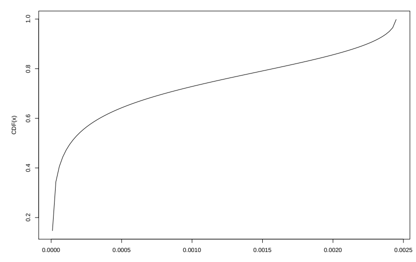
	Johnson Parameters			
	γ	δ	ξ	λ
Cat1	0.7288246	0.3893326	0.00063449	0.02449228
Cat2	6.95E-01	2.40E-01	7.41E-06	2.44E-03
Cat3	7.24E-01	2.44E-01	2.06E-07	6.24E-05
Cat4	7.14E-01	2.39E-01	1.36E-08	6.19E-06
Cat5	4.73E-01	2.69E-01	1.87E-10	5.93E-07
Cat6	4.75E-01	2.73E-01	1.77E-14	8.52E-08

Table 3: NUREG-1829 and fitted Johnson mean, 5th and 95th percentile values

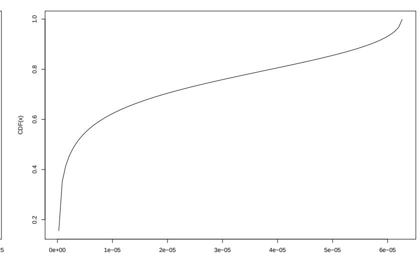
	NUREG-1829			**Fitted Johnson**			**Error**		
	5%	**Mean**	**95%**	**5%**	**Mean**	**95%**	**5%**	**Mean**	**95%**
Cat1	6.90E-04	7.30E-03	2.30E-02	6.89E-04	7.30E-03	2.30E-02	0.08%	0.01%	0.00%
Cat2	7.60E-06	6.40E-04	2.40E-03	7.56E-06	6.42E-04	2.40E-03	0.59%	0.38%	0.04%
Cat3	2.10E-07	1.60E-05	6.10E-05	2.09E-07	1.60E-05	6.12E-05	0.40%	0.23%	0.38%
Cat4	1.40E-08	1.60E-06	6.10E-06	1.40E-08	1.59E-06	6.08E-06	0.26%	0.65%	0.39%
Cat5	4.10E-10	2.00E-07	5.80E-07	4.14E-10	1.98E-07	5.86E-07	0.94%	0.94%	0.98%
Cat6	3.50E-11	2.90E-08	8.10E-08	3.59E-11	2.84E-08	8.41E-08	2.60%	2.03%	3.78%



