Zeszyty Naukowe Wydziału Elektroniki i Informatyki, Nr 3, pp. 133 – 145, 2011

Małgorzata Górska
Akademia Pomorska
w Słupsku
ul. Partyzantów 27, 76-200 Słupsk, Polska
E-mail: gorskamalgorzata16@wp.pl

DO PRIMARY EVENT UNCERTAINTY DISTRIBUTIONS IMPACT TOP EVENT DISTRIBUTION?

Abstract*. The risk analysis is an essential element of planning, production and operation of technical equipment. This paper deals with the fault tree. The fault tree analysis belongs to the most commonly used risk assessment methods. The main aim of the paper is to ask for the question: does the top event uncertainty assessment have a relationship with adopted assumption of primary events distribution? To achieve this aim a computer simulation that involve random numbers, commonly known as the Monte Carlo method, was used. The research makes use of the Beta, Lognormal, Johnson S_J and truncated Normal distribution.*

Keywords*: fault tree, primary events distributions, top event distribution, uncertainty assessment, mode, uncertainty coefficient, reliability structure function, Monte Carlo simulation.*

1. INTRODUCTION

Fault trees are of vital importance for complex systems reliability assessment. They are useful for planning a variety of projects and operations at risk. Figure 1 shows an example of the fault tree created by Patrick O'Connor [3]. What is the structure of the fault tree? The tree comprises two kinds of events: primary and intermediate and consists of "AND" or "OR" gates. The top event is connected with intermediate and primary events by means of branches. Probability of the top event is uncertain since particular probabilities of primary events are uncertain as well. In particular, this paper deals with uncertainty of assessment of top event probability. The main aim is to answer the following question: do primary event uncertainty distributions impact the top event distributions in the fault tree? To answer the question four cases for primary event distributions have been employed: the Beta, Lognormal, Johnson S_J and truncated Normal distribution.

2. CALCULATION ENVIRONMENT

To achieve research aims computer simulation that involve random numbers, commonly known as the Monte Carlo method, was used. Monte Carlo method has been implemented using the Mathcad software package. Mathcad is as versatile and powerful as programming language, yet it is as easy to use as a spreadsheet. This mathematical environment turns out to be an ideal programming and calculation workspace for solving even the complicated problems of reliability engineering. In this paper we present the most important procedures only.

3. THE FUNCTION OF THE FAULT TREE RELIABILITY

The „OR" and „AND" gates correspond to specified events and logical operations. In the „OR" gate the output event is takes place if any of the input events take place. From the point of view of the reliability this situation corresponds to the model of series structure [4].

Let *n* be the number of inputs in the gate and p_i , i=1,2,…,n are fault probabilities.

The reliability structure function for this case is

$$F(p_i)=1-\prod_{i=1}^{n}(1-p_i). \qquad (2.1)$$

In the „AND" gate the output event occurs on condition that all of the input events take place. It corresponds to the model of parallel structure [4]. Appropriately for this case the structure function is

$$F(p_i)=\prod_{i=1}^{n}p_i\ . \qquad (2.2)$$

Figure 1 exemplified the tree of P. O'Connor. The reliability function of this tree is as follows

$$F(p_1,p_2,p_3,p_4,p_5,p_6,p_7,p_8)=1-\prod_{i=1}^{4}p_i\cdot\left[1-\left(1-\prod_{i=5}^{6}p_i\right)\cdot\left(1-\prod_{i=7}^{8}p_i\right)\right]. \qquad (2.3)$$

The Mathcad procedure used to calculate values of the reliability function according to (2.3) is presented below

$$
\mathrm{Proc\_03}(fT3) := \left|\begin{array}{l}
\text{for } i \in 1..10000 \\
\quad \left|\begin{array}{l}
\text{for } j \in 1..8 \\
\quad r_j \leftarrow 1 - fT3_{i,j} \\
R_1 \leftarrow \prod_{k=1}^{4} r_k \\
R_2 \leftarrow \prod_{k=5}^{6} r_k \\
R_3 \leftarrow \prod_{k=7}^{8} r_k \\
R_c \leftarrow R_1 \cdot \left[1-(1-R_2)\cdot(1-R_3)\right] \\
fTte_i \leftarrow 1 - R_c
\end{array}\right. \\
\text{return } fTte
\end{array}\right.
$$

This procedure is easily readable and does not require any additional descriptions or comments.

