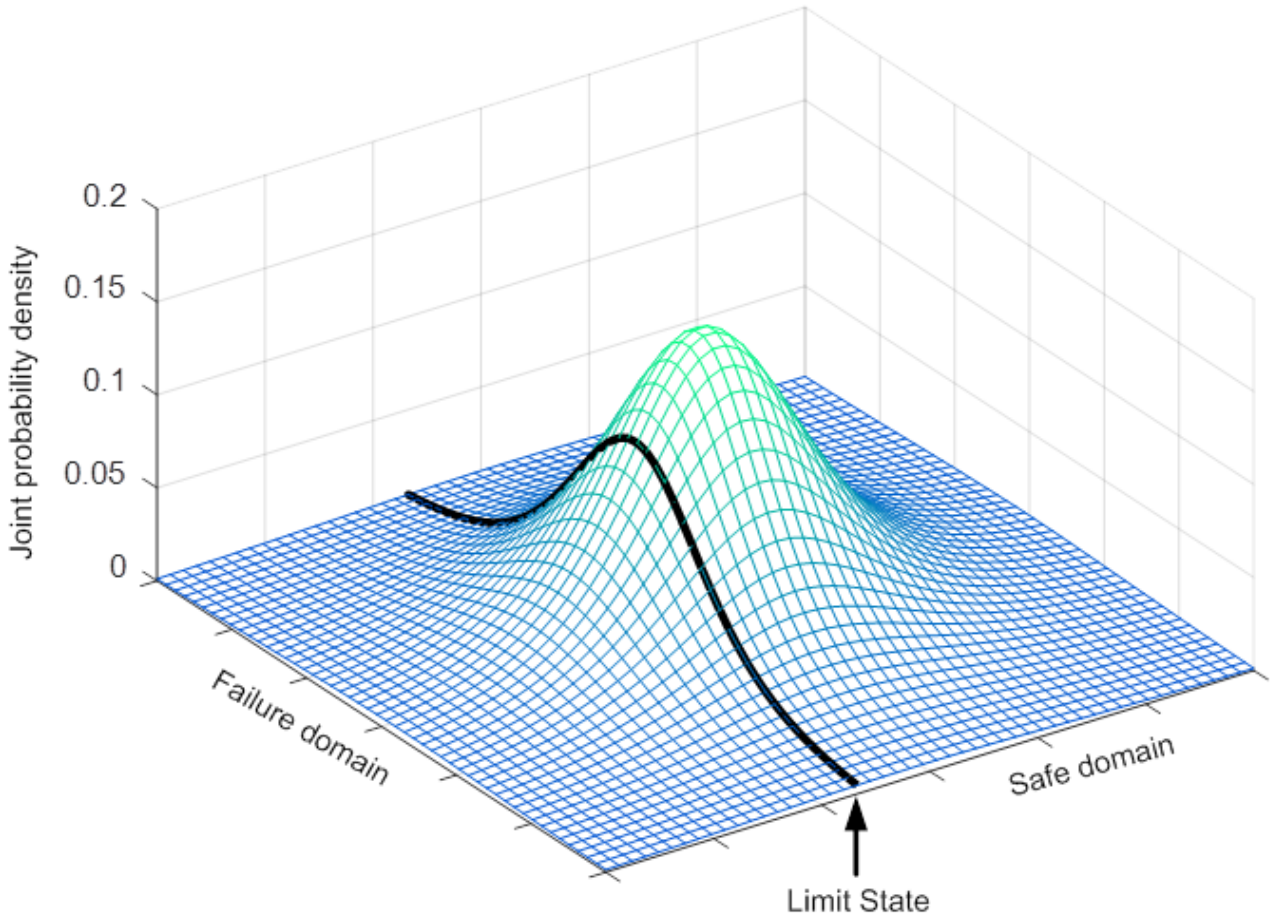
Probabilistic Design: Risk and Reliability Analysis in Civil Engineering

Lecture notes CIE4130

S.N. Jonkman
R.D.J.M. Steenbergen
O. Morales-Nápoles
A.C.W.M. Vrouwenvelder & J.K. Vrijling

Delft University of Technology



TU Delft Delft University of Technology

Challenge the future

PROBABILISTIC DESIGN: RISK AND RELIABILITY ANALYSIS IN CIVIL ENGINEERING

LECTURE NOTES CIE4130

Authors:

S.N. Jonkman
R.D.J.M. Steenbergen
O. Morales-Nápoles
A.C.W.M. Vrouwenvelder
J.K. Vrijling

Edited by:

M.H.G. Baas
M.A. van der Lugt
A. Kosters

With contributions from:

P.J.A.J.M. van Gelder
E.J.C. Dupuits
W.S. Jaeger
V. Vuik
T. Schweckendiek
R.B. Jongejan

Lecture notes (Fourth version November 2017)

Department of Hydraulic Engineering
Faculty of Civil Engineering and Geosciences

Delft University of Technology

TU Delft Delft University of Technology

Table of contents

PART I – GENERAL CONCEPTS

PART II – RELIABILITY ANALYSIS

Part I – General Concepts

Chapter "Introduction"
Author: S.N. Jonkman

Contents

1 Introduction

1.1 General

Almost all activities in life are characterized by some level of risk[1]. Examples of risk-bearing activities are riding a bike or car, boarding an airplane, or living below sea level behind flood defences. Particularly within the field of civil engineering risk and safety are key concepts that need to be taken into account explicitly in the design and management. Failures of systems such as dikes, buildings and other infrastructures are expected to occur with small probability, but can lead to large consequences. On the other hand, many engineering systems, such as transportation infrastructure, flood defences and buildings, provide many benefits to mankind.

In order to determine how safe a structure or system should be, an acceptable level of risk needs to be defined. The eventual decision about acceptable risk is predominantly a political one, but engineers can have an important role in the discussion and decision-making. They can provide information on failure probabilities and consequences (economic, life loss etc.) of a given system and highlight trade-offs between investments in safer systems and risk reduction (see chapter 3 for further details). Risk plays an important role in many current societal discussions. Examples are recent discussions related to the use of nuclear power or shale gas exploration. Both activities bring various benefits (energy generation) but also introduce additional risks to the population and environment. A systematic analysis of risks of (proposed) projects can help to inform the broader societal discussions.

[1] Risk refers to the combination of probability and consequences of undesired events (see chapter 3 for further definitions).

Risks have to be considered in the various stages of a project: during construction, regular use and decommissioning. Every phase has characteristic set of risks. For example, risks of injuries amongst workers and budget overruns during the construction phase, and small probabilities of failure during the regular lifetime.

For standard applications and systems that are frequently constructed, codes are available that define acceptable safety levels. For example, in the Eurocode for structures so-called target values for the failure probability for structures of different safety classes are given. Also, procedures are given to derive design loads and resistances are derived in such a way that the failure probability of a structure can be expected to be sufficiently small (see Figure 1.1).

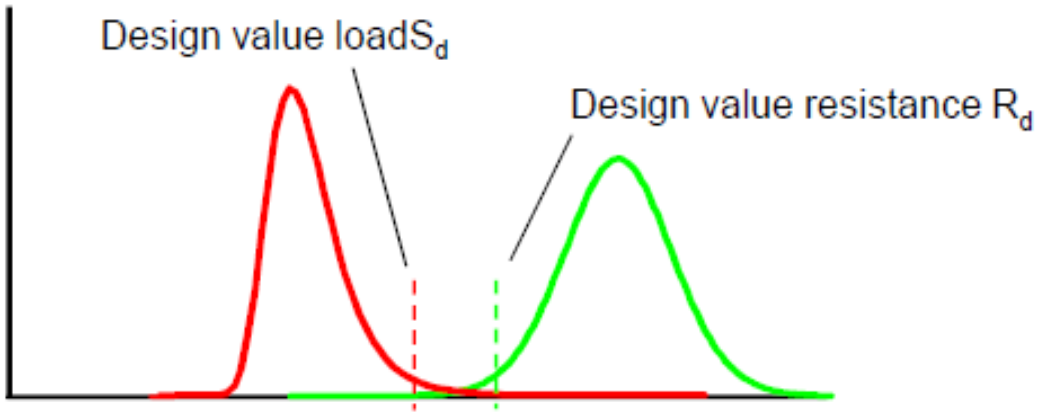


Figure 1.1: Probability density functions showing the variations in load (red, left) and resistance (green, right). The design load and resistance are chosen in such a way that a structure with a sufficiently low probability of failure can be designed. The failure probability is proportional to the overlapping area of the two curves. Codes and guidelines provide information on the load and resistance factors (γ's) that can be used to derive these values – see chapter * on level I techniques for more information.

However, for other applications, e.g. special structures or new applications, no standard codes or guidelines are available and a more explicit analysis of the reliability and risk of the system is required. An example from the past is the design of the Eastern Scheldt barrier. The acceptable probabilities of failure of the structure and non-closure of the gates were determined based on the acceptable risk of flooding of Zeeland. These probability values formed the basis for the so-called probabilistic design of the barrier in the 1970's.

There are various examples of relevant recent developments in the field of risk-based design in civil engineering the Netherlands. A first example is the discussion about the gas extraction in the north of the Netherlands which leads to additional risk for the population. A thorough analysis of the probability of earthquakes, the structural safety of various infrastructures (houses, dikes, hospitals, pipelines) and the resulting level of risk is required and ongoing. A second example, concerns the field of flood management. New safety standards for primary flood defences in the Netherlands have been introduced in the year 2014. These new safety standards are formulated as a tolerable failure probability of dike segments. Future dike reinforcements have to be designed

according to these new standards. This implies that advanced knowledge of probabilistic design of flood defences is needed in all these reinforcement projects.

In light of all these developments, it is crucial that civil engineers are able to understand and apply the concepts of risk and reliability in civil engineering. These lecture notes aim to provide (future) civil engineers with some of the basic techniques and concepts in the field.

1.2 Objectives

These lecture notes are part of the course CIE4130 probabilistic design at TU Delft. After completion of the course, students are expected to be able to:

- Perform a risk analysis of a (simplified) civil engineering system and evaluate the risk of a system – using relevant techniques such as fault and event trees, and criteria to support risk evaluation (e.g. economic optimization, individual and societal risk).
- Perform reliability calculations for at the element level, using relevant techniques, such as level III, II, I and analysis.
- Perform reliability calculations at the system level (using fundamental and Ditlevsen bounds for systems with various levels of correlation).
- Apply the main safety concepts of relevant design codes (Eurocode) and to derive design values for load and strength for civil engineering structures.

1.3 Structure of the lecture notes

These lecture notes have been organized in three parts. An overview of the structure can be found in Figure 1.2. The first part focuses on fundamentals and general principles. It introduces the objectives of the course and lecture notes (this chapter 1) and a general recap of probability theory required for the course (chapter 2). Also, general concepts for systems and risk analysis and risk evaluation are introduced in chapter 3.

Part II summarizes approaches for analysing the reliability of an element. After the introduction of some general concepts (chapter 4), different so-called levels for reliability analysis, level III, II and I and time dependence of loads are treated (chapters 5 - 8). Also methods for the reliability analysis of systems are summarized (chapter 9).

Part III focuses on application and implementation. Chapter 10 describes the implementation of reliability and safety in design codes for structures in civil engineering. The final chapter 11 presents applications of reliability analysis for project planning, maintenance and correctable systems.

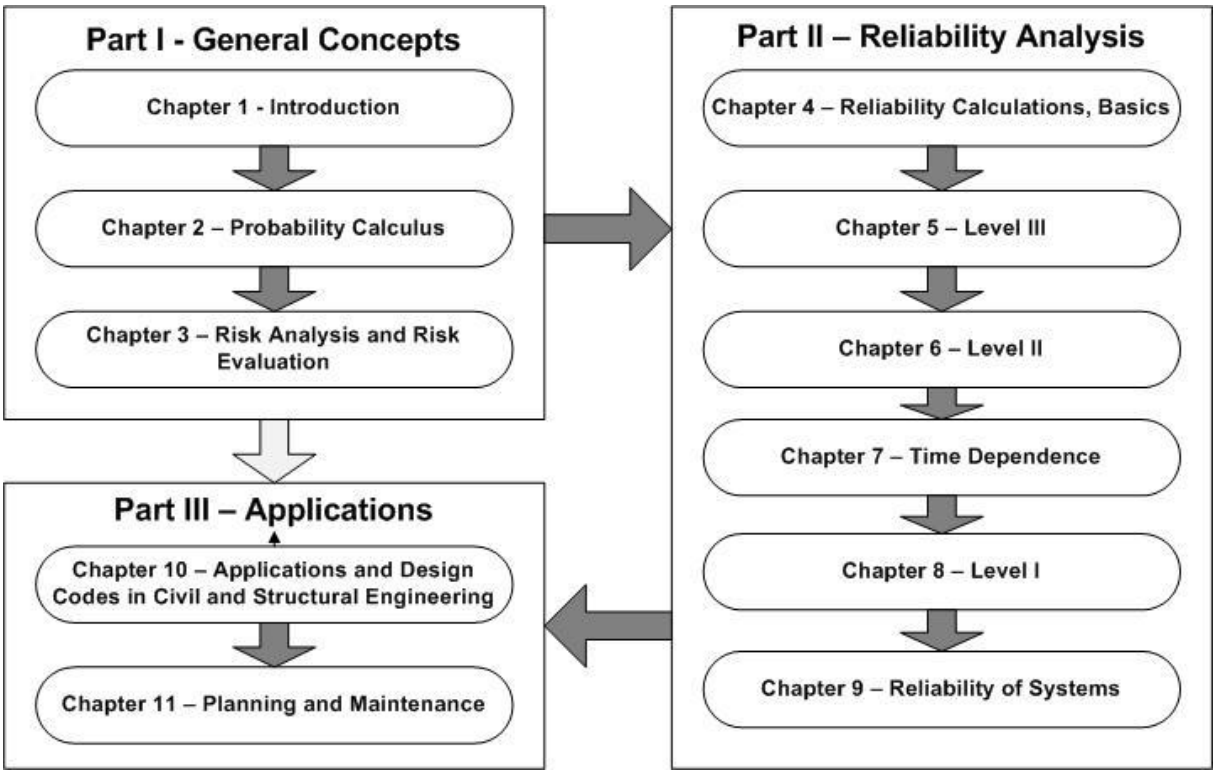


Figure 1.2: Structure of these lecture notes for the course CIE4130.

Previous work& basis of these lectures notes

Parts of these lecture notes are based on previous versions of lecture notes of this course.

- CUR (2015) CUR190: Probability in civil engineering. Version January 2015
- Vrouwenvelder A.C.W.M., Vrijling J.K. (1982) Dictaat b3, probabilistisch ontwerpen, TU Delft (lecture notes b3: Probabilistic design – in Dutch)
- Prof. dr. ir. Luc Taerwe; Prof. dr. ir. Robby Caspeele, Risk Analysis of Construction Processes, Department of Structural Engineering, Ghent University

Some parts of these documents have been used (with permission) as a basis for the current lecture notes. A special word of thank to the professors Taerwe and Caspeele from Ghent University for their kind cooperation.

About the draft version November 2017

These lecture notes are the fourth version of the new lecture notes for the course on probabilistic design (CIE4130) and they have been prepared in the summer and autumn of the year 2017. As such they are an update of the previous version of the lecture notes, i.e. the CUR 190 book. This is work in progress, and the lecture notes will continue to be updated in the coming year. Feedback and suggestions by students and readers is most welcome and can be sent to d.paprotny@tudelft.nl and s.n.jonkman@tudelft.nl.

Further reading

Additional information on various topics introduced in this course can be found in various sources. In every section a reference list is given. The references below cover a broad range of topics.

- Bedford, T. Cooke R. (2001) *Probabilistic risk analysis: foundations and methods*. Cambridge University Press, 2001.
- Benjamin, J. R., & Cornell, C. A. (1970). *Probability, Statistics and Decision for Civil Engineers*. New York: McGraw-Hill.
- Baecher, G. B., & Christian, J. T. (2003). *Reliability and Statistics in Geotechnical Engineering*. West Sussex, UK: Wiley.
- Faber M.H. (2012) Statistics and probability – in pursuit of engineering decision support. Springer Science and business media
- Faber M.H. (2001) lecture notes on risk and safety in civil engineering. ETH Zurich, available at: http://e-collection.library.ethz.ch/eserv/eth:25307/eth-25307-01.pdf
- Vrouwenvelder, A. C. W. M. (1997). The JCSS probabilistic model code. *Structural Safety*, *19*(3), 245–251. doi:10.1016/S0167-4730(97)00008-8
- Also, the previous lecture notes (CUR, 2015; Vrouwenvelder and Vrijling, 1982) provide useful information.

Chapter "Probability Calculus"
Author: O. Morales-Nápoles

Parts of this chapter are based on the old lecture notes B3 "Probabilistisch ontwerpen" (Vrouwenvelder en Vrijling, 1987)

Contents

2 Probability calculus

In this chapter a summary of basic probability theory and statistics which is important for these lectures is given. The axiomatic[1] structure of probability theory is briefly discussed. Also, the methods in which a single random variable is described as well as generalizations to multiple random variables are presented. The lecture notes are consistent with applications in reliability analysis. For more information, regarding probability theory we refer to Hoel et al (1971).

2.1 Axiomatic presentation

Kolmogorov (1956) is recognized as the first to formalize probability theory as an axiomatic discipline similar to algebra, geometry and other branches of mathematics. A formal derivation of probability theory as a mathematical discipline is out of the scope of these notes, however the main results concerning this derivation are presented. The main axioms of probability theory following Kolmogorov (1956) are:

Let Ω be a collection of *elementary events* ξ, ζ, $\eta\ldots$ and $\mathcal{F}$ a set of subsets of Ω; the elements of the set $\mathcal{F}$ will be called random events.

Axiom I:	$\mathcal{F}$ is a field[2] of sets
Axiom II:	$\mathcal{F}$ contains Ω
Axiom III:	To each set A in $\mathcal{F}$ a non-negative real number $P(A)$ is assigned; $P(A)$ is called the probability of event A
Axiom IV:	$P(\Omega) = 1$
Axiom V:	If A and B have no element in common, that is $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

A system of sets $\mathcal{F}$ together with a definite assignment of numbers $P(A)$ satisfying *Axiom I* till *Axiom V* is called a *probability field*.

