PARAMETER ESTIMATION OF COMPONENT RELIABILITY MODELS IN PSA MODEL OF KRŠKO NPP

Romana Jordan Cizelj  
“Jožef Stefan” Institute 
Reactor Engineering Division 
Jamova 39, SI-1000 Ljubljana, Slovenia 
romana.jordan@ijs.si

Ivan Vrbanič  
Nuclear Power Plant Krško 
Vrbina 12, SI-8270 Krško, Slovenia 
ivan.vrbanic@nek.si

ABSTRACT

In the paper, the uncertainty analysis of component reliability models for independent failures is shown. The present approach for parameter estimation of component reliability models in NPP Krško is presented. Mathematical approaches for different types of uncertainty analyses are introduced and used in accordance with some predisposed requirements. Results of the uncertainty analyses are shown in an example for time-related components. As the most appropriate uncertainty analysis proved the Bayesian estimation with the numerical estimation of a posterior, which can be approximated with some appropriate probability distribution, in this paper with lognormal distribution.

1 INTRODUCTION

The basic element of a probabilistic safety assessment (PSA) model is the so-called basic event. In most of the cases, a basic event represents a human error or unavailability of a component. Components can be unavailable due to dependent or independent failures, due to test or maintenance. The term "unavailability" can have in PSA context two meanings. First, it can be used to describe loss of function of a component or second, it can be used as a numerical measure, namely as the probability that a component will not perform the required function in a time point under specified operating conditions [1]. In this paper, the term unavailability is used as a numerical measure (probability).

For calculation of component unavailability, one ought to know the values of some input parameters of component models. Parameter, which is essential in all component models, is failure rate for time-dependent models or probability of failure per demand for demand-related models [2].

The present controlled baseline PSA model of Nuclear Power Plant (NPP) Krško does not contain uncertainty analysis of component parameters. Following the international recommendations, NPP Krško initiated uncertainty analysis for input parameters of level 1 analysis. The results of the analysis should reflect the available knowledge at the plant concerning components’ performance (failures, test and maintenance data) and the up-to-date knowledge about mathematical process.
In the paper, the synergy of theoretical models and project requirements at the plant is presented for an uncertainty analysis of component reliability models. The mathematical bases to perform an uncertainty analysis are shown for time-related and demand-related components. Some results of uncertainty analysis of component failure rates are presented and discussed, and the selected approach for parameters' estimation is presented.

2 THE PRESENT APPROACH FOR PARAMETERS ESTIMATION IN NPP KRŠKO

Component failure rates and probabilities per demand are in present controlled baseline PSA model of NPP Krško estimated with the use of generic data and raw plant-specific data concerning component failures and operation [3]. Three different approaches are used for parameter estimation:

- the use of generic data,
- plant-specific estimation using maximum likelihood approach and
- estimation of parameters using Bayes' theorem.

The selection of the appropriate approach depends on the amount of available data on specific component [4].

Generic data are used, if plant-specific data are not available or if they are too scarce, regarding criteria given in the document [4]. Generic data are gathered from the United States Nuclear Regulatory Commission's (US NRC) documents, the International Atomic Energy Agency's (IAEA) documents and other internationally accepted databases [3].

If a large enough amount of the plant-specific data exist, component parameters are estimated using maximum likelihood approach:

\[ \lambda = \frac{N}{T}, \]  
\[ p = \frac{N}{M}, \]  

where \( \lambda \) is the failure rate for time-related components, \( p \) the probability per demand for demand-related components, \( N \) the number of failures, \( T \) the number of component's operating hours and \( M \) the number of component's demands.

Both parameters \( \lambda \) and \( p \) are in the NPP Krško database modeled with lognormal distribution with error factor of three (EF = 3).

If plant-specific data are too scarce to use maximum likelihood method, but their amount is large enough for some specific assessment, Bayesian analysis is used for updating of generic data with scarce plant-specific data:

\[ f(\lambda|E) = \frac{l(E|\lambda) \cdot h(\lambda)}{\int_{\lambda} l(E|\lambda) \cdot h(\lambda) \cdot d\lambda}, \]

where \( h(\lambda) \) is the prior probability density function of \( \lambda \), gathered from a generic data base, \( l(E|\lambda) \) the likelihood function based on the specific data or the so-called evidence \( E \) and \( f(\lambda|E) \) the posterior probability density function of \( \lambda \) given evidence \( E \).