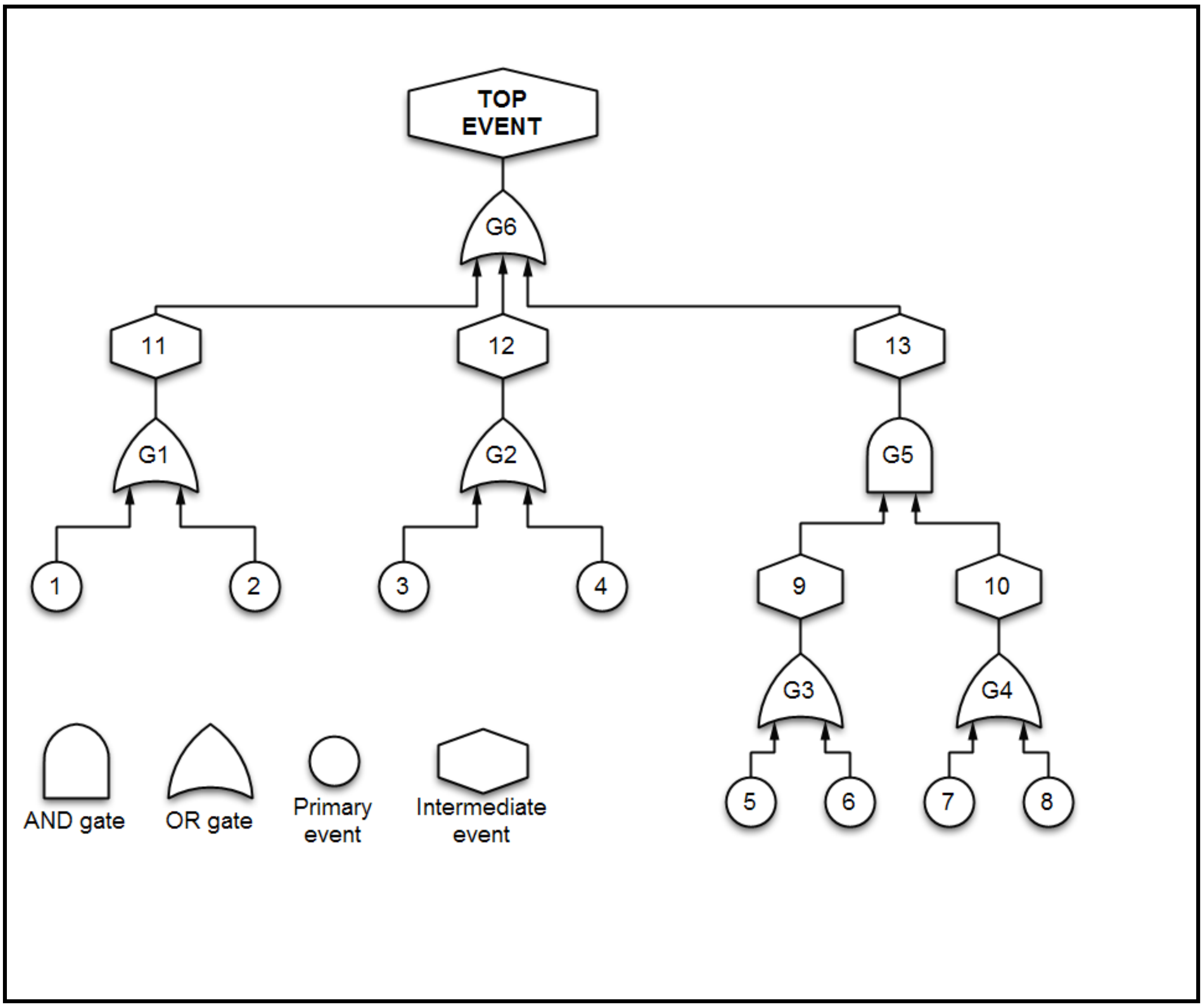

Fig. 1 The fault tree according to O'Connor

4. UNCERTAINTY AND RANDOMNES

In this paper probabilities of primary events (PEPs) are considered random variables not determined constant. It is because PEPs are results of statistical inference and arbitrary engineering assessments procedures. Let n be a number of primary events. A set of input data comprises paired values. The first member of each pair is the most probable PEP value (i.e. pPEP), the second one is PEP's uncertainty coefficient (PEPu).

PEPu is defined as follows (by Apostolakis [1])

$$PEPu = \frac{x_{0.95}}{x_m} \tag{3.1}$$

where $x_{0.95}$ is such a value of x, called 95% quantile, which satisfies the equation $F(x_{0.95})=0.95$ and $x_m = pPEP$ is a maximum of density function called the mode. The mode is understood as the most probable value the r.v. may take. In most cases, just pPEP and PEPu comprise input data.