The first and second axioms will not be discussed. *Axiom III* indicates that the probability P for a random event A is greater than or equal to 0. *Axiom IV* states that the probability of Ω is equal to 1. According to *Axiom V*, the probability of the occurrence of (A or B or both) is equal to the sum of the probabilities of A and B separately, provided that A and B are mutually exclusive. Notice that $\emptyset$ denotes the empty set. For an overview of all the symbols see the list of symbols and Figure 2.1.

Strictly speaking, the given axioms, especially *Axiom V*, are only valid as long as we restrict ourselves to outcomes with a finite number of outcomes. The extension to more general axioms (see Kolmogorov 1956) contains no viewpoints that are important for this course and therefore are excluded.

[1] An axiom is roughly a proposition that is self-evident and is accepted without a proof. In geometry for example, the concept of a "point" and a "line" are accepted as self-evident and are used in order to prove other more elaborate propositions.

[2] A system of sets is called a field if the sum, product and difference of two sets of the system also belong to the same system. Every non-empty field contains the empty set $\varnothing$.

In the above, the probability is introduced as a mathematical quantity. This says little about the interpretation of the notion of probability. This interpretation is essentially a philosophical problem. The main interpretations of probability are:

2.2 Interpretations of probability

Four of the main interpretations of probability are the classical, the logical, the frequentist and the subjective.

1. The classical interpretation of probability is attributed to the French scientist Laplace (Laplace), who defined probability as, "the number of favourable cases divided by the number of equi-possible cases". Examples from coin tossing and dice-throwing were used to illustrate what is meant by "equi-possible". This interpretation however has been criticized regarding the operational definition of "equi-possible" (for insurance companies death cases may not be equi-possible) and hence has fallen into disuse.
2. The logical interpretation was proposed in Keynes (1973). The idea was that conditional probability should be interpreted as partial entailment. This interpretation however is also in disuse and to a large extent forgotten.
3. The frequentist interpretation (Von Misses 1936) introduces probability as limiting relative frequencies in a "collective" or "random sequence". Where, roughly, a random sequence is one which passes all "recursive statistical tests". For example the frequency of 1's in a very large random sequence of 0's and 1's.
4. Finally the subjective interpretation. One of the main proponents is Savage (see Savage 1956) and the interpretation of probability is in terms of degree of belief of a subject. Different subjects can have different degrees of belief for one and the same event. Again, very loosely, the theory would state that if a subject prefers A to B, for rational subjects this would entail that he or she assigns greater probability to A than to B.

2.2.1 Propositions 1 to 7

Starting from the axioms from section 2.1, some propositions can be derived. These propositions may seem quite trivial and their proof superfluous. Still, these proofs, on the basis of the Venn diagrams from Figure 2.1 show the reasoning used in probability theory.

1. $P(\varnothing) = 0$ (2.1)

   Because $P(A \cap \varnothing) = P(\varnothing)$ from *Axiom V*: $P(A \cup \varnothing) = P(A) + P(\varnothing)$.
   Furthermore, $A \cup \varnothing = A$ resulting in $P(A \cup \varnothing) = P(A)$.
   Combining these results gives $P(\varnothing) = 0$.

2. $P(A) + P(\overline{A}) = 1$ (where $\overline{A}$ is the complement of A or not A) (2.2)

   Due to $A \cup \overline{A} = \Omega$ and $A \cap \overline{A} = \varnothing$ (see Figure 2.1e) it follows according to *Axiom IV* & *Axiom V*: $P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1$.

3. $0 \le P(A) \le 1$ (2.3)

   The left inequality follows directly from *Axiom III*. Again, due to *Axiom III* it follows that $P(\overline{A}) \ge 0$. The right inequality follows from proposition 2.

4. If $A \subset B$ then $P(A) \le P(B)$ (2.4)

Since $A \subset B$ is $(B-A) \cup A = B$ and $(B-A) \cap A = \varnothing$ (see Figure 2.1f). Applying *Axiom V*: $P(B) = P((B-A) \cup A) = P(B-A) + P(A)$. With $P(B-A) \geq 0$ (*Axiom III*) it follows $P(B) \geq P(A)$.

5. If $A \subset B$ then $P(A \cup B) = P(B)$ (2.5)

If $A \subset B$ holds then $A \cup B = B$ (see Figure 2.1f), from which the proposition follows directly.

6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (2.6)

$A \cup B = A \cup (B-A)$ is valid (see Figure 2.1g), as long as A and $(B-A)$ are mutually exclusive. Due to *Axiom V*: $P(A \cup B) = P(A) + P(B-A)$. Then $B = (A \cap B) \cup (B-A)$ is valid (see Figure 2.1h), with $(A \cap B)$ and $(B-A)$ also being mutually exclusive.

Due to *Axiom V*: $P(B) = P(A \cap B) + P(B-A)$. The substitution of $P(B-A)$ proves this proposition.

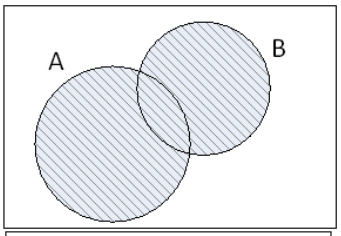
7. $Max\{P(A); P(B)\} \leq P(A \cup B) \leq P(A) + P(B)$ (2.7)

For the left inequality $P(A \cup B) = P(A) + P(B-A)$ is used (see the proof of proposition 6, first step). Since $P(B-A) \geq 0$ (*Axiom III*) it holds that $P(A \cup B) \geq P(A)$. In exactly the same way one may show that $P(A \cup B) \geq P(B)$.

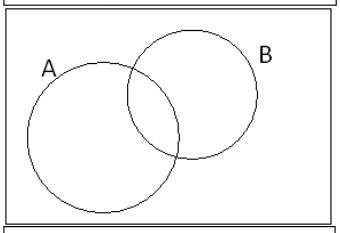
The right inequality follows from proposition (1-6) and the fact that $P(A \cap B) \geq 0$ (from *Axiom III*).

It is now interesting to relate the various propositions regarding $P(A \cup B)$. Proposition 6 gives an exact expression, but has the disadvantage that $P(A \cap B)$ has yet to be determined (see section 2.2). Proposition 7 shows an upper and a lower limit expressed in terms of $P(A)$ and $P(B)$. The upper limit is exact if A and B are mutually exclusive (see *Axiom V* or proposition 6); the lower limit is exact if A implies event B or vice versa (see proposition 5).

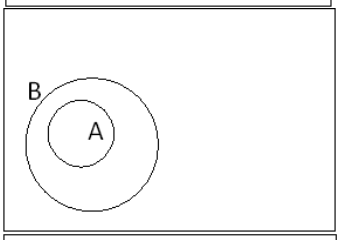
a. Union of event A and B: $A \cup B$;
(A or B)



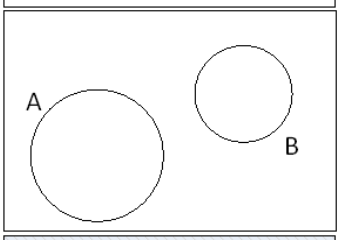
b. Intersection of A and B: $A \cap B$
(A and B)



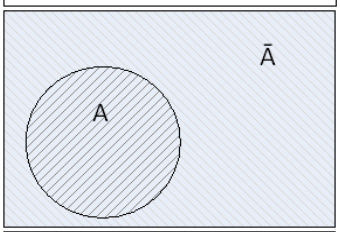
c. A is a subset of B: $A \subset B$
(A part of B or B contains A)



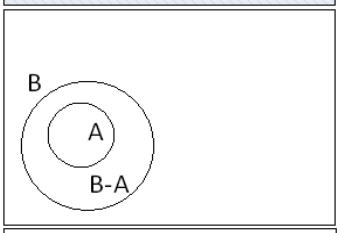
d. A and B are mutually exclusive: $A \cap B = \varnothing$



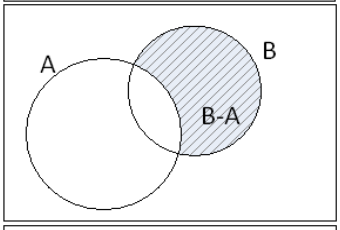
e. $A \cup \overline{A} = \Omega$,
$A \cap \overline{A} = \varnothing$



f. If $A \subset B$ then: $A \cup (B - A) = B$,
$A \cap (B - A) = \varnothing$,
$A \cup B = B$



g. $A \cup (B - A) = A \cup B$,
$A \cap (B - A) = \varnothing$



h. $(B - A) \cup (A \cap B) = B$,
$(B - A) \cap (A \cap B) = \varnothing$

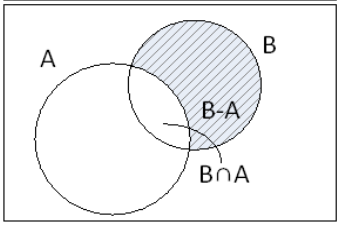


Figure 2.1. Venn Diagrams to use in the clarification of propositions 1 to 7.

Example 2.1

A certain dike would fail if the water level exceeds the height of the dike, or there is a partial structural failure and the dike does not have enough retaining power. Use the following definitions:

F = Failure of the dike

A = The event of a higher water level than the height of the dike

B = The dike observes partial structural failure and has insufficient retaining power.

Assume: $P(A) = 2 \cdot 10^{-4}$ and $P(B) = 3 \cdot 10^{-4}$. For the event of failure of the dike it states:

$$\begin{aligned} &Max\{P(A); P(B)\} \le P(A \cup B) \le P(A) + P(B) \\ &3 \cdot 10^{-4} \le P(F) \le 5 \cdot 10^{-4} \end{aligned} \tag{2.8}$$

Knowing these boundaries is sufficient for many problems. Note that the boundaries can get closer together if one of the probabilities is much larger than the other.

2.2.2 Propositions 8 to 11

Some of the propositions can now be extended to multiple events. A proof is only given for proposition 8; all other propositions can be proven via the same approach.

8. If $A_i \cap A_j = \varnothing$ for every $i \neq j$ then: $P(A_1 \cup A_2 ... \cup A_n) = \mathrm{P}(\mathrm{A}_1) + P(A_2) + ... + P(A_n)$ (2.9)

   This is an extension of *Axiom V*. For $n = 3$ the proposition is proven as follows: $P(A_1 \cup A_2 \cup A_3) = P((A_1 \cup A_2) \cup A_3) = P(A_1 \cup A_2) + P(\mathrm{A}_3) = P(A_1) + P(A_2) + P(A_3)$

   In this proof *Axiom V* is first applied to $(A_1 \cup A_2)$ and A_3 and then again to A_1 and A_2. The extension to $n = 4,\ 5$ etc. continues in the same manner.

9. If $A_i \cap A_j = \varnothing$ for every $i \neq j$ and $(A_1 \cup A_2 \cup ... \cup \mathrm{A}_n) = \Omega$ then: $P(A_1) + P(A_2) + ... + P(A_n) = 1$ (2.10)

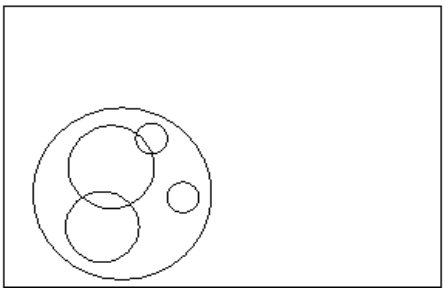
   This is an extension of proposition 2 and a special case of proposition 8.

10. If $A_i \subset A_j$ for every $i \neq j$ as an extension to proposition 5: $P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_j)$ (2.11)

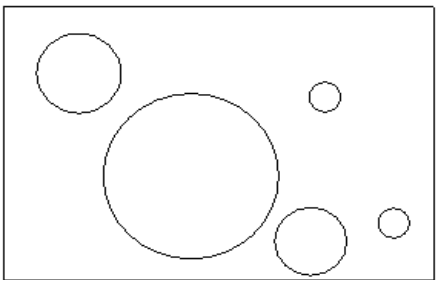
11. For an arbitrary A_i it applies: $Max\{P(A_i)\} \le P(A_1 \cup A_2 \cup ... \cup A_n) \le \sum P(A_i)$ (2.12)

    This is an extension of proposition 7. The probability that at least one of the events A_i occurs is greater than the probability that the most probable event occurs and smaller than the sum of the probabilities of the separate events. The lower bound occurs when one of the events includes every other event. The upper bound occurs when all events are mutually exclusive (see Figure 2.2). A special case is the case when every probability $P(A_i)$ is equal:

$$P(A_i) \le P(A_1 \cup A_2 \cup ... \cup \mathrm{A}_n) \le nP(A_i) \tag{2.13}$$



$P(A_1 \cup A_2 \cup ... \cup A_n) = \max\{P(A_i)\}$



$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum P(A_i)$

Figure 2.2: In proposition 11 the lower bound occurs when one of the events contains all other events, while the upper boundary occurs when all events are mutually exclusive

Example 2.2

Take a six-sided fair die. Let A_i, denote the event that a throw of the die yields the result i. The events A_i are mutually exclusive. Since the die is fair $P(A_i) = P(A_j)$, it follows on the basis of proposition 9 that $P(A_i) = 1/6$ for every i.

Example 2.3

The probability of throwing a 4, 5 or a 6 in one throw is, according to proposition 11: $1/6 \le P(A_4 \cup A_5 \cup A_6) \le 3/6$.

Because the event A_4, A_5 and A_6 are mutually exclusive, we know from proposition 8 that the right equality holds, that is $P(A_4 \cup A_5 \cup A_6) = 3/6$.

Example 2.4

The probability of throwing a 6 in one throw is $1/6$. Because of proposition 11 the probability of throwing at least one 6 in three throws is greater than $1/6$ but smaller than $3/6$. None of these bounds is accurate, why not?

Example 2.5

A construction will be built for a service life of 100 years. Each year the probability of failure is estimated at 10^{-5}. For the probability of failure in 100 years it applies: $10^{-5} \le P(\text{failure in 100 years}) \le 10^{-3}$. Notice that there are two orders of magnitude between these bounds.

In the next section some of the examples will be discussed further.

2.3 Conditional Probabilities – dependence and independence

An important concept in probability theory is the concept of conditional probability defined as follows:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{2.14}$$

Conditional probability (equation (2.14)) is obviously only defined if $P(B) \neq 0$. Before we discuss the interpretation of the conditional probability we first show it is most certainly a probability.
In other words: it will be shown that *Axiom III*, *Axiom IV* and *Axiom V* hold for the definition of conditional probability according to equation (2.14).

Proof.
Axiom III follows naturally. Because $P(A \cap B) > 0$ and $P(B) \neq 0$ it must follow that $P(A \mid B) > 0$.
Consider next *Axiom IV*; the proof that this axiom holds is as follows:

$$P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \tag{2.15}$$

In order to appreciate the statement above see Figure 2.4 (b). Finally it must be shown that equation (2.14) coincides with *Axiom V*. We use the Venn-diagram from Figure 2.3. The mutual exclusive events A_1 and A_2 and event B are shown. It can be concluded that:

$$\{(A_1 \cup A_2) \cap B\} = (A_1 \cap B) \cup (A_2 \cap B) \tag{2.16}$$

Since A_1 and A_2 are mutually exclusive, $(A_1 \cap B)$ and $(A_2 \cap B)$ are also mutually exclusive. This means:

$$\mathrm{P}\{(A_1 \cup A_2) \cap B\} = P(A_1 \cap B) + P(A_2 \cap B) \tag{2.17}$$

If we divide the left part as well as the right part of the equation above by $P(B) \neq 0$:

$$\frac{P\{(A_1 \cup A_2) \cap B\}}{P(B)} = \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} \tag{2.18}$$

Then, using the definition of conditional probability (equation (2.14)) :

$$\mathrm{P}\{(A_1 \cup A_2) \mid B\} = P(A_1 \mid B) + P(A_2 \mid \mathrm{B}) \tag{2.19}$$

This last equation is *Axiom V* for conditional probabilities which concludes the proof.

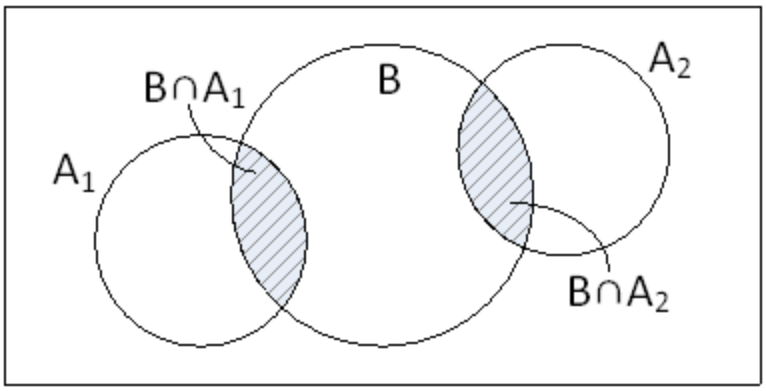
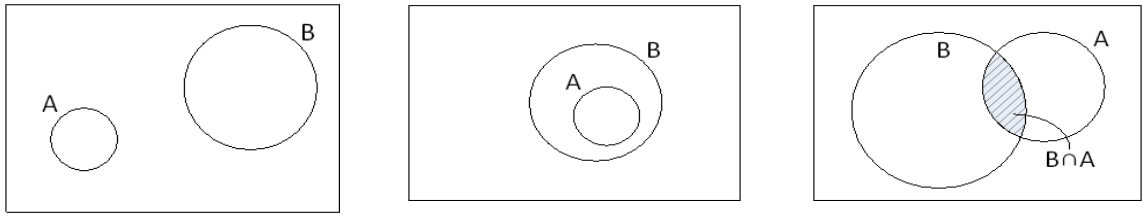


Figure 2.3: If A_1 and A_2 are mutually exclusive, then $(A_1 \cap B)$ and $(A_2 \cap B)$ are also mutually exclusive.



$$P(A|B) = 0 \qquad P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} \qquad P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(a) (b) (c)

Figure 2.4: The different possible computations of A, given B.