Bayesian estimation is performed with a numerical calculation, using discretized lognormal distribution for the prior function. Results of the numerical calculation are the mean value and the variance of a posterior density function.
3 INTERNATIONALLY ACCEPTED METHODOLOGIES

NPP Krško is the western type reactor (Westinghouse), so Slovene activities in the field of nuclear power are strongly connected with the US practice. Slovenia is also a member of IAEA, which promotes development and use of PSA methodologies in the world. These are the reasons to summarize the US NRC and the IAEA guidelines and recommendations for performing an uncertainty analysis.

Two US NRC documents give the basic guidance to the US nuclear power plants, how to perform a PSA analysis [5, 6].

NUREG/CR-2815 recommends performing a maximum likelihood estimation of failure rate, as shown in equation (1) [5]. For an uncertainty description associated with \( \lambda \), a gamma distribution \( g(\lambda) \) should be used as shown in equation (4). The meaning of symbols is the same as in the previous equations.

\[
g(\lambda) = \frac{T \cdot (\lambda \cdot T)^{N-1} \cdot e^{-\lambda T}}{(N-1)!}
\]

If there are no recorded failures of the component \((N = 0)\), Bayesian updating procedure should be used with the Poisson likelihood of having zero failures \( (e^{-\lambda T}) \). A generic database is included to the document for the selection of prior distributions.

In NUREG/CR-2300 two conceptually different methods are recommended: the classical and the Bayesian estimation [6].

Using the classical approach, demand-related components can be modeled by the binomial distribution, and time-related components by the Poisson distribution. Point estimators can be calculated as already shown in equations (1) and (2). But parameters are treated as constants rather than random values. The estimated value depends on the amount of information pertaining to a parameter of interest and can be described with the so-called standard error or statistical confidence interval.

The standard error of a probability per demand \( se(p) \), its upper \( p_U(1-\alpha) \) and lower \( p_L(1-\alpha) \) confidence limits of the \((1-2\cdot\alpha)\)-percent confidence interval are estimated as follows:

\[
se(p) = \sqrt{\frac{p \cdot (1-p)}{N}}, \quad p_U(1-\alpha) = \chi^2(2N + 2;1-\alpha) \cdot 2 \cdot M, \quad p_L(1-\alpha) = \chi^2(2N;\alpha) \cdot 2 \cdot M.
\]

The \( \chi^2(a;b) \) denotes the \((100\cdot b)\)-percentile of the chi-squared distribution with \( a \) degrees of freedom, and the meaning of the other symbols is the same as in the previous equations.

The standard error of a failure rate \( se(\lambda) \), its upper \( \lambda_U(1-\alpha) \) and lower \( \lambda_L(1-\alpha) \) confidence limits of the \((1-2\cdot\alpha)\)-percent confidence interval are estimated as follows:

\[
se(\lambda) = \sqrt{\frac{\lambda}{T}}, \quad \lambda_U(1-\alpha) = \chi^2(2N + 2;1-\alpha) \cdot 2 \cdot T.
\]
The meaning of the symbols is the same as in the previous equations.

In NUREG/CR-2300 is an extensive description of the Bayesian estimation. The basic formula is shown in equation (3). It can be solved analytically using the so-called natural conjugate prior distribution, which has the property that, for a given likelihood function, the posterior and prior distributions are members of the same family of distributions. If the analytical integration is not possible, the numerical techniques should be used.

The likelihood function for demand-related components is the binomial distribution:

$$Pr(N) = \binom{M}{N} \cdot p^N \cdot (1 - p)^{M-N}, \quad (11)$$

where $Pr(N)$ is the probability of $N$ failures given $M$ demands and $p$ the probability of component's failure.

The natural conjugate prior distribution for the binomial likelihood is the beta distribution $b(p)$:

$$b(p) = \frac{\Gamma(M')}{\Gamma(N') \cdot \Gamma(M'-N')} \cdot p^{N'-1} \cdot (1 - p)^{M'-N'-1}, \quad (12)$$

$\Gamma(a)$ is the so-called gamma function defined as:

$$\Gamma(a) = \int_0^\infty x^{a-1} \cdot e^{-x} \cdot dx. \quad (13)$$

If the prior beta distribution has the parameters $M'=M_0$ and $N'=N_0$, then the posterior distribution has parameters $M'=M_1$ and $N'=N_1$, as shown in equation (14), where $N$ is the number of component's failures and $M$ is the number of component's demands.

$$M_1 = M_0 + M, \quad N_1 = N_0 + N. \quad (14)$$

The likelihood function for time-related components is the Poisson distribution:

$$Pr(N) = \frac{(\lambda \cdot T)^N \cdot e^{-\lambda T}}{N!}, \quad (15)$$

where $Pr(N)$ is the probability of $N$ failures in $T$ operating hours and $\lambda$ the component's failure rate.

