TRANSACTIONS OF THE 9TH INTERNATIONAL CONFERENCE ON STRUCTURAL MECHANICS IN REACTOR TECHNOLOGY / LAUSANNE
17-21 AUGUST 1987

Structural Mechanics in Reactor Technology

Editor
FOLKER H.WITTMANN

OFFPRINT

A.A.BALKEMA / ROTTERDAM / BROOKFIELD / 1987

Uncertainty propagation in fault trees using a quantile arithmetic methodology

M.Ragheb & M.Abdelhai
Department of Nuclear Engineering, University of Illinois, Urbana-Champaign, USA

1 INTRODUCTION

A methodology based on Quantile Arithmetic, the probabilistic analog to Interval Analysis (Dempster 1969), is proposed for the computation of uncertainty propagation in Fault Tree Analysis (Apostolakis 1977). The basic events' continuous probability density functions are represented by equivalent discrete distributions through dividing them into a number of quantiles N. Quantile Arithmetic is then used to perform the binary arithmetical operations corresponding to the logical gates in the Boolean expression for the Top Event of a given Fault Tree. The computational characteristics of the proposed methodology as compared with the exact analytical solutions are discussed for the cases of the summation of M normal variables. It is further compared with the Monte Carlo method through the use of the efficiency ratio defined as the product of the labor and error ratios.

2 THEORY OF PROPOSED METHODOLOGY

We consider the approximation of the probability density function f(x) of an absolutely continuous random variable **X** by a discrete distribution $\tilde{\mathbf{X}}$.

Assuming an N-point distribution, the distribution $\tilde{\mathbf{X}}$ can be represented by the following discrete probability density function:

$$(1)\quad g(x_i)=\begin{cases} \alpha & \text{if } x=x_i \quad ,i=1,2,\ \ldots\ ,m-1, \\ & \text{where } P_i=i \\ 1-(N-1)\alpha & \text{if } x=x_m, \\ & \text{where } P_m=1/2 \\ \alpha & \text{if } x=x_{N-i+1} \quad ,i=m-1,\ \ldots\ ,2,1, \\ & \text{where } P_{N-i+1}=1-i \\ 0 & \text{otherwise} \end{cases}$$

In this equation $p_i = P\{X \le x_i\}$, $i=1, 2, \ldots, N$; $\underline{x} \le x_1 \le \ldots \le x_N \le \bar{x}$; $\underline{x}$ is the lower limit of the distribution; $\bar{x}$ is the upper limit of the distribution; x_m is the (1/2)'th quantile or median point; $m=(N+1)/2$ for N odd; x_1 , x_2 , ..., x_{m-1} are the (α)'th, (2 α)'th, ...,

$(m-1)\alpha$'th, quantiles, respectively; and x_{m+1}, ..., x_{N-1}, x_N are the $(1-(m+1)\alpha)$'th, ..., $(1-2\alpha)$'th, $(1-\alpha)$'th quantiles, respectively. The discrete representation underestimates the exact result for points below the median and overestimates it above the median. Clearly, for a larger number of quantiles N, the approximation to the continuous case becomes better.

We consider the two random variables **X**, **Y** , each of which is approximated by an N-point distribution, and let * denotes any one of the four arithmetic operations of addition, subtraction, multiplication, and division. The result of performing the binary operation * on **X** and **Y** is a random variable **Z** having an N-point distribution:

$$f(z) = f(x_i * y_j) = P(\mathbf{X}=x_i, \mathbf{Y}=y_j) . \qquad (2)$$

Since **X** and **Y** are independent, then:

$$g(z_k) = \begin{cases} q_k = p_i p_j & \text{if } z_k = x_i * y_j \quad ,i=1,2,\ \dots\ ,N \\ & \qquad\qquad\qquad ,j=1,2,\ \dots\ ,N \\ & \qquad\qquad\qquad ,k=1,2,\ \dots\ ,N \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where

$$p_i \text{ and } p_j = \begin{cases} & \text{for } i,j = 1,2,\ \dots\ ,m-1 \\ & \quad i,j = m+1,\ \dots\ ,N \\ 1-(N-1) & \text{for } i,j = m \end{cases}$$

and $(x_1, x_2, \dots, x_N)$ and $(y_1, y_2, \dots, y_N)$ are the defining quantile points of **X** and **Y** respectively. The methodology to generate the N quantile points: w_i , i=1 ,2, 3, ..., N corresponding to the operation **X*Y** can then be enunciated. The N^2 numbers z_k are sorted in order of increasing magnitude with their associated probabilities. Then the following rules for generating the N quantile points w_i are used:

For x_i , i=1 , 2, 3, ..., m-1, the points below the median point, take w_i , i=1, 2, 3, ..., m-1 to satisfy:

$$\sum_{k=1}^{r} q_k < i\alpha \le \sum_{k=1}^{r+1} q_k \quad \text{corresponding to the event: } z_{r+1} = w_i \qquad (4)$$

For the median point $x=x_m$, take w_m to satisfy:

$$\sum_{k=1}^{r} q_k < 1/2 \le \sum_{k=1}^{r+1} q_k \quad \text{corresponding to the event: } z_{r+1} = w_m \qquad (5)$$

For x_{N-i+1}, i=(m-1), ..., 3,2,1; those points beyond the median point, take w_{N-i+1}, i=(m-1), ..., 3,2,1 to satisfy:

$$\sum_{k=1}^{r} q_k \le 1-i\alpha < \sum_{k=1}^{r+1} q_k \quad \text{corresponding to the event: } z_{r+1} = w_{N-i+1} \qquad (6)$$

3 NUMERICAL RESULTS

The methodology was implemented into a computer program MQA (Abdelhai 1986), and the numerical results were compared to those of problems possessing exact analytical solutions. Figure 1 shows the relative error

$$e_{QA} = \frac{C - T}{T} \qquad (7)$$

where C is the computed value of a quantile point, and T is the exact value of a quantile point; for the sum of twenty normal distributions. It can be noticed that the exact value of the median is obtained for most values of N.

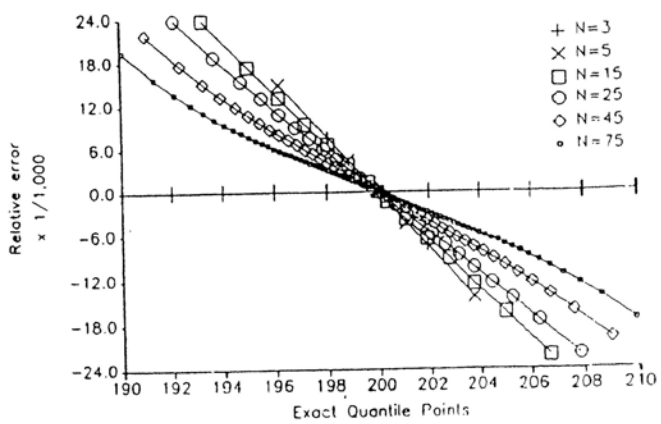


Fig. 1 The relative error for different number of quantiles N for the sum of twenty normal distributions.

Table 1 shows a comparison between the proposed method and the exact analytical solution for the sum of twenty normal variables in [illegible] the relative error as defined by Eq. 7. The results are shown [illegible] the 20'th, 40'th, 60'th and 80'th percentiles, as well as at the median point. The table shows that larger numbers of quantiles are associated with a larger computational load, and a corresponding smaller relative error.

Table 1

Comparison between the results of the sum of twenty normal distributions obtained through exposed method and exact analytical result using different numbers of quantiles N

Percentile Values	Exact Percentile points	Estimated Percentile Points With Different Numbers Of Quantile N									
		N = 5		N = 15		N = 25		N = 45		N = 75	
		CPU=0.052 Sec.		CPU=0.938 Sec.		CPU=6.21 Sec.		CPU=63.512 Sec.		CPU=509.833 Sec.	
		Estimated	C	Estimated	C	Estimated	C	Estimated	C	Estimated	C
20	196.23615	199.15838	0.0149	198.79550	0.0130	198.34743	0.0108	197.79494	0.0079	197.35849	0.0057
40	198.86700	199.74665	0.0044	199.63712	0.0039	199.50639	0.0032	199.34173	0.0024	199.22131	0.0018
50	200.00000	200.00000	0.0000	200.00000	0.0000	200.00000	0.0000	200.00000	0.0000	200.00449	0.00002
60	201.11330	200.25335	-0.0044	200.36288	-0.0038	200.49361	-0.0032	200.66642	-0.0023	200.79306	-0.0017
80	203.76385	200.84162	-0.0143	201.20450	-0.0126	201.65257	-0.0104	202.21321	-0.0076	202.64859	-0.0055

The effectiveness of the proposed methodology can be further assessed when it is compared with Monte Carlo as a widely used method in the intended application. The analysis is carried out by defining the relative efficiency ratio as the product of the labor ratio (T_{MC}/T_{QA}) and the error ratio ($|e_{MC}|/|e_{QA}|$) (Hammersley and Hanscomb 1965) as a figure of merit for the comparison:

$$F = \frac{T_{MC} \mid e_{MC} \mid}{T_{QA} \mid e_{QA} \mid} \tag{8}$$

where F is the efficiency of the quantile arithmetic method relative to the Monte Carlo method, T_{MC} is the units of computing time for the Monte Carlo method, T_{QA} is the units of computing time for the quantile arithmetic method, e_{MC} is the error of the Monte Carlo result relative to the exact analytical result:

$$e_{MC} = \frac{\text{Monte Carlo Result} - \text{Exact Result}}{\text{Monte Carlo Result}},$$

and e_{QA} is the relative error obtained for the quantile arithmetic method as defined by Eq. 7. The results displayed in Table 2 show that the efficiency ratio for M=2 at the 40'th percentile takes the value of 2279 for N=5, 445 for N=25 and 66 for N=45 when the results are compared to those obtained by the Monte Carlo method using 19,200 samples.

