Reliability Calculations

Petersen, Kurt Erling

Publication date:
1986

Document Version
Publisher's PDF, also known as Version of record

Risø-M-2584

RELIABILITY CALCULATIONS

Improvements of methods intended for calculation of reliability of structures and systems

Kurt E. Petersen

Abstract. Risk and reliability analysis is increasingly being used in evaluations of plant safety and plant reliability. The analysis can be performed either during the design process or during the operation time, with the purpose to improve the safety or the reliability.

Due to plant complexity and safety and availability requirements, sophisticated tools, which are flexible and efficient, are needed. Such tools have been developed in the last 20 years and they have to be continuously refined to meet the growing requirements.

Two different areas of application were analysed. In structural reliability probabilistic approaches have been introduced in some cases for the calculation of the reliability of structures or components. A new computer program has been developed based upon numerical integration in several variables.

In systems reliability Monte Carlo simulation programs are used especially in analysis of very complex systems. In order to increase the applicability of the programs variance reduction techniques can be applied to speed up the calculation process. Variance reduction techniques have been studied and procedures for implementation of importance sampling are suggested.

Thesis submitted in partial fulfilment of the requirements for the degree lic. techn. (Danish Ph.D.) at the Technical University of Denmark.

ISBN 87-550-1229-9
ISSN 0418-6435
Grafisk Service, Risø 1986

CONTENTS

Page

1. INTRODUCTION

During this century the industrial world has developed rapidly and a large number of new technologies have been employed. The plants are large and centralized which calls for a high degree of availability leading to complex systems supported by control and regulation systems. This also lead to a higher degree of automation.

In the recent years society are getting more concerned about safety aspects of industrial plants. The society demand that industrial plants are built and operated to a high level of safety, diminishing the consequences of any major accident. The requirements and the public debate on this subject imply the necessity of the availability of analysis methods and tools to be used in assessing the risks inherent in the industrial plant. Such assessments are called risk analyses.

Similarly, the plants, which are complex and large, require large investments which again call upon a high level of availability. This means that also in cases where events will not lead to major accidents, it is important to limit the consequences in order to minimize the shut down time and save money. In such cases analysis methods and tools suited for assessment of the reliability or the availability of the plant are needed.

It is obvious that a large number of methods and tools are common to the two types of assessments described above. Several methods have been developed during the last 20 years within these areas. The aim of this study is to make a review of available methods in two specific areas in order to develop alternative methods or to propose improvements of existing methods.

Chapter 3 deals with the development of a new method to be used in structural reliability - or more specificly in probabilistic fracture mechanics, where the probability of failure of a struc-

ture or component is considered. In chapter 4 a review of available variance reduction techniques to be used in Monte Carlo simulation for systems reliability is presented. Chapter 5 summarizes the results and discuss their applicabilities and limitations.

2. RISK AND RELIABILITY ANALYSIS

To meet the requirements with respect to plant availability and plant safety reliability and risk analysis tools have been developed. The tools can be applied in an analysis of any complex system, for instance in nuclear power industry, air transport, space research, chemical industry, off-shore industry, and traffic systems. The risk analysis does not replace other safety investigations, but it is a supplement to safety codes and standards, quality control and quality assurance, etc.

The following sections will describe the concept of risk and the definitions used in risk and reliability analysis of components and systems.

2.1. Definitions and Descriptions

Risk is a concept used with a variety of meanings and several definitions are given in the literature. In this study the following definition will be used:

> The concept of risk includes both a hazard and its corresponding probability.
>
> A hazard is defined as a situation or a chemical that is potentially harmful to humans or property.

This means that the risk includes an evaluation of unwanted events, their consequences and their probabilities of occurrence:

$$\text{Risk} = F(A, C(A), P(A))$$

where

- A = unwanted events
- $C(A)$ = consequences related to the events
- $P(A)$ = probabilities of the events
- F = some unknown function.

As an example the function F can be defined in the following way:

$$\text{Risk} = \sum_{i=1}^{N} C_i * P_i$$

where

- N = The total number of accidents or situations taken into consideration
- C_i = The consequences related to the accident or situation number i
- p_i = The probability of occurrence of the accident or situation number i

This is a very simple definition of the function F, and its limitations and areas of applications have been heavily debated.

The definition of reliability is the following:

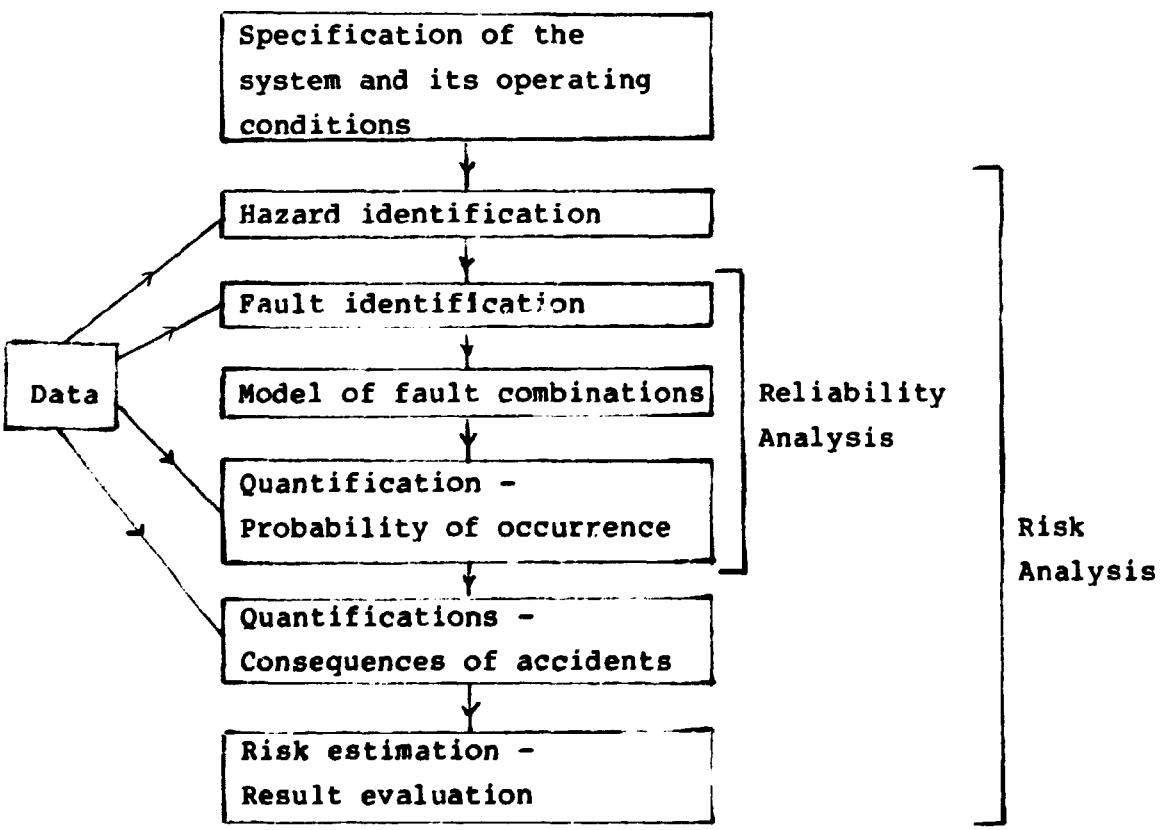
> The reliability of a system (device) or component is the probability that it is performing its purpose adequately for a specified period of time, under the operating conditions encountered.

As it is seen from the definitions of risk and reliability, the analysis of the reliability of a system can be regarded as a part of a risk analysis of the same system. The evaluation of the probability of adequate performance is common to the two types of analysis, whereas the evaluation of the consequences to humans or property is unique in risk (or safety) analysis.

A risk analysis can be described as a set of systematic methods to identify hazards and to quantify their probabilities of occurrence and their consequences. Risk analysis can be used in a broad spectrum of applications with varying purposes, which are shown on the next page. The structure of a risk analysis is also shown.

	approval by authorities
	evaluation of safety measures
RISK ANALYSIS	emergency planning
	selection between alternatives
	reduction of consequences of events
	reduction of probabilities of events
	improvement of availability
RELIABILITY ANALYSIS	planning of operating procedures
	planning of test and maintenance

Purposes of risk and reliability analyses.



The structure of a reliability and risk analysis.

2.2. Reliability and Unavailability of a System

The reliability of a system, which was defined in the previous section, describes the probability that the system is functioning for a specified period of time.

As an example, let us assume a very simple system, consisting of one pump pumping water from one place to another.

The pump has the following characteristics:

- in average constant number of failures per year
- in average one failure every three years
- the failures are detected immediately
- the repair time is five days
- after repair the pump is as good as before.

The system will fail if the pump fails. The reliability R of the system over a year is equal to the probability that the system is functioning adequately over a period of one year.

$$
\begin{aligned}
R &= P(\text{functioning over 1 year}) \\
&= 1 - P(\text{not functioning over 1 year}) \\
&= 1 - (1/3) \\
&= 2/3
\end{aligned}
$$

The unreliability = probability that the system is not functioning properly for a specified period of time
= 1 - reliability.

The steady state availability of a system is the probability that the system is in an acceptable state at any instant of time t, given that the system was fully operative at time t = 0.

If we again consider the example, let us calculate the availability A of the system at time t = 1 year, assuming that the pump was running at time t = 0. The system will in average fail once every three years and it will be down for repair for five days. This means, that

$$A = 1 - (1/3) \cdot (5/356)$$
$$\approx 1 - 0.005 = 0.995$$

The unavailability = 1 - availability.

These two characteristics are important when studying the availability of a system or a whole plant. Further, the characteristics are needed in risk analysis when estimations of accident probabilities are considered.

Tools for this type of calculations are the subject of chapter 4.

2.3. Reliability of a Component

The reliability of a component is in general estimated from experience or from test data describing the behaviour of the component given the operating conditions. Statistical methods are available for the estimation.

In some cases no prior knowledge of the reliability is available. This will be the case if the component is designed using a new technology, new materials or subject to new environmental loads. In such cases it is not possible to use experience and data from components based on other technological design, other materials and other loads. In case that the component in question is a structure an alternative way of estimating its reliability is to treat the component as a system (or a structure) characterized by its strength and the loads imposed. Based on these properties it is possible to estimate the probability that the component will function - the reliability of the component.

Tools for calculation of the reliability of components (structures) are discussed in chapter 3.

3. STRUCTURAL RELIABILITY

This chapter focusses on the problems in structural reliability as mentioned in section 2.3.

3.1. Background

The reliability of a structure can be assessed by using historical data from past failures and non-failures. This is a very simple approach given sufficient data, but it suffers in that only specific structures are considered. The effect of a change in one of the variables which has an influence on the behaviour of the structure, often cannot be predicted as the required data probably do not exist.

An alternative approach is the development and application of engineering models based on an understanding of the failure models and statistical distributions of the variables which have an influence on the behaviour of the structure.

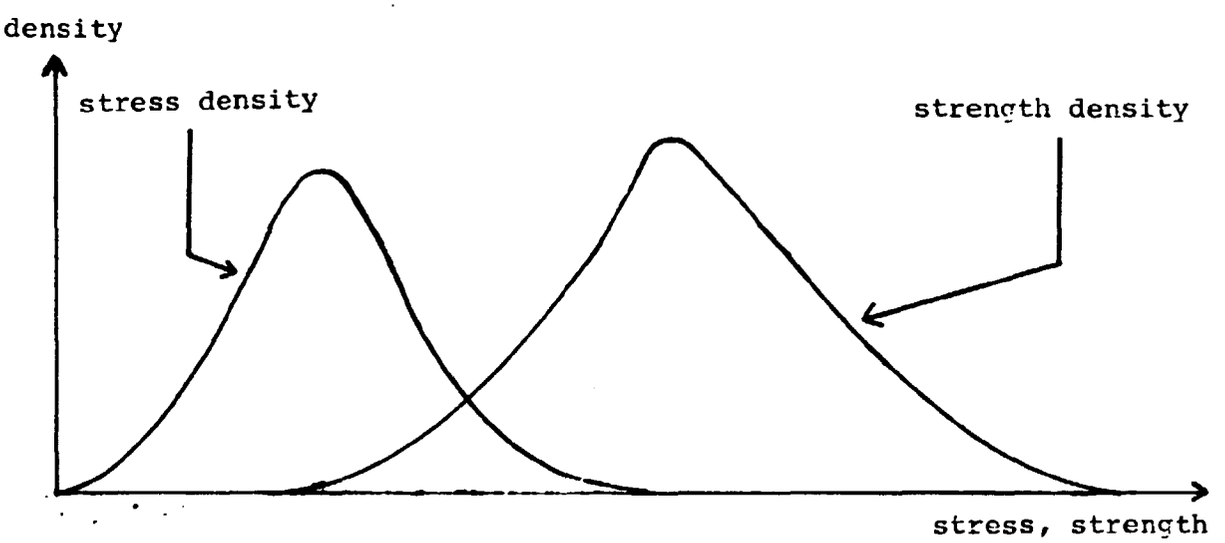
3.2. Stress-Strength Models

One class of engineering models is the stress-strength models. The models are based on the assumption that the structure has a certain strength and that the structure is subject to certain stresses or loads. The strength as well as the stress is defined by an expression in a number of variables. Each variable has an associated distribution. As long as the strength is greater than the stress the structure endures; otherwise the structure fails.

Models of this kind have been applied in probabilistic fracture mechanics especially where high integrity is required, such as pressure components of a nuclear power plant and offshore structures.

Within the limits of application of linear elastic fracture mechanics the relevant quantity for the strength is the fracture toughness, K_{IC}. The relevant quantity for the stress is the stress intensity factor, K_I.

The models are illustrated below where the stress density and the strength density are shown.

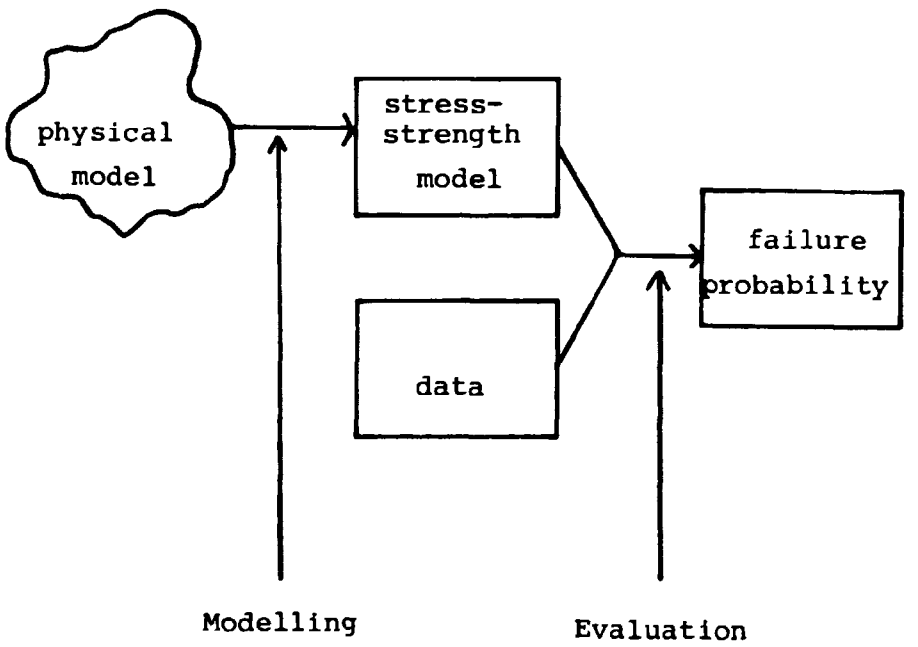


Stress-strength model

For a given value of the stress ,x, the problem is to evaluate the probability that the strength is less than this value, integrated over all values of x.

3.3. Limitations and Conditions

Below is shown a figure which illustrates the scheme for assessment of the reliability of a structure using a stress-strength model



Assessment of the reliability of a structure

This study discusses the evaluation procedure given a stress-strength model and the data for the variables involved.

This report does not discuss the modelling of some specific structures. Neither does it discuss the problem of data which is also a very important task, since the use of engineering models introduces new errors because of the lack of data and the complexity of the approach which requires more input variables to give a detailed description of the behaviour of the structure.

The probability of failure of a structure is calculated by the probability that the stress is greater than the strength.

$$P(\text{stress} > \text{strength})$$

The probability density functions associated with the stress and the strength are given by $f_{stress}(x)$ and $f_{strength}(x)$ respectively. The corresponding cumulative density functions $F_{stress}(x)$ and $F_{strength}(x)$ are given by

$$F_{stress}(x) = \int_{-\infty}^{x} f_{stress}(t)\, dt$$

and

$$F_{strength}(x) = \int_{-\infty}^{x} f_{strength}(t)\, dt$$

Then

$$P(\text{stress} > \text{strength}) = \int_{x=-\infty}^{x=\infty} P(\text{stress} = x \wedge \text{strength} < x)\, dx$$

If stress and strength are statistically independent it follows that

$$P(\text{stress} > \text{strength}) = \int_{x=-\infty}^{x=\infty} P(\text{stress} = x \wedge \text{strength} < x)\, dx$$

$$= \int_{x=-\infty}^{x=\infty} P(\text{stress} = x) \cdot P(\text{strength} < x)\, dx$$

$$= \int_{x=-\infty}^{x=\infty} f_{stress}(x)\, F_{strength}(x)\, dx$$

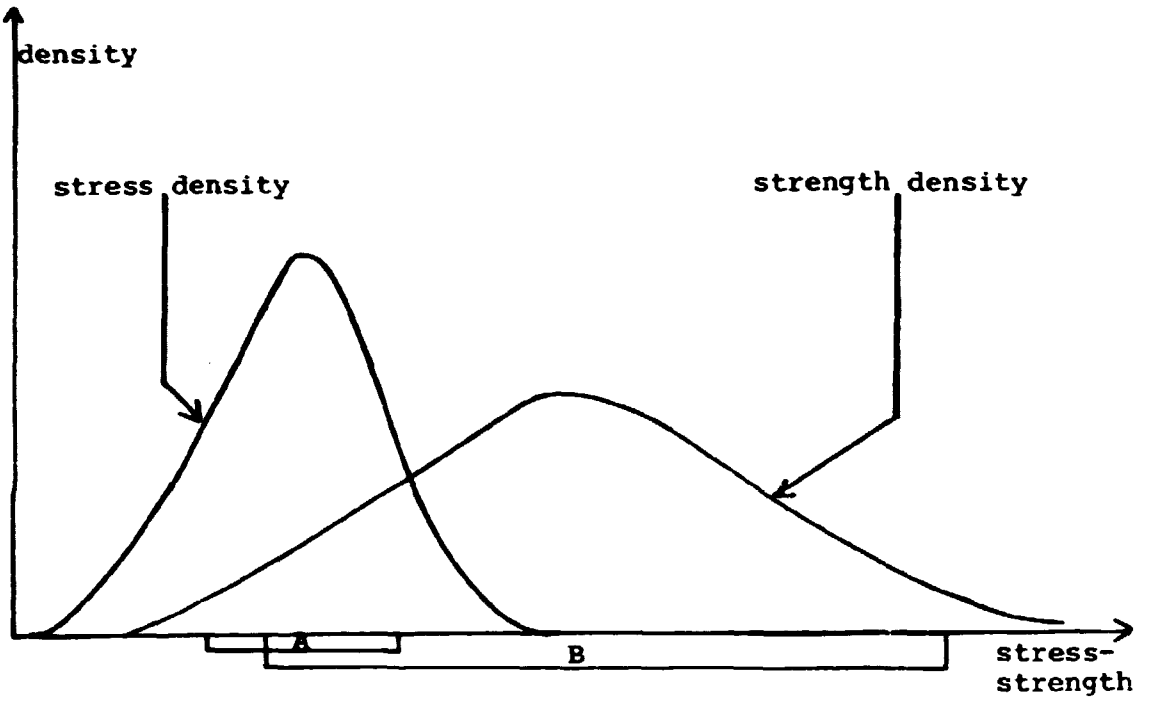
$$= \int_{x=-\infty}^{x=\infty} f_{stress}(x) \int_{y=-\infty}^{y=x} f_{strength}(y)\, dy\, dx$$

$$= \int_{x=-\infty}^{x=\infty} \int_{y=-\infty}^{y=x} f_{stress}(x)\, f_{strength}(y)\, dy\, dx$$

Approximation of the p.d.f.'s - Unreliable structures

If the p.d.f.'s for stress and strength can be approximated by a simple probability density function, i.e. a normal distribution function, the integral can be easily calculated numerically.

This can be utilized, if an unreliable structure is considered. This situation is illustrated below.



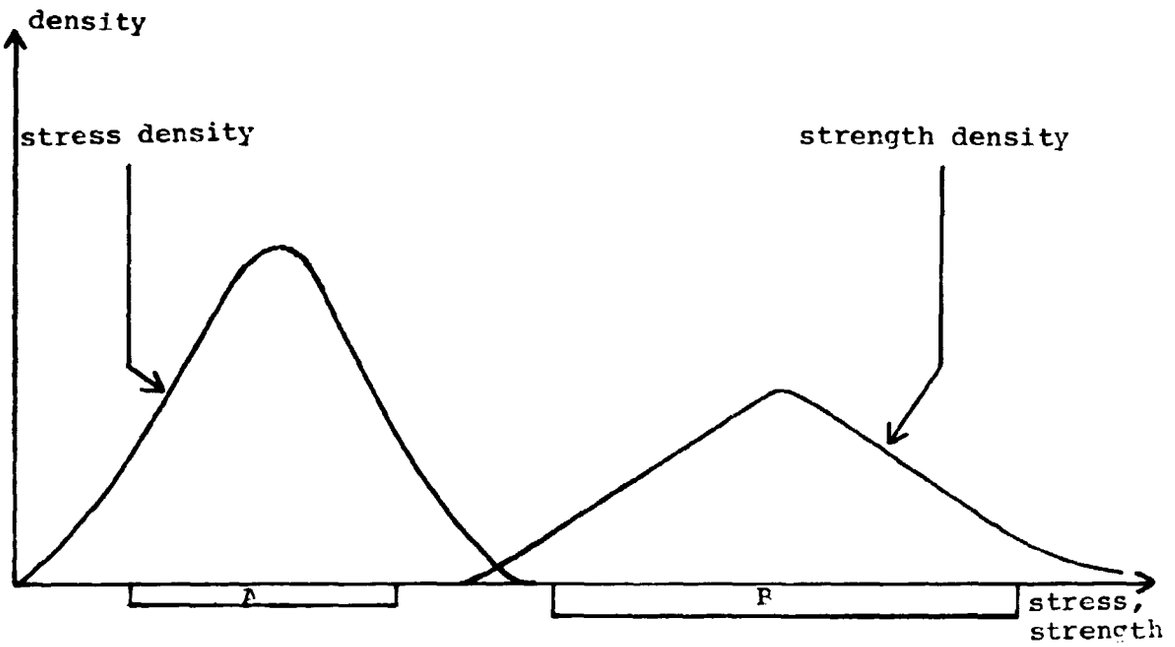
Stress-strength model for an unreliable structure

The areas A and B indicate the areas in which data are available.

In this case the approximation of the p.d.f. by a normal distribution is simple and the values on which the integral is evaluated are well described.

Approximation of the p.d.f.'s - reliable structures

In many cases considering nuclear components and off-shore components, the structure is a highly reliable structure. This situation is illustrated below.