Underneath denoted symbols were used in [2] for the first time:

PEPs - PRIMARY **E**VENT **P**ROBABILITIES,

pPEP - the highest **p**robability of **P**RIMARY **E**VENT **P**ROBABILITIES,

PEPu - **P**RIMARY **E**VENT **P**ROBABILITIES **u**ncertainty coefficient.

5. INPUT DATA

Table 1 contains all the input data related to primary event tree of Fig. 1. There are preferred values of primary events probabilities and uncertainty coefficients i.e. pPEP values and PEPu values.

Primary event	pPEP	PEPu
1	0,01	3
2	0,02	5
3	0,02	5
4	0,02	5
5	0,05	2
6	0,05	2
7	0,05	2
8	0,05	2

Tab. 1 Input data

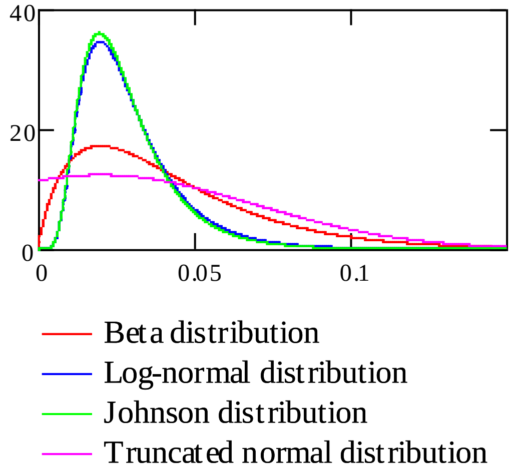

Fig. 2 Density function of the primary event number 2

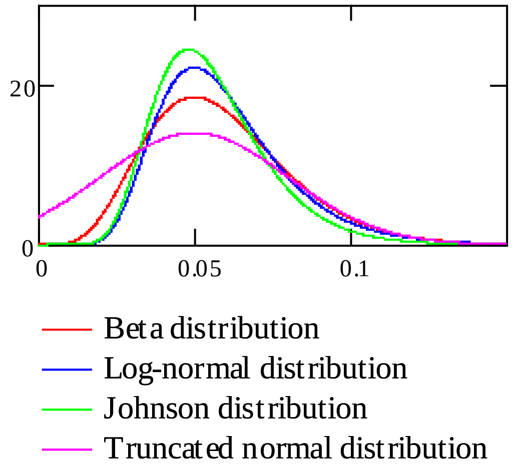

Fig. 3 Density function of the primary event number 8

All the distributions have the same pPEP and PEPu values but they are diversified in respect of the density function shape. The skewness and the kurtosis are different in this case.