From the fact that the conditional probability coincides with the axioms of probability theory it follows that, all propositions for common probabilities can be rewritten for conditional probabilities. For example:

$$P(A|B) + P(\bar{A}|B) = 1 \tag{2.20}$$

$$P(A|B) \leq 1 \tag{2.21}$$

$$P(\varnothing|B) = 0 \tag{2.22}$$

And so on.

It is possible to interpret the conditional probability theory as common probability theory where the total sample space Ω is reduced to subarea B. An important argument for this interpretation is:

$$P(B|B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \tag{2.23}$$

In this case, event B plays the role of the event Ω which has a probability of occurrence of 1. This bring us to the interpretation of B as a "given". For some reason or another the event B can be considered to have occurred.

We now discuss conditional probability in terms of Figure 2.4:

- The conditional probability of an event A given an even B which is mutually exclusive with A is zero (Figure 2.4 a);
- The probability of an event A given B when A implies B is this probability of that event divided by $P(B)$. This procedure normalizes every probability on event A_i with respect to $A_i \subset B$ without compromising their ratios (Figure 2.4 b).
- If an event $(A \cap B) \neq \varnothing$, then A can be broken down in the separate events $(A \cap B)$ and $(\overline{A} \cap B)$. This leads to equation (2.14). The resulting probability $P(A|B)$ is usually called the probability of event A given event B (Figure 2.4 c).

2.4 Dependence and independence

In general $P(A|B)$ will not be the same as $P(A)$. In the special case that this holds we speak of independence. Two events, A and B are independent if:

$$P(A|B) = P(A) \tag{2.24}$$

A is independent of B if the fact that event B occurs does not influence the probability of occurrence of event A.

The rules for AND and OR probabilities in case of independence are as follows:

$$P(A \cap B) = P(A) \cdot P(B) \tag{2.25}$$

$$P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B) \tag{2.26}$$

Mathematically it is better to define (1-20) as a definition by itself instead of a result of independence. This definition is then easily extended to "n" number of events.

The events A_i are independent if:

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2)...P(A_n) \tag{2.27}$$

Calculating with independent events is mathematically trivial though could be computationally expensive. On the other hand the adequate assignment of probabilistic dependence usually proves to be a challenging task. Therefore, there is a tendency to work under the assumption of independence. In practise, independence, or rather dependence, is based on physical attributes of the events or variables under consideration. Assuming independence when in fact the contrary is true may lead to large inaccuracies. Modelling of dependence is thus becoming increasingly important. Decision makers have realized the importance of more accurate probabilistic models and the traditional tools of probability and statistics need to be extended with models for dependence if we want to meet the challenge placed by the questions of interest of decision makers.

Example 2.6

Denote by $A_1,...,A_6$ the event of showing a 1, … , 6 in a throw of an even die. Denote by E the event that the throw results in an even number. Then we have the following:

$$\begin{aligned} &P(A_1 \mid E) = P(A_3 \mid E) = P(A_5 \mid E) = 0 \\ &P(A_2 \mid E) = P(A_4 \mid E) = P(A_6 \mid E) = 1/3 \end{aligned} \tag{2.28}$$

Notice that, for example E and A_6 are not independent because $P(A_6 \mid E) \neq P(\mathrm{A}_6)$.

Example 2.7

Following with the notation from Example 2.6, one may verify that the events E = "throw is even" and A_{56} = "throw is a 5 or 6" are independent because:

$$\begin{aligned} &P(E) = P(A_2 \cup A_4 \cup A_6) = 1/2 \\ &P(A_{56}) = P(A_5 \cup A_6) = 1/3 \\ &P(\mathrm{E} \cap A_{56}) = P(A_6) = 1/6 \end{aligned} \tag{2.29}$$

So it holds that $P(E \cap A_{56}) = P(E)P(A_{56})$.

Example 2.8

In Example 2.4 a fair dice was thrown three times and required was the probability that at least one of those throws is a 6. Now, the assumption is made that those results are independent and the probability can be calculated. Notice that the assumption of independence results from the fact that the dice is fair. It is not true that probability theory indicates that these events should be independent. Assume:

$$\begin{aligned} &B_1 = \text{first throw is a 6} \\ &B_2 = \text{second throw is a 6} \\ &B_3 = \text{third throw is a 6} \\ &B = B_1 \cup B_2 \cup B_3 \end{aligned} \tag{2.30}$$

The example can be solved most easily by flipping the problem statement: What is the probability that none of the throws result in a 6? This can be described as:

$$\overline{B} = \overline{B}_1 \cap \overline{B}_2 \cap \overline{B}_3 \tag{2.31}$$

Because of the independency of $\overline{B}_i$ it holds (see equation (2.27)):

$$
\begin{aligned}
&P(\overline{B}) = P(\overline{B_1})P(\overline{B_2})P(\overline{B_3}) = P(\overline{B_i})^3 \\
&1 - P(B) = \{1 - P(B_i)\}^3 \\
&P(B) = 1 - \{1 - P(B_i)\}^3 \\
&P(B) = 1 - \{1 - 1/6\}^3 \approx 0.42
\end{aligned} \tag{2.32}
$$

See Example 2.4 for the upper and lower bound $1/6 \le P(B) \le 1/2$ and the

Appendix 2.1 Solution to Alternative examples & exercises

Example 2.9

Assume that the event of failure of the same structure as discussed in Example 2.5 is independent over the years. Let the failure in year i be noted as F_i:

$$P(\text{Failure in any year } i,\ i = 1...n) = P(\mathrm{F}_1 \cup \mathrm{F}_2 \cup ... \cup \mathrm{F}_n) \tag{2.33}$$

Similar to the reasoning in Example 2.8:

$$
\begin{aligned}
&P(\text{No failure in any year } i,\ i = 1...n) = \\
&P(\overline{F}_1 \cap \overline{F}_2 \cap ... \cap \overline{F}_n) = \\
&P(\overline{F}_1) \cdot P(\overline{F}_2) \cdot ... \cdot P(\overline{F}_n) = P(\overline{F}_i)^n = \{1 - P(F_i)\}^n \\
&P(\text{failure in any year } i,\ i = 1...n) = 1 - \{1 - P(F_i)\}^n
\end{aligned} \tag{2.34}
$$

With $P(F_i) = 10^{-5}$ and $n = 100$ it follows:

$$P(\text{Failure in year 100}) = 0.0009995 \tag{2.35}$$

Notice the similarity between the result and the upper bound of proposition 11, equation (2.13) $P(\text{Failure in year 100}) \approx 100 \times 10^{-5}$. This will be discussed later on.

2.4.1 Law of total probability

A widely used proposition within probability theory is the so called "Law of Total Probability" given by:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i) \tag{2.36}$$

Where all the B_i are mutually exclusive and together they constitute a so-called *partition* of a sample space (their union corresponds to the total sample space and they are mutually exclusive). To clarify $B_i \cap B_j = \varnothing$ for every $i \neq j$ and $B_i \cup B_2 \cup ... \cup B_n = \Omega$.

The proof lies on the basis of the following statements (see Figure 2.5):

$$
\begin{aligned}
&A = \{(A \cap B_1) \cup (A \cap B_2) \cup ... \cup (A \cap B_n)\} \\
&(A \cap B_i) \cap (A \cap B_j) = \varnothing \text{ for every } i \neq j
\end{aligned} \tag{2.37}
$$

It now follows easily:

$$
\begin{aligned}
&P(A) = P\{(A \cap B_1) \cup (A \cap B_2) \cup ... \cup (A \cap B_n)\} = \\
&P(A) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_n) = \\
&P(A) = \sum P(A \cap B_i) = \\
&P(A) = \sum P(A \mid B_i) P(B_i)
\end{aligned}
\tag{2.38}
$$

The final step follows on the basis of the definition of conditional probability (equation (2.14)).

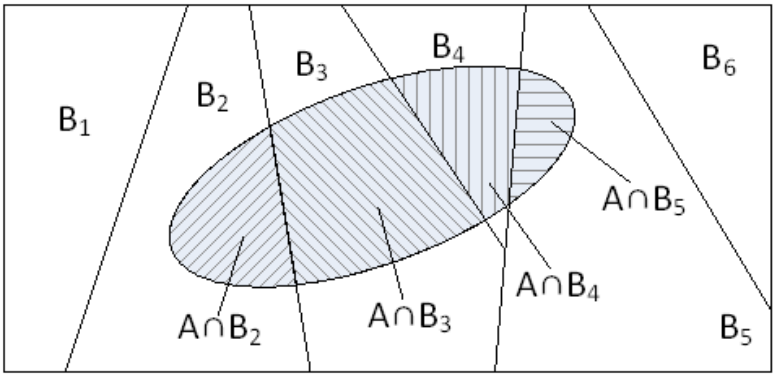


Figure 2.5 The events are $(A \cap B_i)$ mutually exclusive and their union is event A.

Example 2.10

A dice is thrown twice and required is the probability that the sum of two outcomes is equal to 9. Denote by B_i the event that "the outcome of the first throw is i" and by A the event that "the sum of the two outcomes is equal to 9". The *law of total probability* is applied:

$$P(A) = \sum P(A \mid B_i) P(B_i) \tag{2.39}$$

If the first throw is a 1 or a 2, the probability of the sum of the outcomes in each throw is equal to 9 is 0. If the first throw yields a 3, the second throw must be a 6, so $P(A \mid B_3) = 1/6 = P(B_3)$. The same reasoning applies with 4, 5 and 6. Then,

$$
\begin{aligned}
&P(A \mid B_1) = P(A \mid B_2) = 0 \\
&P(A \mid B_3) = P(A \mid B_4) = P(A \mid B_5) = P(A \mid B_6) = 1/6
\end{aligned}
\tag{2.40}
$$

This results in:

$$
\begin{aligned}
&P(A) = \sum P(A \mid B_i) P(B_i) = \\
&P(A) = \sum P(A \mid B_i) \cdot 1/6 = \\
&P(A) = 1/6 \cdot \sum P(A \mid B_i) = 4/36
\end{aligned}
\tag{2.41}
$$

It is of course also possible to calculate this result with other approaches.

2.4.2 Bayes' Theorem

This famous theorem bears the name of reverend Thomas Bayes (1701(?) to 1761) who was credited by Richard Price as the first one to formulate the theorem (see Bayes and Price (1763). Roughly, the theorem provides a technique of calculating a "posterior" distribution on the basis of a "prior". With the same notation used in section 2.4.1 we have:

$$P(A \cap B_i) = P(A) \cdot P(B_i \mid A) \text{ or}$$
$$P(A \cap B_i) = P(B_i) \cdot P(A \mid B_i) \tag{2.42}$$

Which by equality implies that:

$$P(B_i \mid A) = \frac{P(A \mid B_i) \cdot P(B_i)}{P(A)} \tag{2.43}$$

This equation is known as Bayes' Theorem. It can be perceived as a rule for "information processing". The probability of event B_i in the presence of event A is the "prior" probability of event B_i, that is $P(B_i)$, multiplied with the "likelihood" of A (which is $P(A \mid B_i)$) and then normalized by $P(A)$.

Richard Price (Bayes and Price (1763)) but also Laplace (1814) used this theorem to explain the probability that the sun will rise every day. Laplace writes: "Placing the most ancient epoch of history at five thousand years ago, or at 1826213 days, and the sun having risen constantly in the interval at each revolution of twenty-four hours, it is a bet of 1826214 to one that it will rise again tomorrow". Which would end up in a probability of about 0.999999452419348 that the sun will rise the next day. We will come back to this problem in coming sections.

2.5 Random Variables

In previous sections we have already been dealing with random variables. What follows is an extension and classification of these. Intuitively speaking random variables are those that are subject to variation which may be described by a probability distribution function. In other words they are functions that take elements from a sample space and assign them a number in the interval [0,1]. Roughly speaking random variables may be discrete or continuous. In this section we will briefly discuss some of the most important discrete and continuous models describing random variables.

2.5.1 Discrete random variables

These are variables that can take values on a finite outcome space. These have already been discussed in previous sections. Think for example on successive throws of a coin. The outcomes of throwing a coin may be only heads or tails. The outcomes of throwing a die may be $\{1,...,6\}$ and so on. Someone interested in traffic applications may think on counting the number of vehicles passing by a certain point in a particular period of time. Also, someone interested in cyber security may be interested in the number of possible cyber-attacks. In Physics the number of electrons emitted by a radioactive source may be relevant. Notice that these last examples refer also to a discrete outcome which in principle may be very large or even infinite!

2.5.2 Probability mass or density function (pdf) and cumulative distribution function (cdf) for discrete random variables

Discrete random variables may take particular values with positive probability. The function describing this probability is called probability mass function or probability density function (pdf). We will denote it as in equation (2.44).

$$f_X(x) = P(X = x) \tag{2.44}$$

The pdf of discrete random variables has many properties. Two of the most important ones are given next without a proof.

$$\begin{aligned} &0 \le f_X(x) \le 1 \text{ for every } x \\ &\sum_{x \in X} f_X(x) = 1 \end{aligned} \tag{2.45}$$

The cumulative distribution function is the function that described the probability that a random variable takes values less than or equal to a particular value. This function is given by equation (2.46).

$$F_X(x) = P(X \le x) = \sum_{x \in \{X \le x\}} f_X(x) \tag{2.46}$$

Notice that other probabilities of interest may be computed from the pdf or cdf. For example the probability that X lies between two possible outcomes:

$$P(x_1 \le X \le x_2) = \sum_{x \in \{x_1 \le X \le x_2\}} f_X(x) \tag{2.47}$$

Or the probability that X is larger than a particular value X :

$$P(X > x) = 1 - F_X(x) \tag{2.48}$$

2.5.3 Expected Value and Variance for discrete random variables.

Mathematical expectation or simply expectation is an important measure of central tendency for random variables. This measure is sometimes also referred to as the mean value of a random variable. The definition of expectation (when it is finite) is given by equation (2.49).

$$E(X) = \sum_{x \in X} x \cdot f_X(x) \tag{2.49}$$

Mathematical *expectation* has interesting properties that make it also useful as an operator. Some are given next without proof. Let X and Y be two discrete random variables with finite expectation and denote by g some real-valued function of the random variable X and let a and b denote two constants. Then

$$\begin{aligned} &(i) \quad E(g(X)) = \sum_{x \in X} g(x) \cdot f_X(x) \\ &(ii) \quad E(a) = a \\ &(iii) \quad E(aX + bY) = aE(X) + bE(Y) \\ &(iv) \text{ Suppose } P(X \le Y) = 1 \text{ then } E(X) \le E(Y) \text{ with strict equality if and only if } P(X = Y) = 1 \\ &(v) \quad |E(X)| \le E(|X|) \end{aligned} \tag{2.50}$$

Another important summary measure for discrete random variables is the *variance*. The variance is a measure of dispersion around the mean. The variance (just as mathematical expectation) is based on *moments*. Let X be a discrete random variable and let $r \ge 0$ be an integer. We say that X has a moment of order r if X^r has finite expectation and in that case $E(X^r)$ is defined as the r-th moment of X. The variance has to do with the relation between the first and second moments of a random variable.

$$\text{Var}(X) = E\left[\left(X - E(X)\right)^2\right] = E(X^2) - E(X)^2 \tag{2.51}$$

Exercise 2.11 shows that the last equality holds.

The variance of X is often denoted as σ_X^2 or $\sigma^2(X)$. The standard deviation is the squared root of the variance (σ_X or $\sigma(X)$). In these lecture notes we will use both notations for the variance and standard deviation.

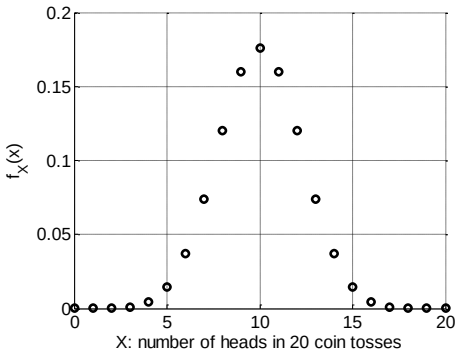
The variance is thus the expectation of the squared deviations between the random variable and its expectation or alternatively the difference between the second moment and the squared expected value a.k.a. mean. The variance also has important properties (like expectation) that makes it attractive as mathematical operator. We will come back to these later.

Example 2.11 The binomial distribution

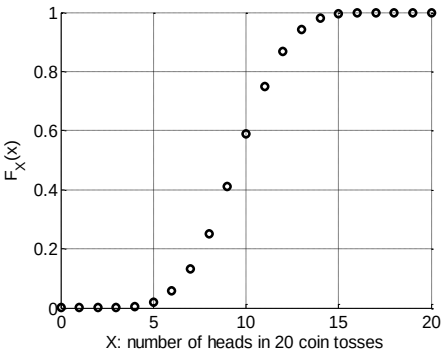
Consider an experiment which consist of n identical trials with two possible outcomes in each trial. The probability of success (let the number 1 denote success) in one trial is p and is constant across trials. Each trial is independent and the variable of interest is X, the number of successful outcomes trials in n trials. The random variable X follows a binomial distribution. The density function of a binomial random variable is:

$$f_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{2.52}$$

An example of such variable is tossing a coin (perhaps not a fair coin) where the probability of "success" (observing heads) is given by p. A possible sequence of 5 tosses is 1,1,1,1,1 that is, all outcomes are a success. Another one is 0,1,1,0,1, that is three successes in 5 tosses. Notice that in general there are $\binom{n}{x}$ possible sequences of x successes in n trials all with probability p.



(a) Probability density or mass function (pdf)



(b) Cumulative distribution function

Figure 2.6 Probability density function and Cumulative Distribution function for a binomial random variable with $p = 0.5$ and $n = 20$.

The mean and variance of a binomial random variable X are given by:

$$E(X) = np$$
$$\text{Var}(X) = np(1-p) \tag{2.53}$$

Example 2.12
Show that the expectation and variance of a binomial random variable are given by (2.53). Observe in Figure 2.6 that the expected number of heads in 20 tosses with a fair coin is 10.

Next, Table 2.1 presents a summary of the most common parametric discrete probability distributions. Next section is concerned with continuous random variables.