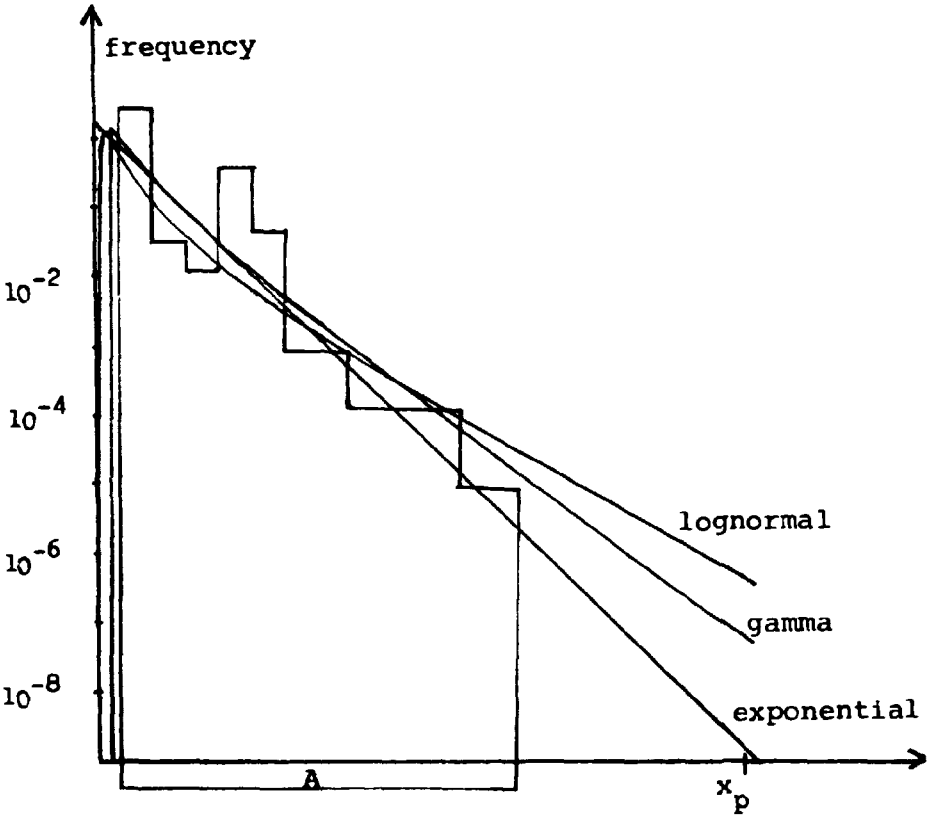
Stress-strength model for a reliable structure

Again the areas A and B indicate the areas in which data are available for approximations of p.d.f.'s for stress and strength. The values on which the evaluation of the integral is based are not well described as they belong to the tails of the distributions, where data are not available.

In this case the result is very sensitive to the tails of the distributions. The following example shows the influence on the result applying different distribution functions as approximations to the data available.

Example

Assume that the data for the stress are as shown in the following figure. The data are fitted to an exponential, a log-normal, and a gamma distribution which are also shown in the figure.



Fitting of different p.d.f.'s to data

In the area A where data are available the exponential, the log-normal, and the gamma distribution fit quite well the data. The tail where no data are available is dependent on the distribution which is chosen. As it is seen the probability associated with the value X_p differs by two or three orders of magnitude depending on the type of the distribution.

The evaluation of the integral is based on a number of values which all belong to the tail (i.e. the situation is similar to that shown above considering the value X_p).

In case of highly reliable structures which are of interest approximations of the p.d.f.'s for stress and strength cannot be used. It is therefore necessary to evaluate the integral directly based on the distribution for each variable describing the stress and the strength.

Application

Within the limits of application of linear elastic fracture mechanics a stress-strength model is described. The theory assumes that the crack is stable, when the stress intensity factor K_I around a crack is smaller than the fracture toughness K_{IC},

$$K_I < K_{IC}.$$

The stress intensity factor, K_I, is given by the relation

$$K_I^2 = K \cdot A \cdot S^2$$

where K is a correction factor
A is the crack depth
S is the applied nominal stress.

The fracture toughness, K_{IC}, is given by the relation

$$K_{IC}^2 = M \cdot S_y \cdot C_v - 0.05 \cdot S_y^2$$

where M is a correction factor
S_y is the yield stress
C_v is the Charpy V notch energy.

A probability density function is associated with each parameter in the expressions.

The probability of a failure is calculated by the probability that the stress intensity factor is greater than the fracture toughness:

$$P(K_I > K_{IC})$$

$$= P(K_I^2 > K_{IC}^2)$$

3.4. Calculation Methods

During the last 10 years some calculation methods have been published. Two of them have been developed and implemented at Risø while the third one is developed at the Joint Research Centre at Ispra in Italy. A brief description of the main ideas behind the methods and the advantages and disadvantages of the methods are given below.

3.4.1. The PEP Code

The PEP706 computer program is described in detail in [2] . The program is based upon Monte Carlo simulation. A simple implementation of the importance sampling technique is available too.

The program simply generates values for each variable involved according to the distributions specified in the input. In each case the quantities, stress and strength are calculated from the actual values simulated and it is tested, if stress is greater than strength. The process is carried out for a large number of trials. The generation of values from each individual distribution is based upon a random number generator which generates numbers uniformly distributed over the interval [0,1].

The number of trials required to obtain an acceptable accuracy is extremely large, thus requiring a large amount of computer time. Especially when the failure probability is very low as in the case of highly reliable structures. A typical example involving four parameters uses about 300 seconds computer time to obtain 10% accuracy with a failure probability of 10^{-7}.

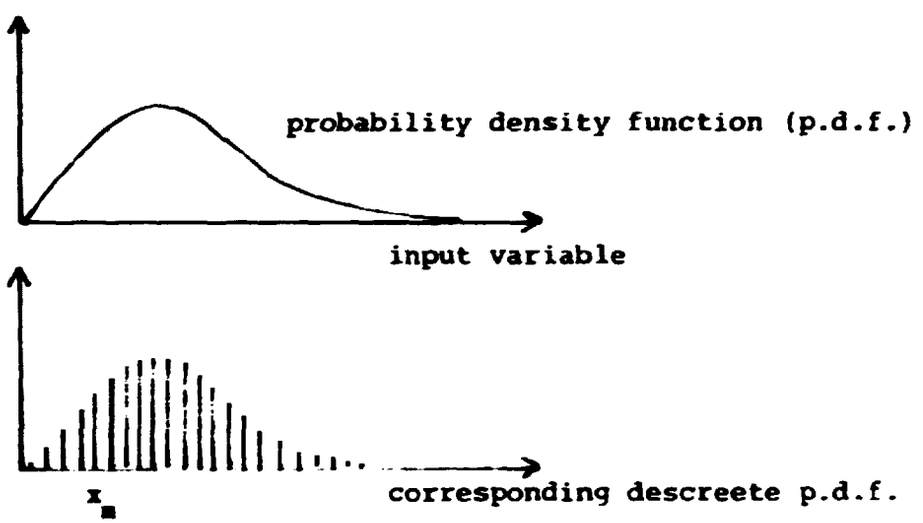
The main advantage of the program is the great flexibility in modelling, where a large number of models can be treated without imposing approximations. Likewise it is not necessary to make any approximations to the input distributions.

The main disadvantage is the large amount of computer time required to obtain an acceptable accuracy.

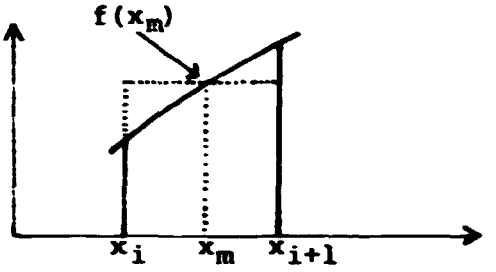
3.4.2. The ANPEP Code

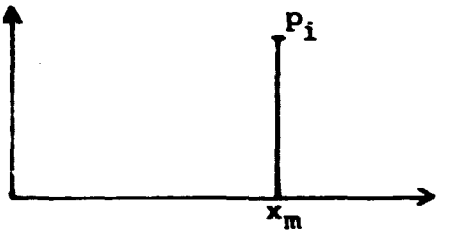
The ANPEP code is described in detail in [3]. In [4] the ANPEP/V2 program is described. This is a further development of the ANPEP code, but it is based upon the same basic idea.

The program is based upon a numerical method to combine random variables. Each distribution specified in the input is represented by a corresponding discrete distribution in the following way:



The values p_i of the discrete density function at a given point X_m is calculated in the following way:





$$x_m = \frac{x_{i+1} + x_i}{2} \qquad p_i = (x_{i+1} - x_i) \cdot f(x_m)$$

The discretization process is carried out for each distribution by dividing the axis into a number of intervals of equal size or of a size which is specified by the user. Now the discretized distributions are combined and the failure criterion, stress greater than strength is checked for any combination. An observed failure contributes to the total failure probability by the product of the corresponding probabilities.

The main advantage, compared to Monte Carlo simulation is the reduction in computer time for smaller problems, i.e. involving four or less variables. The ANPEP code requires approximately 20 seconds for a problem, where PEP706 requires approximately 300 seconds. Secondly, no statistical deviations are introduced, since generation of random numbers is avoided.

The main disadvantages are the reduction in flexibility of the models and the introduction of approximations to the input distributions. From this approximation two problems arise: i) the choice of the number of intervals and ii) the determination of the end point of an unlimited distribution. Both problems are solved empirically. Another problem is that the computer time required is very sensitive to the number of input variables. The problem mentioned above which requires 20 seconds involves four input variables. A similar problem involving six variables requires approximately 2000 seconds.

3.4.3. The COVAL and SCORE Code

The SCORE computer program is developed and implemented at the Joint Research Centre at Ispra in Italy. A description of the program is given in [5] and [6]. This is a further development of the COVAL code.

The program is based upon a numerical method to combine random variables. Each distribution specified in the input is represented by a corresponding histogram of equal probability intervals. Then the program systematically combines the intervals of the random variables.

Let the area of interest for an input variable be divided into N subintervals and let the cumulative density function be given by

$$F(x) = \int_{x_o}^{x} f(t)dt$$

then the endpoint X_{i+1} of subinterval no. i is calculated as the solution to the equation

$$\int_{x_i}^{x_{i+1}} f(t)dt = \frac{1}{N}$$

which is the same as

$$F(x_{i+1}) - F(x_i) = \frac{1}{N}$$

The idea is very similar to that proposed in ANPEP/V2. The main difference is in the approximation of the input distributions. In ANPEP/V2 the intervals are chosen with respect to equal length. In SCORE the intervals are chosen with respect to equal probability.

The advantages and disadvantages compared to Monte Carlo simulation are very similar to those given for the ANPEP/V2 program. The advantages compared to the ANPEP/V2 program are:

- the intervals into which a distribution is subdivided are small, where the density is high and are large, where the density is low. Thus each distribution is reasonably well approximated.

- combinations of equal probability intervals form again equal probability intervals. This prevents the creation of intervals from having probabilities very different one from the other.

Still the problem remains to choose a reasonable number of subintervals and the endpoint for unlimited probability density functions.

3.5. Alternative Methods

In this section different approaches to the evaluation of integrals in several variables are discussed. The general problem is to approximate

$$\int_{R_n} \cdots \int f(x_1,\ldots,x_n)dx_1\ldots dx_n$$

by a sum

$$\sum_{i=1}^{N} A_i \cdot f(v_{i,1},\ldots,v_{i,n})$$

where

N is the number of points in the formula
A_i is the coefficient number i
R_n is a n-dimensional region
$(v_{i,1},\ldots,v_{i,n})$ is integration point number i.

The region R_n in this case is a parallelepiped which can be transformed into the n-cube C_n by an affine transformation. Therefore, the discussion can be restricted to the region C_n.

The approaches can be divided into four groups which are treated separately:

- product formulas
- non-product formulas
- adaptive methods
- Monte Carlo methods

3.5.1. Product Formulas

This topic is discussed in great detail in [8] and [9]. Some of the interesting results from the viewpoint of application in structural reliability are presented below.

Product formulas for C_n are formulas which are constructed by products, or combinations of formulas for regions C_m of lower dimension, $m<n$. Especially, construction of formulas for C_n by products of n formulas for C_1 is of interest.

The integral

$$\int_{C_n} \cdots \int f(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

can be evaluated by a formula constructed in this way.

If a formula exists for a one-dimensional integral

$$\int_{C_1} f(x)\, dx = \sum_{i=1}^{N} A_i \cdot f(v_i),$$

$$C_1 = [0, 1]$$

then

$$\int \ldots \int_{C_n} f(x_1,\ldots,x_n)\, dx_1 \ldots dx_n$$

$$= \sum_{\substack{1 \le i_k \le N \\ k=1,2,\ldots,n}} A_{i_1} \cdot A_{i_2} \ldots A_{i_n} f(v_{i_1},\ldots,v_{i_n}).$$

Formulas constructed in this way are very easy to handle and simple to implement. The main drawback is that the number of integration points increases rapidly with n, the number of dimensions. Suppose that a 8-points formula exists for a one-dimensional integral. The number of points in a formula for a n-dimensional integral is shown for various values of n.

n	number of points
2	64
3	512
4	4.096
5	32.768
6	262.144
10	1.073.741.824

From this table it is obvious that using a 8-points formula for the one-dimensional integral construction of formulas for integrals of a dimension greater than 5 can be time consuming and will therefore in some cases be impractical.

Example

A two-points Gauss formula of degree 3 over $C_1 = [-1, 1]$ is

$$\int_{-1}^{1} f(x)dx = 1 \cdot f(\sqrt{3}/3) + 1 \cdot f(-\sqrt{3}/3)$$

This means, that the four-point product formula over the region $C_2 = [-1, 1] \times [-1, 1]$ based on the above formula is

$$\int_{-1}^{1}\int_{-1}^{1} f(x,y)dxdy = 1 \cdot f(\sqrt{3}/3, \sqrt{3}/3) + 1 \cdot f(-\sqrt{3}/3, \sqrt{3}/3) + 1 \cdot f(\sqrt{3}/3, -\sqrt{3}/3) + 1 \cdot f(-\sqrt{3}/3, -\sqrt{3}/3)$$

3.5.2. Non-Product Formulas

In [9] and especially in [8] other types of formulas are discussed. This is a very difficult subject, since no general theory is available for construction of other types of formulas than the product formulas mentioned above. An existence theorem has been proved [8, page 54], but it includes no guidance for constructing a formula with given characteristics. In [8] a very comprehensive discussion on this subject is presented, including a large collection of special results and formulas.

The main advantage of a non-product formula compared to a product formula is that fewer integration points are needed for a formula of a given degree. If a formula of degree d with m points exists for a given region R, the product formula of degree d over R^n has m^n points. In some cases it is possible to construct a non-product formula of degree d with less than m^n points.

Example

The region is $C_n = [-1, 1]^n$

The product Gauss formula of degree 5 has 3^n points.

A non-product formula has been constructed by Stroud and Hammer with $2n^2+1$points

n	3^n	$2n^2+1$	Reduction (in percent)
2	9	9	0
3	27	19	30
4	81	33	59
5	243	51	79

This table illustrates the advantage of an effective non-product formula compared to a similar product formula. It is also known that for a given degree d, that reduction in the number of point is increasing with the dimension. Unfortunately, the difficulties in constructing non-product formulas for higher degrees are also increasing. Therefore, the reduction i· the number of points needed in a formula does not always com ,sate the difficulties in the construction.

The difficulties arise from the lacking theory about orthogonal polynomials in several variables. In one variable the theory is fairly simple. Further, the zeroes of orthogonal polynomials can be used in construction of formulas. The difficulties in several variables are of two types:

- the one-dimensional space is very simple compared to higher-dimensional spaces. Since all line segments are equivalent under an affine transformation there is essentially only one bounded connected region in the Euclidean space. In higher-dimensional spaces these simple rules do not apply. For example, the square, the circle, and the triangle are regions in the two-dimensional space which are not equivalent under an affine transformation.

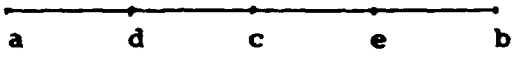
- the theory of orthogonal polynomials is well-known in one variable and simply related to integration formulas. In several variables the theory is much more complicated and not completely described. Until now no formulas have been constructed *in n dimensions, n>2, using* the theory of orthogonal polynomials in n variables.

In practice a couple of other difficulties in constructing non-product formulas arise. The weights in the formulas are not always positive, which should be a natural property for a formula. Furthermore, it is not certain that all integration points belong to the region, which also should be a natural property.

3.5.3. Adaptive Methods

Adaptive methods are methods where the calculation is performed with stepwise refinements within important subregions. In each step the calculation is continued within the subregion, where the highest benefit of a refined calculation is expected. The idea can be illustrated in this way:

f(x) is to be integrated over [a,b]. We choose a formula which is used over [a,c] and ⌊c,b], where c = (a+b)/2.



The calculated integrals are I_1 and I_2 respectively. Then the integrals I_{11}, I_{12}, I_{21}, and I_{22} over [a,d], [d,c], [c,e], and [e,b] are calculated. The process is continued with refined calculations within the subregion with the highest difference, D_1 or D_2, where $D_1 = I_1 - (I_{11} + I_{12})$ and $D_2 = I_2 - (I_{21} + I_{22})$.

It is very advantageous if the integration points already used can be reused in all subsequent refinements. Formulas exist with this property, but in general this characteristic is not common.

3.5.4. Monte Carlo Methods

Other approaches to the problem of evaluating integrals in several variables are the Monte Carlo method or number theoretical methods. The Monte Carlo methods are very simple to construct, but their main drawback is the large number of points needed to achieve a reasonable accuracy.

The Monte Carlo method for approximation of an integral in several variables is

$$\int\cdots\int_{R_n} f(x_1,\cdots,x_n)dx_1\cdots dx_n$$

$$\approx (V/N)\sum_{i=1}^{N} f(\underline{v}_i)$$

where V is the volume of R_n
N the number of points
$\underline{v}_i$ point number i in R_n

The points $\underline{v}_i$ are chosen at random uniformly distributed in R_n. In general the points are not generated randomly, but quasirandomly, by a deterministic method. Some common methods are based on the linear congruential method

$x_{i=1} = a\cdot x_i+b$, where a and b are fixed.

A measure of the error in the above formula is given by $c/\sqrt{N}$, and this quantity does not decrease very rapidly as N increases, leading to very large computing times.

Other methods are known - number theoretical methods - where the points are found, so that the error is decreased. Such methods are discussed in [8].

3.5.5. Conclusions

Based on the study of the various methods of numerical integration in several variables, it is concluded that simple product formulas constructed on the basis of simple Gauss formulas in one dimension are recommended. The reason is that Gauss formulas are effective formulas and formulas exist which are specialized for specific regions, such as $[-1,1]$, $[0,\infty[$, $]-\infty,\infty[$.

Secondly, the drawbacks of non-product formulas are too large to be compensated by their reduction in the number of points. In particular, the fact that several formulas have negative weights and integration points outside the integration region make them useless in practical applications.

Finally, it has been found that Monte Carlo methods are too time consuming in practice. In [9] it is shown that Monte Carlo methods are reasonable only in cases where the number of dimensions is greater than 10. Furthermore, they introduce a statistical error in the calculation by nature, which is avoided in other types of approaches.

It is therefore decided to develope a tool to evaluate an integral in several variables using a product formula based upon Gauss formulas in one variable.

3.6. The NUMPEP Code

A computer program, NUMPEP, is developed based on numerical integration in several variables using product formulas.

The probability of failure is given by

$$\int_{x=-\infty}^{x=\infty} \int_{y=-\infty}^{y=x} f_{K_I^2}(x) \cdot f_{K_{IC}^2}(y)\, dy\, dx$$

where the probability density functions associated with K_I^2 and K_{IC}^2 are described.

The program is intended for the class of problems within probabilistic fracture mechanics as described.

The program is prepared for other relationships between the variables which means that it is possible to solve problems within other fields of structural reliability.

Furthermore, the program is modularized, so it is easy to extend the number of variables involved. This extention requires some further work on programming.

The program is written in Fortran for the Burroughs B7800 computer.

A detailed description of the principles is given and the input and the output from the program are shown.

3.6.1. Construction

The requirements for the method used for evaluation of the integral are the following:

- the formula should be simple
- the method should be fast
- the accuracy of the result should be within the first two digits
- the number of variables allowed should be at least six
- the program should be prepared for extension of the number of variables
- the program should be designed to solve the problem: calculate the probability of failure within probabilistic fracture mechanics in two cases:

a. $K_I^2 = K \cdot A \cdot S^2$

where

K is the constant k

A is given by an exponential distribution with parameters λ and x_o,

$f(x) = \lambda \cdot e^{-\lambda(x-x_o)}$

S is given by a normal distribution with parameters x_m and σ,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(x-x_m)^2}{2\sigma^2}}$$

$K_{IC}^2 = M \cdot C_v \cdot S_y$

where

M is the constant m

C_v is given by a Weibull distribution with parameters x_o, m and k,

$$f(x) = k \cdot (x-x_o)^m \cdot e^{-\frac{k(x-x_o)^{m+1}}{m+1}}$$

S_y is given by a Weibull distribution with parameters x_o, m and k,

$$f(x) = k \cdot (x-x_o)^m \cdot e^{-\frac{k(x-x_o)^{m+1}}{m+1}}$$

b. This case is equivalent to case a except that S_y is here given by a normal distribution with parameters x_m and σ,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-x_m)^2}{2\sigma^2}}$$

- the program should be flexible and extensible in order to take into account:
  - other distributions for the variables than those shown above
  - the modelling of $K_{IC}^2 = M \cdot (C_v \cdot S_y - 0.05 \cdot S_y^2)$.

Integration limits

The limits of integration $-\infty$ and $+\infty$ have no physical interpretation. In each case it is possible to specify limits which are meaningful, i.e. if we look at the crack depth, a value less than zero or greater than the wall thickness is meaningless.

The integral to be evaluated is therefore:

$$\int_{x=a}^{x=b} \int_{y=a}^{y=x} f_{K_I}^2(x)\, f_{K_{IC}}^2(y)\, dy\, dx$$

Using Dirichlet's formula, see [7], we find

$$\int_\alpha^\beta \int_\alpha^x \cdots \int_\alpha^x f(x)dx^{k+1} = \frac{1}{k!} \int_\alpha^\beta (\beta-x)^k f(x)dx$$

For k = 1

$$\int_{\alpha}^{\beta} \int_{\alpha}^{x} f(x)\, dx^2 = \int_{\alpha}^{\beta} (\beta - x)\, f(x)\, dx$$

Therefore

$$\int_{a}^{b} \int_{a}^{x} f_{K_I^2}(x) \cdot f_{K_{IC}^2}(y)\, dy\, dx$$

is equal to

$$\int_{a}^{b} (b-x) \cdot f_{K_I^2}(x)\, f_{K_{IC}^2}(x)\, dx$$

Given p.d.f.'s for each variable in the expressions for K_I and K_{IC} it is possible to construct p.d.f.'s for K_I and K_{IC}. These are unfortunately not well-known p.d.f.'s, so it is necessary to construct the density functions explicitely using the theory of transformation of variables.

Transformation of variables

In this section a short description of the formulas for transformation of variables is given.