6. THE SCHEME OF MONTE CARLO SIMULATION

The simulation method was applied in four steps.

Step 1: Calculating *a, b* parameters values for primary events distribution.
Step 2: Generating x_{ij} input random numbers, i=1,2,…,10 000, j=1,2…,8.
Step 3: Calculating reliability function values (resulting vector comprises 10 000 values).
Step 4: Calculating values pPEP, PEPu and x_{95} quantile for output distribution.

Every calculation was done for four different distributions: beta, lognormal, Johnson and truncated normal.

7. THE BETA DISTRIBUTION EMPLOYED

The first of two major characteristics of probability distribution is the probability density function denoted *f(x)*. In the case of beta distribution we have

$$f(x)=\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}; \qquad (7.1)$$

where x denotes the random variable; a, b are beta distribution parameters (a>0, b>0) and $B(a,b)$ is the beta function.

By the definition

$$B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}dt\,. \qquad (7.2)$$

For a>1 and b>1 the probability density function has a mode

$$x_m=\frac{a-1}{a+b-2}\,. \qquad (7.3)$$

The second characteristics is the cumulative probability function defined as follows

$$F(t)=\frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt\,. \qquad (7.4)$$

In particular:

$$\frac{1}{B(a,b)}\int_0^{x_{0,95}} t^{a-1}(1-t)^{b-1}dt=0{,}95. \qquad (7.5)$$

The parameters a and b values we obtain after having solved system of equation (7.3) and (7.5) for values x_m (pPEP) and $x_{0,95}$ $(x_{0,95}=pPEP\cdot PEPu)$.

The next step is the generation of random variables data and now we can calculate reliability structure function values (Proc_03).

The final calculation step is to find the values of pPEP, $x_{0,95}$ and PEPu for output distribution. To achieve this aim the following Mathcad procedure is applied. Particular rows of the procedure are commented in detail. Numbers in brackets are intended as reference numbers.

$$
\mathrm{Proc\_11}(fX) :=
\left|
\begin{array}{l}
nn \leftarrow \mathrm{rows}(fX) \\
fX \leftarrow \mathrm{sort}(fX) \\
x_d \leftarrow fX_{2000} \\
x_g \leftarrow fX_{nn-2000} \\
\delta \leftarrow 10^{-6} \\
x_s \leftarrow \dfrac{x_g + x_d}{2} \\
m \leftarrow 0 \\
\text{while } 2 > 1 \\
\quad \left|
\begin{array}{l}
f_d \leftarrow \mathrm{Proc\_10a}(x_d, fX) \\
f_g \leftarrow \mathrm{Proc\_10a}(x_g, fX) \\
x_d \leftarrow x_s - \delta \ \text{ if } \ f_g > f_d \\
x_g \leftarrow x_s + \delta \ \text{ if } \ f_g < f_d \\
x_s \leftarrow \dfrac{x_g + x_d}{2} \\
m \leftarrow m + 1 \\
TWP_{m,1} \leftarrow x_d \\
TWP_{m,2} \leftarrow x_s \\
TWP_{m,3} \leftarrow x_g \\
TWP_{m,4} \leftarrow \dfrac{f_g}{f_d} - 1 \\
\text{break } \ \text{if } \left| \dfrac{f_g}{f_d} - 1 \right| \le 0.001 \vee \ m \ge 100
\end{array}
\right. \\
\begin{pmatrix}
x_s \\
fX_{9500} \\
\dfrac{fX_{9500}}{x_s}
\end{pmatrix}
\end{array}
\right.
$$

The description of Proc_11procedure

[1] - Ascribing the number of fX vector rows to nn variable.

[2] - Sorting the rows of fX vector in ascending order.

[3],[4] - Defining of the initial mode interval; ascribe values to x_d and x_g.

[5] - Ascribing the values to δ variable.