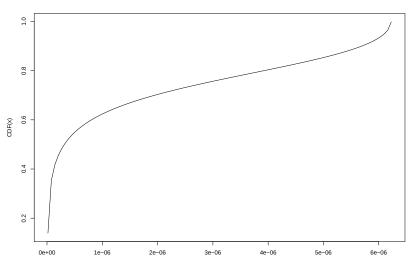
(a) Cat1



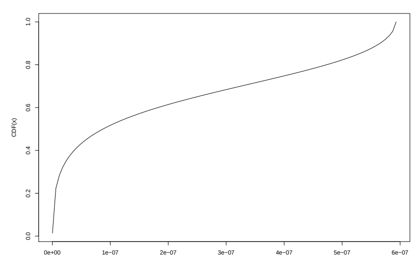
(b) Cat2



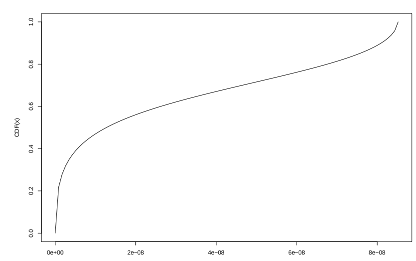
(c) Cat3



(d) Cat4

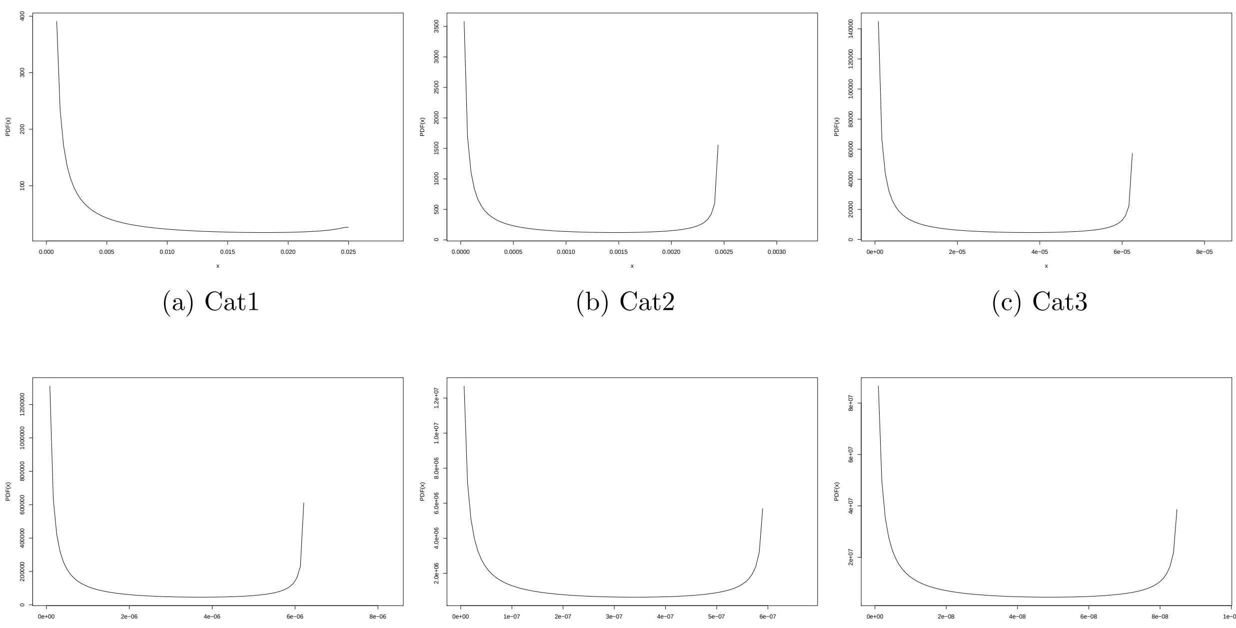


(e) Cat5



(f) Cat6

Figure 1: Johnson CDFs for each category



(a) Cat1 (b) Cat2 (c) Cat3

(d) Cat4 (e) Cat5 (f) Cat6

Figure 2: Johnson PDFs for each category

Once the best fit is found, we sample the LOCA frequencies for each category to get $Frequency[cat_j]$ - realization of LOCA frequency for category j.

2.2 Distribution of LOCA frequencies to different weld locations

We first convert LOCA frequencies to probabilities using

$$P[cat_j] = Frequency[cat_j] / \sum_{l \in J} Frequency[cat_l],$$

where

- $P[cat_j]$ - probability of observing a break that falls into category j given a break was observed
- $Frequency[cat_j]$ - frequency of failure of type j where $j \in J$
- $J = \{cat_1, cat_2, cat_3, ..., cat_B\}$ - set of possible break types (categories)

Given $P[cat_j]$ the next step is to distribute that probability among all welds that can experience a break from the same category. We compute $P[cat_j \text{ at } location_i]$ - the probability that weld i will experience break of type j using $P[cat_j \text{ at } location_i] = P[cat_j]/M_j$, where $M_j = |I_j|$ - number of welds that can experience category j breaks, $i \in I_j, I_j$ - set of welds that can experience break category j. Here we assume that every weld that can experience a break of category j has equal probability of actually experiencing it.