Table 2.1 Summary Parametric Discrete Probability Distributions

Distribution	Density function $f_X(x) = P(X = x)$	$E(X)$	$\text{Var}(X)$	Description
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0,1,...,n$	np	$np(1-p)$	Number of successful outcomes in n identical and independent trials with probability of success p
Geometric	$p(1-p)^{x-1}$ $x = 1,2,3,...$	$1/p$	$(1-p)/p^2$	Number of trial in which the first success is observed. The probability of success is p
Hypergeometric	$\dfrac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$ $x = 0,1,...,n$ if $n \le r$, $x = 0,1,...,\text{r}$ if $n > r$	$\dfrac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	Number of successes in a sample size of n from a population of size N with r possible successes in the population
Poisson	$\dfrac{\lambda^x e^{-x}}{x!}$ $x = 0,1,2,...$	λ	λ	Number of successes appearing in a fixed interval of time or space.
Negative Binomial	$\binom{x-1}{r-1} p^r (1-p)^{x-r}$ $x = r, r+1,...$	r/p	$r(1-p)/p^2$	Number of trial in which the r-th success is observed. The probability of success is p

2.5.4 Continuous random variables

Up to now we have considered random variables that can take a finite number of values. For example the number of successes in a certain number of trials or the number of objects observed in a certain time or space interval. However there are many situations (theoretical and applied) in which the variables under investigation take values in a continuous space. In civil engineering and related fields this is usually the case. Variables often have units of length, time, or mass for example.

2.5.5 Probability density function (pdf) and cumulative distribution function (cdf) for continuous random variables

Just like a discrete random variable, continuous random variable can be described by a distribution function that, by definition, describes the probability of observing different values of the random variable of interest:

$$F_X(x) = P(X \le x) \tag{2.54}$$

The distribution function is monotonic and non-decreasing from $F_X = 0$ at $X = -\infty$ until $F_X = 1$ at $X = +\infty$ (see Figure 2.7).

By differentiating the distribution function the probability density function is obtained:

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2.55}$$

To interpret the probability density function it is necessary to notice that the two events $X \le x$ and $x < X \le x + dx$ are mutually exclusive:

$$P(X \le x + dx) = P(X \le x) + P(x < X \le x + dx) \tag{2.56}$$

It follows that:

$$\begin{aligned} P(x < X \le x + dx) &= P(X \le x + dx) - P(X \le x) \\ P(x < X \le x + dx) &= F_X(x + dx) - F_X(x) = f_X(x)dx \end{aligned} \tag{2.57}$$

The probability density function multiplied with an infinitesimal interval yields the probability that the stochastic variable will take on a value within the interval (see Figure 2.7). Notice however that the density function itself for continuous variables is not a probability! For continuous random variables $P(X = x) = 0$

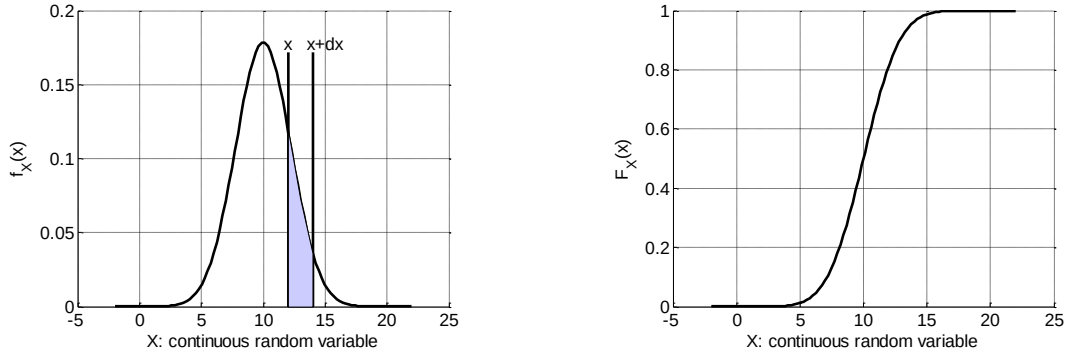
Some properties of the probability density function are presented without proof:

$$f_X(x) > 0 \text{ for all } x \tag{2.58}$$

$$P(X \le a) = \int_{-\infty}^{a} f_X(x)dx = F_X(a) \tag{2.59}$$

$$P(X \le \infty) = \int_{-\infty}^{+\infty} f_X(x)dx = F_X(\infty) = 1 \tag{2.60}$$

$$P(X \in A) = \int_A f_X(x)dx \tag{2.61}$$



(a) pdf (b) cdf

Figure 2.7: Probability density function (pdf) and cumulative distribution function (cdf) for an arbitrary X

2.5.6 Expected Value and Variance of continuous random variables

Equation (2.49) and the properties of the expected value shown in (2.50) have their counterpart for continuous random variables where summation is replaced by integration. Notice that the strict equality (iv) in (2.50) does not hold for continuous random variables. More generally if X is a random variable and $g(x)$ is a function of x, then the expected value of $g(x)$ is defined as:

$$E\{g(x)\} = \int_{-\infty}^{+\infty} g(x) f_X(x) dx \tag{2.62}$$

It can be easily seen that $E\{g(x)\}$ has the following properties:

$$\begin{aligned}
&(i) \quad E(a) = a \\
&(ii) \quad E(a \cdot g(\mathrm{x})) = a \cdot E[\mathrm{g(x)}] \\
&(iii) \quad E\{g(x) + h(x)\} = E\{g(x)\} + E\{h(x)\} \\
&(iv) \quad \text{Suppose } P(X \le Y) = 1 \text{ then } E(X) \le E(Y) \\
&(v) \quad |E(X)| \le E(|X|)
\end{aligned} \tag{2.63}$$

The most important examples of expected values are the mean and the variance of continuous random variables. Notice in equation (2.65) that the variance is the expected value of the "squared deviations" of the random variable from their mean.

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) dx \tag{2.64}$$

$$\mathrm{Var}(X) = E[\{X - E(X)\}^2] = \int_{-\infty}^{+\infty} \{x - E(X)\}^2 f_X(x) dx \tag{2.65}$$

Similarly as with discrete random variables, the mean or expectation of X is a measure for the "central location" of a random variable. The square root of the variance, the standard deviation $\sigma(X)$ or σ_X, is a measure for the dispersion of a random variable. The coefficient of variation $V(X)$ or V_X, is also widely used:

$$V(X) = \frac{\sigma(X)}{E(X)} \tag{2.66}$$

The coefficient of variation is a relative measure for the dispersion of the variable around the mean. Notice that the mean and variance are expected values but the standard deviation and the coefficient of variation are not.

2.5.7 Linear Transformations

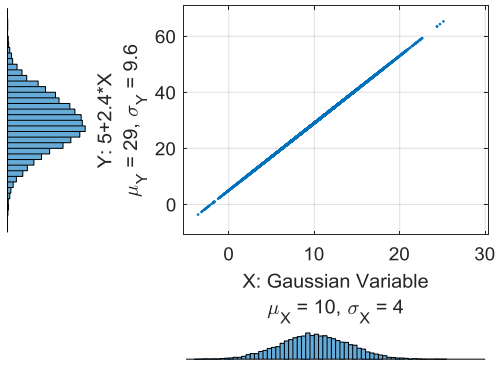
If X is a random variable, then $Y = g(X)$ is also a random variable. Given de probability density function of X the probability density function of Y can be calculated. In most cases it is only necessary to determine the mean and standard deviation of Y. For a linear function $g(X)$ it can be done exactly, for a nonlinear function it can be estimated. For a linear function of X ($Y = a \cdot X + b$) we have:

$$E(Y) = E(a \cdot X + b) = a \cdot E(X) + b \tag{2.67}$$

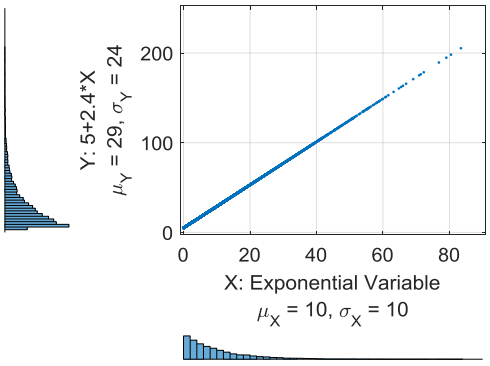
After some algebraic calculations one may show that:

$$\mathrm{Var}(Y) = a^2 \mathrm{Var}(X) \tag{2.68}$$

We will come back to (2.68) later when discussing dependence. Notice that these results are obtained solely by making use of the definition and properties of the expected values. Examples are given in Figure 2.8. Linear transformation of Gaussian random variables are still Gaussian (Figure 2.8a). This holds in general for different distributions. In Figure 2.8b a linear transformation of an exponential variable is presented. For the exponential distribution the mean and standard deviation are equal (see table). In this case the mean and standard deviation of Y are different because Y becomes a two parameter exponential distribution (shifted by the intercept term in the linear transformation equal to 5).



a) Linear transformation of Gaussian variable



b) Linear transformation of exponential random variable

Figure 2.8 Linear transformations of random variables.

2.5.8 Non-Linear Transformations

An example of a non-linear transformation of a log normal random variable is given in Figure 2.9.

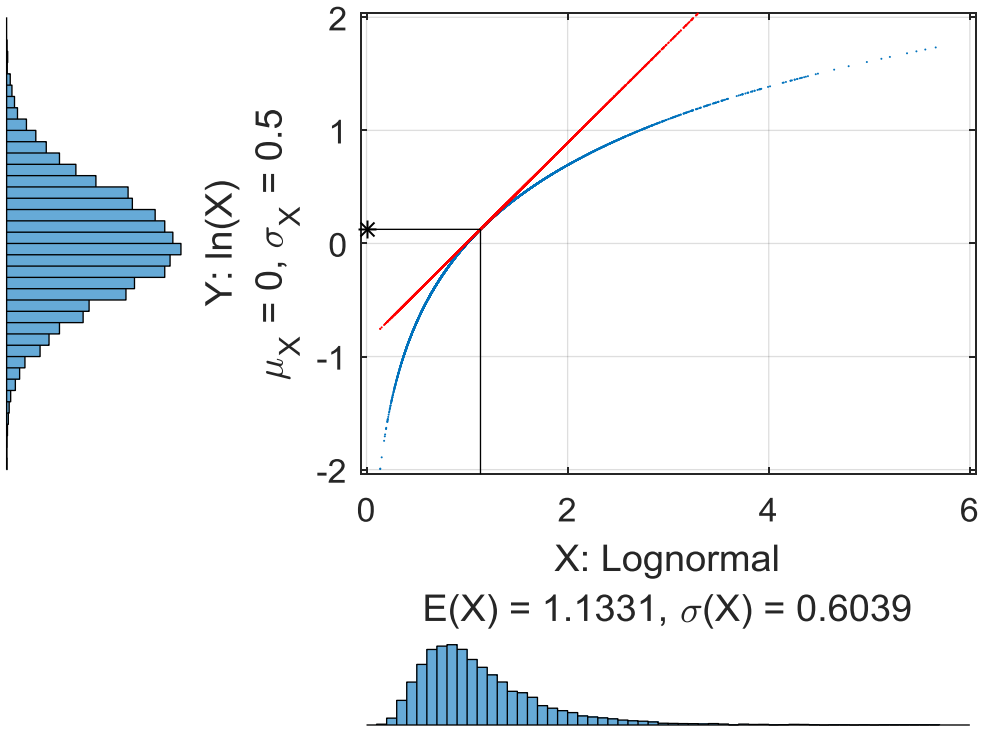


Figure 2.9 Non-linear transformation of a lognormal random variable.

In Figure 2.9 we observe that due to the non-linearity of ln(X) the parametric family of the density of X is not preserved. This means that in practice knowing the distribution function of X (for example concrete strength, wind speed, etc) does not immediately gives us an idea of $Y = g(X)$ (for example a bending moment following from a model).

One common method to approximate non-linear functions is through Taylor-polynomial. A tailor expansion evaluated in the point x_0 is presented in equation (2.69) where $\frac{\partial^n g(x_0)}{\partial x^n}$ denotes the n-th order derivative of g with respect to x:

$$g(\mathrm{X}) \approx g(x_0) + \frac{\partial g(x_0)}{\partial x}(\mathrm{X}\text{-}x_0) + \frac{1}{2!}\frac{\partial^2 g(x_0)}{\partial x^2}(\mathrm{X}\text{-}x_0)^2 + \ldots + \frac{1}{n!}\frac{\partial^n g(x_0)}{\partial x^n}(\mathrm{X}-x_0)^n \tag{2.69}$$

When a function is approximated by the first two terms of the Taylor-polynomial one speaks of a linear function. It is of particular interest where to choose exactly the point x_0 in the linearization. In Figure 2.9 the linearization point $x_0 = E(X)$ is chosen. This is called the mean value approximation. In the case of $Y = \ln(X)$ we have the following expression for the mean value approximation:

$$\mathrm{Y}(X) \approx \ln\big(E(X)\big) + \frac{1}{X}\big(X - E(X)\big) \tag{2.70}$$

Which corresponds to the line in Figure 2.9. Observe that while linearizing in the mean value gives a good approximation of the non-linear function ln(X) around the mean and therefore

$$E(Y) = 0 \approx \ln\big(E(X)\big) + \frac{1}{E(X)}\big(E(X) - E(X)\big) = \ln(1.1331) = 0.125 \tag{2.71}$$

However with the mean value approximation, the "tail of the distribution" is badly approximated. The line representing $g(X)$ deviates significantly from ln(X) for values of X away from the mean; only in the case of a very small $\sigma(X)$ is the mean value approximation more accurate. If we are particularly interested to get a good representation of Y in the tales (e.g. for calculating failure probabilities) we should use a linearization in x_0 located in the tail region. We will come back on this in the Level II reliability calculations. By a similar argument as in (2.68) it may be shown that $\mathrm{Var(Y)} \approx \mathrm{Var}(g(X)) \approx \left(\frac{\partial g(E(X))}{\partial x}\right)^2 \sigma_X^2$ which for the case of Y = ln(X) with the example from Figure 2.9 gives $\mathrm{Var(Y)} = 0.5 \approx \left(\frac{1}{E(X)}\right)^2 \sigma_X^2 = \left(\frac{1}{1.1331}\right)^2 0.6039^2 = 0.2840$.

2.6 Continuous Parametric Distributions

Describing a random variable through a parametric distribution may have many advantages. For example, investigating probabilities not observed in a sample becomes possible. In many cases operations with random variables become traceable analytically when parametric distribution describe them. In this section some of the commonly used parametric distributions used in civil engineering are introduced.

2.6.1 The Gaussian or Normal Distribution

One of the most widely used distribution functions is the Gaussian (Normal) distribution. Figure 2.7 is in fact an example of a Gaussian distribution. The probability density function is described by:

$$f_X(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}} \tag{2.72}$$

Here μ denotes the mean and σ denotes the standard deviation. A Normally distributed variable with a mean of 0 and a standard deviation of 1 is the *Standard Normal* distributed variable, usually denoted by Z. An arbitrary variable X can be rewritten as:

$$X = \mu_X + Z \cdot \sigma_X \tag{2.73}$$

The distribution function of a Gaussian distribution has no closed form, however it has been studied extensively. Almost every computer software and scientific calculator may return values for a Normal Distribution. Another option is to approximate the probabilities of interest by searching in tables (see for example

Table 2.2). For small probabilities it is usually sufficient to use the following approximation:

$$P(Z < a) = \Phi_N(z) = \frac{1}{-z\sqrt{2\pi}}e^{\left\{-\frac{z^2}{2}\right\}} \tag{2.74}$$

This approximation holds for $z < -2$.

Example 2.13

A certain material has a mean strength of $\mu(x) = 30MPa$ with a coefficient of variation of 13.3%. What is probability that the strength of the material is less or equal to $20\ MPa$ when assuming the variables are Gaussian distributed?

The standard deviation is $\sigma(X) = \mu(X) \cdot V(X) \approx 4MPa$ so it holds from equation (2.73):

$$P(X \le 20) = P(\mu(X) + Z\sigma(X) \le 20) \quad \text{and}$$

$$P(X \le 20) = P(30 + 4Z \le 20) = P(Z \le -2.5) \approx 0.62 \times 10^{-2} \tag{2.75}$$

This result is obtained using table 2.1.

Approximation equation (2.74) yields: $P(z \le -2.5) = 0.70 \times 10^{-2}$. For some applications this approximation is sufficient. If the probability decreases, the approximation error also decreases.

The Gaussian distribution emerges when a large number of independent random variables, from which none dominates any other, are created disregarding the output distributions of these variables (see Figure 2.12 for an example). This result is known as the central limit theorem. A consequence of this proposition is that the sum of two Gaussian distributed variables is also a Gaussian distributed variable.

2.6.2 The uniform distribution

The uniform distributions assigns equal density to all outcomes within an interval. This distribution is often used in applications to generate random numbers from other distributions. It is also useful as a "first guess" if no other information about a random variable *X* is known, other than that it is in [*a*, *b*]. The density and cumulative distribution functions of a uniform random variable are respectively:

$$f_X(x) \begin{cases} \dfrac{1}{b-a} & a \le x < b \\ 0 & otherwise \end{cases} \tag{2.76}$$

And:

$$F_X(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases} \tag{2.77}$$

Plots of the pdf and cdf for a uniform [0,1] random variable are presented in Figure 2.10.