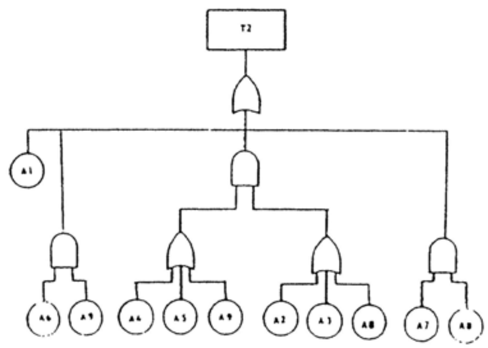
Table 2

Comparison of the quantile methodology and the Monte Carlo methods in terms of the efficiency ratio (F)

(M = 2)

Percentile Values	Number of Samples (S)	Efficiency Ratio (F) — Number (N) of Quantiles.				
		N=5	N=15	N=25	N=45	N=95
20	4800	77.84	47.29	10.95	2.17	0.23
	19200	163.20	99.15	22.54	4.56	0.47
40	4800	72.46	99.75	14.14	2.11	0.24
	19200	2278.63	3136.82	444.82	66.28	7.55
50	4800	∞	∞	∞	∞	∞
	19200	∞	∞	∞	∞	∞
60	4800	92.31	137.15	18.94	2.90	0.30
	19200	1161.23	1725.25	238.29	36.45	3.78
80	4800	5.82	3.56	0.80	0.17	0.02
	19200	274.47	168.04	37.91	7.81	0.7[illegible]

We further compare the results obtained by the proposed methodology to those obtained through using the Monte Carlo program SAMPLE used by Rasmussen et al.(1975). We here consider the Fault Tree shown in Fig. 2 that was earlier analyzed by Senna et al.(1982), with the corresponding failure probabilities. We consider the distributions of the basic events to be normal.

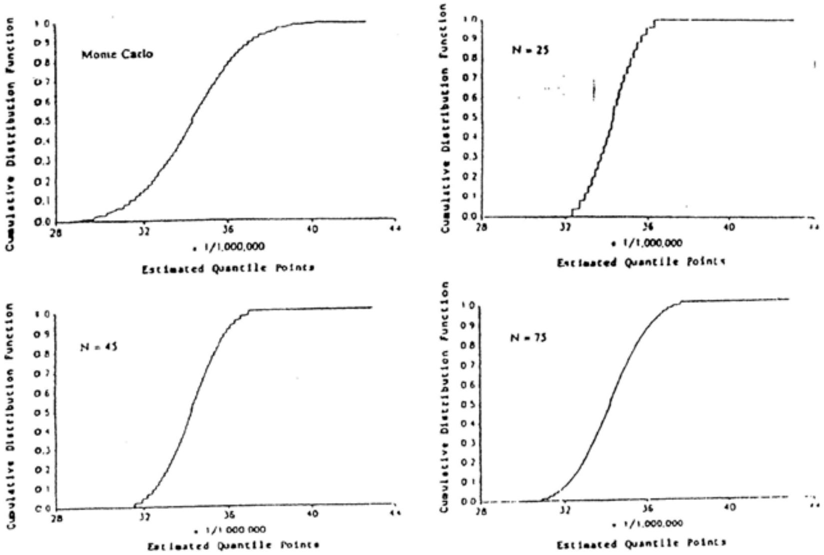


Component Failure	Mean	Standard Deviation	90% Error Spread
A1	1.7E-05*	1.7E-06	2.79650E-06
A2	3.6E-04	3.6E-05	5.92200E-05
A3	1.0E-03	1.0E-04	1.64500E-04
A4	1.0E-03	1.0E-04	1.64500E-04
A5	3.6E-04	3.6E-05	5.92200E-05
A6	6.1E-03	6.1E-04	1.00345E-03
A7	6.1E-03	6.1E-04	1.00345E-03
A8	9.7E-04	9.7E-05	1.59565E-04
A9	9.7E-04	9.7E-05	1.59565E-04

* Read as 1.7×10^{-5}

Fig. 2 Representative Fault Tree used by Senna et al. and associated components failure data.