2.3 Sampling of the break size

The final step is to sample the actual break size conditioned on the break category. Here we assume that the break size has a uniform distribution within a given category, formally

$$breakSize_j^i \sim U[minBreak_j^i, maxBreak_j^i], j \in J, i \in I_j,$$

where

- $U[minBreak_j^i, maxBreak_j^i]$ is the Uniform distribution bounded by $minBreak_j^i$ and $maxBreak_j^i$
- $minBreak_j^i = cat_j^{minBreak}$
- $maxBreak_j^i = \min[cat_j^{maxBreak}, weld_i^{size}]$
- $cat_j^{minBreak}$ - minimum break size that would put it into category j
- $cat_j^{maxBreak}$ - maximum break size that would put it into category j
- $weld_i^{size}$ - actual weld size (it can not experience break size larger than it's diameter).

2.4 Methodology summary

This methodology will require two sampling loops in our simulator CASA Grande, Letellier (2011). We need one sampling loop for the break size within each category and a second loop that samples LOCA frequencies from their fitted distributions. Below is a step-by-step description of that procedure:

1. Set N - number of LOCA frequency samples and S - number of break size samples to generate

2. Sample LOCA frequencies $Frequency[cat_j]$ from the fitted Johnson distributions, see Section 2.1
3. Distribute frequency across plant specific welds as described in Section 2.2
4. Sample actual break size for each possible weld / break category combination as described in Section 2.3
5. Estimate performance measures, store them
6. If we ran S break sizes samples go to the next step, otherwise go to step 4
7. Compute the performance measures summary, store them
8. If we ran N LOCA frequencies samples go to the next step, otherwise, go to step 2
9. Make aggregated performance measures summary

3 Illustrative example

We illustrate our approach described in the first four steps from Section 2.4 using the following example, see Figure 3. Assume we have a total of six welds and these are the only locations where a break can occur. Three of them (welds 1, 2 and 3) are small and have a size of 2.5 inches and hence can experience only small breaks (category1 and category2). Two of those six (welds 4 and 5) are of medium size and have a size of 4 inches and thus can have small and medium breaks (category1, category2 and category3 only,

they can't experience category 4 break). The last weld (weld 6) is large and has a size of 35 inches and can have all three types of breaks - small, medium and large (category1, category2, category3, category4, category5 and category6).