Table 2.2: Distribution of a Gaussian (Normal) distribution

z	$\Phi_N(z)$	z	$\Phi_N(z)$	z	$\Phi_N(z)$
0.0	0.5				
-0.1	0.46	-2.1	0.018	-4.1	$0.21\cdot10^{-4}$
-0.2	0.42	-2.2	0.014	-4.2	$0.13\cdot10^{-4}$
-0.3	0.38	-2.3	0.011	-4.3	$0.85\cdot10^{-5}$
-0.4	0.34	-2.4	0.0082	-4.4	$0.54\cdot10^{-5}$
-0.5	0.31	-2.5	0.0062	-4.5	$0.34\cdot10^{-5}$
-0.6	0.27	-2.6	0.0047	-4.6	$0.21\cdot10^{-5}$
-0.7	0.24	-2.7	0.0035	-4.7	$0.13\cdot10^{-5}$
-0.8	0.21	-2.8	0.0026	-4.8	$0.79\cdot10^{-6}$
-0.9	0.18	-2.9	0.0019	-4.9	$0.48\cdot10^{-6}$
-1.0	0.16	-3.0	0.0013	-5.0	$0.29\cdot10^{-6}$
-1.1	0.14	-3.1	0.00097	-5.1	$0.17\cdot10^{-6}$
-1.2	0.12	-3.2	0.00069	-5.2	$0.10\cdot10^{-6}$
-1.3	0.10	-3.3	0.00048	-5.3	$0.58\cdot10^{-7}$
-1.4	0.080	-3.4	0.00034	-5.4	$0.33\cdot10^{-7}$
-1.5	0.067	-3.5	0.00023	-5.5	$0.19\cdot10^{-7}$
-1.6	0.055	-3.6	0.00016	-5.6	$0.11\cdot10^{-7}$
-1.7	0.045	-3.7	0.00011	-5.7	$0.60\cdot10^{-8}$
-1.8	0.036	-3.8	$0.72\cdot10^{-4}$	-5.8	$0.33\cdot10^{-8}$
-1.9	0.029	-3.9	$0.48\cdot10^{-4}$	-5.9	$0.18\cdot10^{-8}$
-2.0	0.023	-4.0	$0.32\cdot10^{-4}$	-6.0	$0.99\cdot10^{-9}$

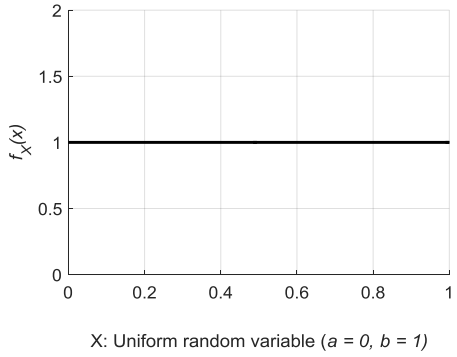
$$\Phi_N(z) = P(u < z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{2.78}$$

For $z > 0.0$ it follows: $\Phi_N(z) = 1 - \Phi_N(-z)$

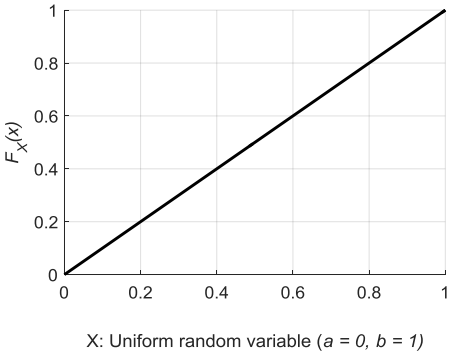
Example 2.14

X is Gaussian distributed with $\mu = 10$ $\sigma = 2$; what is the probability of $X < 14$?

$$P(X < 14) = P(10 + 2Z < 14) = P(Z < 2) = \Phi_N(2) =$$
$$P(X < 14) = 1 - \Phi(-2) = 0.977 \tag{2.79}$$



(c) pdf



(d) cdf

Figure 2.10: Probability density function (pdf) and cumulative distribution function (cdf) for an uniformly distributed random variable in [0,1]

The expectation and variance of this random variable are given by:

$$E(X) = \frac{a+b}{2} \tag{2.80}$$

$$\mathrm{Var}(X) = \frac{(b-a)^2}{12} \tag{2.81}$$

2.6.3 The triangular distribution

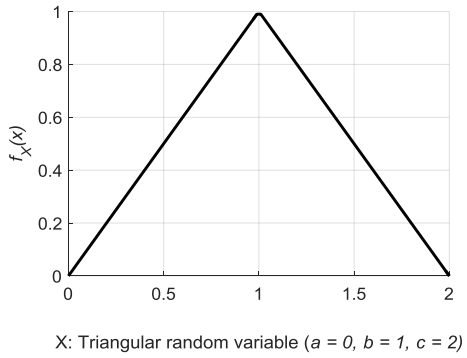
The triangular distribution is a three parameter continuous distribution. Like the uniform distribution it is defined on a fix interval [*a*, *c*], however it has an extra parameter *b* which is the single mode (value where the density is highest) of the distribution. The density and cumulative distribution functions are given by:

$$f_X(x)\begin{cases} \dfrac{2(x-a)}{(c-a)(b-a)} & a \le x \le b \\ \dfrac{2(\mathrm{c}-x)}{(c-a)(\mathrm{c}-\mathrm{b})} & b < x \le c \\ 0 & \text{otherwise} \end{cases} \tag{2.82}$$

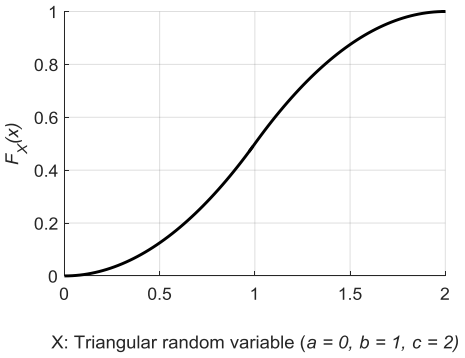
And:

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{(x-a)^2}{(c-a)(b-a)} & a \le x \le b \\ 1 - \dfrac{(c-x)^2}{(c-a)(\mathrm{c}-b)} & b < x \le c \\ 1 & x > c \end{cases} \tag{2.83}$$

Plots of the pdf and cdf for a triangular random variable with parameters $a = 0$, $b = 1$ and $c = 2$ are presented in Figure 2.11.



(a) pdf



(b) cdf

Figure 2.11: Probability density function (pdf) and cumulative distribution function (cdf) for an arbitrary triangular random variable

The expectation and variance of this random variable are given by:

$$E(X) = \frac{a+b+c}{3} \tag{2.84}$$

$$\mathrm{Var}(X) = \frac{a^2+b^2+c^2-ab-ac-bc}{18} \tag{2.85}$$

2.6.4 Sums of continuous random variables

Let X and Y be two independent continuous random variables with density functions $f(x)$ and $g(y)$, respectively. Assume that both $f(x)$ and $g(y)$ are defined for all real numbers. Then the convolution $(f \otimes g)(z)$ of f and g is the function given by

$$\begin{aligned} f \otimes g &= \int f(z-y)g(y)dy \\ &= \int f(y)g(z-y)dy \end{aligned} \tag{2.86}$$

Example 2.15

Let *X* and *Y* be uniformly distributed random variables on [0,1]. Required is the distribution of *Z = X + Y* their sum. Recall that the density of a uniform random variable on [*a, b*] is given by $1/(a-b)$ Then we have:

$$f_X(x) = f_Y(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \textit{otherwise} \end{cases} \tag{2.87}$$

The density of the sum is $f_Z(z) = (f_X \otimes f_Y)(z) = \int f_X(z-y) f_Y(y) dy$ where $f_Y(y) = 1$ for $y \in [0,1]$ and hence $f_Z(z) = \int f_X(z-y) dy$. The integrand in the last expression is 1 if $0 \le z-y \le 1$. If $0 \le z \le 1$ the integrand has value 1 on the set $0 \le y \le z$ and hence $f_Z(z) = \int_0^z dy = z$. If $1 \le z \le 2$ the integrand has value 1 on the set $z-1 \le y \le 1$ and hence $f_Z(z) = \int_{z-1}^1 dy = 1-(z-1) = 2-z$. The density is thus:

$$f_Z(z) = \begin{cases} z, & 0 \le z \le 1 \\ 2-z, & 1 \le z \le 2 \\ 0, & \textit{otherwise} \end{cases} \tag{2.88}$$

To visually appreciate this example, see Figure 2.12. Notice that after summing up 3 uniform [0,1] variables the distribution of the sum starts converging to a Gaussian distribution.

2.6.5 Lognormal Distribution

If *X* has a lognormal distribution then $Y = \ln(X)$ is Gaussian distributed. In other words, $Y = e^X$ with *Y* being Gaussian distributed. The lognormal distribution is the result of a nonlinear transformation of a Gaussian distribution. If the mean and standard deviation of *Y* are known, the corresponding values for *X* can be determined:

$$E(X) = \exp\{E(Y) + \frac{1}{2}\sigma^2(Y)\} \approx \exp\{E(Y)\} \tag{2.89}$$

$$\sigma(X) = E(X)\sqrt{\exp\{\sigma^2(Y)\} - 1} \approx E(X)\sigma(Y) \tag{2.90}$$

The approximation equations (2.89) and (2.90) hold for values of the standard deviation much smaller (less) than the mean $(\sigma(Y) << E(Y))$. The probability density function is then defined as:

$$f_X(x) = \frac{1}{\sigma_Y x\sqrt{2\pi}} e^{\left\{-\frac{(\ln(x)-\mu_Y)^2}{2\sigma_Y{}^2}\right\}} \tag{2.91}$$

Where $\mu_Y = E(Y)$. The pdf and cdf of a lognormal distribution with $\mu_Y = 0$ and $\sigma_Y^2 = 1$ are presented in Figure 2.13. The cumulative distribution function cannot be written explicitly but has to be determined on the basis of the tables for the Gaussian distribution or computed numerically. The lognormal distribution is usually used for variables that, because of physical limitations,

cannot take on negative values. The Gaussian distribution is in such cases less applicable although with small coefficient of variation the difference is marginal.

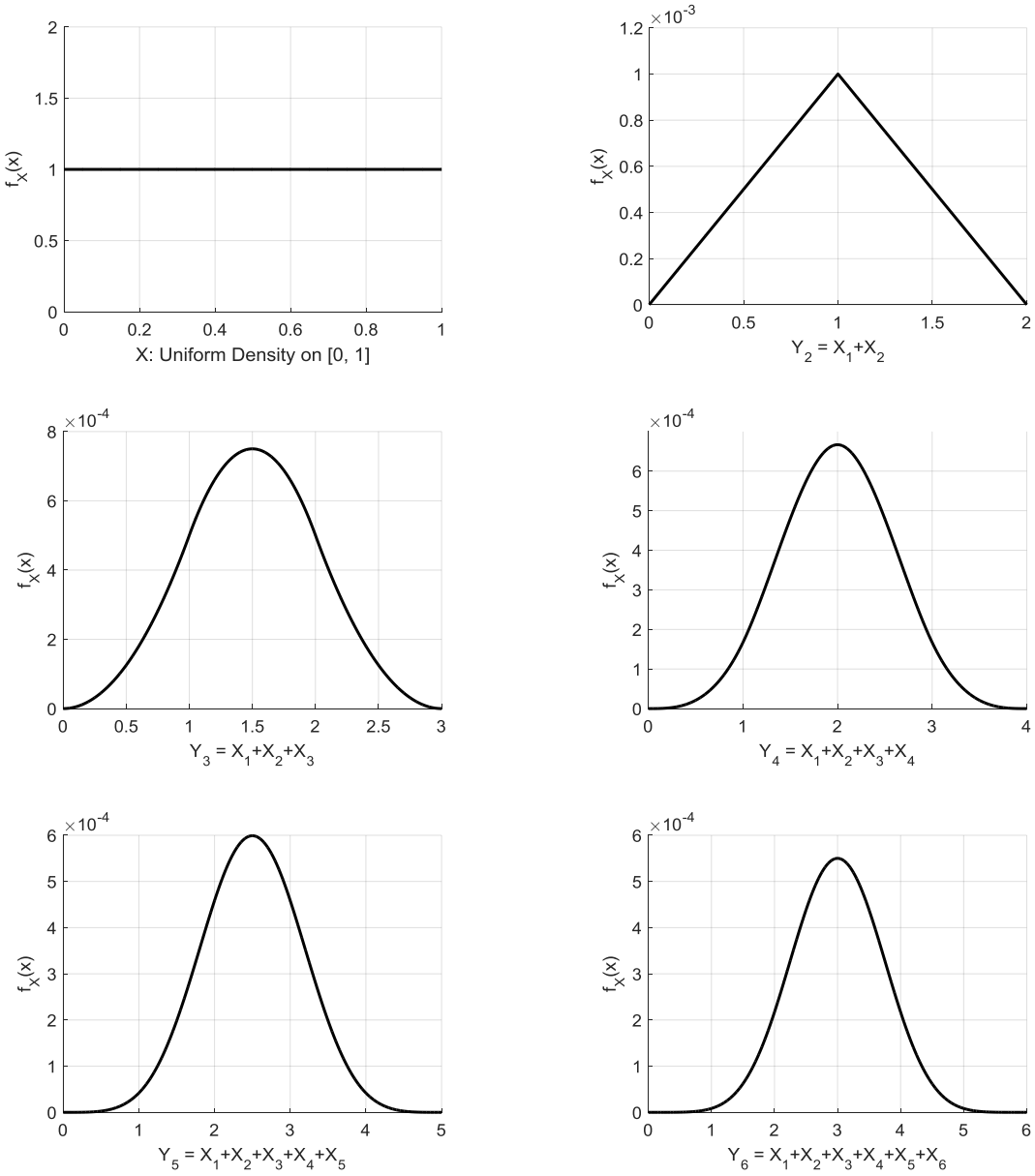


Figure 2.12: Illustration of the central limit theorem: The variables X_i are independent and have a uniform distribution on [0,1]. The sum of only four of these variables already gives a distribution (except from the tails) that is similar to the Gaussian distribution.

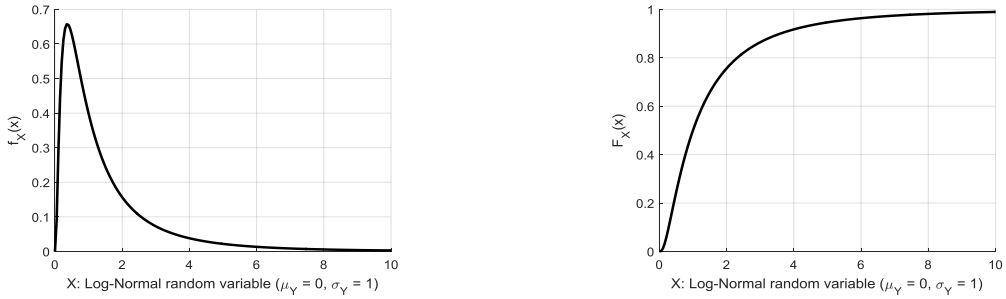


Figure 2.13: Lognormal pdf and cdf.

Finally, in the same way that the Gaussian distribution is the result of the sum of a large amount of variables, the lognormal distribution is the result of a multiplication of a large number of random variables.

Example 2.16

Consider the same problem in example 2.13, now with X being lognormally distributed. The material has a strength of $\mu(x) = 30MPa$ with a coefficient of variation of 13.3%. What is probability that the strength of the material is less than or equal to $20\ MPa$?

Given are $E(X)$ and $\sigma(X)$ from which $E(Y)$ and $\sigma(Y)$ may be determined by applying the inverse of equations (2.89) and (2.90).

First, the approximation equations yield $E(\mathrm{Y}) \approx \ln(30) \approx 3.40$ and $\sigma(\mathrm{Y}) \approx 4/30 \approx 0.133$. Substituting in the exact equations yield:

$$\begin{aligned} \mathrm{Var}(Y) &= \ln\left\{1 + \frac{\mathrm{Var}(X)}{E(X)^2}\right\} \approx (0.133)^2 \\ E(Y) &= \ln(E(X)) - \frac{1}{2}\mathrm{Var}(Y) \approx 3.39 \end{aligned} \tag{2.92}$$

The approximations are in this case close enough. Now, the probability of exceedance is calculated:

$$\begin{aligned} &P(X < 20) = P\{\exp(Y) < 20\} \\ &P(X < 20) = P\{Y < \ln(20)\} \\ &P(X < 20) = P\{3.39 + 0.133Z < \ln(20)\} \\ &P(X < 20) = P\{Z < (3.00 - 3.39)/0.13\} \\ &P(X < 20) = P\{Z < -3.0\} = 0.0013 \end{aligned} \tag{2.93}$$

Note that the probability of exceedance with the Gaussian distribution in example 2.13 was almost five times as high.

Example 2.17

In general, a lognormal distribution provides a good description of the number of load changes Y until failure in a fatigue test. The results of fatigue tests on a series of welds indicate a mean $\mu_X = 430000$ cycles and a standard deviation $\sigma_X = 215000$ cycles. Determine the pdf and cdf of the distribution of Y and calculate the 5^{th} -percentile.

The parameters of the lognormal distribution are calculated similarly as in Example 2.16:

$$Var(Y) = \ln\{1+\frac{Var(X)}{E(X)^2}\} = \ln\{1+\frac{215^2}{430^2}\} \approx 0.2231 \approx 0.4724^2$$
$$E(Y) = \ln(\mathrm{E(X)}) - \frac{1}{2}Var(Y) = \ln(430000) - \frac{1}{2}\cdot 0.2231 \approx 12.86 \quad (2.94)$$

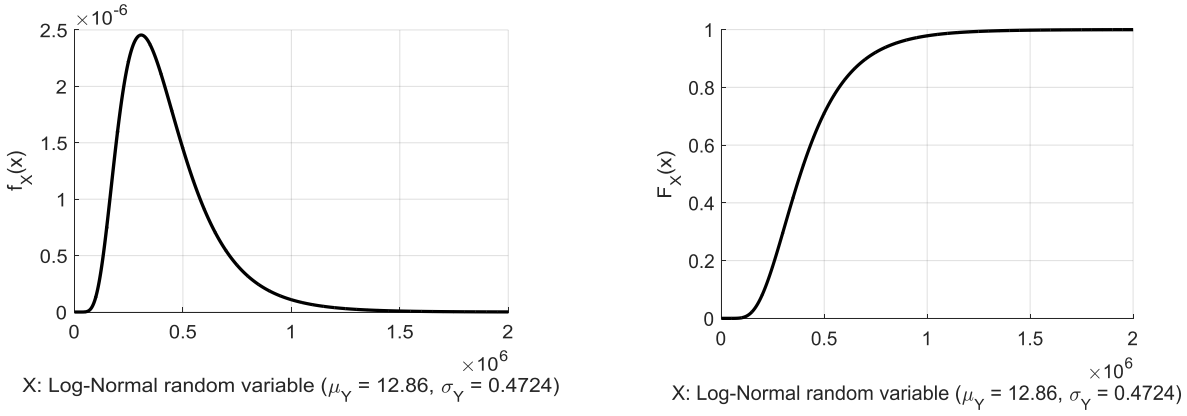


Figure 2.14: Lognormal pdf and cdf from Example 2.17

2.6.6 Extreme Value Distributions

In many applications, the most interesting values of a large group of random variables are the largest, or smallest values. For example, values of a big wave, the maximum wind speed or the lowest strength. The distribution of maxima or minima of a number of variables moves to the so called extreme value distributions. This convergence is not absolute and is a fairly slow process; much slower than for example the convergence of the sum of variables to a Gaussian distribution.