One variable

The formula for transformation of one variable is given:

$$h(u) = \left|\frac{dx}{du}\right| \; f(x(u))$$

where

x is the original variable
u is the transformed variable
f(x) is the probability density function of x
h(u) is the probability density function of u
x(u) is the expression relating x to u
$\left|\frac{dx}{du}\right|$ is the absolute value of the derivative of x with respect to u

Example

Given X with p.d.f. f(x)

We want the p.d.f., h(u), related to the transformation $u = a + x$, where $a \in R$

$$x = u-a \qquad \frac{dx}{du} = 1$$

Then $h(u) = \left|\frac{dx}{du}\right| \cdot f(x(u))$

$= f(u-a)$

Several variables

The formula for changing several variables is given similarly by

$$\Psi(u,v) = \phi(x(u,v),y(u,v))|Det(J)|$$

$$Det(J) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

where

x,y are the original variables
u,v are the transformed variables
$\phi(x,y)$ is the original joint density function of x and y
$\Psi(u,v)$ is the transformed joint density function of
u and v
u(x,y) are the expressions relating u and v
v(x,y) to x and y
Det(J) the determinant of the Jacobian

Normally u is the desired transformation and v a convenient dummy variable. In this case we want the marginal density function for u associated with the joint density function $\Psi(u,v)$ which is given by:

$$h(u) = \int_v \Psi(u,v)\, dv.$$

If x and y are statistically independent the equation simplifies to

$$\Psi(u,v) = f(x(u,v)) \cdot g(y(u,v)) \cdot |Det(J)|$$

If v(x,y) = x, which is often convenient, and correspondingly x(u,v) = v, then

$$Det(J) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = -\frac{\partial y}{\partial u}$$

The marginal density function, h(u), is then given by

$$h(u) = \int_v \left|\frac{dy}{du}\right| . f(v) \cdot g(y(u,v)) \, dv.$$

Example

Given x and y with p.d.f.'s f(x) and g(y). We want the p.d.f., h(u), related to the transformation given by

$u = x + y$
$v = x$

Then

$x = v$
$y = u-v$ and $\frac{\partial y}{\partial u} = 1$

$$h(u) = \int_v |1| \cdot f(v) \cdot g(u-v) \, dv$$

$$= \int_v f(v) \cdot g(u-v) \, dv$$

(which is the well-known convolution integral).

Using these formulas for transformation of variables it is possible to express the p.d.f.'s related to the stress intensity factor and the fracture toughness.

The stress intensity factor

The stress intensity factor is given by

$$K_I^2 = K \cdot A \cdot S^2$$

K is often specified by a constant. In other cases by a p.d.f., h(x).

A is specified by a p.d.f., f(x) which is often an exponential distribution.

S is specified by a p.d.f., g(x) which is often a normal distribution.

Two variables

The most simple case is when K is given by a constant value.

Using the formulas of transformation of variables given above we find h(u) associated with S^2 as

$$h(u) = \frac{1}{2\sqrt{u}} \left(g(\sqrt{u}) + g(-\sqrt{u})\right)$$

Associated with $A \cdot S^2$ we get,

when $U = A \cdot S^2$ $\qquad A = V$
$V = A$ $\qquad S^2 = \frac{U}{V}$

and

$$\left|\frac{\partial S^2}{\partial U}\right| = \left|\frac{1}{v}\right|$$

$$h(u) = \int_v \left|\frac{1}{v}\right| \cdot f(v) \cdot \frac{1}{2\sqrt{\frac{u}{v}}} \cdot g\left(\sqrt{\frac{u}{v}}\right) + g\left(-\sqrt{\frac{u}{v}}\right) dv$$

Associated with $K \cdot A \cdot S^2$ we find (K takes the value k)

$$h(u) = \int_v \frac{1}{k} \cdot \left|\frac{1}{v}\right| \cdot f(v) \cdot \frac{1}{2\sqrt{\frac{u}{kv}}} \left(g\left(\sqrt{\frac{u}{kv}}\right) + g\left(-\sqrt{\frac{u}{k \cdot v}}\right)\right) dv.$$

Three variables

As in the case of two variables we find associated with $A \cdot S^2$

$$1(u) = \int_v \frac{1}{|v|} \cdot f(v) \cdot \frac{1}{2\sqrt{\frac{u}{v}}} \cdot \left(g\left(\sqrt{\frac{u}{v}}\right) + g\left(-\sqrt{\frac{u}{v}}\right)\right) dv$$

The p.d.f. for K is given by h(x).

Associated with $K \cdot A \cdot S^2$ we find, when

$$U = K \cdot A \cdot S^2 \qquad K = \frac{U}{V}$$

$$V = A \cdot S^2 \qquad A \cdot S^2 = V$$

and

$$\left|\frac{\partial K}{\partial U}\right| = \left|\frac{1}{V}\right|$$

$$m(u) = \int_w \left|\frac{1}{w}\right| \cdot h(w) \cdot 1\left(\frac{u}{v}\right)\, dw$$

$$= \int_w \int_v \left|\frac{1}{vw}\right| \cdot h(w) \cdot f(v) \cdot \frac{1}{2\sqrt{\frac{u}{vw}}} \cdot \left(g\left(\sqrt{\frac{u}{vw}}\right)+g\left(-\sqrt{\frac{u}{vw}}\right)\right)\, dvdw$$

The fracture toughness

The fracture toughness is given by

$$K_{IC}^2 = M \cdot (C_v \cdot S_y - 0.05 \cdot S_y^2).$$

M is often specified by a constant. In other cases by a p.d.f., h(x).

C_v is specified by a p.d.f., f(x) which is often a normal or a Weibull distribution.

S_y is specified by a p.d.f., g(x) which is often a Weibull distribution.

In some cases the last term $M \cdot S_y^2 \cdot 0.05$ is negligible and the expression is

$$K_{IC}^2 = M \cdot C_v \cdot S_y$$

Using the formulas of transformation of variables given above we find h(u) associated with $C_v \cdot S_y$ as

$$h(u) = \int_v \left|\frac{1}{v}\right| \cdot f(v) \cdot g\left(\frac{u}{v}\right) dv$$

Two variables

If M is specified by a constant value m the p.d.f. associated with $M \cdot C_v \cdot S_y$ is

$$h(u) = \int_v \left|\frac{1}{v \cdot m}\right| \cdot f(v) \cdot g\left(\frac{u}{v \cdot m}\right) dv.$$

Three variables

If M is specified by a p.d.f., h(x), the p.d.f. associated with $M \cdot C_v \cdot S_y$ is

$$l(u) = \int_w \left|\frac{1}{w}\right| \cdot h(w) \cdot \int_v \left|\frac{1}{v}\right| \cdot f(v) \cdot g\left(\frac{u}{v \cdot w}\right) dv\, dw$$

$$= \int_w \int_v \left|\frac{1}{v.w}\right| \cdot h(w) \cdot f(v) \cdot g\left(\frac{u}{vw}\right) dv\, dw.$$

Conclusion

We will evaluate an integral of the form

$$\int_a^b (b-x) f_1(x) f_2(x) dx$$

where $f_1(x)$ and $f_2(x)$ are given by one of the following two types:

$$f(x) = \int_{v=x_0}^{v=\infty} g(x,v)dv$$

or

$$f(x) = \int_{w=x_0}^{w=\infty} \int_{v=x_1}^{v=\infty} g(x,v,w)dv\, dw.$$

A Gauss-Legendre formula is well-suited for the integration

$$\int_a^b f(x)dx.$$

In principle is it possible to construct a product formula based on this type of formula. It has been found that in the cases of interest both $f_1(x)$ and $f_2(x)$ can be rewritten using a simple affine transformation.

Then

$$f_1(x) = \int_{v'=0}^{v'=\infty} e^{-v'} g'(x,v')dv'$$

or

$$f_1(x) = \int_{v'=0}^{v'=\infty} \int_{w'=0}^{w'=\infty} e^{-v'} e^{-w'} g'(x,v',w')dv'dw'$$

In evaluation of integrals of this type Gauss-Laguerre formulas are well-suited.

The difference between the application of a Gauss-Legendre and a Gauss-Laguerre formula is shown in the example below.

Example

$$\int_a^b f(x)dx$$

where

$$f(x) = e^{-x}/(x+1),\ a = o,\quad b = \infty,\ x \geq o$$

A Gauss-Laguerre formula is applied for evaluating the integral of $1/(x+1)$ with the weighting function e^{-x}.

A Gauss-Legendre formula is applied for evaluating the integral of $e^{-x}/(x+1)$ with the weighting function 1.

In the latter case the upper integration limit has been chosen in the following way:

Since $1/(x+1) \leq 1$ for $x \geq o$
then $e^{-x}/(x+1) \leq e^{-x}$ for $x \geq o$

If the function is cut off at some value N

$$\int_o^\infty e^{-x}/(x+1)dx = \int_o^N e^{-x}/(x+1)dx + \int_N^\infty e^{-x}/(x+1)dx$$

The last term

$$\int_N^\infty e^{-x}/(x+1)dx \leq \int_N^\infty e^{-x}dx = e^{-N}$$

Choose N, so that $e^{-N} < \delta$.

In this example $\delta = 10^{-4}$ is chosen.

$$e^{-N} = 10^{-4}$$

or

$$N = 9.210340$$

An upper limit N = 10 is chosen, which gives an error less than $e^{-10} = 4.54 \cdot 10^{-5}$.

Another possibility would have been to transform the integral over [o, ∞[into an integral over [a,b].

Let $y = 1/(x+1)$

or $x = (1-y)/y$

then $o \leq x < \infty$ or $o < y \leq 1$

and $dx/dy = -1/y^2$

So

$$\int_o^{\infty} e^{-x}/(x+1)dx = \int_o^1 (e^{-(1-y)/y})/y \, dy$$

A Gauss-Legendre formula can now be used directly without any cutting on the new function.

Using a Gauss-Legendre formula on [0.10] and a Gauss-Laguerre formula, the following results are obtained:

Number of points	Gauss-Legendre	Gauss-Laguerre
2	0.194292	0.571429
4	0.541168	0.593301
8	0.595932	0.595867

Using 30 points the result is 0.595278.

It is therefore decided, that the Gauss-Laguerre formulas shall be used as the basis for constructing the product formulas in combination with the Gauss-Legendre formula.

These formulas are described in [10].

3.6.2 Input and Output

The input as well as the output from the program are described below including a check facility which is available.

Input

The input consists of five categories of data:

1. The type of the distributions
2. The value of the parameters in each distribution
3. The number of integration points
4. The integration limits
5. Indicator for the check facility

The functional relationship used in the model has to be expressed explicitly in the program.

Type

The following four types of distributions are allowed in the program:

1. Constant value
2. Normal distribution
3. Exponential distribution
4. Weibull distribution

The sequence of indicators for the type of the six variables used in the model defines the model in connection with the functional relationship between the variables specified in the program.

Example

The sequence 1-1-1-1-1-1 indicates a model, where all six variables are given by a constant value.

The sequence 1-3-2-1-4-4 indicates the model which is discussed in the main example (see 3.6.3.).

Parameter Values

The values of the parameters are specified, depending on the type of the distribution, in the following form:

type = 1 (constant)	: the value of the constant
type = 2 (normal distribution)	: the mean value x_m and the standard deviation σ
type = 3 (exponential distribution)	: the origin x_o and the scale parameter λ
type = 4 (Weibull distribution):	the origin x_o, the scale parameter k and the form parameter m

Number of integration points

Two values are required. The first one, N_1, defines the number of points used in the Gauss-Legendre formula

$$\int_a^b (b-x) f_{K_I}{}^2(x) f_{K_{IC}}{}^2(x)\, dx = \sum_{i=1}^{N_1} w_i (b-x_i) f_{K_I}{}^2(x_i) f_{K_{IC}}{}^2(x_i)$$

where w_i specifies the weight.

The other value, N_2, defines the number of integration points used in evaluation of $f_{K_I}{}^2(x)$ and $f_{K_{IC}}{}^2(x)$ by the Gauss-Laguerre formula

$$f_{K_I}{}^2(x_k) = \int_o^\infty e^{-w} \cdot f_1(w, x_k)\, dw$$

$$= \sum_{i=1}^{N_2} f_1(w_i, x_k) \cdot v_i$$

where v_i is the weight.

And correspondingly

$$f_{K_{IC}^2}(x_k) = \int_0^\infty e^{-w} \cdot f_2(w, x_k)\, dw$$

$$= \sum_{i=1}^{N2} f_2(w_i, x_k) \cdot v_i$$

where v_i is the weight.

The integration limits

The limits a and b in the formula

$$\int_a^b (b-x) f_{K_I^2}(x) f_{K_{IC}^2}(x)\, dx = \sum_{i=1}^{N1} w_i (b-x_i) f_{K_I^2}(x_i) f_{K_{IC}^2}(x_i)$$

should be specified.

Check facility

It is possible to print intermediate results from the integration process. For each integration point x_i the following values will be printed

$$x_i,\ (b-x_i),\ f_{K_I^2}(x_i),\ f_{K_{IC}^2}(x_i).$$

Output

Execution of the program will give the following output:

1. Identification of the computer run
2. Input for the program

3. Intermediate check result (optional, see above)
4. The resulting probability of failure.

In section 3.6.3. two examples are shown with the corresponding printer output in appendix 1.

3.6.3. Examples

Two examples of the use of the program NUMPEP are given. In appendix 1 the output from the computer runs are given.

Example 1

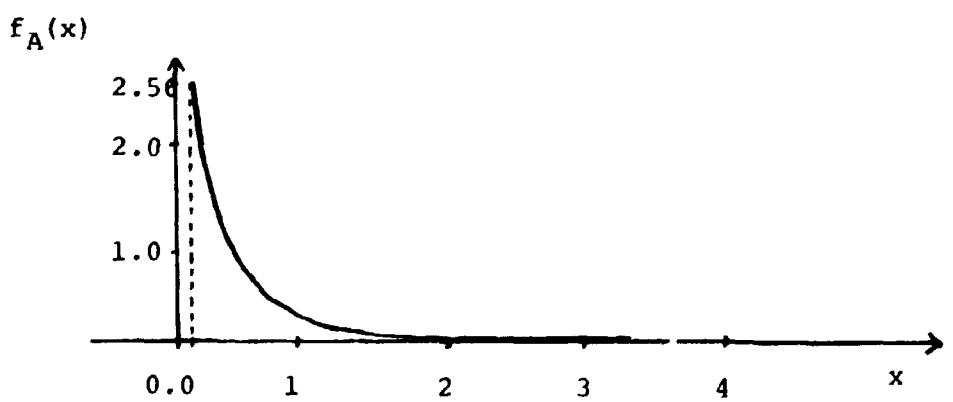
The data originate from [1].

The model is

$$K_I^2 = K \cdot A \cdot S^2$$

where

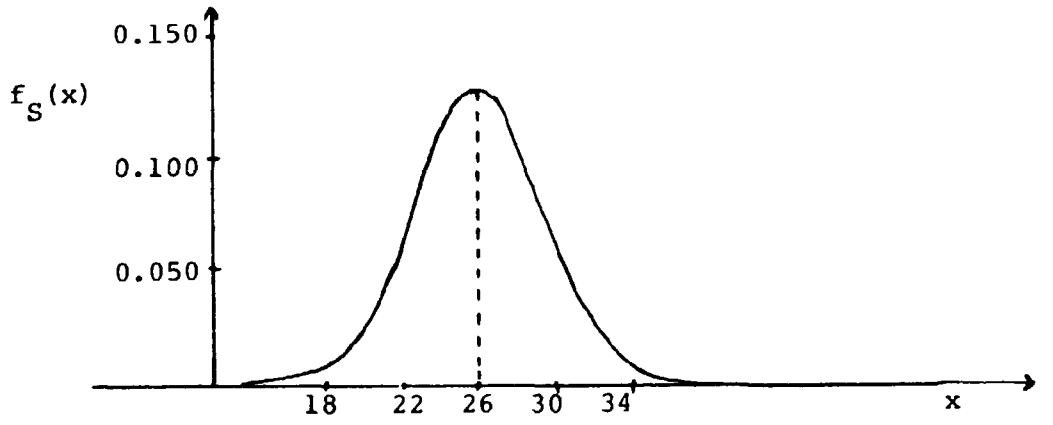
K is the constant 3.8
A is given by the exponential distribution
$f_A(x) = \lambda \cdot \exp(-\lambda(x-x_o))$ with $\lambda = 2.56$ and $x_o = 0.1$



S is given by the normal distribution

$$f_S(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-x_m)^2}{2\sigma^2}\right)$$

with $x_m = 26.0$ and $\sigma = 3.2$



$$K_{IC}^2 = M \cdot C_v \cdot S_y$$

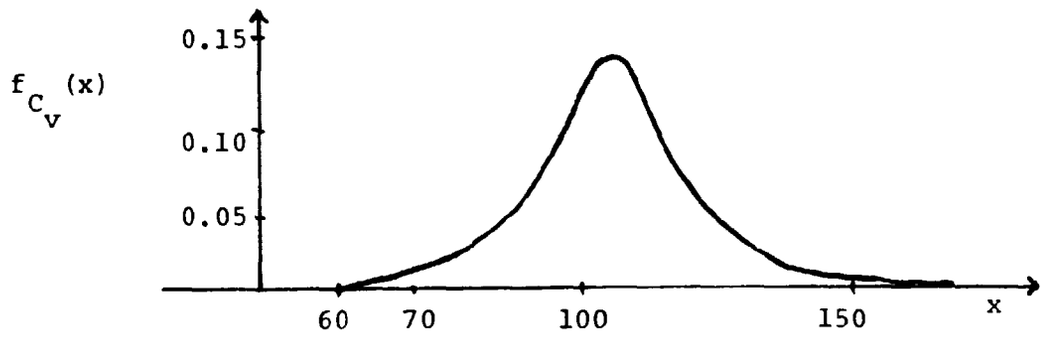
where

M is the constant 5.0

C_v is given by the Weibull distribution

$$f_{C_v}(x) = k \cdot (x-x_o)^m \cdot \exp\left(\frac{-k(x-x_o)^{m+1}}{m+1}\right)$$

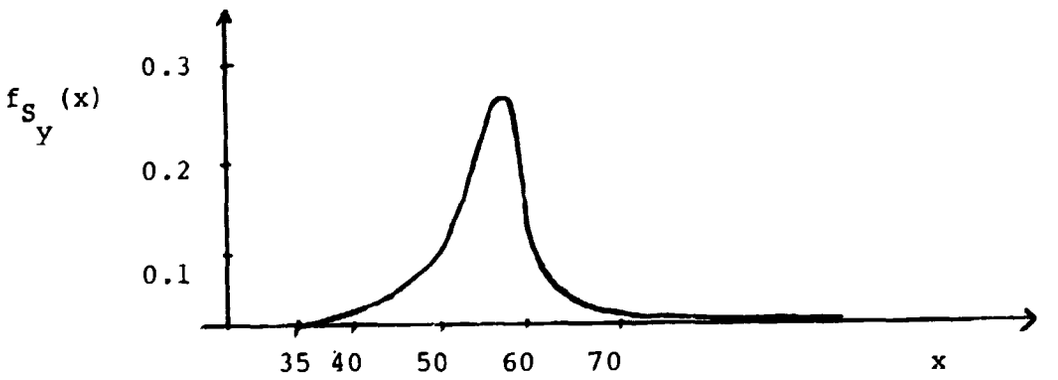
with $k = 1.57 \cdot 10^{-12}$, $m = 8.2$ and $x_o = 35.0$



S_y is given by the Weibull distribution

$$f_{S_y}(x) = k \cdot (x-x_o)^m \cdot \exp\left(\frac{-k(x-x_o)^{m+1}}{m+1}\right)$$

with $k = 2.53 \cdot 10^{-6}$, $m = 2.65$, $x_o = 60.0$



K_{IC}^2 takes the minimum value

$$5.0 \cdot 35.0 \cdot 60.0 = 10500$$

and therefore the lower integration limit is 10500.

K_I^2 takes no upper limit value, but due to the physical existence of a maximum crack depth and a maximum value of the nominal stress a maximum of 60000 is reasonable.

The probability of failure is in this case

$$\int_{10500}^{60000} (60000-x) \cdot f_{K_I^2}(x) \cdot f_{K_{IC}^2}(x)\, dx$$

This integral is evaluated by a Gauss-Legendre formula with 32 points. The evaluation of $f_{K_I^2}$ and $f_{K_{IC}^2}$ is performed by applying a 30 points Gauss-Laguerre formula.

The result is $0.2826 \cdot 10^{-6}$.

With a 8 points Gauss-Legendre formula and a 8 points Gauss-Laguerre formula for $f_{K_I^2}$ and $f_{K_{IC}^2}$ the resulting probability of failure would be

$$0.2678 \cdot 10^{-6}$$

The latter result is obtained in 0.9 seconds CPU-time, while the former requires 4.7 seconds in CPU-time.

Example 2

The data originate from the same reference as in example 1, ref. [1].

The model is similar to that in example 1

$$K_I^2 = K \cdot A \cdot S^2$$

where

K is the constant 3.8
A is given by the exponential distribution
$f_A(x) = \lambda \cdot \exp(-\lambda(x-x_o))$ with $\lambda = 2.56$ and $x_o = 0.1$
S is given by the normal distribution

$$f_S(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-x_m)^2}{2\sigma^2}\right)$$

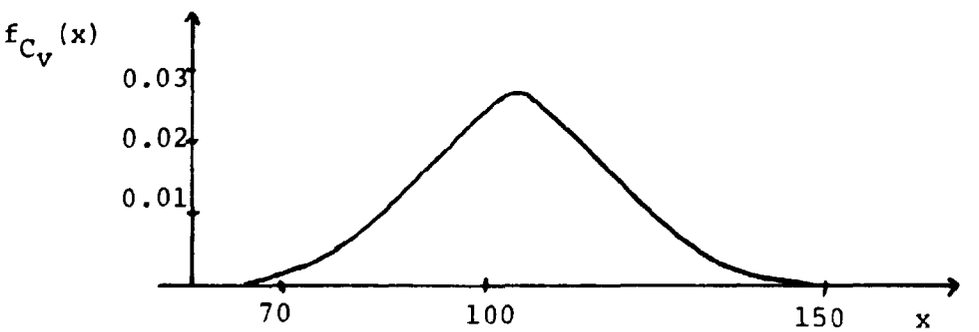
with $x_m = 26.0$ and $\sigma = 3.2$

$$K_{IC}^2 = M \cdot C_v \cdot S_y$$

where

M is the constant 5.0
C_v is given by the normal distribution

$$f_{C_v}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-x_m)^2}{2\sigma^2}\right) \text{ with } x_m = 105.0 \text{ and } \sigma = 15.0$$



Sy is given by the Weibull distribution

$$f_{S_y}(x) = k(x-x_o)^m \exp\left(\frac{-k(x-x_o)^{m+1}}{m+1}\right)$$

with $k = 2.53 \cdot 10^{-6}$, $m = 2.65$, $x_o = 60.0$

Reasonable limits in this case are 10500 and 60000. Therefore, the probability of failure is:

$$\int_{10500}^{60000} (60000-x) \cdot f_{K_I^2}(x) \cdot f_{K_{IC}^2}(x)\, dx$$

This integral is evaluated by a Gauss-Legendre formula with 8 points and $f_{K_I^2}$ and $f_{K_{IC}^2}$ are evaluated by a Gauss-Laguerre formula with 8 points. The resulting probability of failure is $0.6360 \cdot 10^{-6}$ using 1.0 second of CPU-time.

3.6.4. Experience and Application

Applications of the developed computer program show that a result is obtained within a very short time for the type of problem dealt with within probabilistic fracture mechanics. It is possible to specify the number of integration points needed to achieve a reasonable accuracy. As it was described in section 3.6.1. it is normally not needed to evaluate the integral with a high accuracy due to the nature of the problem.

In general relevant data are lacking which means that the uncertainty inherent in the problem is large. This means, that in stead of having a very accurate result of one calculation, it is likely that several evaluations are needed with slightly different data in a detailed sensitivity study. In this type of application a fairly simple and fairly accurate evaluation is important. It has been seen that the NUMPEP code can fulfill these requirements.