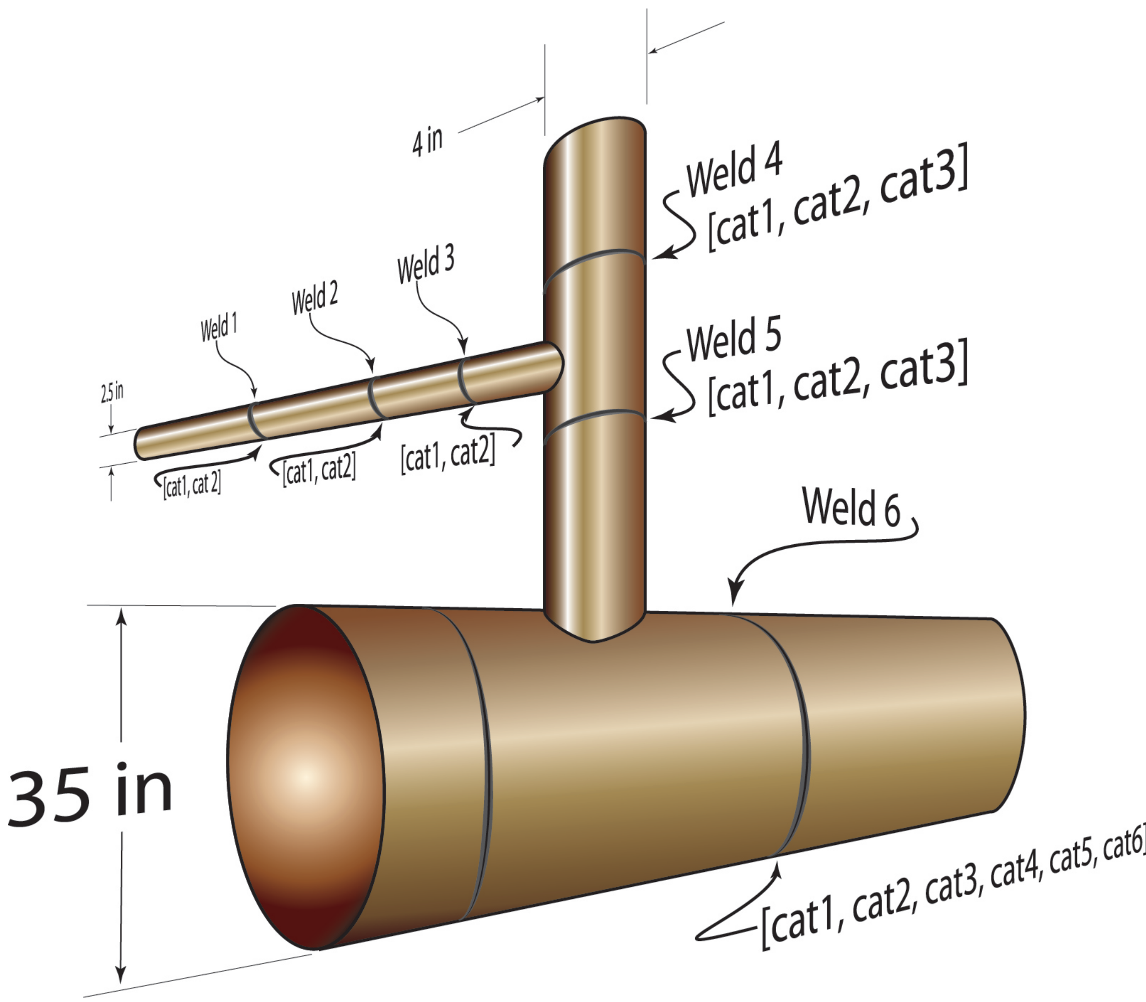


Figure 3: Illustrative Example

1. Assume $S = 1, N = 1$.

2. The sampled LOCA frequencies (using the fitted Johnson distributions) are given in Table 4

Table 4: Sampled LOCA frequencies and corresponding probabilities

Failure Type	Category	Break Size	Frequency	Probability
small	1	0.5-1.625	3.9E-03	9.64E-01
small	2	1.625-3	1.4E-04	3.46E-02
medium	3	3-7	3.4E-06	8.41E-04
medium	4	7-14	3.1E-07	7.67E-05
large	5	14-31	1.2E-08	2.97E-06
large	6	31-41	1.2E-09	2.97E-07

3. We have

$J = \{cat_1, cat_2, cat_3, cat_4, cat_5, cat_6\}$,

$I_{cat1} = \{weld_1, weld_2, weld_3, weld_4, weld_5, weld_6\}$,
$I_{cat2} = \{weld_1, weld_2, weld_3, weld_4, weld_5, weld_6\}$,
$I_{cat3} = \{weld_4, weld_5, weld_6\}$,
$I_{cat4} = \{weld_6\}$, $I_{cat5} = \{weld_6\}$, $I_{cat6} = \{weld_6\}$.

M_{cat_1}=6, M_{cat_2}=6, M_{cat_3}=3, M_{cat_4}=1, M_{cat_5}=1, M_{cat_6}=1.