The first book devoted to the theory of asymptotic extreme value theory seems to be the book by Gumbel (1958). In his book Gumbel cites page 141 of the President's Water Resources Policy Commission (1950) regarding floods. The citation is famous in extreme value analysis and especially in hydrology and reads: "*However big floods get, there will always be a bigger one coming; so says the theory of extremes, and experience suggests it is true*".

Extreme value distributions, maxima as well as minima, are classified in three types: Type I, II and III. Table 2.3 provides an overview of the most important formulas regarding extreme value distributions. The most important difference between the various types is that type I is defined for

the range of $(-\infty,+\infty)$; type II is bounded for maxima by a lower bound and for the minima by an upper bound; type III is precisely in an opposite relation.

Type I for maxima is usually classified as a Gumbel-distribution, type II for maxima as a Frèchet distribution and type III for minima as a Weibull distribution. An interesting property of the extreme value distributions is (of course) that the maximum of two or more extreme value distributions (regardless whether they are from the same type or not) is also an extreme value distribution. We will illustrate this on the basis of an example.

Notes:

1. The type II and III distributions are given with lower (upper) boundary 0; an arbitrary different boundary can be introduced by means of parameter *u*.
2. The gamma function is defined as: $\Gamma(r)=\int_0^\infty t^{r-1}\cdot e^{-t}dt$ with $\Gamma(r)=(r-1)!$ (for $r\geq 0$). Furthermore, the following applies to $r:\Gamma(r+1)=1\cdot 2\cdot 3\cdot\ldots\cdot r$, the factorial function.
3.

Example 2.18

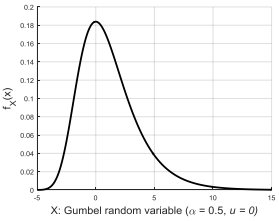
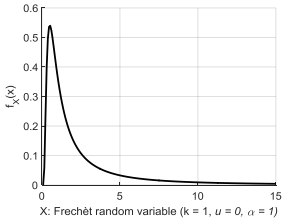
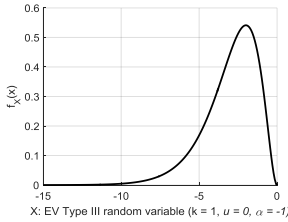
Let $Y_1,...,Y_n$ be independent identically distributed Frèchet (extreme value distributed of a type II maximum) random variables. Define $X=\max(Y_1,...,Y_n)$. What is the distribution of *X*?

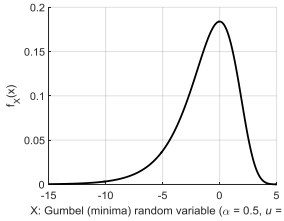
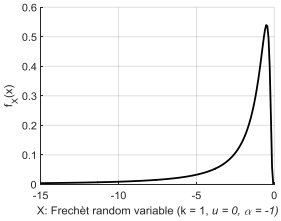
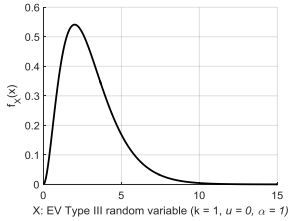
$$\begin{aligned} F_X(x) &= P(X<x)=P(\max(Y_1,...,Y_n)<x) \\ &= P(Y_1<x \text{ and } Y_2<x \text{ and } ... \text{ and } Y_n<x) \\ &= P(Y_1<x)\ P(Y_2<x)...\ P(Y_n<x) \\ &= F_Y{}^n(x) \\ &= \{\exp-(x/\alpha)^{-k}\}^n \\ &= \exp-\left\{n(x/\alpha)^{-k}\right\} \\ &= \exp-\left\{x/(\alpha\cdot n^{1/k})\right\}^{-k} \end{aligned} \qquad (2.95)$$

Conclusion: X also has an extreme value distribution of a type II maximum with parameters k and $\alpha\cdot n^{1/k}$.

When looking at the means and standard deviations it follows that $E(X)=E(Y)\cdot n^{1/k}$ and $V(X)=V(Y)$. The mean increases but the coefficient of variation remains the same.

Table 2.3: Overview Extreme Value Distributions

	Type I maxima (Gumbel)	Type II maxima (Frèchet)	Type III maxima
$F_X(x)$	$\exp[-e^{-\alpha(x-u)}]$	$\exp\left\{\left(-(x-u)/\alpha\right)^{-k}\right\}$	$\exp\left\{\left(-(x-u)/\alpha\right)^{k}\right\}$
$f_X(x)$	$\alpha\cdot\exp[-\alpha(x-u)-e^{-\alpha(x-u)}]$	$\left(\frac{k}{\alpha}\right)\left(\frac{x-u}{\alpha}\right)^{-k-1}\exp\left\{\left(\frac{-(x-u)}{\alpha}\right)^{-k}\right\}$	$-\left(\frac{k}{\alpha}\right)\left(\frac{x-u}{\alpha}\right)^{k+1}\exp\left\{\left(\frac{-(x-u)}{\alpha}\right)^{k}\right\}$
Range	$x,u\in(-\infty,\infty),\ \alpha>0$	$x,\ \alpha,\ k>0$	$u,\ x<0,\ k>0$
μ_x	$\mu=u+0{,}577/\alpha$	$\mu=u\cdot\Gamma(1-1/\mathrm{k})\ (\mathrm{k}>1)$	$\mu=u\cdot\Gamma(1+1/\mathrm{k})$
σ_x	$\sigma=\pi/\alpha\sqrt{6}$	$\sigma^2+\mu^2=u^2\cdot\Gamma(1-2/\mathrm{k})\ (\mathrm{k}>2)$	$\sigma^2+\mu^2=u^2\cdot\Gamma(1+2/k)$
$X=\max y_i$ $i=1\ldots n$	$\alpha_x=\alpha_y,$ $u_x=u_y+\{\ln(n)/\alpha\}$	$k_x=k_y,\ u_x=u_y,\ n^{1/k_y}$	$k_x=k_y,\ u_x=u_y,\ n^{-1/k_y}$
			

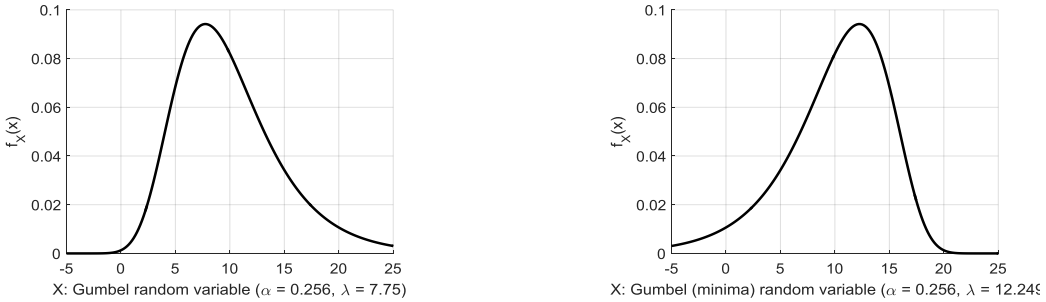
	Type I minima	Type II minima	Type III minima (Weibull)
$F_X(x)$	$1-\exp[-e^{+\alpha(x-u)}]$	$1-\exp\left\{\left(-(x-u)/\alpha\right)^{-k}\right\}$	$1-\exp\left\{\left(-(x-u)/\alpha\right)^{k}\right\}$
$f_X(x)$	$\alpha\cdot\exp[\alpha(x-u)-e^{\alpha(x-u)}]$	$-\left(\frac{k}{\alpha}\right)\left(\frac{x-u}{\alpha}\right)^{-k-1}\exp\left\{\left(\frac{-(x-u)}{\alpha}\right)^{-k}\right\}$	$-\left(\frac{k}{\alpha}\right)\left(\frac{x-u}{\alpha}\right)^{k+1}\exp\left\{\left(\frac{-(x-u)}{\alpha}\right)^{k}\right\}$
Range	$-\infty<x,$ $-\infty<u<+\infty,\ \alpha>0$	$x,\ u<0,\ k>0$	$u,\ x,\ k>0$
μ_x	$\mu=u-0{,}577/\alpha$	$\mu=u\cdot\Gamma(1-1/\mathrm{k})\ (\mathrm{k}>1)$	$\mu=u\cdot\Gamma(1+1/\mathrm{k})$
σ_x	$\sigma=\pi/\alpha\sqrt{6}$	$\sigma^2+\mu^2=u^2\cdot\Gamma(1-2/\mathrm{k})\ (\mathrm{k}>2)$	$\sigma^2+\mu^2=u^2\cdot\Gamma(1+2/k)$
$X=\max y_i$ $i=1\ldots n$	$\alpha_x=\alpha_y,$ $u_x=u_y-\{\ln(n)/\alpha\}$	$k_x=k_y,\ u_x=u_y\cdot n^{1/k_y}$	$k_x=k_y,\ u_x=u_y\cdot n^{-1/k_y}$
			

Example 2.19

Experimental measurements of the annual maximum flow rate (X) in a river over a long time frame have resulted in a mean $\mu_X = 10$ m³/s and $\sigma_X = 5$ m³/s. Experience and previous statistical analysis suggests that the distribution of X is well approximated by a Type I (Gumbel) distribution so no further statistical analysis is required. The parameters of the distribution of interest according to Table 2.3 are:

$$\alpha = \frac{\pi}{\sigma_X \sqrt{6}} = \frac{\pi}{5\sqrt{6}} = 0.2565 \text{ while } u = E(X) - \frac{0.577}{\alpha} \approx 7.7505 \tag{2.96}$$

The probability density function is visualized in Figure 2.15a.



a b

Figure 2.15 Extreme value distributions of type I

The probability that the maximum flow rate exceeds 20 m³/s is calculated as:

$$P(X > 20) = 1 - F_X(20) = 1 - \exp\left\{-e^{-0.256(20-7.75)}\right\} \approx 0.043 \tag{2.97}$$

This flow rate has a return period of 1/0.043 = 23.3 years. The 100 year flow rate equals 25.77 m³/s.

Now assume that the mean and standard deviation from above relate to the minimum flow rate in the river. In this case α remains the same and $\lambda = E(X) + \frac{0.577}{\alpha} \approx 12.2495$

The probability density function is visualized in Figure 2.15b. This graph clearly indicates that there is a significant probability of obtaining negative minimum flow rates, which is off course impossible. The model based on minima is thus not recommended.

2.6.7 The Gamma distribution

The gamma distribution is very common in applications. It is of special interest in the mathematical theory of reliability presented for example in Barlow and Proschan (1965). The gamma distribution is the foundation of the *gamma process* to be briefly discussed later in Chapter 10. The gamma density is defined as:

$$f_X(x) = \frac{1}{a\Gamma(k)}\left(\frac{x}{a}\right)^{k-1} e^{\left(-\frac{x}{a}\right)} \tag{2.98}$$

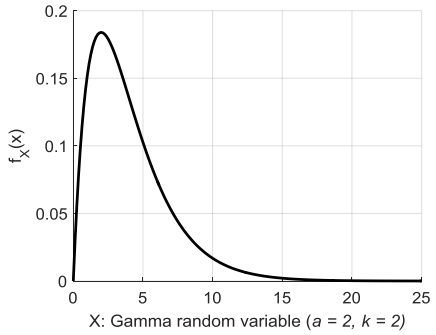
Where $\Gamma(a) = \int_{z=0}^{\infty} z^{a-1}e^{-z}dz$ with. The mean and standard deviation are defined as:

$$E(X) = a \cdot k, \qquad \sigma(X) = a \cdot \sqrt{k} \tag{2.99}$$

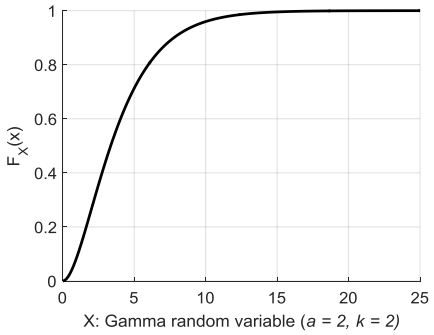
Pictures for a Gamma density and cumulative distribution function are provided in Figure 2.16. For even numbers of k it holds that (2.98) is de density of:

$$X = \frac{a}{2}\sum_{i=1}^{k/2} Z_i \tag{2.100}$$

With Z_i being independent standard normal random variables. For $a = 2$ and $v = k/2$, the gamma distribution is equal to a chi-squared distribution with v degrees of freedom. A summary of other important continuous distribution functions is provided in .



a



b

Figure 2.16 Gamma pdf and cdf

Distribution	Density function $f_X(x) = P(X = x)$	$E(X)$	Var(X)	Description
Uniform	$\frac{1}{b-a}$ $a \le x \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	Equal probability all outcomes in an interval
Exponential	$\lambda e^{-\lambda x}$ $x > 0$	$1/\lambda$	$1/\lambda^2$	Continuous counterpart of the geometric distribution. Time between events in a Poisson process
Chi squared (χ^2)	$\frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1}e^{-\frac{x}{2}}$	v	$2v$	Sum of the squares of v independent standard normal random variables

Beta	$\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]\cdot x^{a-1}(1-x)^{b-1}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	Model for random variables limited to intervals of finite length.

Table 2.4 Summary Parametric Continuous Probability Distributions

Having revised some of the most common continuous and discrete models for random variables, we come back to the sun rise problem. This is repeated below in Laplace's words: "Placing the most ancient epoch of history at five thousand years ago, or at 182623 days, and the sun having risen constantly in the interval at each revolution of twenty-four hours, it is a bet of 1826214 to one that it will rise again tomorrow". The solution according to Richard Price starts with letting the "sun rising" define a success in a series of Bernoulli experiments. The probability of observing x successes in n trials is given by a Binomial distribution $\binom{n}{x}p^x(1-p)^{n-x}$ assume that p has a "prior" distribution which is uniform on [0,1]. The probability of observing x successes in n trials considering this uniform prior is then $\int_0^1\binom{n}{x}p^x(1-p)^{n-x}dp$. We are interested in the probability that $p\in[a,b]$ given the sun has risen x times in n days. This conditional probability would be then:

$$P(a\le p\le b\,|\,x \text{ successes in } n \text{ trials})=\frac{\int_a^b\binom{n}{x}p^x(1-p)^{n-x}dp}{\int_0^1\binom{n}{x}p^x(1-p)^{n-x}dp}=\frac{\int_a^b p^x(1-p)^{n-x}dp}{\int_0^1 p^x(1-p)^{n-x}dp} \tag{2.101}$$

The difficult part for Bayes was to evaluate the integrals. Bayes showed that the denominator equals $1/(n+1)$ regardless of the value of x. For the numerator the case where the successes are equal to the number of trials is considered:

$$P(a\le p\le b\,|\,x \text{ successes in } x \text{ trials})=\frac{\int_a^b p^x dp}{1/(x+1)}=b^{x+1}-a^{x+1} \tag{2.102}$$

This quantity approaches 1 for any interval that includes p=1 and zero for any other interval. Thus as we experience more "sunrises" the probability of success should approach 1 more and more though never with absolute certainty.

Laplace's solution is commonly referred to as the "rule of succession" and follows a similar derivation. In modern statistical terms the likelihood function is: $L(p\,|\,n)=\binom{n}{n}p^n(1-p)^{n-n}=p^n$ the prior is still uniform in [0,1]. The posterior distribution is $P(p\,|\,n)\propto L(p\,|\,n)f_p\propto p^n$ which is

the kernel of a Beta distribution with $a = n+1$ and $b = 1$. The expected value is thus (see): $E(p \mid n \text{ successes in } n \text{ trials}) = \frac{n+1}{n+2}$. From which the odds mentioned by Laplace follow.

2.7 Two random variables *X* and *Y*

2.7.1 The two dimensional probability density function

To describe two random variables X and Y the two dimensional probability density function is introduced. It is defined as:

$$f_{XY}(x,y)dxdy = P(x < X < x+dx \text{ and } y < Y < y+dy) \tag{2.103}$$

Properties of this function are:

$$f_{XY}(x,y) \geq 0 \tag{2.104}$$

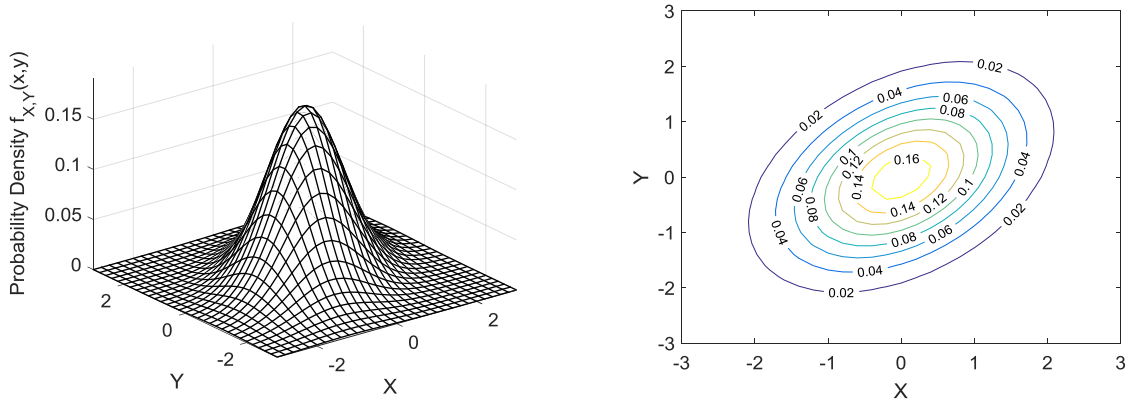
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)dxdy = 1 \tag{2.105}$$

$$P((X,Y) \in A) = \iint_{\{X,Y\} \in A} f_{XY}(x,y)dxdy \tag{2.106}$$

A special case arises in the case that the variables x and y are independent:

$$\begin{aligned} f_{XY}(x,y)dxdy &= P(x < X < x+dx \text{ and } y < Y < y+dy) \\ &= P(x < X < x+dx)P(y < Y < y+dy) \\ &= f_X(x)\, f_Y(y)dxdy \end{aligned} \tag{2.107}$$

The two dimensional probability density function is the product of two single dimension probability density functions. With dependant variables this is not the case and, in general, not enough information is provided by the separate one dimensional probability density functions to shape the two dimensional probability density function.