In order to check the validity of the result the program can be applied by dividing the region into subregions with subsequent calculation of the integral in each subregion succeeded by a summation. If this refined calculation satisfies the requirement to the accuracy, the execution of the program is stopped, otherwise the procedure can be continued by further subdivision.

3.6.5. Error Analysis

The integral

$$I_n = \int\cdots\int w(x_1,\cdots,x_n) f(x_1,\cdots,x_n) dx_1\cdots x_n$$

can be evaluated using a numerical integration formula

$$S_N = \sum_{i=1}^{N} A_i f(v_{i,1},\cdots, v_{i,n})$$

The error $E[f]$ is given by

$$I_n = S_N + E[f].$$

Estimation of the error is in general a very difficult task. In [8] a thorough discussion is presented, giving estimates proposed by Sard and Barnhill, respectively.

In the first case the theory is based on a generalization of the Peano error estimates for one variable. The estimates are given for functions f(x,y) which have a certain type of Taylor series expansion.

In the second case the theory is based on a generalization of the estimates of P.J. Davis for one variable. The estimates are given for functions of two complex variables, which are analytic for all points in a certain region containing R_2.

SARD-ESTIMATES

The derivatives of a function $f(x,y)$ over $[a,b]x[c,d]$ are defined as

$$f^{(i,j)}(x,y) = \frac{\partial^{i+j}}{\partial x^i \partial y^j} f(x,y) \qquad i \geq o,\ j \geq o$$

If the derivatives are defined and continuous, then $f(x,y)$ has a Taylor's expansion.

$B_{p,q}$ is the set of all functions with the following properties:

1. $f^{(p,q)}(x,y)$ is Riemann integrable on R_2
2. $f^{(m-j,j)}(x,c)$, $j<q$ is Riemann integrable on $a \leq x \leq b$
3. $f^{(i,m-i)}(a,y)$, $i<p$ is Riemann integrable on $c \leq y \leq d$
4. $f^{(i,j)}(a,c)$, $i+j<m$ exists
5. Taylor expansion holds for $f(x,y)$ for all (x,y) in R_2.

Kernel theorem, [8], states that if an integration formula is exact, when $f(x,y)$ is a polynomial of degree $<m$ in x and y, then there exist functions

$K_{p,q}(x,y)$ $\qquad$ (x,y) in R_2

$K_{m-j,j}(x)$ $\qquad$ x in $[a,b]$, $j<q$

$K_{i,m-i}(y)$ $\qquad$ y in $[c,d]$, $i<p$

so that

$$E[f] = \sum_{j<q} \int_a^b K_{m-j,j}(u) f^{(m-j,j)}(u,c)\, du$$

$$+ \sum_{i<p} \int_c^d K_{i,m-i}(v)\ f^{(i,m-i)}(a,v)\, dv$$

$$+ \int_a^b \int_c^d K_{p,q}(u,v) f^{(p,q)}(u,v)\, dudv$$

whenever $f(x,y)$ belongs to $B_{p,q}$.

Furthermore,

$$K_{p,q}(u,v) = E\left[\frac{(x-u)^{p-1}(y-v)^{q-1}}{(p-1)!\quad(q-1)!}\right]$$

$$K_{m-j,j}(u) = E\left[\frac{(x-u)^{m-j-1}(y-c)^{j}}{(m-j-1)!\qquad j!}\right]$$

$$K_{i,m-i}(v) = E\left[\frac{(x-a)^{i}(y-v)^{m-i-1}}{i!\quad(m-i-1)!}\right]$$

which are the errors using the formula on the expressions in the brackets.

Several estimates are given of which the most common is:

$$|E[f]| \leq \sum_{j<q} e_{m-j,j:x}\, M_{m-j,j:x} + \sum_{i<p} e_{i,m-i:y}\, M_{i,m-i:y} + e_{p,q}\, M_{p,q}$$

where $e_{m-j,j:x} = \int_a^b |K_{m-j,j}(u)|\, du$

$$e_{i,m-i:y} = \int_c^d |K_{i,m-i}(v)|\, dv$$

$$e_{p,q} = \int_a^b \int_c^d |K_{p,q}(u,v)|\, dudv$$

$$M_{m-j,j:x} = \sup_{u\in[a,b]} |f^{(m-j,j)}(u,c_o)|$$

$$M_{i,m-i:y} = \sup_{v\in[c,d]} |f^{(i,m-i)}(a_o,v)|$$

$$M_{p,q} = \sup_{R_2} |f^{(p,q)}(u,v)|$$

The first three e-values are only dependent on the formula, while the three M-values depend on the integrand. The Kernel functions K defined above are unfortunately not simple functions, and this leads to calculational problems. Some examples of Kernel functions from [8] are shown on the next page.

Example 1

$f(x,y) = \sqrt{(3+x+y)}$ is an element of $B_{2,4}$

The estimate for $E[f]$ is

$$|E[f]| \leq \sum_{j<q} e_{m-j,j:x}\, M_{m-j,j:x} + \sum_{i<p} e_{i,m-i:y}\, M_{i,m-i:y} + e_{p,q}\, M_{p,q}$$

Here $p = 2,\ q = 4,\ p+q = 6 = m \qquad (a_o, c_o) = (0,0)$
$(a,b) = (c,d) = (-1,1)$

$$|E[f]| \leq e_{6,0:x}\, M_{6,0:x} + e_{5,1:x}\, M_{5,1:x} + e_{4,2:x}\, M_{4,2:x} + e_{3,3:x}\, M_{3,3:x} + e_{0,6:y}\, M_{0,6:y} + e_{1,5:y}\, M_{1,5:y} + e_{2,4}\, M_{2,4}$$

$$f^{(1,0)}(x,y) = 1/2 \cdot (3+x+y)^{-1/2}$$

$$f^{(2,0)}(x,y) = -1/4 \cdot (3+x+y)^{-3/2}$$

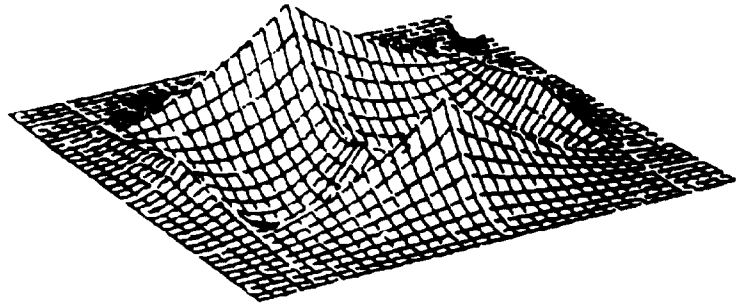
$$f^{(3,0)}(x,y) = 3/8 \cdot (3+x+y)^{-5/2}$$

$$f^{(4,0)}(x,y) = -15/16 \cdot (3+x+y)^{-7/2}$$

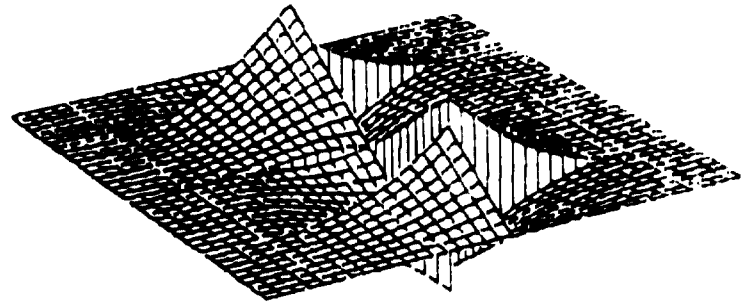
$$f^{(5,0)}(x,y) = 105/32 \cdot (3+x+y)^{-9/2}$$

$$f^{(6,0)}(x,y) = -945/64 \cdot (3+x+y)^{-11/2}$$

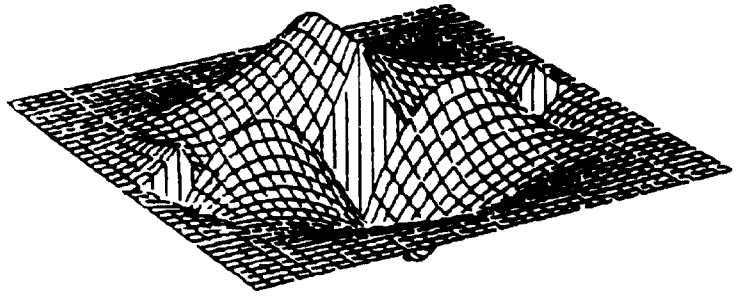
$$M_{6,0:x} = \sup_{u \text{ in } [-1,1]} \left| -945/64\ (3+u)^{-11/2} \right|$$



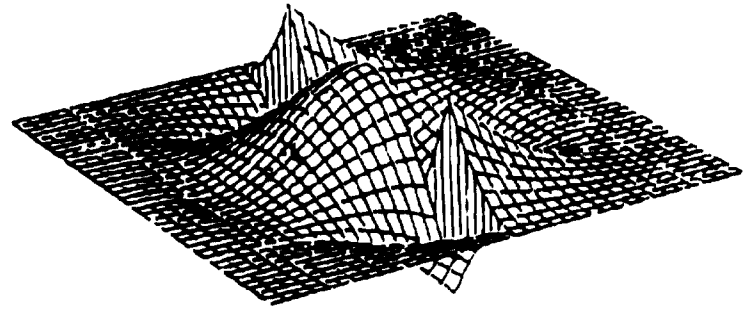
$(p, q) = (2, 2)$ $K_{max} = 0.178$



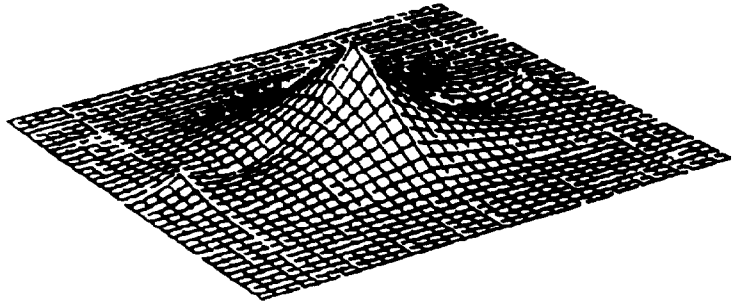
$(p, q) = (3, 1)$ $K_{max} = 0.576$



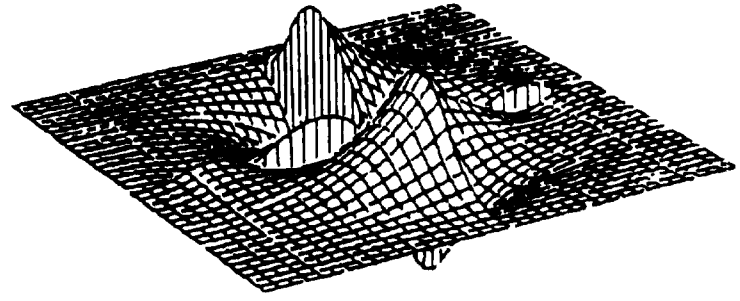
$(p, q) = (2, 3)$ $K_{max} = 0.144$



$(p, q) = (3, 2)$ $K_{max} = 0.176$



$(p, q) = (2, 4)$ $K_{max} = 0.337$



$(p, q) = (3, 3)$ $K_{max} = 0.121$

maximum is reached for u = -1

$$M_{6,0:x=} \mid -945/64(2)^{-11/2} \mid = 0.326$$

$$M_{5,1:x} = M_{4,2:x} = M_{3,3:x} = M_{6,0:x}$$

$$M_{0,6:y} = M_{1,5:y} = M_{6,0:x}$$

$$M_{2,4} = \sup_{(u,v)\text{in } [-1,1] \times [-1,1]} |f^{(2,4)}(u,v)|$$

$$= \sup_{(u,v) \text{ in } [-1,1] \times [-1,1]} |-945/64(3+u+v)^{-11/2}|$$

maximum is reached for u = v = -1

$$M_{2,4} = |-945/64(1)^{-11/2}| = 14.8$$

Using the seven-point fifth-degree formula C_2:5-1 [8] the e-values are tabulated:

$$e_{6,0:x} = 0.000126$$

$$e_{4,2:x} = 0.000302$$

$$e_{0,6:y} = 0.0000377$$

$$e_{2,4} = 0.00246$$

$$e_{5,1:x} = e_{3,3:x} = e_{1,5:y} = 0$$

$$\text{Thus } |E[f]| \leq (0.000126 \cdot 0.326) + (0.000302 \cdot 0.326) + (0.0000377 \cdot 0.326) + (0.00246 \cdot 14.8)$$

$$= 0.0356$$

To see how close the estimate of $|E[f]|$ is to the actual error a couple of examples are shown below.

Example 2

Let $f(x,y) = e^{x+y}$

Integrate f(x,y) over $[-1,1]$ x $[-1,1]$ using two different formulas C_2:5-1 and C_2:5-2 from [8].

Formula	Integral	Estimated Error	Actual Error
C_2:5-1	5.521576985	$0.1948 \cdot 10^{-1}$	$0.281439583 \cdot 10^{-2}$
C_2:5-2	5.52275698	$0.2765 \cdot 10^{-1}$	$0.211568298 \cdot 10^{-2}$
Exact	5.524091381		

The ratio between the estimated error and the actual error is

C_2:5-1 : 6.9
C_2:5-2 : 13.1

which shows that the error is overestimated using SARC-estimates.

Example 3

Let $f(x,y) = x^4y^2$

Integrate f(x,y) over $[-1,1]$ x $[-1,1]$ using the two different formulas C_2:5-1 and C_2:5-2 from [8].

Formula	Integral	Estimated Error	Actual Error
C_2:5-1	0.266666667	0.0145	$0.363797881 \cdot 10^{-11}$
C_2:5-2	0.207407407	0.0593	$0.592592593 \cdot 10^{-1}$
Exact	0.266666667		

The ratio between the estimated error and the actual error is

C_2:5-1 : $4.0 \cdot 10^9$
C_2:5-2 : 1.0

This shows that using C_2:5-1 the estimated error is useless. Furthermore, the estimates provided by SARD cannot in general be improved, since the use of the formula C_2:5-2 shows that the error limits cannot be sharpened, since a function and a formula exist where the error estimate is equal to the actual error.

BARNHILL-ESTIMATES

Barnhill error estimates are given in [8] for functions of two complex variables, which are analytic for all points in a certain region containing R_2.

$$|E[f]| \leq ||E|| \cdot ||f||$$

where

$||E||$ depends on the formula

$||f||$ depends on the integrand

$$||f||^2 = \iiint_{E\rho xE\rho} |f(z,w)|^2 \, dxdudydv$$

where $E\rho xE\rho$ are ellipses

$$||f||^2 \leq \max |f(z,w)| \cdot \pi ab$$

$$||E||^2 = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} |E[w_r^* w_s^*]|^2$$

where w_r^* and w_s^* are Chebyshev polynomials. This means, that $||E||^2$ expresses the error using the formula on the specified Chebyshev polynomials.

Example 1

Let $f(x,y) = e^{x+y}$
Integrate $f(x,y)$ over $[-1,1] \times [-1,1]$ using C_2:5-1 from [8] (see also example 2, SARD).

In [8] values for $||E||$ and $||f||$ are given.

The error estimate is 0.3417, which can be compared with the similar SARD-estimate.

Formula C_2:5-1 $\quad f(x,y) = e^{x+y}$

Errors/Ratio	Sard	Barnhill	Actual
C_2:5-1	0.01948	0.3417	0.002814396
Ratio between errors	6.9	121.4	

The results of this example are typical for Barnhill-estimates. They are very complex to calculate and they overestimate the error dramatically.

It can be concluded that the error estimates in general are useless for practical purposes. The reasons are that

1. the Kernel functions are very difficult to treat, even in very simple cases

2. in most cases the estimated errors are much greater than the actual error.

In stead it is proposed to use recalculation of the integral using a larger number of integration points and compare the results. If the difference between subsequent calculations is acceptable then the computation is stopped, otherwise it is continued with an increased number of points.

3.6.6. Conclusions and Comparisons

The NUMPEP code has been thoroughly tested and the results of the calculations from the examples given in section 3.6.3. have been compared to the results obtained by some of the codes described in section 3.4.

Example 1

Using the NUMPEP code the probability of failure of the structure is $0.2826 \cdot 10^{-6}$. Product formulas have been used based upon a Gauss-Legendre and a Gauss-Laguerre formula. The number of points in each formula was $32 \cdot 30 = 960$, and the CPU-time was 4.7 seconds. A similar calculation gave the probability of failure $0.2678 \cdot 10^{-6}$, using $8 \cdot 8 = 64$ points with a CPU-time of 1.1 seconds.

The NUMPEP code is compared to the PEP code [2] and the ANPEP/V2 code [4] , which are described in section 3.4. The results are given in the table below.

Code	Probability of failure	CPU-time (seconds)
PEP	$0.2867 \cdot 10^{-6}$	300.0
ANPEP/V2	$0.2843 \cdot 10^{-6}$	12.5
NUMPEP	$0.2826 \cdot 10^{-6}$	4.7

As it is seen there is an agreement between the results of all three codes taking into account the requirements to the accuracy. Furthermore, it is obvious that NUMPEP has obtained the result using a shorter CPU-time than the other codes.

Example 2

Using the NUMPEP code the probability of failure of the structure is $0.6360 \cdot 10^{-6}$. Product formulas have been used based upon a Gauss-Legendre and a Gauss-Laguerre formula. The number of points in each formula was $8 \cdot 8 = 64$, while the CPU-time was

1.0 second. The NUMPEP code is compared to the PEP code [2] and the ANPEP/V2 code [4], which are described in section 3.4. The results are given in the table below.

Code	Probability of failure	CPU-time (seconds)
PEP	$0.6257 \cdot 10^{-6}$	300.0
ANPEP/V2	$0.6392 \cdot 10^{-6}$	14.3
NUMPEP	$0.6360 \cdot 10^{-6}$	1.0

Again, the results are in good agreement and the reduction in computing time using the NUMPEP code is significant.

Number of points

As described in section 3.5.3. Monte Carlo methods are time consuming, since the error in the formula is given by $c/\sqrt{N}$, where N is the number of points. This means, that an increase in the number of points by a factor 100 will only reduce the error by a factor 10. It is shown in [3] that the ANPEP code is much more efficient than Monte Carlo methods.

In ANPEP and ANPEP/V2 each distribution is decretized and combined subsequently. In general 100 points are recommended for each decretization. In the examples discussed above this leads to $2 \cdot (100 \cdot 100) = 20000$ points.

The product formula in NUMPEP requires in total $N = N_1 \cdot 2N_2 = 2N_1N_2$ points, where N_1 is the number of points in the Gauss-Legendre formula and N_2 is the number of points in the Gauss-Laguerre formula.

In the examples above the following number of points were used:

example 1: $N_1 = 32$, $N_2 = 30$
$N = 1920$

example 2: $N_1 = 8$, $N_2 = 8$
$N = 128$

These differences between the number of points in the methods explain the differences in the observed CPU-times.

On the contrary if the ANPEP/V2 code is used with 8 points in each discretization, the same number of integration points are needed as for the NUMPEP code in example 2. The result of the calculation using ANPEP/V2 code is $0.4931 \cdot 10^{-6}$, which shows, not surprisingly, that the choice of integration points and weights in the Gauss formulas are optimized compared to the strategy used in the ANPEP/V2 code.

4. SYSTEMS RELIABILITY

In order to optimize the design of industrial systems, analysis of their performance with respect to safety, reliability and availability is of increasing importance. Methods and tools for such analyses have been developed during the past 20 years using the power of computers to solve complex questions within a reasonable time.

4.1. Background

A variety of computerized methods are available for reliability and safety analysis of complex systems. A comprehensive presentation of available methods and tools is given in [11].

Within the Nordic countries a couple of research projects have been carried out with the aim to describe the state of the art of the techniques in industrial applications. The SCRATCH project was finished in 1982 issuing a summary report [12]. It contains a description of the usefulness of risk and reliability analysis techniques in any industrial application. In the period 1981-1985 a new Nordic project was carried out with the aim to study probabilistic risk assessment (PRA) and licensing in nuclear applications. The main results are summarized in [13]. These results have subsequently been reviewed with respect to non-nuclear applications [14].

One of the common results of the above research projects is the necessity of methods for detailed description of complex systems with subsequent reliability or availability calculation. Fault trees or reliability block diagrams are the most commonly used methods for representation of specification of the failure conditions of complex systems. They both allow a quantitative assessment of system reliability and availability. Analytical as well as Monte Carlo simulation programs are available for this part.

One of the main drawbacks of Monte Carlo simulation programs is the amount of CPU-time needed to perform a calculation with reasonable accuracy on the results. This problem can in some cases be overcome by applying a variance reduction technique. In this study variance reduction techniques are analyzed with respect to implementation in an existing Monte Carlo simulation program.

Section 4.2.-4.4. describes the definitions and ideas behind system representation methods and quantification methods. In 4.5. the existing Monte Carlo simulation program MOCARE is briefly presented. Section 4.6. gives an outline of variance reduction techniques. The most commonly used and flexible method, the importance sampling method, is analyzed in section 4.7. Finally, section 4.8. and 4.9. present the proposal for the implementation of importance sampling technique in the MOCARE program.

4.2. Fault Trees and Block Diagrams

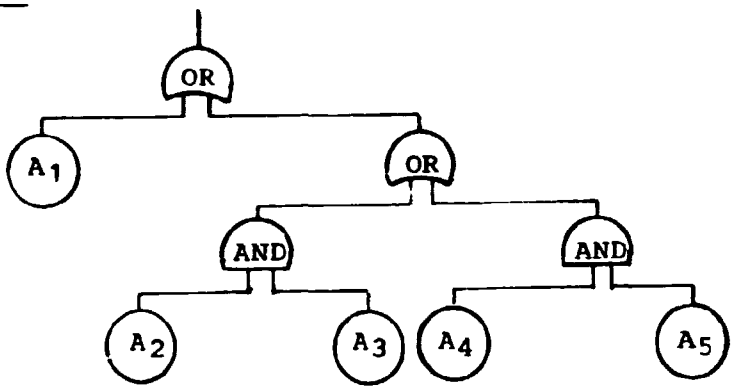
Fault trees and reliability block diagrams can be regarded as dual methods. A fault tree represents all combinations of basic events which lead to failure of the system. Basic events can be either component failures, human errors or failure conditions. A reliability block diagram on the contrary represents all combinations of components which have to be functioning to assure system functioning. The fault tree representation is most widely used.

Fault Trees

The basic concepts and the techniques of fault tree analysis are given in [15]. Furthermore, its advantages and limitations are discussed in [11], [12] and [13].

A fault tree consists of basic events linked into a description of the system using logical OR-gates and AND-gates.

Example



A_1-A_5 are basic events, such as pump fails to start on demand.

The system will fail, if
- A_1 occurs
- A_2 and A_3 occur simultaneously
- A_4 and A_5 occur simultaneously.

Fault trees are useful diagrams well-suited for presentation of system failure conditions to be used in communication between the designers and the operating staff and between the designers and the authorities. Further, it is prepared for a subsequent quantification.

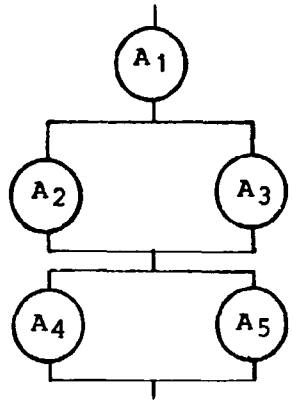
Block Diagrams

Reliability block diagrams are discussed in [13] and a computer program RELVEC applying this methodology is described in [16].