$BreakSize_{cat_1}^{weld_1} \sim U[0.5, 1.625]$
$BreakSize_{cat_1}^{weld_2} \sim U[0.5, 1.625]$
$BreakSize_{cat_1}^{weld_3} \sim U[0.5, 1.625]$
$BreakSize_{cat_1}^{weld_4} \sim U[0.5, 1.625]$

$$BreakSize^{weld_5}_{cat_1} \sim U[0.5, 1.625]$$
$$BreakSize^{weld_6}_{cat_1} \sim U[0.5, 1.625]$$

$$BreakSize^{weld_1}_{cat_2} \sim U[1.625, 2.5]$$
$$BreakSize^{weld_2}_{cat_2} \sim U[1.625, 2.5]$$
$$BreakSize^{weld_3}_{cat_2} \sim U[1.625, 2.5]$$
$$BreakSize^{weld_4}_{cat_2} \sim U[1.625, 3]$$
$$BreakSize^{weld_5}_{cat_2} \sim U[1.625, 3]$$
$$BreakSize^{weld_6}_{cat_2} \sim U[1.625, 3]$$

$$BreakSize^{weld_4}_{cat_3} \sim U[3, 4]$$
$$BreakSize^{weld_5}_{cat_3} \sim U[3, 4]$$
$$BreakSize^{weld_6}_{cat_3} \sim U[3, 7]$$

$$BreakSize^{weld_6}_{cat_4} \sim U[7, 14]$$

$$BreakSize^{weld_6}_{cat_5} \sim U[14, 31]$$

$$BreakSize^{weld_6}_{cat_6} \sim U[31, 35]$$

Using the formula for $P[cat_j\ at\ location_i]$ we compute probabilities for each weld. The results are given in Table 5. We see that the sum of the distributed probabilities and the targeted probabilities are the same.

4. We simulate break sizes for each weld within each category using uni-

form distribution with the specified above parameters. The sample is shown in Table 6.

We find it is worth mentioning that our assumptions lead to a piecewise linear CDF function of the break size distribution for a given weld. For example, consider weld 6. The CDF of the Break size for that weld will have 6 break points with the slopes determined by the $P[cat_j \text{ at } location_{weld_6}]$ values and break points at $cat_j^{maxBreak}$ values (1.625, 3, 7, 14, 31), see Figure 4.

Table 5: Distributed LOCA probabilities among all welds

Weld	1	2	3	4	5	6	Actual	Target
Cat1	1.61E-01	1.61E-01	1.61E-01	1.61E-01	1.61E-01	1.61E-01	9.64E-01	9.64E-01
Cat2	5.77E-03	5.77E-03	5.77E-03	5.77E-03	5.77E-03	5.77E-03	3.46E-02	3.46E-02
Cat3				2.80E-04	2.80E-04	2.80E-04	8.41E-04	8.41E-04
Cat4						7.67E-05	7.67E-05	7.67E-05
Cat5						2.97E-06	2.97E-06	2.97E-06
Cat6						2.97E-07	2.97E-07	2.97E-07

Table 6: Sampled break sizes for all welds within each break category

Weld	1	2	3	4	5	6
Cat1	1.1	0.6	0.87	1.34	0.79	1.23
Cat2	2.4	1.9	2.1	2.9	1.75	2.36
Cat3				4.56	6.54	5.97
Cat4						9.67
Cat5						25.68
Cat6						32.67