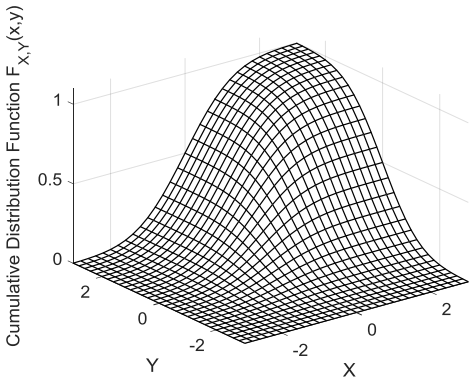


Figure 2.17: The two dimensional probability density function. At the right, the height contours are visualized.

In Figure 2.17 an example of a two dimensional probability density function is presented. In many cases only the contour plot is provided. The cumulative distribution function is also presented in Figure 2.17.

2.7.2 Expected Value, Mean, Variance

Like with the one dimensional distribution the expected value can also be defined for the two dimensional distribution:

$$E\{g(x,y)\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\cdot f_{XY}(x,y)dxdy \tag{2.108}$$

The most important properties are:

$$E(a) = a \tag{2.109}$$

$$E(a\cdot g) = a\cdot E(g) \tag{2.110}$$

$$E(g+h) = E(g)+E(h) \tag{2.111}$$

If x and y are independent it also holds:

$$E\{g(x)\cdot h(y)\} = E\{g(x)\}\cdot E\{h(y)\} \tag{2.112}$$

The proof of the above follows:

$$\begin{aligned} E\{g(x)\cdot h(y)\} &= \int\int g(x)\cdot h(y)\cdot f_{XY}(x,y)dxdy \\ &= \int\int g(x)\cdot h(y)\cdot f_x(X)\cdot f_y(Y)dxdy \\ &= \int g(x)\cdot f_X(x)dx\cdot\int h(y)\cdot f_Y(y)dy \\ &= E\{g(x)\}\cdot E\{h(y)\} \end{aligned} \tag{2.113}$$

To characterise the shared probability density function of X and Y the properties that usual summary measures (expectations and variances) are still used.

$$\mu(X) = E(X);\ \sigma^2(X) = E[\{X-\mu(X)\}^2] \tag{2.114}$$

$$\mu(Y) = E(Y);\ \sigma^2(Y) = E[\{Y - \mu(Y)\}^2] \tag{2.115}$$

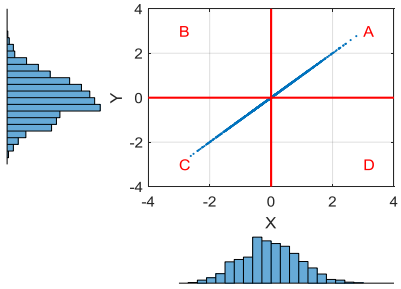
These definitions are fully compatible with the earlier given, one-dimensional, definitions (2.64) and (2.65).

2.7.3 Covariance

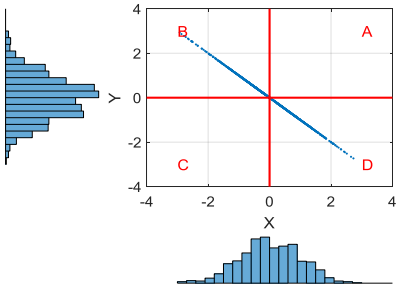
Given the definitions for expected value and variance, it becomes obvious to define a third statistical measure. This third measure is the expected value of the product of the deviations of the random variables with respect to their corresponding mean:

$$\text{cov}(X,Y) = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] \tag{2.116}$$

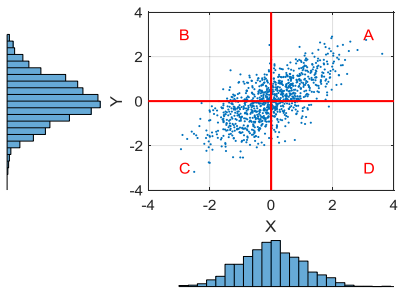
The covariance is a measure of linear dependence and is closely related to the bivariate Gaussian distribution. For perfect linear dependence the covariance would be positive and large. This is illustrated in Figure 2.18 a). Both random variables are normally distributed with $E(X) = E(Y) = 0$ and take an arbitrary value of $X > 0$, it corresponds to values of $Y > 0$, that is a value in the A quadrant in Figure 2.18a). In this case $\left(X - E(X)\right) > 0$ and $\left(Y - E(Y)\right) > 0$, hence $E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] > 0$. Similarly, take an arbitrary value of $X < 0$ (quadrant C). This value would correspond to values of $Y < 0$ and hence $E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] > 0$. Thus the covariance will be positive and large.



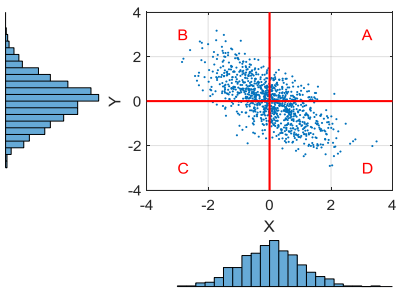
a) Perfect positive linear dependence



b) Perfect negative linear dependence



c) Positive dependence



d) Negative dependence

Figure 2.18: Dependence patterns examples for bivariate Normal distribution.

For negative dependence a similar analysis may be performed to verify that the Covariance in this case would be Negative and Large (see Figure 2.18b).

In case there is not perfect dependence, for example for Figure 2.18c), it may be seen that $X > 0$ in quadrant D would correspond to $Y < 0$. In this case for the samples in quadrant D $\big(X - E(X)\big) > 0$ and $\big(Y - E(Y)\big) < 0$, hence $\big(X - E(X)\big)\big(Y - E(Y)\big) < 0$ for these samples. Conversely, for samples $X < 0$ in quadrant B $\big(X - E(X)\big) < 0$ and $\big(Y - E(Y)\big) > 0$, hence $\big(X - E(X)\big)\big(Y - E(Y)\big) < 0$. The overall expected value however $E\Big[\big(X - E(X)\big)\big(Y - E(Y)\big)\Big] > 0$ still because there are more samples in quadrants A and C than in quadrants B and D. The magnitude of the covariance would however be smaller in the case of Figure 2.18c) than in Figure 2.18a) because of the negative contribution of samples in quadrants B and D. An analogue situation may be verified for the case of Figure 2.18d) concerning negative covariance.

A concept directly related to the covariance is the Pearson's product moment correlation-coefficient $\rho_{XY} = \rho(X,Y)$ (both notations are widely used), which is defined as follows:

$$\rho_{XY} = \frac{\mathrm{cov}(XY)}{\sigma(X)\sigma(Y)} = E\left\{\frac{X - E(X)}{\sigma(X)} \cdot \frac{Y - E(Y)}{\sigma(Y)}\right\} \tag{2.117}$$

The correlation coefficient has the following properties:

If X and Y are independent: $\rho_{XY} = 0$ (2.118)

If X and Y are completely linearly dependant: $\rho_{XY} = \pm 1$ (2.119)

In general: $-1 \le \rho_{XY} \le +1$ (2.120)

The correlation coefficient is a nominal measure for the mutual dependency of two stochastic variables. The proof for the given properties is given below.

a. If x and y are independent it holds (see (2.112) and (2.116)):

$$\begin{aligned} \mathrm{cov}(X,Y) &= E\big[\{X - E(X)\}\{y - E(Y)\}\big] \\ &= E\{X - E(X)\} \cdot E\{Y - E(Y)\} \\ &= 0 \cdot 0 = 0 \end{aligned} \tag{2.121}$$

With this result it also holds that $\rho_{XY} = 0$; the fact that $E\{X - E(X)\}$ follows directly from the definition given in (2.64).

b. Complete linear dependency means that Y can be written as a linear function of X with known coefficients: $Y = aX + b$. For the expected value and standard deviation is holds: $E(Y) = aE(X) + \mathrm{b}$ and $\sigma(Y) = |a|\sigma(X)$.

For the covariance of X and Y if follows:

$$\begin{aligned} \mathrm{cov}(X,Y) &= E\big[\{X - E(X)\}\{Y - E(Y)\}\big] \\ &= E\big[\{X - E(X)\}\{aX + b - aE(X) - \mathrm{b}\}\big] \\ &= aE\{X - E(X)\}^2 \\ &= a \cdot \sigma^2(x) \end{aligned} \tag{2.122}$$

The correlation coefficient then is:

$$\rho_{XY} = \frac{\mathrm{cov}(XY)}{\sigma(X)\sigma(Y)} = \frac{a\sigma^2(X)}{\sigma(X)\{|\mathrm{a}|\sigma(X)\}} = \frac{a}{|a|} = \pm 1 \tag{2.123}$$

c. Consider the inequality:

$$E\left[\left\{\frac{X-E(X)}{\sigma(X)}-\frac{Y-E(Y)}{\sigma(Y)}\right\}^2\right]>0 \tag{2.124}$$

This inequality is based on the fact that the square (of real numbers) is always positive. This also holds for the expected value of a square. The result of this, with the help of the properties of the expected value and others, is:

$$E\left[\left\{\frac{X-E(X)}{\sigma(X)}\right\}^2-2\left\{\frac{X-E(X)}{\sigma(X)}\right\}.\left\{\frac{Y-E(Y)}{\sigma(Y)}\right\}+\left\{\frac{Y-E(Y)}{\sigma(Y)}\right\}^2\right]>0$$

$$1-2\rho_{XY}+1>0 \tag{2.125}$$

$$\rho_{XY}<1$$

The inequality $-1<\rho_{XY}$ can be proven by assuming the square with a positive sign.

Note: if X and Y are independent it follows that $\rho_{XY}=0$, but from $\rho_{XY}=0$ it cannot be concluded that X and Y are independent; some other non-linear dependencies can be in play. If X and Y are Gaussian distributed it can be determent that if $\rho(X,Y)=0$, X and Y are independent.

Example 2.20

If the variables x and y are given as:

$$\begin{aligned}&E(X)=E(X)=0\\&\sigma(X)=1,\ \sigma(Y)=2,\ \mathrm{cov(XY)}=1\end{aligned} \tag{2.126}$$

Asked are the expected values, variances and covariances of $S=X+Y$ and $T=2X-Y$.

One can directly notice that $E(S)=E(T)=0$. The variance of s follows:

$$\begin{aligned}\sigma^2(\mathrm{S})&=\mathrm{E}[\{\mathrm{S}-\mathrm{E(S)}\}^2]\\&=\mathrm{E}[\{\mathrm{X}+\mathrm{Y}\}^2]\\&=\mathrm{E}(\mathrm{X}^2)+2E(\mathrm{XY})+\mathrm{E}(\mathrm{Y}^2)\\&=\sigma^2(X)+2\,\mathrm{cov(XY)}+\sigma^2(Y)\\&=1+2+4=7\end{aligned} \tag{2.127}$$

Determine yourself that $\sigma^2(T)=4-4+4=4$.

Finally, the covariance is determined as follows:

Determine yourself that $\sigma^2(T)=4-4+4=4$.

Finally the covariance is determined as follows:

$$\begin{aligned}
&\operatorname{cov}(ST) = E\left[\{S-\mu(S)\}\{T-\mu(T)\}\right] = \\
&\operatorname{cov}(ST) = E\left[(X+Y)(2X-Y)\right] = \\
&\operatorname{cov}(\mathrm{ST}) = E\left(2X^2\right) + E\left(XY\right) - E\left(Y^2\right) = \\
&\operatorname{cov}(ST) = 2+1-4 = -1
\end{aligned} \tag{2.129}$$

The correlation coefficient $\rho_{XY} = 0.5$ and $\rho_{ST} = -0.2$.

Example 2.21

Given are two independent variables X and Y with $\mu(X) = \mu(Y) = 0$ and $\sigma(X) = \sigma(Y) = 1$. If $Z = aX + bY$, then $\mu(Z) = 0$ and $\sigma(Z)$ follows from:

$$\begin{aligned}
&\sigma^2(\mathrm{Z}) = \mathrm{E}\left(aX+bY\right)^2 \\
&\sigma^2(Z) = a^2\,\mathrm{E}(X^2) + 2\,\mathrm{abE}(XY) + b^2\,\mathrm{E}(Y^2) = a^2 + b^2
\end{aligned} \tag{2.130}$$

The covariance of X and Z is determined next:

$$\begin{aligned}
&\operatorname{cov}(XZ) = E\left[\{X-E(X)\}\{Z-E(Z)\}\right] = \\
&\operatorname{cov}(XZ) = E\{(X)(aX+bY)\} = \\
&\operatorname{cov}(XZ) = aE(X^2) + \mathrm{bE}(XY) = a\sigma^2(X) + b\operatorname{cov}(XY) = a\sigma^2(X) = a
\end{aligned} \tag{2.131}$$

It follows that $\rho(XZ)$ is:

$$\rho(XZ) = \frac{\operatorname{cov}(XZ)}{\sigma(X)\sigma(Z)} = \frac{a}{\sqrt{a^2+b^2}} \tag{2.132}$$

Assume $a^2 + b^2 = 1$ leading to $\sigma(Z) = 1$ and $\rho(XZ) = a$. If $a \approx 1$ then X and $Z = aX + Y\sqrt{1-a^2}$ have much in common and $\rho(XZ)$ is high; if $|a|$ is small then X and Z do not have much in common and $\rho(XZ) \approx 0$.

2.7.4 Function of two variables

Consider the function $Z = g(X,Y)$. If the mutual probability density function of X and Y is known then the probability density function of Z can be calculated. For many application it is sufficient to determine the expected value and the standard deviation.

1. Z is the sum of X and Y: $Z = X + Y$.
   In that case the expected value is determined by:

$$E(Z) = E(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y) \tag{2.133}$$

   The variance is determined as follows:

$$\sigma^2(Z) = E\left[\left\{Z - E(Z)\right\}^2\right] =$$
$$\sigma^2(Z) = E\left[\left\{X + Y - E(X) - E(Y)\right\}^2\right] =$$
$$\sigma^2(Z) = E\left[\left\{X - E(X)\right\}^2 + 2\left\{X - E(X)\right\}\left\{Y - E(Y)\right\} + \left\{Y - E(Y)\right\}^2\right] = \tag{2.134}$$
$$\sigma^2(Z) = \sigma^2(X) + 2\,\mathrm{cov}(XY) + \sigma^2(Y) =$$
$$\sigma^2(Z) = \sigma^2(X) + 2\rho\sigma(X)\sigma(Y) + \sigma^2(Y) =$$

   If $\rho = 0$ this simplifies to $\sigma^2(Z) = \sigma^2(X) + \sigma^2(Y)$.

2. An arbitrary linear function $Z = a + bX + cY$
   In the same manner as case 1. It can be determined that:

$$E(Z) = \mathrm{a} + \mathrm{b}\,E(X) + cE(Y) \tag{2.135}$$

$$\sigma^2(Z) = b^2\sigma^2(X) + 2bc\,\mathrm{cov}(XY) + c^2\sigma^2(Y) \tag{2.136}$$

3. Approximation of a nonlinear function
   If $Z = g(XY)$ is a nonlinear function in many cases the following approximation, on the basis of linearization, can be used:

$$Z = g(X_0Y_0) + (X - X_0)g'_X(X_0Y_0) + (Y - Y_0)g'_Y(X_0Y_0) \tag{2.137}$$

$$E(Z) = g(X_0Y_0) + (E(X) - X_0)g'_X(X_0Y_0) + (E(Y) - Y_0)g'_Y(X_0Y_0) \tag{2.138}$$

$$\sigma^2(Z) = \left\{g'_X(X_0Y_0)\sigma(X)\right\}^2 + 2g'_X(X_0Y_0)g'_Y(X_0Y_0)\,\mathrm{cov}(XY) + \left\{g'_Y(X_0Y_0)\sigma(Y)\right\}^2 \tag{2.139}$$

In the definitions above, g'_X and g'_Y are respectively the partial derivative $\delta g/\delta X$ and $\delta g/\delta Y$. (X_0Y_0) is the chosen linearization point. Sometimes this point is taken as $(E(X), E(Y))$, the so called mean value approximation.

Example 2.22

Examine $Z = XY$ with X and Y being independent. We choose $X_0 = E(X)$ and $Y_0 = E(Y)$:

$$\begin{aligned} Z &= E(X)E(Y) + \{X - E(X)\}E(Y) + \{Y - E(Y)\}E(X) \\ E(Z) &= E(X)E(Y) \\ \sigma^2(Z) &= \{E(Y)\sigma(X)\}^2 + \{E(X)\sigma(Y)\}^2 \end{aligned} \qquad (2.140)$$

The variance coefficient can be determined:

$$V^2(Z) = \frac{\sigma^2(Z)}{E^2(Z)} = \frac{E(Y)^2\sigma(X)^2 + E(X)^2\sigma(Y)^2}{E^2(X)E^2(Y)} = V^2(X) + V^2(Y) \qquad (2.141)$$

This result is quoted many times in the literature.

Interesting to notice is that the exact solution of this example is known: $E(Z) = E(X)E(Y)$ and $\sigma^2(Z) = E^2(Y)\sigma^2(X) + E^2(X)\sigma^2(Y) + \sigma^2(X)\sigma^2(Y)$. Thus the approximation holds if $V(X)$ and $V(Y)$ are small.