The interpretation of a reliability block diagram, which focuses on success in stead of failure of the components, is: a system is considered functioning if a path through the blocks exists, that are functioning. If no such paths exists, the system is failed.

Example

The reliability block diagram which represents the system given by the fault tree above is:



A_1-A_5 are basic events, such as pump is running.

The system is functioning, if
- A_1, A_2 and A_4 are functioning
- A_1, A_2 and A_5 are functioning
- A_1, A_3 and A_4 are functioning
- A_1, A_3 and A_5 are functioning.

Remarks

The two approaches will give identical system failure conditions, but in practical use as an aid in the identification process, some differences are found.

It has been argued, that reliability block diagrams are easier to use than fault trees due to their close relationship to the flow schemes of the systems. Furthermore, it is more natural to plant personnel to think in terms of success than in terms of failure, since they are responsible for a continuous operation of the plant - also in cases of disturbances.

On the other hand, it has been found that the searching for system failure conditions in practice assure a higher degree of completeness than searching for system function conditions.

4.3. Cut Sets and Path Sets

Given a system of components, a cut set is a set of components, which imply a system failure if all components in the set are failed simultaneously. A minimal cut set is a cut set, where any change of a component state from failed to nonfailed will imply that the system is functioning. This means that all components need to be failed simultaneously before a system failure occurs. In the example in section 4.2. the minimal cut sets are: A_1, (A_2, A_3), (A_4, A_5).

Similarly, given a system of components, a path set is a set of components which assure the functioning of a system if all components in the set are functioning. A minimal path set is a path set, where any change of a component state from functioning to failed will imply that the system fails.

4.4. Reliability and Availability Calculations

The definitions of the reliability and the availability of a system are given in chapter 2. In general each component in a system has a specific time-to-failure distribution, of which the exponential distribution plays an important role.

$$f(x) = \lambda \exp(-\lambda(x-x_o)) \quad \text{for } x \geq x_o$$

$$F(x) = \int_{x_o}^{x} \lambda \exp(-\lambda(t-x_o))\, dt$$

$$= 1 - \exp(-\lambda(x-x_o))$$

The hazard rate is defined as

$$h(x) = f(x)/(1-F(x))$$

where 1-F(x) is called the reliability at time x. In case of an exponential distribution the hazard rate

$$h(x) = f(x)/(1-F(x))$$
$$= \lambda\exp(-\lambda(x-x_o))/(1-(1-\exp(-\lambda(x-x_o))))$$
$$= \lambda$$

i.e. a constant hazard rate, independent of the time.

The reliability over time t given an exponential distribution is

$$R(t) = 1-F(t) = \exp(-\lambda(t-x_o))$$

which is the probability that the component is functioning at time t given it was functioning at time x_o.

Parallel system

Given a parallel system, S, consisting of two components, C_1 and C_2, (which corresponds to an OR gate)

$$R_s(t) = R_1(t) + R_2(t) - R_1(t)R_2(t)$$

if C_1 and C_2 are independent.

Series system

Given a series system, S, consisting of two components, C_1 and C_2, (which corresponds to an AND gate)

$$R_s(t) = R_1(t) \cdot R_2(t)$$

if C_1 and C_2 are independent.

Similar expressions can be derived taking into account maintained components, where a repair time is specified. Such expressions are given in [17].

Availability

From [17] the availability of a repairable component at time t is

$$A(t) = \mu/(\lambda+\mu) + (\lambda/(\lambda+\mu))\exp(-(\lambda+\mu)t)$$

given exponential failure distribution with failure rate λ and exponential repair distribution with repair rate μ.

The average availability for a period of time [o,t] is the avarage uptime

$$\overline{A}(t) = (1/t) \int_o^t A(s)ds$$

$$= \mu/(\lambda+\mu) + \lambda/(\lambda+\mu)^2 t - (\lambda/(\lambda+\mu)^2 t)\exp(-(\lambda+\mu)t)$$

If $t \to \infty$ the availability becomes $\mu/(\lambda+\mu)$, the steady-state availability.

Expressions can be derived for series and parallel systems, but the expressions are rather complex.

In order to achieve a result within a reasonable time, two different approaches have been applied

- approximate, analytical methods
- simulation methods.

4.4.1. Analytical Programs

Computer programs have been developed to calculate the system reliability or system availability given component data and a representation of the system failure conditions.

One such program has been developed at Risø, FAUNET, [18], using the fault tree representation. The program searches for minimal sets (cut sets or path sets) utilizing advanced techniques. Then the reliability or the availability is calculated based on the minimal sets and component data. The program accepts components with either constant failure probabilities or exponentially distributed failure times. Repair times are either constant or exponentially distributed. Furthermore, constant test intervals can be specified. The basic assumption is that the components are statistically independent.

The well known expressions for the unavailability as a function of time t are used (see [18]).

The program performs a calculation of the system unavailability from the minimal cut sets, using the rules:

OR-gate: $P = P_1+P_2-P_1P_2$
AND-gate: $P = P_1 \cdot P_2$

Further, to calculate the probability of the union of the minimal cut sets, only the first terms in the expansion is included. It is possible to write the expression explicitly, but due to calculational difficulties only the first terms are included.

4.4.2. Monte Carlo Simulation

Several methods based on Monte Carlo simulation have been developed. One such program is the MOCARE program [19] described in section 4.5.

In this type of program each component is assigned a time where it is failed based upon the input describing the p.d.f. for the time-to-failure and repair. The assignment is performed using a random number generator, which is available on any computer.

In many cases, the random number generator is using the linear congruential method:

$$x_{n+1} = (x_n \cdot a+b) \text{MOD } N, \; a, \; b \text{ fixed}$$
$$R = x_{n+1}/N$$

where N is the largest integer value.

The user specifies a start number x_o and new numbers are generated successively using the formula. Using this method numbers which are uniformly distributed are generated on $[0, 1[$.

Assume a component with specified time-to-failure distribution F(x), with density function f(x). If R is a random number uniformly distributed over 0 to 1, the following equation gives a value x, following the specified time-to-failure distribution:

$$R = F(x)$$

or

$$x = F^{-1}(R)$$

Example

Assume an exponential density function

$$f(x) = \lambda \exp(-\lambda(x-x_o)) \text{ and } F(x) = 1-\exp(-\lambda(x-x_o))$$

$$R = 1-\exp(-\lambda(x-x_o))$$
$$1-R = \exp(-\lambda(x-x_o))$$
$$-\lambda(x-x_o) = \ln(1-R)$$
$$x = x_o-(1/\lambda)\cdot\ln(1-R)$$

If R is uniformly distributed over 0 to 1, also 1-R will be uniformly distributed over 0 to 1.

To save one algebraic operation x can be found from

$$x = x_o-(1/\lambda)\cdot\ln(R)$$

The main advantages of Monte Carlo simulation programs are their flexibility to model complex systems and the userfriendliness. The main drawback is the computing time required to obtain a given accuracy of the result.

4.5. The MOCARE Code

The MOCARE program which is based on direct Monte Carlo simulation, is developed at Risø [19].

4.5.1. Principles

The program accepts either a fault tree or a reliability block diagram as the part of the input which contains the specification of the system failure condistions. Furthermore, minimal cut sets from a previous calculation by the FAUNET program are also accepted as input. For each component five different p.d.f. types for time-to-failure are available:

- exponential distribution
- Weibull distribution
- normal distribution
- log-normal distribution
- constant time

and similarly for the repair times.

The program carries out a number N of trials over a specified period of time T. The process is illustrated in the example below. The history of each fault is shown over the period [o,T]. The presence of a fault is symbolized by a hatched field on the time axis. The status of the system is evaluated every time a change in a component state occurs, i.e. if a failure occurs or if a fault has been repaired. If the system has failed, it is registered together with the duration of the system failure for further treatment.

Finally, the mean reliability over the period [0,T] is calculated using:

$$R = NS/N$$

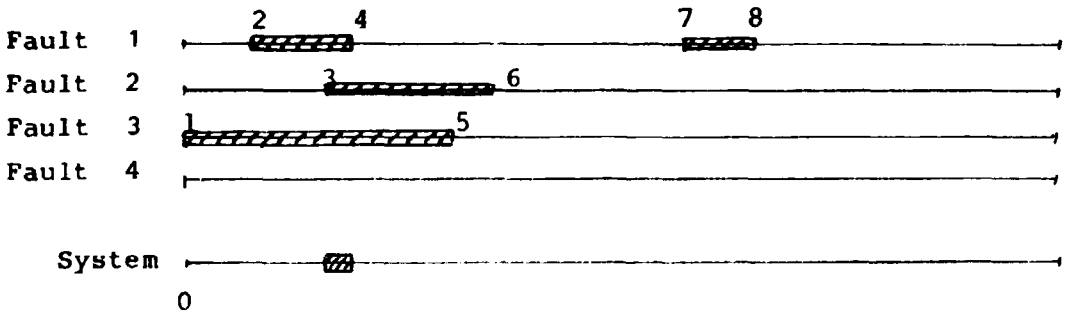
where NS is the number of trials with no system failures
N is the number of trials

The mean unavailability is calculated using:

$$UA = \sum_{i=1}^{NF} D_i/(N \cdot T)$$

where NF is the number of trials with system failure (NF=N-NS)
N is the number of trials
T is the observation period
D_i is the system downtime for system failure no.i.

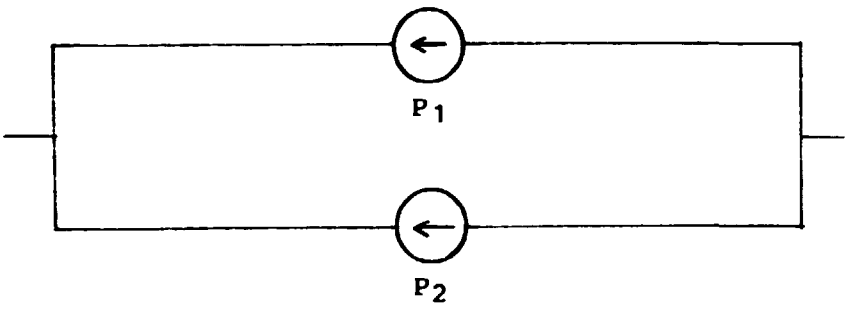
Example



The status of the system is evaluated 8 times, and only in one case, at time point 3, the system has failed. The downtime of the system is from time point 3 to time point 4, where fault no. 1 is repaired. Furthermore, it is seen that fault no. 4 does not occur in this trial. The histories can be presented on graphical displays which has been found very valuable both as a check of the correctness of the model and as an illustration of how the system performs.

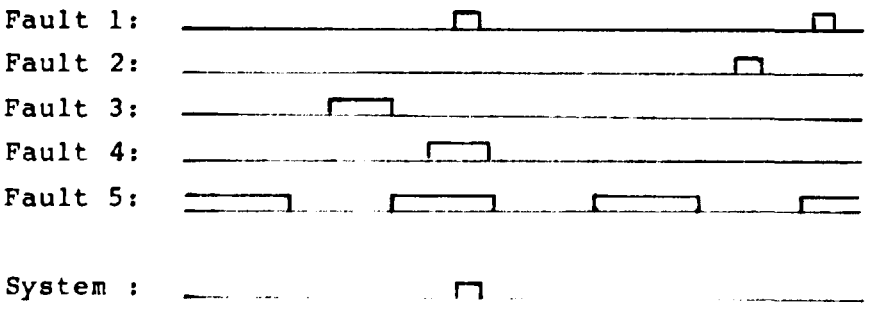
4.5.2. Example

Assume a simple system consisting of two identical pumps, P_1 and P_2, in parallel each with 100% capacity.



One pump is running at a time with the other pump in stand-by mode. The pumps are switched every one week. In case of a failure of a running pump an attempt to start the other pump is initiated. The pumps are assumed to fail either during operation or during stand-by following certain distributions. In the latter case it is assumed that the failure occurred during stand-by is detected only when a start is attempted.

A history of the faults might look like this:



where

Fault 1: failure of P_1 during operation
Fault 2: failure of P_2 during operation
Fault 3: failure of P_1 during stand-by
Fault 4: failure of P_2 during stand-by
Fault 5: dummy fault specifying the scheduled operation times for P_1 and P_2 - if fault no. 5 is in a failed state P_1 is in operation, otherwise P_2 is in operation. The pump which is not in operation, is in stand-by.

A failure of the system occurs in the situation where P_1 is running and P_2 is in stand-by. A failure of P_2 occurs which has no immediate effect, since P_1 is running. Some time later a failure of P_1 occurs. An attempt to start P_2 is initiated, but without success due to the failure.

4.5.3. Experience

The MOCARE program has been applied in several analyses of complex systems and has proved to be a very flexible tool. This is accomplished by using a flexible input and by using subsystems for the specification of conditions for the occurrence of basic faults and system failures. Secondly, the large number of special facilities that are available, makes a detailed analysis of the system possible. The program can handle a series of different types of fault:

- faults with various probability density functions for the time to failure and the repair
- faults having a constant probability of failure per period of observation
- consequential faults, occurring with a specified probability of failure per event, that can be defined as the failure of a specified subsystem
- faults which can only occur under certain circumstances, that can be specified by means of subsystems.

The application of subsystems is very useful in specification of different modes and in specification of the transition rules between the states.

In analysis of very large and complex systems MOCARE has been used and reasonable results have been obtained. One such analysis is performed within the Nordic research project on probabilistic methods in nuclear applications [20]. A reliability analysis of the feedwater systems ability to assure feedwater to the reactor core in case of a loss of off-site power was performed. In the analysis the normal feedwater system, the auxiliary feedwater system and the electrical power supply systems were taken into account. The analyzed system was fairly large and very complex. The total number of cut sets was approximately 410 mill. Several methods were applied in the calculation of the unavailability of the feedwater system. MOCARE was able to get a reasonable result using about 1/2 hour CPU-time on a Burroughs B7800 computer. It was not possible to get a result using the FAUNET code due to combinatorial problems. Other methods applying cutting of the fault tree, if the probabilities were below some specified value were also able to give reasonable results. In this case one is never sure that the cutting is allowable and that the results are correct.

The main drawbacks of the MOCARE program are the considerable amount of computing time required to obtain reasonable results and the statistical nature of the program which will give answers with a related statistical uncertainty. This means that there is a certain probability that you have obtained an incorrect result.

While the latter problem is insolubly due to the nature of the program, some proposals are given in order to overcome the former drawback using variance reduction techniques.

4.6. Variance Reduction Techniques

A series of variance reduction techniques have been developed mainly within queueing theory and radiation transport calculations. But some of them are also useful methods within reliability calculations. A very comprehensive presentation of the concepts of variance reduction and descriptions of the various methods is given in [21]. Below is shortly summarized some known results of variance reduction used in reliability calculations and an outline of the methods available.

4.6.1. Experiences and Knowledge

Variance reduction is concerned with increasing the accuracy of Monte Carlo estimates of parameters. In direct Monte Carlo simulation you try to describe the performance of the system as close to reality as possible. In case of simulation of rare events as is the general case in reliability analysis, this lead to long computing times. Variance reduction techniques attempt to increase the effectiveness of the Monte Carlo method by:

- modifying the simulation process
- utilization of approximate information
- studying the system within a different context.

In reliability analysis it is the experience that the importance sampling method in general is the most favorable method. Some other methods have been applied, but with less effective results with respect to variance reduction. Importance sampling is studied in more detail below.

Furthermore, it has been shown that unfortunately no single method exists, that in general will lead to an optimal variance reduction in an analysis of a system.

Finally, it has been emphasized [21], that any kind of prior knowledge of the problem should be utilized in selection of methods and parameters.

4.6.2. Outline of Methods

A comprehensive outline of methods for variance reduction is presented in [21]. The following classification of available methods is proposed:

- Modification of the simulation process
  - Importance sampling
  - Russian Roulette and splitting
  - Systematic sampling
  - Stratified sampling

- Use of approximate information
  - Expected values
  - Statistical estimation
  - Correlated sampling
  - History reanalysis
  - Control variates
  - Antithetic variates
  - Regression

- Study of the system within a different context
  - Sequential sampling
  - Adjoint formulation
  - Transformations
  - Orthonormal functions
  - Conditional Monte Carlo.

The four methods classified as methods based on modification of the simulation process are emphasizing the same strategy: to select some areas of interest of each random variable and concentrate the simulation process on these areas.

In importance sampling the simulation process is modified by a choice of a more relevant density function $f^*(x)$ in stead of $f(x)$. To compensate for the modification the result is weighted by $f(x)/f^*(x)$.

In Russian roulette a choice is made in each stage to evaluate the importance of the state. If it is of interest the number of simulations starting in this state is increased and if not the simulation process is stopped with a certain probability.

In systematic sampling the sample space is divided into subareas in the way that the generated uniformly distributed random numbers are scaled into intervals. This means, that you are sure that some of the simulations will treat the area of interest. Stratified sampling is similar to systematic sampling where the simulation is directed to regions of special interest. In this sense stratified sampling combines systematic sampling and importance sampling.

All four methods seek as the basic idea to concentrate the sampling upon regions of interest. The importance sampling method is the most flexible and therefore this method has been chosen in reliability analysis.

4.7. Importance Sampling

In this section the importance sampling technique is described and the most important results concerning its characteristics and applicability are outlined.

4.7.1. Method

The importance sampling technique can be illustrated by considering a Monte Carlo estimate of a parameter I, where

$$I = E(g(x)) = \int g(x)\, f(x)\, dx$$

where E is the expectation.

The direct Monte Carlo simulation process would be:

- select a random sample $X_1,\ldots,X_N$ from the distribution with probability density function f(x)
- estimate I as $\hat{I} = (1/N) \sum_{i=1}^{N} g(X_i)$

The sample variance for the estimate is given by

$$S^2 = 1/(N-1) \sum_{i=1}^{N} (g(X_i)-\hat{I})^2$$

$$= N/(N-1)((1/N) \sum_{i=1}^{N} g^2(X_i)-\hat{I}^2)$$

If the sampling is from another probability distribution function $f^*(x)$, then

$$I = \int (g(x)f(x)/f^*(x))\, f^*(x)\, dx$$

The modified procedure is now:

- select a random sample $X_1,\ldots,X_N$ from the distribution with probability density function $f^*(x)$
- estimate I as $\hat{I}_1 = (1/N) \sum_{i=1}^{N} g(X_i)f(X_i)/f^*(X_i)$

This modification should be compensated. The result is weighted by $f(X_i)/f^*(X_i)$ in the final calculation.

The sample variance for the new estimate is given by

$$S_1^2 = 1/(N-1) \sum_{i=1}^{N} (g(X_i)f(X_i)/f^*(X_i)-\hat{I}_1)^2$$

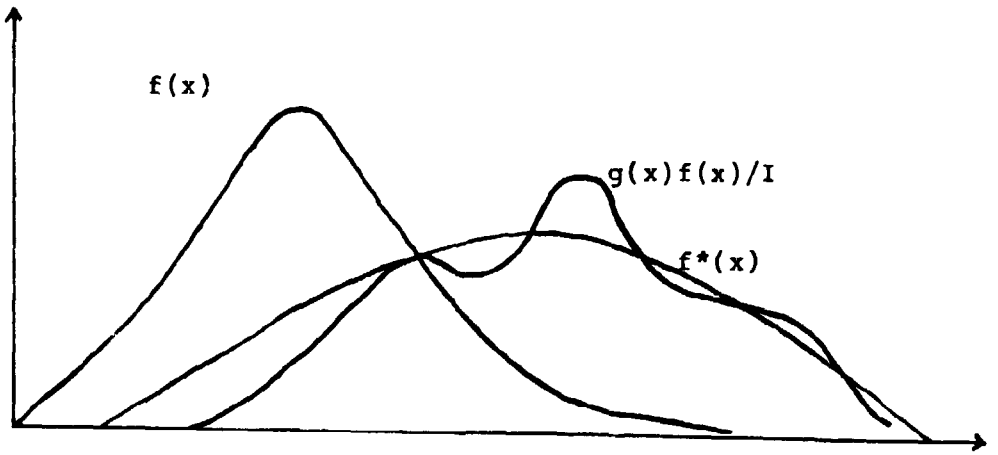
$$= N/(N-1)\ (1/N \sum_{i=1}^{N} (g(X_i)f(X_i)/f^*(X_i))^2 - \hat{I}_1^2)$$

The difference between I and $\hat{I}_1$

$$E((\hat{I}_1 - I)^2) = E((1/N \sum_{i=1}^{N} g(X_i)f(X_i)/f^*(X_i) - I)^2)$$

$$= (1/N)(\int (g(x)f(x)/f^*(x))^2 f^*(x)dx - I^2)$$

If $f^*(x) = g(x)f(x)/I$ then $E((\hat{I}_1 - I)^2) = 0$. This means, that if the answer is known, a sampling plan can be constructed with expected zero variance. This is of course of no practical interest, but it illustrates that some function of the form $g(x)f(x)/I$ will be preferable. Further, it is convenient that $f^*(x)$ is a simple, well known function which is easy to use, which can be a conflicting requirement to having $f^*(x)$ as close to $f(x)g(x)/I$ as possible. This situation is illustrated below:



The variance reduction can be calculated by

$$E(S^2 - S_1^2)$$
$$= E(S^2) - E(S_1^2)$$
$$= \int g^2(x)\ (1-f(x)/f^*(x))f(x)\ dx$$

This shows, that a careful selection of $f^*(x)$ is necessary to avoid a worse result than by direct simulation, namely if the contents of the parenthesis is negative.

Example

Assume that we shall throw two dice and estimate the probability of getting a sum equal to 3, given

$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$

for both dice.

We know that the correct result is $2/36 = 1/18$.

This means, that using direct simulation the result of a simulation will be of no interest in 17 out of 18 cases.

We can modify the sampling distribution in this way:

$P(1) = P(2) = 1/2 \qquad P(3) = P(4) = P(5) = P(6) = 0$

Then we can estimate the probability of getting the sum equal to 3 as $2/4 = 1/2$. This must be modified in the final result by $f(x)/f^*(x)$, which is $(1/6)/(1/2) = 1/3$. Both dice are modified which lead to the following correction of the final result:

$$(1/2)\ (1/3)\ (1/3) = 1/18$$

which is the correct result.

In the latter case you will only get useless results in 2 out of 4 simulations.

Further, to optimize the sampling plan, we can propose this modification of the sampling distribution:

1. Throw the first die with density function
   $P(1) = P(2) = 1/2 \qquad P(3) = P(4) = P(5) = P(6) = 0$
2. If die number one becomes a 1, throw the second die with density function
   $P(2) = 1 \qquad P(1) = P(3) = P(4) = P(5) = P(6) = 0$

   If die number one becomes a 2, throw the second die with density function

   $P(1) = 1 \qquad P(2) = P(3) = P(4) = P(5) = P(6) = 0$

The estimate is then 1/1 = 1. The result must be modified by $f(x)/f^*(x)$, which is

$(1/6)/(1/2) = 1/3$
$(1/6)/(1/1) = 1/6$

The estimate is then 1 (1/3) (1/6) = 1/18 which again is the correct result.

This simple example also shows the importance of utilization of prior knowledge of the problem considered.

4.7.2. Examples and Applications

The method of importance sampling has been applied as a variance reduction technique in some Monte Carlo simulation programs. The first example was presented by P. Nagel, [22], where the application of importance sampling yields a reduction in computing time by a factor of 10-100 depending on the complexity of the system, the failure conditions and the time of observation.