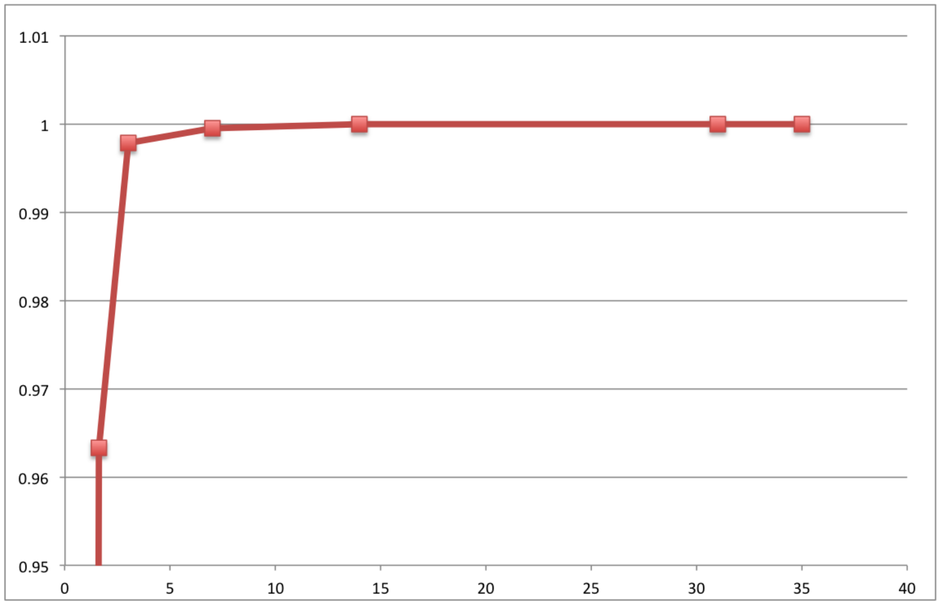


Figure 4: CDF function of break size distribution for weld 6

Conclusion

In this document we are presenting solutions to three problems:

1. How to distribute the NUREG 1829 LOCA frequencies to different locations (welds) in a nuclear plant. We make the simple assumption that small breaks are equally likely to occur in small, medium, or large welds; medium breaks are equally likely to occur in medium and large welds; large breaks can occur only in large welds. This allows us to preserve the NUREG 1829 LOCA frequencies.

2. The six break size categories (columns in NUREG-1829 Table 1) are ranges bounded by six discrete points. For a particular weld we need to be able to sample from the continuous range of break size values. We

propose to use the linear interpolation which is equivalent to assigning equally likely probabilities within each break size category.

3. How to model the distribution of the LOCA frequencies - we propose and fit the Johnson distributions.

We believe that the first problem is the most important that we need to agree on its solution. The other two can be modeled with different distributions. We have already implemented the Gamma distributions fit from the NUREG/CR 6928 and working on fitting a set of Beta distributions. This will give a portfolio of options to apply.

In this document we do not discuss the different sampling techniques needed. Popova and Galenko (2011) describe all the sampling methodologies that we implement.

References

Crenshaw, J. W. (2012, January). South Texas Project Units 1 and 2 Docket Nos. STN 50-499, Summary of the South Texas Project Risk-Informed Approach to Resolve Generic Safety Issue (GSI-191). Letter from John W. Crenshaw to USNRC.

EPRI (1999). Revised Risk-Informed In-Service Inspection Procedure. TR 112657 Revision B-A, Electric Power Research Institute, Palo Alto, CA.

Fleming, K. N., B. O. Lydell, and D. Chrun (2011, July). Development of LOCA Initiating Event Frequencies for South Texas Project GSI-191. Technical report, KnF Consulting Services, LLC, Spokance, WA.

Johnson, N. (1949). Systems of frequency curves generated by methods of translations. *Biometrika 36*, 149–176.

Letellier, B. (2011). Risk-Informed Resolution of GSI-191 at South Texas Project. Technical Report Revision 0, South Texas Project, Wadsworth, TX.

Mosleh, A. (2011, October). Technical Review of STP LOCA Frequency Estimation Methodology. Letter Report Revision 0, University of Maryland, College Park, MA.

Popova, E. and A. Galenko (2011, Deecember). Uncertainty Quantification (UQ) Methods, Strategies, and Illustrative Examples Used for Resolving the GSI-191 Problem at South Texas Project. Technical Report Revision 0, The University of Texas at Austin, Austin, TX.

Tregoning, R., P. Scott, and A. Csontos (2008, April). Estimating Loss-of-Coolant Accident (LOCA) Frequencies Through the Elicitation Process: Main Report (NUREG-1829). NUREG 1829, NRC, Washington, DC.