[6] - Ascribing the values to x_s variable.

[7] - Initialization of m variable and ascribing zero to m.

[8] - Started the while loop.

[9], [10] - Ascribing the density function values for x_d and x_g arguments to f_d and f_g variable, respectively. The Parzen's method (Proc_10a) was employed to estimate the density function of the top event distribution.

[11],[12] - Definition of the new limits of the mode interval; conditional ascribing the values to x_d and x_g variable.

[13] - Ascribing the value to x_s variable.

[14] - Incrementing the m value.

[15], [16], [17], [18] - Ascribing the values to columns of TWP matrix.

[19] - The exit condition of the while loop.

[20] – Projection the vector of results.

We obtain the following result in the beta case

pPEP	0,128
$x_{0,95}$	0,228
PEPu	1,776

8. THE LOGNORMAL DISTRIBUTION EMPLOYED

The lognormal distribution was derived from the well-known normal distribution. It is said that a given random variable follows the lognormal distribution if its natural logarithm follows the normal distribution.

In the case of lognormal distribution we have

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot b \cdot x} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{\ln x - a}{b} \right)^2 \right], \tag{8.1}$$

where x denotes the random variable and a, b denote parameters.

The probability density function has a mode located at

$$x_m = pPEP = \exp\left(a - b^2\right). \tag{8.2}$$

The cumulative failure function takes the form

$$F(x)=\frac{1}{\sqrt{2\cdot\pi}\cdot b}\int_0^x\frac{1}{u}\cdot\exp\left[-\frac{1}{2}\cdot\left(\frac{\ln u-a}{b}\right)^2\right]du\ . \tag{8.3}$$

If x follows the lognormal distribution then random variable

$$z=\frac{\ln(x)-a}{b} \tag{8.4}$$

follows the normal N(0,1) distribution. It can be easily read-out from tables of the N(0,1) distribution that corresponding 95% quantile is equal to 1.64. Consequently, we get

$$\frac{\ln(x_{0.95})-a}{b}=1.645\Rightarrow x_{0.95}=\exp(a+1.645\cdot b)\ . \tag{8.5}$$

.

Substituting (2b) and (5) into (6) we get

$$PEPu=\exp\left(b^2+1.645\cdot b\right). \tag{8.7}$$

After transformation of formulas (2b) and (7) we obtain the following equations for *a* and *b* parameters

$$\begin{aligned} b&=\sqrt{\ln(PEPu)+0.676}-0.822\\ a&=\ln(pPEP)+b^2 \end{aligned} \tag{8.8}$$

The Mathcad procedure named Proc_11 returns the following result vector:

pPEP	0,101923
$x_{0,95}$	0,159174
PEPu	1,561699

9. THE JOHNSON DISTRIBUTION EMPLOYED

The Johnson distribution of S_B type has the following probability density function

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot b\cdot x\cdot(1-x)}\cdot\exp\left[\frac{-1}{2}\cdot\left(\frac{\ln\left(\frac{x}{1-x}\right)-a}{b}\right)^2\right], \tag{9.1}$$

where x denote r.v. and *a*, *b* are parameters.

The cumulative distribution function

$$F(x)=\frac{1}{\sqrt{2\pi}\cdot b}\cdot\int_0^x\frac{1}{x\cdot(1-x)}\cdot\exp\left[-\frac{1}{2}\cdot\left(\frac{\ln\left(\frac{x}{1-x}\right)-a}{b}\right)\right]^2. \tag{9.2}$$

If x follows the Johnson distribution the random variable

$$z = \frac{\ln\left(\frac{x}{1-x}\right) - a}{b} \tag{9.3}$$

follows the normal N(0,1) distribution.
In particular

$$\frac{\ln\left(\frac{x_{0,95}}{1-x_{0,95}}\right) - a}{b} = 1{,}645\,. \tag{9.4}$$