Figure 3 shows that the proposed methodology gives a result that is close to the cumulative distribution function of the top event T_2 as generated by the Monte Carlo method. The results are shown for N=25, 45 and 75 quantiles.



.3 Comparison of the results obtained from the quantile [illegible] approach using different quantiles to those obtained from the

A comparison on a pointwise basis of the results of the generated distribution is attempted by defining the relative difference betwe[illegible] the two approaches as:

$$\varnothing = \frac{\text{Quantile arithmetic Result} - \text{Monte Carlo Result}}{\text{Monte Carlo Result}}$$

is shown in Table 3. The comparison to a Monte Carlo sample of 120[illegible] histories shows that the relative difference between the two results at the median point ranges from 0.06 to 0.03 percent. It can be

Table 3

Comparison of the results for the 20'th, 40'th, 50'th, 60'th, and 80'th percentiles points obtained through Quantile Arithmetic method and the Monte Carlo method with 1200 samples for the Fault Tree

Percentile Values	Monte Carlo Estimated Percentile Points	MQA Estimated Percentile Points					
		N = 25		N = 45		N = 75	
		Points	♦	Points	♦	Points	♦
20	.32432341E-04	.33295000E-04	0.0265987	.33063113E-04	0.0194494	.32896134E-04	0.0143002
40	.33706592E-04	.33959000E-04	0.0074884	.33895800E-04	0.0056134	.33841400E-04	0.0039995
50	.34268734E-04	.34258000E-04	-0.0003131	.34248300E-04	-0.0005962	.34248700E-04	-0.0005845
60	.34803898E-04	.34533600E-04	-0.0077661	.34601400E-04	-0.0058180	.34662300E-04	-0.0040682
80	.36042228E-04	.35227800E-04	-0.0225962	.35444600E-04	-0.0165811	.35624200E-04	-0.0115980

observed that the relative difference decreases as we move toward the median point. At the 20'th percentile, the relative difference decreases from 2.7 percent to 1.4 percent as the number of quantiles N is increased from 25 to 75.

4 SUMMARY AND CONCLUSIONS

A methodology based on Quantile Arithmetic is proposed for the computation of uncertainties propagation in Fault Trees. The basic events' continuous probability density functions are represented by equivalent discrete distributions through dividing them into a number of quantiles N. Quantile Arithmetic is then applied to perform the binary arithmetic operations correpending to the logical gates in the Boolean expression of the Top Event expression of a given Fault Tree.

A general purpose computer code has been developed for the application of the methodology to Fault Trees. A comparison is carried out with the exact results for the sum of M normal variables. A Fault Tree from the work of Senna et al.(1982) is used for comparison of the proposed methodology to the Monte Carlo results.

The proposed methodology, due to its potential advantages, is expected to contribute to the effort in analyzing nuclear systems from the probabilistic perspective. Due to its competitiveness with the Monte Carlo method and the Interval Analysis methods, particularly at the median point, larger systems can be analyzed within a reasonable computational time-frame. Work is pursued in applying this methodology to the estimation of the propagation of the uncertainty in goal trees used for decision making under uncertainty in Production Rules Analysis systems (Ragheb and Gvillo 1986) in the field of Knowledge Engineering.

REFERENCES

Abdelhai M. I. 1986. A Quantile Arithmetic Methodology for Uncertainty Propagation in Fault Trees. Thesis, Dept. of Nuclear Engineering, Univ. of Illinois at Urbana-Champaign.

Apostolakis G and Lee Y. T. 1977. Methods for the Estimation of Confidence Bounds for the Top-Event Unavailability of Fault Trees, Nucl. Eng. Des. 41, p 411.

Dempster M. 1969. Distributions in Intervals and Linear Programming, in: E. Hansen, Ed., "Topics in Interval Analysis," Oxford University Press, pp. 107-127, Ely House, London.

Hammersley J. M. and Handscomb D. C. 1965. Monte Carlo Methods, Methuen and Co. Ltd., London, 2nd. sd., p.51.

Ragheb M. and Gvillo D. 1986. Development of Model-Based Fault-Identification Systems on Microcomputers, in: J. F. Gilmore, Ed., "Applications of Artificial Intelligence III," Vol. 635, pp.268-275.

Rasmussen N. et al. 1975. Reactor Safety Study: An Assessment of Accident Risk in US Commercial Nuclear Power Plants, WASH-1400, USNRC, Washington D.C., Oct.

Senna V., Milidiu R. L., Fleming P. V., Salles M. R. and Oliveira L. F. S. 1982. A Fast Analytical Method for the Addition of Random Variables, Report NUREG/CP-0027, "Proceedings of the International Meeting on Thermal Nuclear Reactor Safety," Chicago, IL, Vol. 1, pp.245-252.