A thorough discussion of the application of importance sampling to fault tree analysis is provided by Kamarinopoulos, [23] and [24]. The results are discussed based on the following assumptions:

- only exponential probability density functions are considered
  $f(x) = \lambda \exp(-\lambda(x-x_0))$ parameter λ
- the importance functions $f^*(x)$ are exponential probability density functions of the form
  $f^*(x) = B\lambda \exp(-B\lambda(x-x_0))$ parameter $B\lambda$
  where $B > 1$
- the components are statistically independent.

In section 4.4.2. it is shown that the fictitious lifetimes given an exponential probability density function can be calculated using a random number generator, by

$$x = x_0 - 1/(\lambda)\ln(R)$$

where R is uniformly distributed over 0 to 1.

Sampling from $f^*(x)$ the lifetimes are calculated by

$$x = x_0 - (1/(\lambda B))\ln(R)$$

which shows that the lifetimes become shorter, since $B > 1$, and the probability of system failure is artifically increased. In the calculation of the system reliability, this modification should be compensated by a weighting factor, which gives:

$$S = (1/N) \sum_{k=1}^{N} \delta_k \left(\prod_{i=k}^{M} \delta_{k,i} (f_i(z_{k,i})/f_i^*(z_{k,i})) \right)$$

where

- N is the number of simulation trials
- M is the number of components
- δ_k is an indicator with $\delta_k = 1$ if the system fails in trial no. k, and 0 otherwise
- $\delta_{k,i}$ is an indicator with $\delta_{k,i} = 1$ if component i has failed in trial k, and 0 otherwise
- $z_{k,i}$ is the lifetime calculated using the random number generator
  $z_{k,i} = x_{i,0} - (1/(\lambda_i B))\ln(R)$
  where $x_{i,0}$, λ_i and B are the parameters of the exponential distribution
- f_i is the original exponential distribution specified for component i
- f_i^* is the importance function specified for component i
- S is the system unreliability.

It has been shown, [23] and [24], that the estimation of S does not converge to the true value of S in the case of OR-gates. The contributions from unfailed components are missing. A correction shall be added:

$$\hat{C} = (1/N) \sum_{k=1}^{N} \delta_k \left(\prod_{i=1}^{M} (1-\delta_{k,i}) (f(c_k)/f_i^*(c_k))B \right)$$

where $c_k = \max \{ z_{k,i} \mid \text{component } i \text{ has failed}, i=1,\ldots,M \}$

Similarly, in the expression of the variance a correction term is needed. The variance is calculated by:

$$(1/N) \sum_{k=1}^{N} \delta_k \left(\prod_{i=1}^{M} (\delta_{k,i} (f_i(z_{k,i})/f_i^*(z_{k,i})))^2 + \right.$$

$$\left. \prod_{i=1}^{M} ((1-\delta_{k,i}) (f_i(c_k)/f_i^*(c_k))B)^2 \right) - S^2.$$

The main result is however, that no general rules about how to choose B to make the variance as small as possible exists. The first product decreases with increasing B, while the second product increases at the same time. This means, that the variance exhibits a minimum.

It is found in [23] that choices of B with $1 < B < 5$ may be reasonable, but no general rules exist, which means that B has to be found by trial and error.

If other importance functions and other probability density functions are chosen, no results or rules exist for the calculation of a proper importance sampling function. In this case the importance function has to be found by trial and error.

4.7.3. Conclusions

It has been generally agreed that importance sampling is the most suitable method to be used as variance reduction technique within reliability analysis. Furthermore, it has been verified, that in case of exponential probability density functions the choice of an importance function of the same type with parameter λB with $1 < B < 5$ is reasonable. The choice of B is to be made by trial and error.

The expressions for the variance taking into account the correction terms which compensate the modified sampling plans are given in [23]. Detailed descriptions are given regarding:

Non-Repairable systems
- one component
- parallel systems
- series systems
- general systems with active components

Repairable systems
- one component
- parallel systems
- general systems with active components.

It is recommended to use these rules in cases where the components are specified by exponential probability density functions. One of the advantages of the MOCARE code is its flexibility with respect to the choice of probability density functions. This means, that to preserve this advantage in a new program utilizing importance sampling, other and perhaps less effective implementations are needed.

It is worth mentioning, that in cases with other choices of the importance function $f^*(x)$ similar expressions of the variance as obtained in [23] are valid. Finally, the results are only valid for systems with active components. In [26] it is shown how to compensate in cases involving stand-by systems. The compensation is much more complex and time consuming.

4.8. Simulation using Importance Sampling

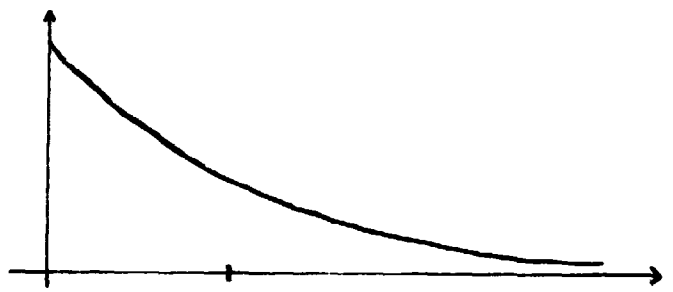
In this section the possible implementations of importance sampling using the MOCARE program is discussed.

As described in section 4.7. in case of an exponential probability density function the results obtained by Kamarinopoulos are the best guidelines in constructing a modified density function and the corresponding correction terms are given to compensate the modification. The main drawback is that only exponential density functions are allowed. This contradicts one of the advantages of the MOCARE program, namely the flexibility in the choice of probability density functions.

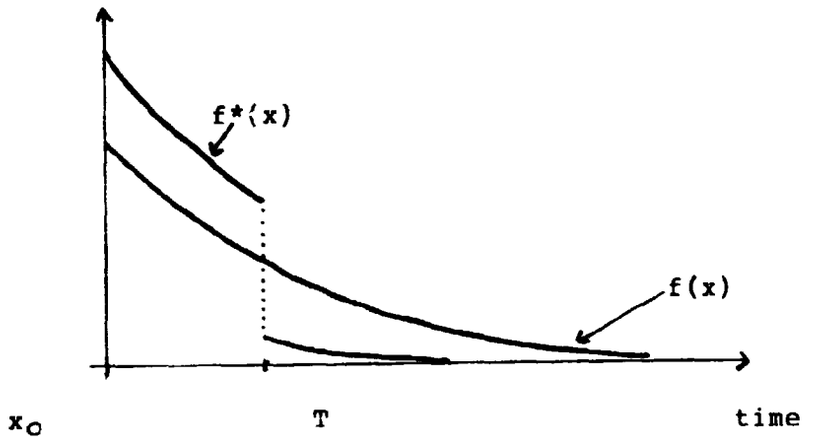
In the general case another implementation can be applied based on the idea given below, which was presented in [25].

4.8.1. Method

Assume a probability density function f(x)



You are only interested in lifetimes less than a given value T. T can be either the observation time, the maximum lifetime of the component or some other value based on knowledge to the problem. In stead another density function f*(x) emphasizing the interval between x_o and T is chosen.



In this case f*(x) can be constructed in this way:

$$\text{let } f^*(x) = \begin{cases} af(x) & x_o \leq x \leq T \\ bf(x) & T < x \end{cases}$$

Find a and b, so you get more lifetimes between x_o and T than using f(x).

As f*(x) is a probability density function

$$\int_{x_o}^{\infty} f^*(x)\,dx = 1$$

$$\int_{x_o}^{\infty} f^*(x)\,dx = \int_{x_o}^{T} af(x)\,dx + \int_{T}^{\infty} bf(x)\,dx$$

Exponential distribution

If $f(x) = \lambda \exp(-\lambda(x-x_o))$, $x_o \geq o$, $\lambda > o$

$$F(x) = \int_{x_o}^{x} \lambda \exp(-\lambda(t-x_o))\,dt$$

$$= 1 - \exp(-\lambda(x-x_o))$$

$$\int_{x_o}^{\infty} f^*(x)$$

$$= \int_{x_o}^{T} af(x)\,dx + \int_{T}^{\infty} bf(x)\,dx$$

$$= aF(T) + b\,(1-F(T))$$

$$= a(1-\exp(-\lambda(T-x_o))) + b(1-(1-\exp(-\lambda(T-x_o))))$$

$$= a - a\exp(-\lambda(T-x_o)) + b\exp(-\lambda(T-x_o))$$

$$= a - (a-b)\exp(-\lambda(T-x_o))$$

Now

$$a - (a-b)\exp(-\lambda(T-x_o)) = 1$$

or

$$b = a + (1-a)/\exp(-\lambda(T-x_o))$$

If a is chosen, such that

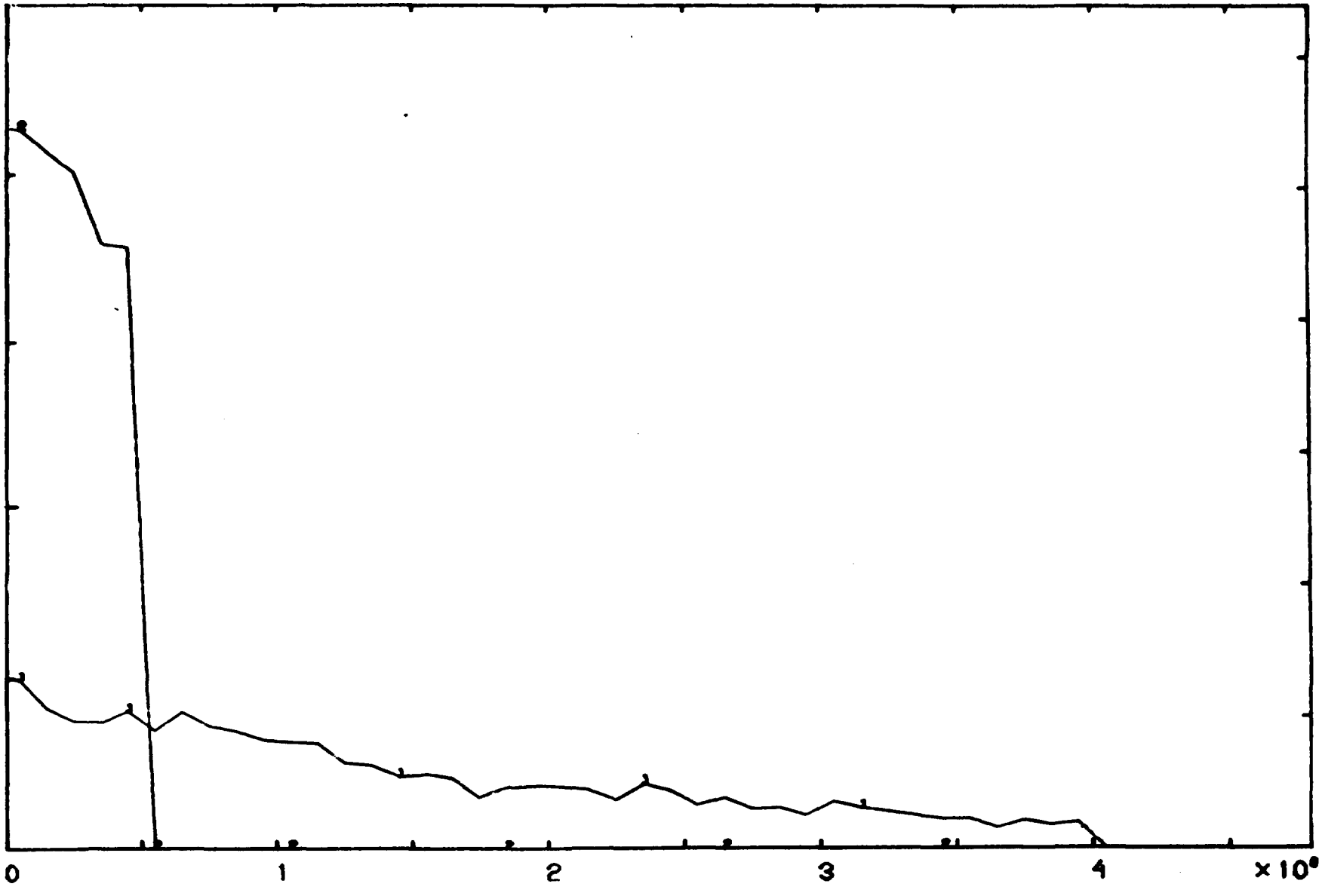
$$(1-a)/\exp(-\lambda(T-x_o)) = -a$$

Then b = o.

It is not allowable to modify the function in the way that you exclude some lifetimes, since $f^*(x)$ is then zero for $x > T$ and the modification factor $f(x)/f^*(x)$ is undefined. This is an example of overbiasing.

An example is shown on the next page using a developed program, EXTEP. This program will plot the results to give the user an outline of the results of the biasing process.

In the example an exponential distribution with $\lambda = 0.5$ and $x_o = 0.0$ is chosen (curve 1) and lifetimes less than 0.5 are emphasized. a is calculated using the above formula, giving $a = 4.5$ and $b = 5 \cdot 10^{-3}$ (curve 2) to avoid overbiasing. The lifetimes of interest are generated 4.5 times as frequent using $f^*(x)$ than using $f(x)$.


EXPONENTIAL
0
1
2
3
4
x10⁸

In direct simulation you generate lifetimes using a random number generator

$$x = x_o - (1/\lambda) \cdot \ln(R)$$

In the modified simulation you generate R as a random number uniformly distributed over 0 to 1.

If $R < aF(T)$

then the expression $f^*(x) = af(x)$ is used, otherwise $f^*(x) = bf(x)$.

In the former case

$$R = aF(x) \qquad x_o \leq x \leq T$$

$$= a(1-\exp(-\lambda(x-x_o)))$$

which gives

$$x = x_o - \ln(1-R/a)/\lambda$$

In the latter case

$$R = aF(T) + b(F(x)-F(T)) \qquad x>T$$

We know

$$aF(T) + b(1-F(T)) = 1$$

or

$$a(F(T))-b(F(T)) = 1-b$$

which means that

$$R = aF(T) + b(F(x)-F(T))$$

$$= 1-b + bF(x)$$
$$= 1-b + b(1-\exp(-\lambda(x-x_o)))$$
$$= 1-b\exp(-\lambda(x-x_o))$$

$$R = 1-b\exp(-\lambda(x-x_o))$$

or

$$x = x_o - \ln((1-R)/b)/\lambda$$

Weibull distribution

The Weibull distribution is given by

$$f(x) = k(x-x_o)^m \exp(-k/(m+1)(x-x_o)^{m+1})$$

$$F(x) = \int_{x_o}^{x} f(t)dt$$

$$= 1-\exp(-k/(m+1)(x-x_o)^{m+1})$$

Using direct simulation and a random number R, we find

$$R = 1-\exp(-k/(m+1)(x-x_o)^{m+1})$$

The lifetime x is found by

$$x = x_o+(-(m+1)/k \ln R)^{(1/(m+1))}$$

If again

$$f^*(x) = \begin{cases} af(x) & x_o \leq x \leq T \\ bf(x) & x>T \end{cases}$$

the lifetimes are found using the modified function.

1. $x_o \leq x \leq T$

$$R = \int_{x_o}^{x} af(t)dt$$

or

$$x = x_o + (-(m+1)/k \ln(1-R/a))^{(1/(m+1))}$$

2. $x>T$

$$R = aF(T) + b(F(x)-F(T))$$

We know that

$$a(F(T)) + b(1-F(T)) = 1$$

or

$$a(F(T)) - b(F(T)) = 1-b$$

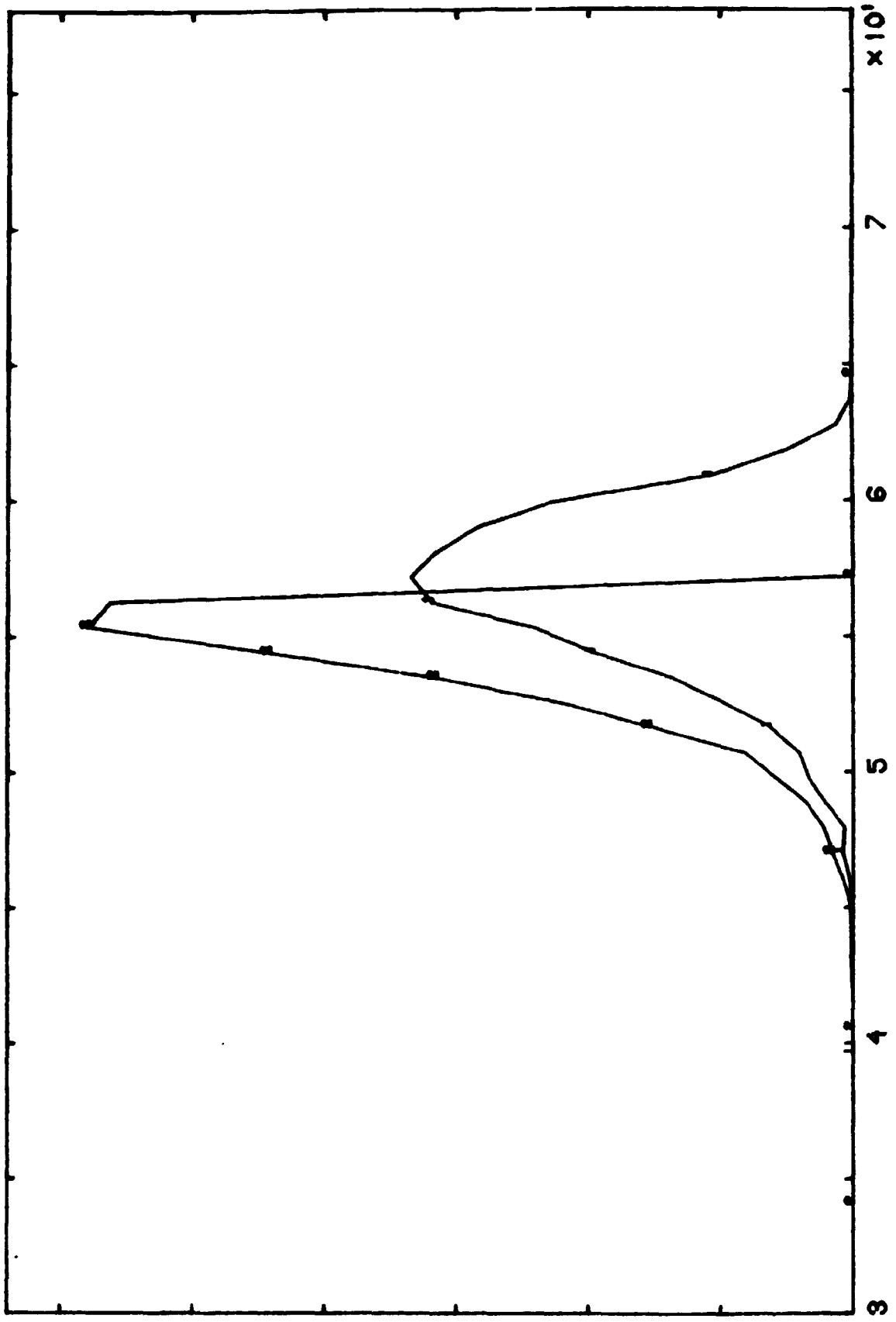
Then

$$R = 1-b + bF(x)$$

or

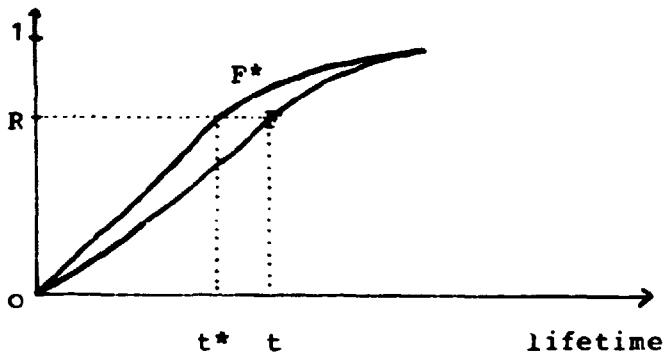
$$x = x_o + (-(m+1)/k \ln((1-R)/b))^{(1/(m+1))}$$

An example is shown on the next page using the EXTEP program. A Weibull distribution with $k = 1.57 \cdot 10^{-12}$, $m = 8.2$ and $x_o = 35$ is chosen. Lifetimes between x_o and 54.12 are emphasized, giving $a = 1.6$ and $b = 10^{-3}$. As in the example with the exponential distribution the lifetimes of interest are emphasized. The value of a depends on the length of the interval of interest. If the interval is shorter than [35.0,54.12] higher values of a are allowed.


x 10¹
7
6
5
4
3

4.8.2. Modified Random Numbers

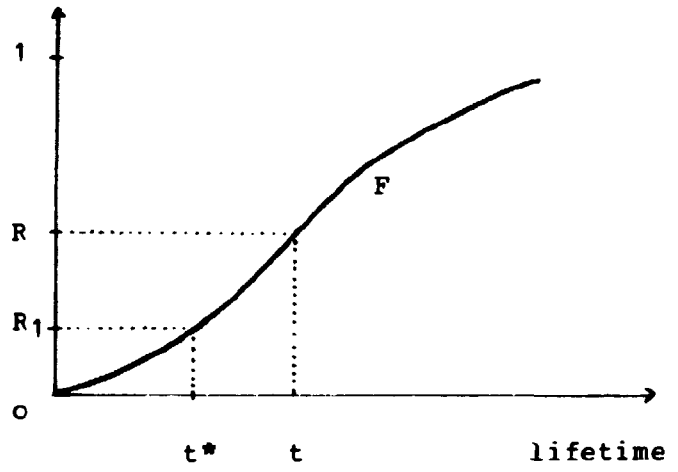
The methods of importance sampling described so far are based on the idea, that a modified probability density function is chosen, so that shorter lifetimes are generated. The modification is compensated in the final results. The idea can be illustrated this way:



The cumulative density function F is shown and a random number R uniformly distributed over o to 1 is generated. The corresponding lifetime is t. We choose another probability density function with cumulative density function F*(x) assuring shorter lifetimes. Given the same random number R, the lifetime t* is used, where $t^* < t$.

Another method to be used in importance sampling is based on a modification of the random number generator and use of the original probability density function.

Choose random numbers which are distributed over o to 1 according to some other distribution than the uniform distribution. The choice should emphasize small values of random numbers. This can be illustrated in the following way:



Again using direct simulation you generate R following a uniform distribution with a corresponding l. .etime t. If a large sample is generated,

$$P(\text{generated values between o and } R_1) = R_1$$

which corresponds to generation of lifetimes between o and t*.

A weighting is made by the function

$$g(x) = \begin{cases} 2 & o \leq x \leq R_1, \quad R_1 < 0.5 \\ (1-2R_1)/(1-R_1) & R_1 < x \leq 1 \end{cases}$$

where $\int_o^1 g(x)dx = 2 \cdot R_1 + ((1-2R_1)/(1-R_1))(1-R1) = 1$

This means, that lifetimes between o and t* are generated twice as frequent as when using direct sampling.

No general results exists on how to find an optimal weighting function. It has to be found by trial and error.

A computer program IMPSAMPLOT is developed which can treat this type of importance sampling with plotting facilities available to give the user an outline of the process. A series of examples giving all details of the calculation are presented in appendix 2.

Example A

A Weibull distribution is chosen with

$k = 1.57 \cdot 10^{-12}$

$m = 8.2$

$x_o = 35.0$

The lifetimes of interest are lifetimes between 35.0 and 54.12. If direct simulation is used the probability of getting lifetimes between 35.0 and 54.12 is 0.1. This means, that in a sample of size 9000 we would expect 900 lifetimes of interest. In the example we actually found 893. In appendix 2, example A, you will find the results of the calculation, together with a table specifying the distribution of the random numbers, which generated the lifetimes of interest.