The natural conjugate prior distribution for the Poisson likelihood is the gamma distribution $g(\lambda)$:

$$g(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha')} \cdot \lambda^{\alpha'-1} \cdot e^{-\beta \cdot \lambda}. \quad (16)$$
If the prior gamma distribution has the parameters $\alpha' = \alpha_0$ and $\beta' = \beta_0$, then the posterior distribution has the parameters $\alpha' = \alpha_1$ and $\beta' = \beta_1$, as shown in equation (17), where $N$ is the number of component's failures and $T$ is component operating time:

$$
\alpha_1 = \alpha_0 + N, \\
\beta_1 = \beta_0 + T.
$$

The IAEA recommendations about PSA level 1 methodology for nuclear power plants are given in the IAEA Safety Series document [7]. Similarly to the US NRC guidelines, the maximum likelihood estimation is recommended as the point estimation for failure rates and probabilities of failure per demand, and the Bayesian approach is recommended for an assessment of the uncertainties.

4 UNcertainty Analysis

The goal of the uncertainty analysis described in this paper is to assign appropriate probability distributions to component failure rates and probabilities per demand, which reflect the specific knowledge about a component, such as component failures, operating time or number of demands. Some additional project requirements are [8]:

A. The existing database already contains a generic distribution for each component [3]. It was our presumption that the generic distributions were appropriately selected.

B. Component models should be in accordance with the basic event reliability models in the computer code, used at NPP Krško. Unavailability due to time-related failure modes is modeled by means of the exponential distribution. Unavailability due to demand-related failure modes is modeled by means of fixed probability, assuming that demand-related components are modeled with the binomial distribution.

C. For posterior, one of distributions defined in the computer code used at Krško NPP should be used. It includes the following distributions: lognormal, beta, gamma, normal, uniform, log-uniform and discrete. It was decided to avoid the use of discrete distribution.

The initial information, available for the uncertainty analysis of each parameter, is the generic distribution of the parameter, the number of component's failures, and the number of component's demands or component's operating time.

Taking into account the above project requirements and available specific and generic data, it was decided to perform five different types of uncertainty analysis. From the results a judgement can be made about epistemic uncertainties, connected with the lack of knowledge, which approach could be most appropriate to be used. Below, the analysis for a time-related failure mode of a specific component is described.

Input data for the analysis are:

- Component: Residual heat removal (RHR) pump
- Failure mode: Failure to operate (run)
- Generic distribution: Lognormal; mean = 3,00E-5 [1/hours]; variance = 5,48E-9
- Number of failures, $N$: 1
- Operating time, $T$: 20.858 [hours]

4.1 Classical estimation

Classical estimation is performed in accordance with the guidelines in NUREG/CR-2300 [6]. The point estimate is calculated using equation (1), and the upper and the lower
confidence limits of 90% confidence interval are calculated using equations (9) and (10). For a further analysis, this type of calculation is marked with the label Classical.

Because the failure rate in this type of estimation is not treated as a random variable, it cannot be directly compared with the failure rates resulting from other estimations. Here this calculation is used as a statistical estimate on the amount of specific information.

4.2 Gamma distribution

The component failure rate is approximated with the gamma distribution according to the NUREG/CR-2815 guidelines, see equation (4) [5]. For the further analysis, this type of calculation is marked with the label Gamma.

4.3 Bayesian estimation

Three types of Bayesian calculations are performed.

The first type is the numerical calculation, labeled B-numer. For the prior, the generic lognormal distribution is used, with the parameters mean and variance as given with input data (see the introduction of section 4). For the likelihood, the Poisson distribution is used. The posterior is calculated as a discrete distribution, so the numerical calculation of distribution parameters, such as mean, median and 5\textsuperscript{th} and 95\textsuperscript{th} percentiles, is performed.

The second type is labeled B-ln2. This one consists of the identical Bayesian calculation as B-numer. However, upon completion of numerical Bayesian calculation, a resultant discrete posterior distribution is approximated by the lognormal distribution.