Consequently

$$b = \frac{1}{1{,}645} \cdot \left[\ln\left(\frac{x_{0,95}}{1-x_{0,95}}\right) - a\right]. \tag{9.5}$$

For $x_{0,95}$

$$\frac{df(x)}{dx} = 0\,. \tag{9.6}$$

Taking (9.5) and (9.6) into consideration we obtain the distribution parameters *a* and *b*. The Mathcad procedure named Proc_11 returns the following result vector.

pPEP	0,096
$x_{0,95}$	0,151
PEPu	1,581

The Beta, Lognormal and Johnson S_J distributions are a skewed to the right. The right tail is longer than the left side and the bulk of the values lie to the left on the mean. The truncated normal distribution is right-skewed as well; nonetheless the skewness is not such great in this case. This distribution presents more optimistic situation from the reliability point of view.

10. THE TRUNCATED NORMAL DISTRIBUTION

The probability density function of the truncated normal distribution takes the following form

$$f_c(x) = \begin{cases} \frac{1}{c} \cdot f(x) \text{ gdy } x \geq 0 \\ 0 \text{ gdy } x < 0 \end{cases}, \tag{10.1}$$

where *f(x)* is probability density function of normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot b} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{x-a}{b}\right)^2\right] \tag{10.2}$$

and $$c=\int_0^{\infty} f(x)dx \quad (c<1) \tag{10.3}$$

is constant.

The cumulative probability function

$$F_c(x)=\begin{cases}\frac{1}{c}\cdot\int_0^{\infty} f(x)\text{ gdy } x\geq 0\\ 0 \text{ gdy } x<0\end{cases}. \tag{10.4}$$

Left side cut-off normal distribution (10.1) and normal distribution (10.2) have the same values of x_m and $x_{0,95}$.

The appropriately Mathcad procedure gives the following result.

pPEP	0,148
$x_{0,95}$	0,249
PEPu	1,679

11. CONCLUSION

Since Apostolakis [1] has originated uncertainty analysis the lognormal and Johnson S_B distributions have gained popularity as basic uncertainty distributions. Moreover, one may treat the lognormal distribution as a special case of the Johnson S_B distribution when $x<<1$.

In ranges of x we deal with in uncertainty assessment these distribution are very similar.

In Reality nothing follows the lognormal, Johnson or any other distribution known in probability theory. The distributions are only mathematical models of Reality made for engineering purposes. This is so called “engineering judgment”.

It may happen, despite this common believe, that actual uncertainty distribution of primary events will considerably differ from the Lognormal and Johnson distributions. If so, a question is sure to be asked: In what degree does this fact impact uncertainty distribution of the top event?

Table 2 is intended to compare output distributions in terms of pPEP and PEPu.

	TYPE OF DISTRIBUTION			
	BETA	**LOG-NORMAL**	**JOHNSON**	**TRUNCATED NORMAL**
pPEP	0,128	0,102	0,096	0,148
PEPu	1,776	1,562	1,581	1,679

Tab. 2 Mode (pPEP) and uncertainty coefficient (PEPu) of output distributions

Taking into account that simulations comprised 10000 repetitions one may conclude that mode and uncertainty coefficient of the top event distribution does not depend on the primary event distributions.

In other words: there is no significant relation between assumptions relating to primary events distribution and uncertainty assessment of output distribution.

Bibliography

1. Apostolakis G.,Lee Y.T.: *Methods* for the estimation of confidence bounds for the top event unavailability of fault trees. Nuclear Engineering and Design. 41 (1977) , pp 411-419.
2. Drapella A., Górska M.: Uncertainty Propagation in Fault Tree. A Monte Carlo Approach. Network Centric Warfare. Naval Academy Gdynia 2009.
3. O'Connor P. D.T.: Practical Reliability Engineering. Wiley 2002.
4. PN-EN 61078 (U), Techniki Analizy Niezawodności. KT 009 ds. Niezawodności 2002.