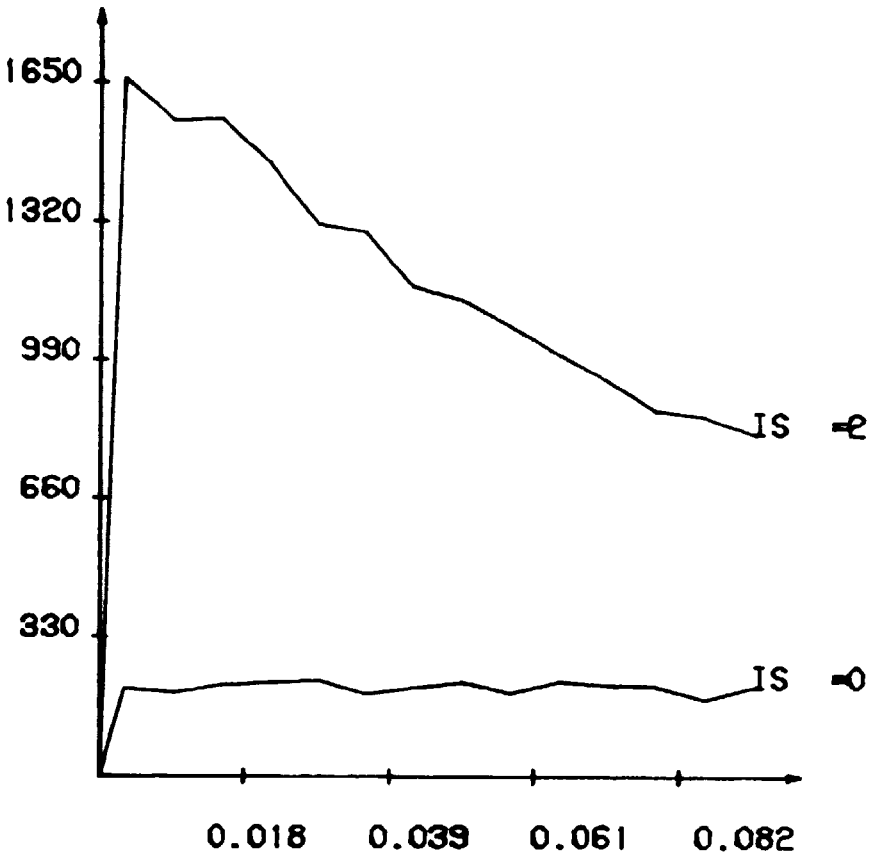
An exponential weighting function is specified with parameter 8.1. In this case, you find 4995 lifetimes between 35.0 and 54.12 in the sample of size 9000. These results are also shown in appendix 2, example A.

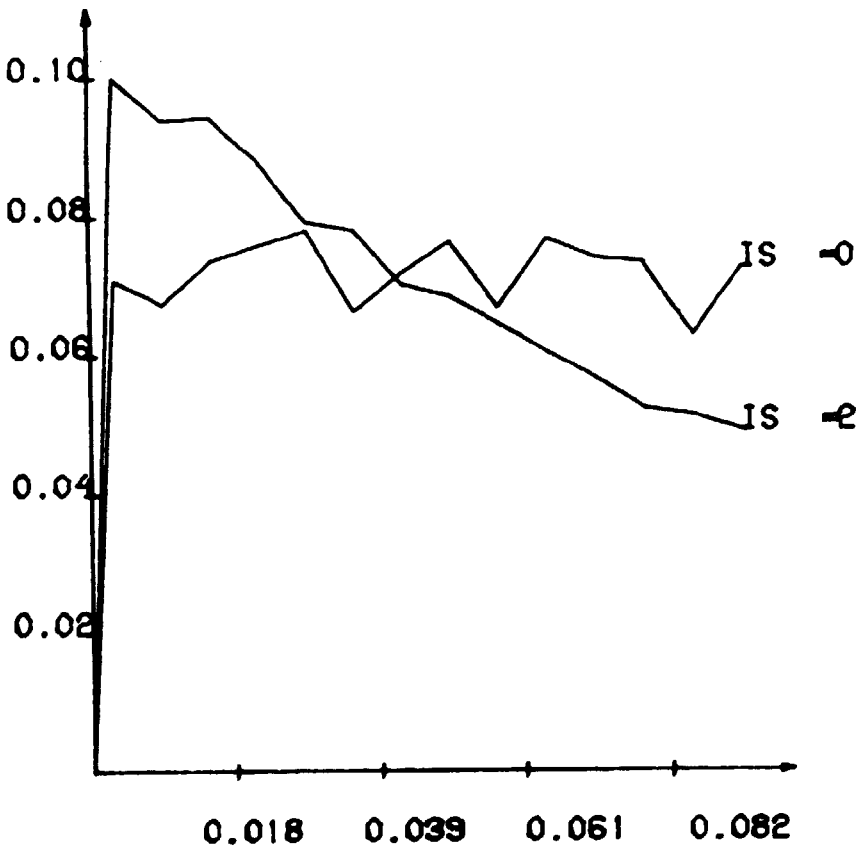
Example B

The situation is similar to that given in example A. A sample of size of 20.000 is generated using direct sampling, weighting by a pointwise defined density function and weighting by the exponential distribution specified in example A.

Example C

The situation is similar to that given in example A. A sample size of 30.000 is generated using direct sampling and weighting by the exponential distribution specified in example A. A plot of the results is shown on the next page. On the upper part you can see the difference without weighting (IS=0) and with weighting (IS=2). On the lower part the difference is shown when the functions are normalized, showing that the smaller values are emphasized using weighting.


1650
1320
990
660
330
IS =2
IS =0
0.018
0.039
0.061
0.082


0.10
0.08
0.06
0.04
0.02
IS =0
IS =2
0.018
0.039
0.061
0.082

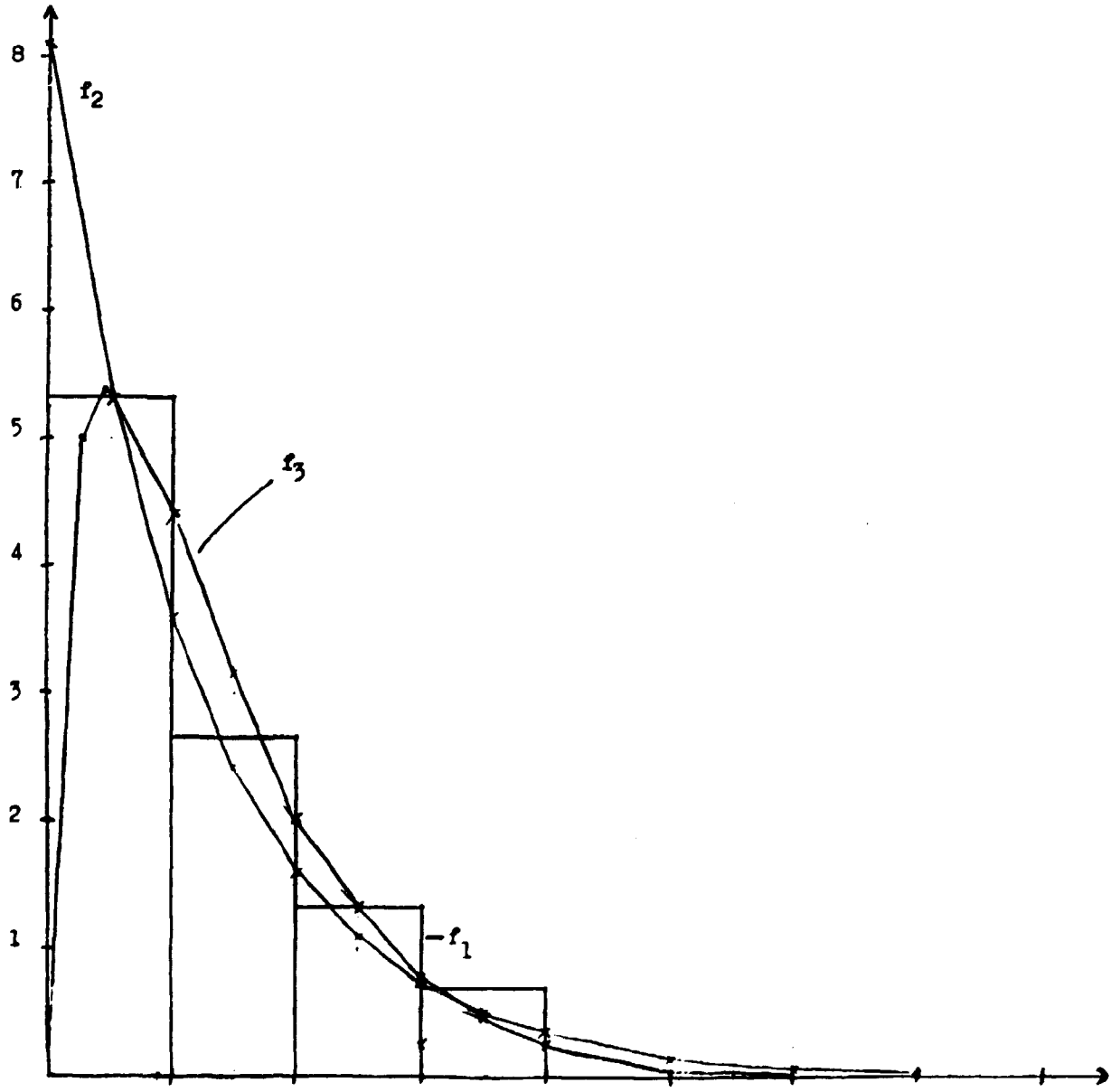
Example D

The situation is similar to that given in appendix A. A sample of size 20.000 is generated using

- no weighting
- weighting by a pointwise defined function (f_1)
- weighting by an exponential distribution with $\lambda = 8.1$ (f_2)
- weighting by a Weibull distribution with (f_3)

$k = 17.0$ $m = 0.3$ $x_o = 0.0$

The three weighting functions are shown below.



As it is seen they are very close to each other. The results of the calculation show that the number of generated lifetimes of interest were:

without weighting	3
pointwise defined distribution	11
exponential distribution	15
Weibull distribution	0

This shows, that even if the distributions are similar they do not have similar effect on the results.

In general the use of Weibull distributions or other distributions with a zero is not recommended, since it is not possible to emphasize values close to zero.

4.8.3. Possible Implementations

The two types of application of importance sampling discussed above can be used in connection with the MOCARE program. The problems are:

- which method to be used
- which type of weighting distribution to be used
- which parameters to be used.

Furthermore, another problem arise: time is lost on administration of the weighting process, if the number of components is large.

To solve the problems, where no general rules exists, a procedure is proposed in two cases:

- some prior knowledge
- no prior knowledge.

4.8.4. Prior Knowledge

If some knowledge is available about which of the components are important with respect to system failure, only these components are subject to weighting. The choice of method, weighting func-

tion and parameters of the function is decided on a trial and error basis. Some guidelines based on the experience with the EXTEP and IMPSAMPLOT programs can be given.

Using the method described in section 4.8.1. an exponential or a Weibull distribution is preferable. The choice of parameters is very difficult due to the problem of overbiasing. Overbiasing means, that you exclude important lifetimes by the modification.

In the simple example with the two dice discussed in section 4.7. overbiasing is done in the following case:

Assume the first die is modified, so

$$P(1) = 1 \qquad P(2) = P(3) = P(4) = P(5) = P(6) = o$$

The second die is modified, so

$$P(2) = 1 \qquad P(1) = P(3) = P(4) = P(5) = P(6) = o$$

The probability that you get the sum equal to 3 is now 1 and the compensation factor is $((1/6)/(1/1))^2 = 1/36$. The result 1/36 is wrong due to the fact, that the possibilities that die 1 shows 2 and die 2 shows 1 are excluded.

One reasonable choice of the maximum lifetimes of interest might be the total observation time or a time based on some knowledge of the system.

4.8.5. Adaptive Methods

If no prior knowledge exists another procedure is proposed:

1. Use direct simulation and stop after N system failures

2. Select the most important components contained in the cut sets generated

3. Use a weighting function for each of the selected components after the principles discussed in section 4.8.4.

4. Continue the simulation and stop after N system failures

5. Continue with item 2.

One way to make the selection of the most important components under item 2 could be to select those which are in cut sets with the largest contribution to the system unreliability.

However, it must be emphasized that using this procedure there is another risk of overbiasing, namely that some components and hereby some cut sets are generated more and more frequently, excluding the appearance of other cut sets, which might be important.

4.9. Conclusions

The application of importance sampling as a variance reduction technique in Monte Carlo simulation programs is very limited due to lack of rules for constructing the weighting functions. Mostly, use of the trial and error method is recommended. Even in simple cases as discussed by Kamarinopoulos, with exponential distributions and weighting functions of the same type, the selection of parameters must be based on trial and error. Furthermore, for systems with stand-by components the simple modification factors are not valid. Much more complicated modification factors are needed, see [26].

However, the two developed programs, EXTEP and IMPSAMPLOT, have shown to be very flexible tools for the trial and error process. They can treat the same density functions as accepted by MOCARE and in this respect they are very useful in the preparation of relevant weighting functions.

5. SUMMARY OF CONCLUSIONS

In this chapter the purpose of the project and a summary of the main results are summarized.

5.1. General

In risk and reliability analysis several methods are available. They can be applied in the various disciplines within risk and reliability analysis.

One class is the methods within structural reliability, where the probability of failure of a structure or a component is calculated. In linear fracture mechanics some probabilistic methods have been developed. One purpose of the project was to review existing methods and to make proposals for refinements or new methods.

Fault tree methods have been very extensively used and very fast methods for quantification are available. One class of methods is the Monte Carlo simulation methods of which MOCARE is one program. The other purpose of the project was to review existing methods for variance reduction with respect to implementation within the MOCARE program.

5.2. Structural Reliability

In probabilistic linear fracture mechanics methods exist based on

- Monte Carlo simulation (PEP 706)
- integration using discretized probability density functions (ANPEP, ANPEP/V2)
- combination of discretized probability density functions (COVAL, SCORE).

A review of the methods and some comparisons between the methods show, that the Monte Carlo approach is a very flexible tool, but unfortunately the computing time required to obtain a reasonable accuracy is large.

ANPEP, ANPEP/V2, COVAL and SCORE require less computing time to obtain the same accuracy as the Monte Carlo program. The main limitation of the programs is that increasing the number of variables will dramatically increase the computing time. It has been shown, that in cases of up to 4 variables, the programs are well functioning, while an increase to 6 variables lead to unsatisfactory computing times.

A new program, NUMPEP, based on numerical integration in several variables using product formulas of Gauss-Legendre and Gauss-Laguerre type has been developed and compared to the other programs. It has been found that NUMPEP is faster than the other programs for problems involving up to 6 variables. It is fair to mention that it is possible to increase the number of variables, but it is expected that if it exeeds 10, the computing time will exceed a reasonable value.

The program can also be used within other areas than linear fracture mechanics, if the problem can be formulated in terms of multidimensional integrals using probability density functions. Furthermore, it is modularly constructed which means that it can be easily extended to more that 6 variables.

Finally, to prepare the program for a new combination of the variables more effort is required than to prepare the Monte Carlo program, due to the favorable flexibility of the latter program. Since the results are promising it is recommended to apply the NUMPEP program.

5.3. Systems Reliability

The MOCARE program which utilizes direct Monte Carlo simulation, is a very flexible tool in quantification of fault trees. The

flexibilities are with respect to

- choice of probability density functions
- modelling the system taking into account complex test and maintenance schedules, specification of switches between redundant equipment, detailed description of component or subsystem performance.

In some cases a considerable amount of computing time is required to get a reasonable estimate of the reliability of a highly reliable system.

A review of existing variance reduction techniques in Monte Carlo simulation programs for reliability analysis has been performed. It is found that the importance sampling method is useful in these applications. Furthermore, it is found that no general results exists for constructing an optimal weighting function.

For special cases where all distributions are exponential, some results are available, if the weighting function or importance function is chosen as an exponential distribution with a parameter which is multiplied by B, $B>1$. The corresponding correction terms are specified for a general system, but the choice of B has to be based on trial and error.

In the general case no results is available, but importance sampling can be used with advantage. Two implementations have been proposed and studied. A very simple modification of the original probability density function and a modification of the random number generator, so that some skewed distribution is used in stead of the uniform distribution. The correction terms are of course similar to those given in the exponential case, but the specifications of which method, which weighting distribution and which parameters to be used must be based on trial and error and knowledge of the problem considered. Two computer programs are developed which can be used with the purpose of guidance in selection of weighting function and parameters.

As no general results is derived for constructing the weighting function the large computing time to run the MOCARE program remains a problem.

Therefore, in general other programs such as FAUNET, is recommended in reliability analysis. They are faster and they do not introduce a statistically uncertainty as in MOCARE. Furthermore, it is recommended to use MOCARE in special cases where FAUNET does not apply:

- systems where time dependencies are important
- systems with subsystems, where detailed modelling of a complex performance is required.

In these cases MOCARE is an efficient supplement to the FAUNET program.

It is also recommended to use the developed importance sampling methods within the MOCARE program in order to speed up the calculations. Even if an optimal weighting function is not constructed, a less efficient choice of function can be of importance. It is important to emphasize that the application of the importance sampling methods must be controlled very carefully by the user due to the risk of overbiasing, which will lead to incorrect results, and due to the lack of general results.

5.4. Applications

The NUMPEP program is applicable within probabilistic fracture mechanics. It can be used to determine the failure probability of the pressure vessel in nuclear power plants. It has also been suggested to apply this approach to estimate the probability of failure of the pipework in nuclear power plants. Obviously, also non-nuclear problems are within the range of applications of the program.

The MOCARE program has been used in particular within nuclear applications in calculating the reliability or the unavailabil-

ity of complex systems. Worldwide this type of approach has been substituted fully or partly by other methods, such as the FAUNET program. Its advantage is its flexibility which makes it a very useful tool for some parts of an analysis requiring very detailed modelling of the performance of complex systems. Obviously, it can be applied as well on non-nuclear problems. I expect, that detailed modelling of non-nuclear system will also be required in the future, as the systems get more complex.

5.5. Future Areas of Development

The developed programs NUMPEP, EXTEP and IMPSAMPLOT can be further developed.

Structural Reliability

The NUMPEP code should be extended to take into account

- other combinations of variables
- other probability density functions
- more than six variables
- dependencies between the variables.

Such developments will increase the area of applicability.

Systems Reliability

The EXTEP and IMPSAMPLOT programs can be further developed, if other probability density functions are allowed. The most essential continuation will be to use the programs within a large number of analyses to get experience with various choices of weighting function, since the selection is based on trial and error.

6. ACKOWLEDGEMENTS

The author would like to thank all colleagues for their help and patience during the period of writing. Specifically many thanks are due to Professor Thøger Busk, The Technical University of Denmark, for many inspiring and fruitful discussions during the project time. Further, special thanks are also due to Mr. Hans Erik Kongsø for his interest and encouragement.

Finally, I would like to take the opportunity to thank Gytha Egelund and Anette Grohnheit for the excellent typing of the manuscript.

7. REFERENCES

1. BECHER, P.E. and PEDERSEN, A. (1973). Application of Statistical Elastic Fracture Mechanics to Pressure Vessel Reliability Analysis. Risø-M-1650. (Roskilde, Denmark) 13 p.

2. BECHER, P.E. (1975). PEP 706. SRE-4-75. 33 p. (Internal report prepared for Department of Reactor Technology, Risø National Laboratory)

3. KONGSØ, H.E. and PETERSEN, K.E. (1976). ANPEP: A Computer Program for Calculation of the Probability of Failure of Structures: Program Description and Manual. RISØ-M-1855. (Roskilde, Denmark) 38 p.

4. KONGSØ, H.E. and PETERSEN, K.E. (1977). ANPEP/V2: A Computer Program for Calculation of the Probability of Failure of Structures: Program Description and Manual. RISØ-M-1918. 78 p. (Internal report prepared for Department of Reactor Technology, Risø National Laboratory)

5. COLOMBO A.G. and JAARSMA, R.J. (1980). A Powerful Numerical Method to Combine Random Variables. IEEE Transactions on Reliability, R-29, 126-129.

6. COLOMBO, A.G. and JAARSMA, R.J. (1980). SCORE: A Computer Program for the Systematic Combination of Random Variables. EUR-6819. (Bruxelles) 43 p.

7. STEFFENSEN, J.F. (1925). Interpolationslære. (Gad, København) 232 p.

8. STROUD, A.H. (1971). Approximate Calculations of Multiple Integrals. (Prentice-Hall, Englewood Cliffs, N.J.) 431 p.

9. DAVIS, P.J. and RABINOWITZ, P. (1975). Methods of Numerical Integration. (Academic Press, New York) 459 p.

10. STROUD, A.H. and SECREST, D. (1966). Gaussian Quadrature Formulas. (Prentice Hall, Englewood Cliffs, N.J,) 374 p.

11. TAYLOR, J.R. (1979). A Background to Risk Analysis. (Risø National Laboratory, Roskilde) 4 vols.

12. NORDFORSK (1984). Sikkerhetsanalyse som beslutningsunderlag. Slutrapport från SCRATCH-programmet. (Yrkeslitteratur, Oslo) 2 vols.

13. PRA Uses and Techniques: a Nordic Perspective: summary report of the NKA project SÄK-1 (1985). Edited by S. Dinsmore. Nord-Series. (NKA, [s.l.]) 121 p.

14. PETERSEN, K.E. (1986). Risk Analysis Uses and Techniques in the Non-Nuclear Field: a Nordic Perspective. (NKA, [s.l.]) 34 p.

15. FUSSELL, J.B. (1974). Fault Tree Analysis - Concepts and Techniques. [In:] Generic Techniques in Systems Reliability Assessment. Edited by E.J. Henley and J.W. Lynn. (Nordhoff, Leiden)

16. NIEMELÄ, I. (1985). RELVEC Manual, SAH 10/85, SAH 11/85. (Technical Research Centre of Finland, Espoo) 2 vols.

17. MANN, N.R., Schafer, R.E., and Singpurwalla, N.D. (1974). Methods for Statistical Analysis of Reliability and Life Data. (Wiley and Sons, New York) 564 p.

18. PLATZ, O. and OLSEN, J.V. (1976). FAUNET: A Program Package for Evaluation of Fault Trees and Network. Risø Report No. 348. 49 p.

19. KONGSØ, H.E., LARSEN, H.L., NILSSON, K., and PETERSEN, K.E. (1978). MOCARE, A Computer Program for System Reliability Analysis by Monte Carlo Simulation: Program Description and Manual. Risø-M-2109. 54 p. (Internal report)

20. PETERSEN, K.E. and KONGSØ, H.E. (1983). Notes on the computational Results. NKA/SAK-1 Benchmark Study II. SAK-1-D(83)5. (Internal paper, not published)

21. MCGRATH, E.J. and IRVING, D.C. (1975). Techniques for Efficient Monte Carlo Simulation, Volume III: Variance Reduction. ORNL-RSIC-38 (Oak Ridge, Tenn.) 161 p.

22. NAGEL, P. (1966). Importance Sampling in Systems Simulation. Fifth Annual Conference on Reliability and Maintainability, Held in New York, July 1966. (American Institute of Aeronautics and Astronautics, New York) 330-337

23. KAMARINOPOULOS, L. Direkte und Gewichtete Simulationsmetoden zur Zuverlässigkeitsuntersuchung Technischer Systeme. D 83, Techniche Universität, Berlin, 1972. (Dissertation)

24. KAMARINOPOULOS, L. Anwendung von Monte-Carlo-Verfahren zur Ermittlung von Zuverlässigkeitsmerkmalen Technischer Systeme. ILS Bericht 14, Technische Universität, Berlin, 1976.