The third type, labeled B-gamma, is a Bayesian calculation, using a natural conjugate distribution. The lognormal distribution, characterized with the parameters mean and variance as given with input data, is approximated with a gamma distribution. The parameters of the posterior are determined using equation (17).

4.4 Comparison of results

In Figure 1, the probability density functions (PDF), which are results of the uncertainty analyses, are shown together with generic PDF. Please note logarithmic scale on abscissa.

\[ \lambda \ [1/h] \]

\[ 1,E-08 \quad 1,E-07 \quad 1,E-06 \quad 1,E-05 \quad 1,E-04 \quad 1,E-03 \]

\[ 0 \quad 10000 \quad 20000 \quad 30000 \quad 40000 \quad 50000 \quad 60000 \quad 70000 \]

\[ 1,E-08 \quad 1,E-07 \quad 1,E-06 \quad 1,E-05 \quad 1,E-04 \quad 1,E-03 \]

Figure 1: Probability density functions
The specific PDFs are narrower than the generic PDF, as specific information about the component reduce the uncertainty about component unavailability. This could be seen better in Figure 2, where the cumulative density functions (CDF) of generic distribution and specific distributions are shown. The specific distributions have larger gradient around distribution mean than the generic distribution.

![Figure 2: Cumulative density functions](image)

The often-used measures of dispersion or variation of random variable are the 5th and 95th percentiles of the 90% probability interval. Very useful is also information about the distribution variance, which is a different measure of dispersion or variation of random variable about its mean. In Table 1 values of the 5th ($\lambda_{5\%}$) and 95th ($\lambda_{95\%}$) percentiles, median ($\lambda_{50\%}$), mean and variance of the generic distribution and the estimated uncertainty distributions are shown. For example, variance of the probability distribution assessed with the Bayesian estimation (B-numer or B-ln2) is about six-times (6x) smaller than variance of the generic distribution. This is in accordance with the discussion above.

![Table 1: Characteristic values of probability distributions](image)

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{5%}$</th>
<th>$\lambda_{50%}$</th>
<th>$\lambda_{95%}$</th>
<th>mean</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>1.10E-06</td>
<td>1.00E-05</td>
<td>1.43E-04</td>
<td>3.00E-05</td>
<td>5.48E-09</td>
</tr>
<tr>
<td>Classical</td>
<td>2.46E-06</td>
<td>2.27E-04</td>
<td>4.79E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-numer</td>
<td>3.80E-06</td>
<td>2.20E-05</td>
<td>9.20E-05</td>
<td>3.26E-05</td>
<td>9.42E-10</td>
</tr>
<tr>
<td>B-ln2</td>
<td>6.30E-06</td>
<td>2.30E-05</td>
<td>9.70E-05</td>
<td>3.24E-05</td>
<td>9.45E-10</td>
</tr>
<tr>
<td>B-gamma</td>
<td>3.20E-06</td>
<td>3.20E-05</td>
<td>1.25E-04</td>
<td>4.42E-05</td>
<td>1.63E-09</td>
</tr>
<tr>
<td>Gamma</td>
<td>2.40E-06</td>
<td>3.30E-05</td>
<td>1.43E-04</td>
<td>4.79E-05</td>
<td>2.30E-09</td>
</tr>
</tbody>
</table>

Figure 3 is a visualization of the values shown in Table 1. Note that the meaning of the results of the classical evaluation is conceptually different than the meaning of the results of the other uncertainty analyses. They are shown in Figure 3 only to complement results of the uncertainty analysis.
To conclude, the assessment proved the numerical Bayesian estimation as a suitable tool for the uncertainty analysis even if the quantity of specific data is small (in the upper example only 1 failure in 20.858 hours). Additionally, approximation of the discrete posterior distribution with the lognormal distribution seems to give appropriate results to satisfy given project requirements (see the introduction in section I).

ACKNOWLEDGMENTS

The financial support of the Ministry of Education, Science and Sport of the Republic of Slovenia for the project "Identification and Treating of Uncertainties in Probabilistic Safety Assessment" (Z2-3534) is gratefully acknowledged.

Additionally, the authors wish to thank the NPP Krško for permission to publish some of specific data, used for uncertainty analysis, performed under the project "NEK Level 1 IIE PSA Parameter Uncertainty Analysis".

REFERENCES