25. PETERSEN, K.E. (1980). An Implementation of Importance Sampling in a Monte Carlo Simulation Program for Analysis of Reliability of Systems. Paper presented at COMPSTAT 1980. 4. symposium held in Edinburgh. (Not published)

26. KONGSØ, H.E. (1972). RELY 4: A Monte Carlo Computer Program for Systems Reliability Analysis. Risø-M-1500, (Roskilde, Denmark) 76 p.

DANSK RESUME

PÅLIDELIGHEDSBEREGNINGER

Forbedringer af metoder til beregring af pålideligheden af strukturer og systemer.

Risiko- og pålidelighedsanalyse anvendes i stigende grad ved evalueringer af anlægssikkerhed og anlægspålidelighed. Analyser kan udføres enten i designfasen eller efter idrifttagning, men de foreslåede designændringer eller modforanstaltninger mod designsvagheder er selvfølgelig forskellige i de to tilfælde. Analyserne kan udføres som en del af den dokumentation, der er nødvendig for myndighedsgodkendelse, eller med det formål at forbedre sikkerheden eller pålideligheden af anlægget.

Anlæg, som anvendes i dag, er komplekse på grund af krav til sikkerhed og tilgængelighed. Dette medfører behov for avancerede analyseværktøjer, som er fleksible og effektive, til brug i evalueringer af sikkerheden eller pålideligheden af anlægget. Sådanne værktøjer er blevet udviklet i de seneste 20 år, og de må løbende forbedres, for at kunne møde de øgede krav med hensyn til kompleksitet, fuldstændighed og nøjagtighed.

To forskellige anvendelsesområder er blevet analyseret i dette projekt. I strukturel pålidelighed er sandsynlighedsbaserede metoder blevet anvendt i visse beregninger af pålideligheden af strukturer eller komponenter. Et regnemaskineprogram er blevet udviklet og sammenlignet med eksisterende programmer. Dette program, som er baseret på numerisk integration i flere variable, har vist sig at være meget hurtigt sammenlignet med eksisterende programmer. Dette er vigtigt, da mangel på data og usikkerhederne relateret til problemet kræver omfattende brug af følsomhedsanalyser, der medfører en mængde gennemregninger.

I systempålidelighed er Monte Carlo programmer brugt specielt ved analyse af meget komplekse systemer, der består af flere redundante delsystemer, procedurer for skift mellem delsystemerne og komplekse test- og vedligeholdsprocedurer. I sådanne tilfælde er simuleringsprogrammer meget fleksible, hvilket muliggør denne type af vanskelige og detaljerede modelleringer. For at øge anvendeligheden af programmerne, kan variansreduktionsmetoder anvendes til at hurtiggøre beregningsprocessen.

Variansreduktionsmetoder er blevet analyseret, og procedurer for implementering af "importance sampling" er foreslået, både i tilfælde med et vist forhåndskendskab til problemstillingen, og i tilfælde uden et sådant forhåndskendskab. Det er konkluderet, at selv med brug af variansreduktionsmetoder anbefales brug af analytiske programmer i forhold til simuleringsprogrammer, mens variansreduktionsteknikker kan forbedre anvendeligheden af simuleringsprogrammer i de specielle situationer, der er beskrevet ovenfor.

APPENDIX 1

PROGRAM EXAMPLES

STRUCTURAL RELIABILITY

EXAMPLE 1

```
**********************************************************************

          EVALUATION OF FAILURE-PROBABILITY
                    USING
          NUMERICAL INTEGRATION

**********************************************************************
```

INPUT DATA

EXPRESSION FOR K1**2:

VARIABLE NO. 1:	CONSTANT	
	C =	.3800E+01
VARIABLE NO. 2:	EXPONENTIAL DISTRIBUTED	
	ZERO =	.1000E+00
	PARAMETER =	.2560E+01
VARIABLE NO. 3:	NORMAL DISTRIBUTED	
	MEAN VALUE =	.2600E+02
	STANDARD DEVIATION =	.3200E+01

EXPRESSION FOR K1C**2:

VARIABLE NO. 4:	CONSTANT	
	C =	.5000E+01
VARIABLE NO. 5:	WEIBULL DISTRIBUTED	
	ZERO =	.3500E+02
	SCALE PARAMETER =	.1570E-11
	FORM PARAMETER =	.8200E+01
VARIABLE NO. 6:	WEIBULL DISTRIBUTED	
	ZERO =	.6000E+02
	SCALE PARAMETER =	.2530E-05
	FORM PARAMETER =	.2650E+01

INTEGRATION: P(K1>K1C)

INTEGRATION LIMITS: 10500.00 - 60000.00

NUMBER OF INTEGRATION POINTS = 32

NUMBER OF INTEGRATION POINTS USED IN EVALUATION OF EACH OF THE P.D.F.'S FOR K1 AND K1C = 30

PROBABILITY OF FAILURE = .2826E-06

Example 2

```
**********************************************************************
                EVALUATION OF FAILURE-PROBABILITY
                        USING
                NUMERICAL INTEGRATION
**********************************************************************
```

INPUT DATA

EXPRESSION FOR K1**2:

VARIABLE NO. 1:	CONSTANT	
	C =	.3800E+01
VARIABLE NO. 2:	EXSPONENTIAL DISTRIBUTED	
	ZERO =	.1000E+00
	PARAMETER =	.2560E+01
VARIABLE NO. 3:	NORMAL DISTRIBUTED	
	MEAN VALUE =	.2600E+02
	STANDARD DEVIATION =	.3200E+01

EXPRESSION FOR KIC**2:

VARIABLE NO. 4:	CONSTANT	
	C =	.5000E+01
VARIABLE NO. 5:	WEIBULL DISTRIBUTED	
	ZERO =	.3500E+02
	SCALE PARAMETER =	.1570E-11
	FORM PARAMETER =	.8200E+01
VARIABLE NO. 6:	NORMAL DISTRIBUTED	
	MEAN VALUE =	.1050E+03
	STANDARD DEVIATION =	.1500E+02

INTEGRATION: P(R1><1C)

INTEGRATION LIMITS: 10500.00 - 60000.00

NUMBER OF INTEGRATION POINTS = 8

NUMBER OF INTEGRATION POINTS USED IN EVALUATION OF EACH OF THE P.C.F.'S FOR R1 AND R1C = 8

PROBABILITY OF FAILURE = .6360E-05

APPENDIX 2

PROGRAM EXAMPLES

SYSTEMS RELIABILITY

EXAMPLE A

WEIGHTING FUNCTION
NO IMPORTANCE SAMPLING
IS =

INTEGRATION LIMIT
CHECK = .5612419960E+02

FP= 0.00922222 (993) (FAILURE PROBABILITY)
STD= 0.00315117 (STANDARD DEVIATION)

MIN= 0.00005023
MAX= 0.00964754
NINT= 18 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.00	-	0.00	66
0.00	-	0.01	76
0.01	-	0.02	52
0.02	-	0.03	79
0.03	-	0.03	76
0.03	-	0.04	70
0.04	-	0.05	71
0.05	-	0.05	99
0.05	-	0.06	93
0.06	-	0.07	77
0.07	-	0.08	71
0.08	-	0.09	66
0.09	-	0.09	75

TOTAL CPU-TIME= 15.8 SEC
TIME PR. TRIAL= 1.8 MSEC

DATE OF EXECUTION: 15. 8.80.

SAMPLING FUNCTION

SAMPLING FROM DEFAULT DISTRIBUTION
ITYP =
P =
M =
X =

SPECIFIC NUMBERS

NUMBER OF TRIALS =
RANDOM INITIAL VALUE =

DESIGNING FUNCTION

IMPORTANCE SAMPLING
IS =
EXPONENTIAL DISTRIBUTION
LAMBDA=

INTEGRATION LIMIT

CHECK =

FP=
FPI=
STD=
STDSUM=
SDSUM2=

(CORRECTED FAILURE PROBABILITY)
(UNCORRECTED)
(ST.DEV. USING FPI - SQRT(FP*(1-FP)/N))
(ST.DEV. USING FP - SQRT(FP*(1-FP)/N))
(ST.DEV. USING DIRECT CALCULATION)

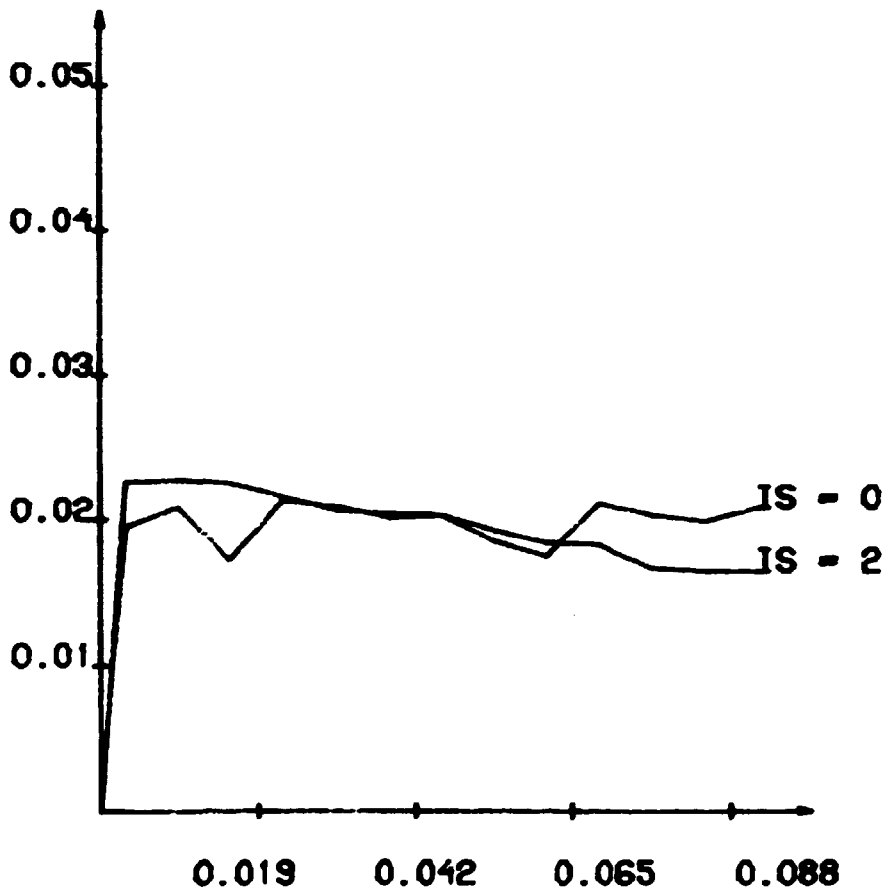
MIN=
MAX=
NINT= 11

(NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000011711	-	0.007733647	506
0.007733641	-	0.015334775	511
0.015334775	-	0.022935908	497
0.022935908	-	0.030777041	457
0.030777041	-	0.038464372	411
0.038464372	-	0.046153705	432
0.046153705	-	0.053851037	397
0.053851037	-	0.061542370	356
0.061542370	-	0.069233702	324
0.069233702	-	0.076925034	323
0.076925034	-	0.084616367	272
0.084616367	-	0.092307699	295
0.092307699	-	0.099999031	270

TOTAL CPU-TIME= 30.1 SEC

TIME PR. TRIAL= 3.3 MSEC


0.05
0.04
0.03
0.02
0.01
IS - 0
IS - 2
0.019
0.042
0.065
0.088

DATE OF EXECUTION: 7. 3.CU.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
K = .157E-11
M = 8.20
XO = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

NO IMPORTANCE SAMPLING
IS = 0

CHECKPOINT

CHECK = .54124199630E+02

EP= 0.087600000 (1948) (FAILURE PROBABILITY)
STD= 0.002896583 (STANDARD DEVIATION)

MIN= 0.000005023
MAX= 0.099937093
NINT= 11 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000005023	-	0.009089757	167
0.009089757	-	0.018174490	177
0.018174490	-	0.027259224	150
0.027259224	-	0.036343958	174
0.036343958	-	0.045428691	194
0.045428691	-	0.054513425	174
0.054513425	-	0.063598159	162
0.063598159	-	0.072682892	167
0.072682892	-	0.081767626	183
0.081767626	-	0.090852359	184
0.090852359	-	0.099937093	186

DATE OF EXECUTION: 7. 1.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
A = .157E-11
M = 8.20
XO = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 1
POINTWISE DEFINED FUNCTIONS

XF			FXF	NUMBER
0.000000	-	0.100000	5.333333	10704
0.100000	-	0.200000	2.666667	0
0.200000	-	0.300000	1.333333	0
0.300000	-	0.400000	0.666667	0

XC	FXC
0.000000	0.000000
0.100000	0.533333
0.200000	0.800000
0.300000	0.933333
0.400000	1.000000

CHECKPOINT

CHECK = .54124199600E+02

FP=	0.10035000 (10704)	(CORRECTED FAILURE PROBABILITY)
FP1=	0.53520000	(UNCORRECTED ")
STD=	0.00352676 2	(ST.DEV. USING FP1 - SQRT(P*Q/N))
STDSUM=	0.00212461 6	(ST.DEV. USING FP - SQRT(P*Q/N))
SDSUM2=	0.00006612 68	(ST.DEV. USING DIRECT CALCULATION)
MIN=	0.00000094 2	
MAX=	0.09599797 8	
NINT=	14	(NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000942	-	0.007143587	743
0.007143587	-	0.014286233	730
0.014286233	-	0.021428878	756
0.021428878	-	0.028571524	818
0.028571524	-	0.035714169	802
0.035714169	-	0.042856814	719
0.042856814	-	0.049999460	749
0.049999460	-	0.057142105	772
0.057142105	-	0.064284751	755
0.064284751	-	0.071427396	737
0.071427396	-	0.078570041	815
0.078570041	-	0.085712687	754
0.085712687	-	0.092855332	770
0.092855332	-	0.099997978	774

DATE OF EXECUTION: 7. 3.80.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
K = .157E-11
M = 8.20
XO = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 2
EXPONENTIAL DISTRIBUTION
LAMBDA= 8.10

CHECKPOINT

CHECK = .54126199600E+02

FP= .120795717 (11165) (CORRECTED FAILURE PROBABILITY)
FP1= .558453030 (UNCORRECTED)
STD= .003511293 (ST.DEV. USING FP1 - SQRT(P*Q/N))
STDSUM= .002128802 (ST.DEV. USING FP - SQRT(P*Q/N))
SDSUM2= .000672735 (ST.DEV. USING DIRECT CALCULATION)

MIN= 0.000000620
MAX= 0.099999031
NINT= 14 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000620	-	0.007143364	1100
0.007143364	-	0.014286107	1038
0.014286107	-	0.021428851	1072
0.021428851	-	0.028571595	958
0.028571595	-	0.035714338	832
0.035714338	-	0.042857082	872
0.042857082	-	0.049999826	766
0.049999826	-	0.057142569	748
0.057142569	-	0.064285313	728
0.064285313	-	0.071428057	697
0.071428057	-	0.078570800	643
0.078570800	-	0.085713544	591
0.085713544	-	0.092856288	552
0.092856288	-	0.099999031	566

TOTAL CPU-TIME= 59.8 SEC
TIME PR. TRIAL= 5.4 MSEC

EXAMPLE C

WEIGHTING FUNCTIONS

NO IMPORTANCE SAMPLING
IS = 0

CHECKPOINT

CHECK = .5612419600E+02

FP= 0.098466667 (2954) (FAILURE PROBABILITY)
STD= 0.001720184 (STANDARD DEVIATION)

MIN= 0.000016470
MAX= 0.099944400
NINT= 14 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000620	-	0.007143764	208
0.007143364	-	0.014286107	198
0.014286107	-	0.021428851	216
0.021428851	-	0.028571595	223
0.028571595	-	0.035714338	229
0.035714338	-	0.042857082	195
0.042857082	-	0.049999826	211
0.049999826	-	0.057142569	224
0.057142569	-	0.064285313	196
0.064285313	-	0.071428057	225
0.071428057	-	0.078570800	217
0.078570800	-	0.085713544	215
0.085713544	-	0.092856288	184
0.092856288	-	0.099999031	213

TOTAL CPU-TIME= 41.2 SEC

TIME PR. TRIAL= 1.4 MSEC

DATE OF EXECUTION: 2. 7.80.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
K = .157E-11
M = 8.20
XO = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 30000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 2
EXPONENTIAL DISTRIBUTION
LAMBDA= 8.19

CHECKPOINT

CHECK = .5412419960OE+02

FP= 6.100198604 (16688) (CORRECTED FAILURE PROBABILITY)
FP1= 6.356266667 (UNCORRECTED - -)
STD= 6.002868415 (ST.DEV. USING FP1 - SQRT(P*Q/N))
STDSUM= 6.001733579 (ST.DEV. USING FP - SQRT(P*Q/N))
SOSUM2= 6.000548198 (ST.DEV. USING DIRECT CALCULATION)

MIN= 6.000000620
MAX= 6.099999031
NINT= 14 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000620	-	0.007143364	1658
0.007143364	-	0.014286107	1558
0.014286107	-	0.021428851	1564
0.021428851	-	0.028571595	1458
0.028571595	-	0.035714338	1311
0.035714338	-	0.042857082	1294
0.042857082	-	0.049999826	1165
0.049999826	-	0.057142569	1134
0.057142569	-	0.064285313	1071
0.064285313	-	0.071428057	1002
0.071428057	-	0.078570800	940
0.078570800	-	0.085713544	868
0.085713544	-	0.092856288	851
0.092856288	-	0.099999031	814

TOTAL CPU-TIME = 87.3 SEC
TIME PR. TRIAL= 2.9 MSEC

EXAMPLE D

DATE OF EXECUTION: 7. 3.80.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
A = .157E-11
M = 8.20
X0 = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

NO IMPORTANCE SAMPLING
IS = 0

CHECKPOINT

CHECK = .439749840206+02

FP= 0.000150000 (3) (FAILURE PROBABILITY)
STD= 0.000086596 (STANDARD DEVIATION)

MIN= 0.000005023
MAX= 0.000094881
NINT= 2 (NUMBER OF INTERVALS)

SUBINTERVALS		NUMBER
0.000005023 =	0.000047952	2
0.000047952 =	0.000094881	1

TOTAL CPU-TIME= 29.9 SEC
TIME PR. TRIAL= 9955.6 MSEC

DATE OF EXECUTION: 6. 1.80.

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
A = .157E-11
M = 8.20
X0 = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE = 5E2911

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 1
POINTWISE DEFINED FUNCTIONS

XF			FXF	NUMBER
0.000000	-	0.100000	5.333333	11
0.100000	-	0.200000	2.666667	0
0.200000	-	0.300000	1.333333	0
0.300000	-	0.400000	0.666667	0

XC	FXC
0.000000	0.000000
0.100000	0.533333
0.200000	0.800000
0.300000	0.933333
0.400000	1.000000

CHECKPOINT

CHECK = .439749840205+02

FP= 0.000103125 (11) (CORRECTED FAILURE PROBABILITY)
FPI= 0.000550000 (UNCORRECTED)
STD= 0.000165786 (ST.DEV. USING FPI - SQRT(P*Q/N))
STDSUM= 0.000071803 (ST.DEV. USING FP - SQRT(P*Q/N))
SDSUM2= 0.000031085 (ST.DEV. USING DIRECT CALCULATION)

MIN= 0.000000942
MAX= 0.000080668
NINT= 4 (NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000942	-	0.000020873	3
0.000020873	-	0.000040805	1
0.000040805	-	0.000060736	2
0.000060736	-	0.000080668	5

DATE OF CALCULATION 08 JUL 80

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
R = .1575-11
M = 8.20
X0 = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 20000
RANDOM INITIAL VALUE= 504711

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 2
EXPONENTIAL DISTRIBUTION
LAMBDA= 8.10

CHECKPOINT

CHECK = .43974984020E+02

FP=	0.000292537 (15)	(CORRECTED FAILURE PROBABILITY)
FP1=	0.000750000	(UNCORRECTED)
STD=	0.000193577	(ST.DEV. USING FP1 - SQRT(P*Q/N))
STDSUM=	0.000068040	(ST.DEV. USING FP - SQRT(P*Q/N))
SDSUM2=	0.000023899	(ST.DEV. USING DIRECT CALCULATION)
MIN=	0.000000620	
MAX=	0.000086200	
NINT=	4	(NUMBER OF INTERVALS)

SUBINTERVALS			NUMBER
0.000000620	-	0.000022015	4
0.000022015	-	0.000043410	4
0.000043410	-	0.000064805	3
0.000064805	-	0.000086200	4

SAMPLING FUNCTION

SAMPLING FROM WEIBULL DISTRIBUTION
ITYP = 1
A = .157E-11
M = 8.20
XO = 35.00

SPECIFIC NUMBERS

NUMBER OF TRIALS = 10000
RANDOM INITIAL VALUE = 504711

WEIGHTING FUNCTIONS

IMPORTANCE SAMPLING
IS = 3
WEIBULL DISTRIBUTION
A = 17.00
M = 0.30

CHECKPOINT

CHECK = .43974984020E+02

FP= 6.020000000 (0) (CORRECTED FAILURE PROBABILITY)
FP1= 6.000500000 (UNCORRECTED)
STD= 0.000000000 (ST.DEV. USING FP1 - SQRT(P*Q/N))
STDSUM= 0.000000000 (ST.DEV. USING FP - SQRT(P*Q/N))
SDSUM2= 0.000000000 (ST.DEV. USING DIRECT CALCULATION)

Risø National Laboratory **Risø - M - 2584**

RELIABILITY CALCULATIONS

Kurt E. Petersen

Date: March 1986

Department or group: Risk Analysis Group

Groups own registration number(s):

Project/contract no.:

Pages: 141 Tables: Illustrations: References: 26

ISBN: 87-550-1229-9

Abstract (Max. 2000 char.):

Risk and reliability analysis is increasingly being used in evaluations of plant safety and plant reliability. The analysis can be performed either during the design process or during the operation time, with the purpose to improve the safety or the reliability.

Due to plant complexity and safety and availability requirements, sophisticated tools, which are flexible and efficient, are needed. Such tools have been developed in the last 20 years and they have to be continuously refined to meet the growing requirements.

Two different areas of application were analysed. In structural reliability probabilistic approaches have been introduced in some cases for the calculation of the reliability of structures or components. A new computer program has been developed based upon numerical integration in several variables.

In systems reliability Monte Carlo simulation programs are used especially in analysis of very complex systems. In order to increase the applicability of the programs variance reduction techniques can be applied to speed up the calculation process. Variance reduction techniques have been studied and procedures for implementation of importance sampling are suggested.

Descriptors

A CODES; C CODES; COMPARATIVE EVALUATIONS; COMPUTERIZED SIMULATION; FAULT TREE ANALYSIS; FRACTURE MECHANICS; INDUSTRIAL PLANTS; M CODES; MONTE CARLO METHOD; N CODES; NUCLEAR POWER PLANTS; P CODES; PROBABILITY; RELIABILITY; RISK ASSESSMENT; S CODES; SAFETY ANALYSIS

Available on request from Risø Library, Risø National Laboratory, (Risø Bibliotek, Forskningscenter Risø), P.O. Box 49, DK-4000 Roskilde, Denmark.
Telephone 02 37 12 12, ext. 2262. Telex: 43116, Telefax: 02 36 06 09

RR 5000 99