Example 2.23

The bivariate distribution

The density function of the bivariate normal distribution is given by:

$$\phi_\rho(\mathrm{x},\mathrm{y}) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}{}^2}} e^{-\frac{(\frac{x-\mu_X}{\sigma_X})^2 - (\frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}) + (\frac{y-\mu_Y}{\sigma_Y})^2}{2(1-\rho_{XY}{}^2)}} \qquad (2.141)$$

Notice that the bivariate normal distribution is parameterized by the individual means and variances and also by ρ_{XY}. The corresponding cumulative distribution is thus:

$$\Phi_\rho(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y}\phi_\rho(s,t)dsdt \qquad (2.142)$$

Figure 2.17 represents in fact a bivariate normal distribution with correlation $\rho_{XY} = 0.4$. The samples presented in Figure 2.18 correspond to $\rho_{XY} = 1$ (a), $\rho_{XY} = -1$ (b), $\rho_{XY} = 0.7$ (c) and $\rho_{XY} = -0.7$ (d).

2.7.5 Spearman's Rank correlation coefficient

So far we have considered the covariance and Pearson's product moment correlation coefficient as measures of statistical dependence. These measures are taken in the original units of variables. Spearman's Rank correlation coefficient is a way to extend the concept of "linear dependence" to "monotonic dependence". Spearman's correlation coefficient $r_{X,Y} = r(X,Y)$ is:

$$r(X,Y) = \rho(F_X(X), F_Y(Y)) \qquad (2.143)$$

Spearman's coefficient is thus, Pearson's product moment correlation computed with the "ranks" of variables *X* and *Y*. Notice that $F_X(X)$ and $F_Y(Y)$ would bring values of *X* and *Y* to the [0,1] interval. In practice F_X may be a parametric estimate (for example from a particular family discussed in section 2.6) or an empirical estimate:

$$\hat{F}_X(x) = \frac{\#\text{ samples } \le x}{N+1} \tag{2.144}$$

where *N* denotes the total number of samples and *N*+1 in the denominator is used to avoid values "too close" to 0 and 1 in small samples. Important properties of Spearman's correlation coefficient are:

1. $r(X,Y) = r(G(X),Y)$ for a strictly increasing function *G*
2. $r(X,Y) = -r(G(X),Y)$ for a strictly decreasing function *G*
3. If $r(X,Y) = 1$ then there exists a strictly increasing function *G* such that *X* = *G*(*Y*)

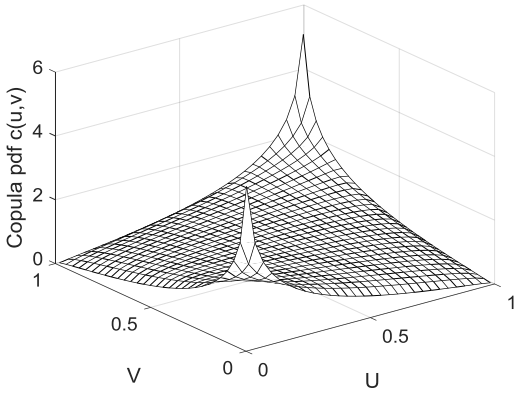
2.8 Bivariate copulas

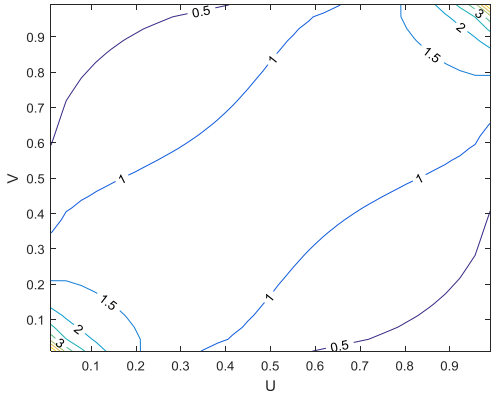
Modern statistics is shifting from the study of bivariate distributions to the study of copulas. Roughly speaking, bivariate copulas are the bivariate distribution corresponding to the "ranks" of the original variables. In this sense, the notion of 'copula' was introduced to separate the effect of dependence from the effect of marginal distributions in a joint distribution. And this property is precisely what is making the study of copulas the standard in modern statistics.

More formally Sklar's (1959) theorem states that random variables *X* and *Y* are joined by a copula *C* if their joint distribution can be written

$$F_{X,Y}(x,y) = C\left(F_X(x), F_Y(y)\right) \tag{2.145}$$

Every continuous bivariate distribution $F_{X,Y}$ can be represented in terms of a copula. Moreover, we can always find a unique copula *C* that corresponds to given continuous joint distribution. One of the most common copulas is the Normal or Gaussian Copula.





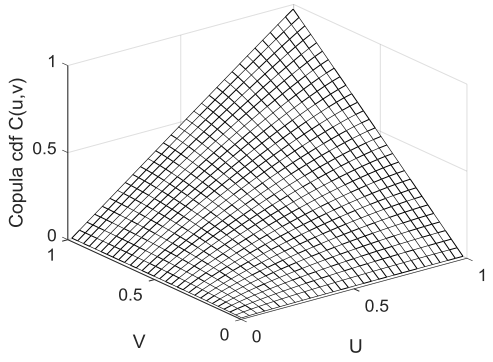


Figure 2.19: Two dimensional Normal copula.

Example 2.24

For maintenance or installation operations of offshore wind turbines to be started, environmental variables need to be at certain levels. In particular variables X_1: mean wind speed (m/s) and X_2: significant wave height (cm) are of interest for a particular operation. It has been found that both these variables may be well approximated by a Gumbel distribution:

$$F_{X_i}(x) = e^{-e^{(-\alpha_i(x-\gamma_i))}} \tag{2.146}$$

The parameters for wind speed and significant wave height are according to the following table:

x_i	α_i	γ_i
mean wind speed (m/s)	0.35	8
significant waveheight (m)	0.016	155

If the wind speed is $>11(m/s)$ or the wave height is $>220(cm)$ the mission may not ship.

1. If the two variables wave height and windspeed are assumed to be independent, the probability of not being able to ship can be calculated through standard probability relations.

$$\begin{aligned} P(X_1 > 11 \text{ OR } X_2 > 220) &= P(X_1 > 11) + P(X_2 > 220) - P(X_1 > 11 \text{ AND } X_2 > 220) \\ &= P(X_1 > 11) + P(X_2 > 220) - P(X_1 > 11) \cdot P(X_2 > 220) \end{aligned} \tag{2.147}$$

.

Using the Gumbel distribution with parameters as indicated in the table, we find:

$$\begin{aligned} P(X_1 > 11) &= 1 - e^{-e^{-0.35(11-8)}} = 0.295 \\ P(X_2 > 220) &= 1 - e^{-e^{-0.016(220-155)}} = 0.298 \end{aligned} \tag{2.148}$$

The chance of occurence of wheather conditions that hinder the offshore operations is:

$$P(X_1 > 11 \text{ OR } X_2 > 220) = 0.295 + 0.298 - 0.295 \cdot 0.298 = 0.505 \tag{2.149}$$

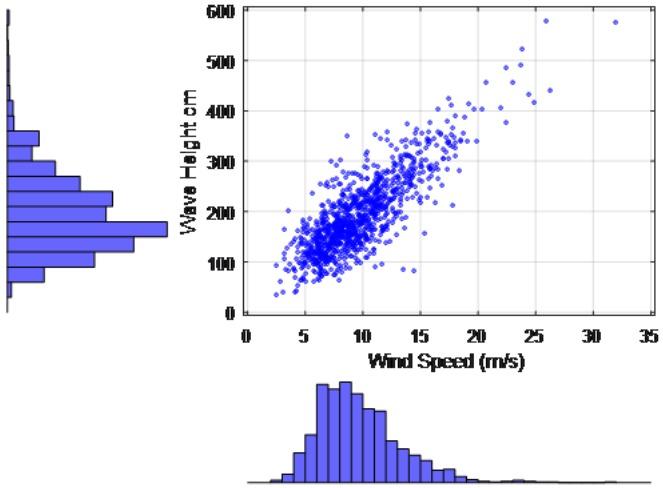


Figure 20 Scatter plot of realizations taken from the joint distribution for wave height and wind speed.

2. As can be seen in the scatterplot, it appears that the wind speed and the wave height are correlated. If it is known that Spearman's rank correlation coefficient between wave height and wind speed is $r(X_1, X_2) \approx 0.8$, then the joint probability of exceedance of $x_1 > 11$ AND $x_2 > 0.22$ can no longer be calculated from the marginal distributions only.

The third term of equation (2.147) must then be expanded as:

$$P(X_1 > 11 \text{ AND } X_2 > 220) \quad = \quad P(X_1 > 11) \cdot P(X_2 > 220 \mid X_1 > 11) \tag{2.150}$$

We also know that:

$$\begin{aligned} P(X_1 > 11 \text{ AND } X_2 > \quad 220 \quad) = & \\ = \quad & 1 - \left[P(X_1 < 11 \text{ OR } X_2 < 220) \right] \\ = \quad & 1 - \left[P(X_1 < 11) + P(X_2 < 220) - P(X_1 < 11 \text{ AND } X_2 < 220) \right] \\ = \quad & 1 - \left[F_{X_1}(11) + F_{X_2}(220) - H_{X_1,X_2}(11, 220) \right] \end{aligned} \tag{2.151}$$

In general, we can denote

$$\begin{aligned} F_{X_1}(x_1) &= u \\ F_{X_2}(x_2) &= v \end{aligned} \tag{2.152}$$

We know from (2.145) that the joint probability distribution can be obtained from substitution of the marginal distributions in the expression for a copula:

$$H_{X_1,X_2}(x_1,x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)) = C(u,v) \tag{2.153}$$

Such that the second term in (2.150) can be rewritten as:

$$P(X_2 > x_2 \mid X_1 > x_1) \quad = \quad \frac{1-\left[u+v-C(u,v)\right]}{(1-u)} \tag{2.154}$$

3. We make use of the Gumbel copula to combine the two marginal distribution functions into a joint distribution function, as in (2.145).
   The Gumbel copula distribution function is given by:

$$C(u,v;\theta) = e^{-\left[(-\ln(u))^{\theta}+(-\ln(v))^{\theta}\right]^{\frac{1}{\theta}}} \tag{2.155}$$

The parameter θ can be related to Spearman's rank correlation as $\theta(r=0.8)=2.5$, as can be seen from the graph below:

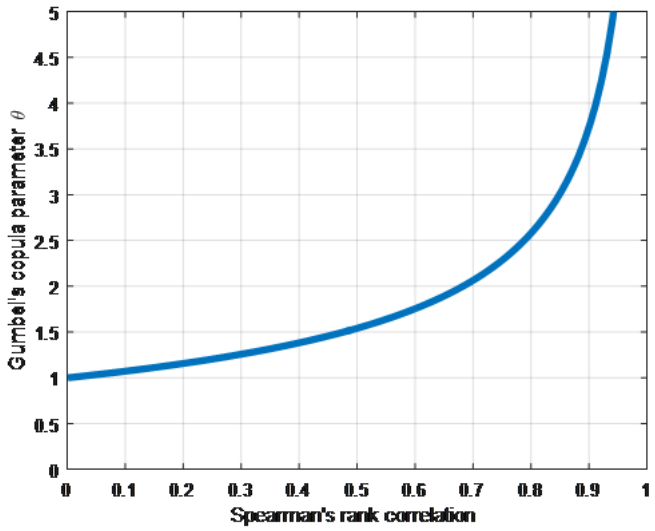


Figure 21 The Gumbel copula parameter as a function of Spearman's rank correlation coefficient.

With an expression for the copula distribution, the joint probability of exceedance can be calculated from (2.151):

$$P(X_1 > 11 \text{ OR } X_2 > 220) = (1-u)+(1-v)-\left[1-u-v+C(u,v)\right] \tag{2.156}$$

With $u = P(X_1 < 11) = 0.705$, $v = P(X_2 < 220) = 0.702$ and $C(0.705, 0.702, 2.5) = 0.629$ this becomes:

$$P(X_1 > 11 \text{ OR } X_2 > 220) = (1-0.705)+(1-0.702)-\left[1-0.705-0.702+0.629\right] = 0.371 \quad (2.157)$$

We can therefore conclude that the correlation between the two weather states positively influences the workability at sea, since the probability of occurrence of weather conditions that fall within the acceptable range for a mission to take place has increased from $(1-0.505)=0.495$ to $(1-0.371)=0.629$.

An algebraic expression for the copula distribution function, $C(u,v)$ joining two marginal distributions, $F_X(x)$ and $F_Y(y)$ could be obtained also from the joint probability density function, $H_{X,Y}(x,y)$ if this one is known. The relationship is given by:

$$C(u,v) = H\left(F^{-1}(u), G^{-1}(v)\right) \quad (2.159)$$

in which $u,v \in \left[0\ 1\right]$ and F^{-1} and G^{-1} are the inverses of F and G, which are the marginal cumulative distribution functions. When the joint cumulative probability distribution function is given by $H_{X,Y}(x,y)$, the marginal distribution functions can be retrieved as:

$$\begin{aligned} F_X(x) &= \lim_{y \to \sup(Y)} H_{X,Y}(x,y) \\ G_Y(y) &= \lim_{x \to \sup(X)} H_{X,Y}(x,y) \end{aligned} \quad (2.160)$$

Where $\sup(X)$ is the supremum of X, which on a closed interval is equal to $\max(X)$ and on an open interval is equal to the upper interval boundary, or, if there is no boundary, $\sup(X)=\infty$.

To clarify the procedure, an example will be given.

Example 2.25

Consider two random variables X and Y with a joint probability density function given by

$$h_{X,Y}(x,y) = \frac{4}{5a \cdot b}\left(\frac{x}{a}+\frac{y}{b}+\frac{xy}{a \cdot b}\right) \quad (2.161)$$

with $0 \le x \le a$ and $0 \le y \le b$.

First, the cumulative distribution function needs to be found.

$$H_{X,Y}(x,y) = \int_0^a \int_0^b f_{X,Y}(x,y)\,dx\,dy = \frac{x \cdot y}{5a^2 \cdot b^2}\left(2a \cdot y + 2b \cdot x + x \cdot y\right) \quad (2.162)$$

Then the marginal distribution functions can be found by obtaining limits of equation (2.162) according to (2.160):

$$F_X(x) = H_{X,Y}(x,b) = \frac{x}{5a^2 \cdot b}\left(2a \cdot b + 3b \cdot x\right)$$
$$G_Y(y) = H_{X,Y}(a,y) = \frac{y}{5a \cdot b^2}\left(2a \cdot b + 3a \cdot y\right) \tag{2.163}$$

Now the inverses of F and G need to be found. For this example, it is possible to find a closed form expression for the inverse on its domain.

$$F_U^{-1}(u) = \frac{a}{6}\left(5\sqrt{\frac{12x}{5} + \frac{4}{25}} - 2\right)$$
$$G_V^{-1}(v) = \frac{b}{6}\left(5\sqrt{\frac{12y}{5} + \frac{4}{25}} - 2\right) \tag{2.164}$$

According to (2.159), These expressions in (2.164) must then be substituted for the arguments x and y of (2.162) to obtain the copula:

$$C_{U,V}(u,v) = \frac{1}{6480}((2\sqrt{15u+1}-2)(2\sqrt{15v+1}-2)(4\sqrt{15u+1}\sqrt{15v+1} + 20\sqrt{15u+1} + 20\sqrt{15v+1} - 44)) \tag{2.165}$$

The copula distribution function can be differentiated twice to obtain the copula density function:

$$c_{U,V}(u,v) = \frac{\partial^2 C_{U,V}(u,v)}{\partial u \partial v} \tag{2.166}$$

The copula density is plotted in Figure 22.

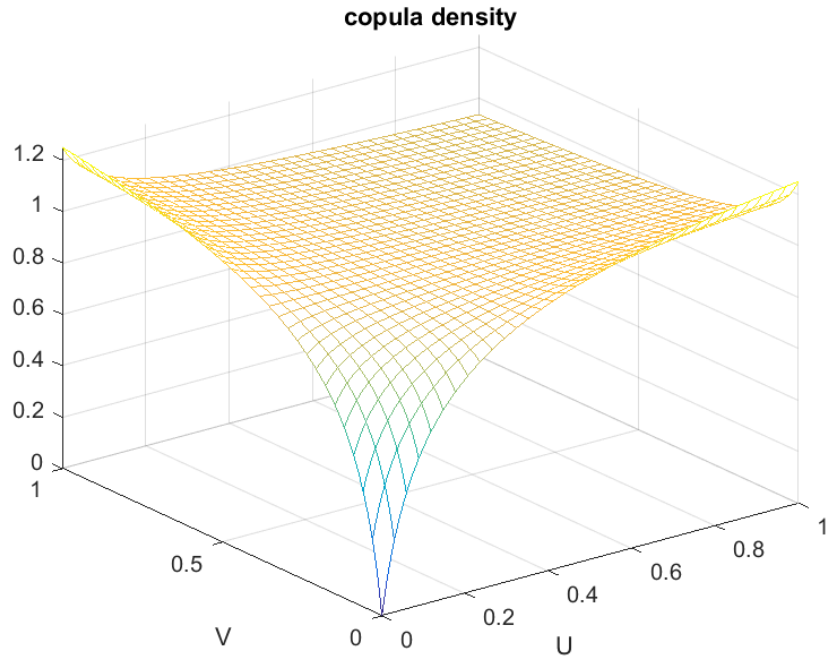


Figure 22 The copula density function derived from the joint probability density

function given by (2.161).

Even though one might expect there to be a correlation between the two variables due to the mixed term in equation (2.162), the copula shows that this correlation is very weak, because the density function is for the largest part a horizontal surface parallel to the U,V plane. This indicates that all pairs of outcomes (u,v) are equally likely. Only on the borders of the intervals of interest, i.e. u nearly zero while v approximates 1 or v nearly zero while u approximates 1, does the copula show a slightly higher probability of occurence. The steep drop of the surface to the origin $(u,v)=(0,0)$ indicates that there are very few, if none, realizations possible in this neighbourhood.
Indeed, Spearman's correlation coefficient is given by $\rho_{X,Y} \approx -0.06$, so the correlation between u and v is very weak.

2.9 "n" random variables

The extension from 2 to "n" variables holds, in principle, no new elements. The n-dimensional probability density function for the variables $\underline{x} = \{x_1 .... x_n\}$ is defined as:

$$f_{\underline{X}}(\underline{x})dx_1...dx_n = P(x_1 < X_1 < x_1 + dx_1 \text{ and } ... \text{ and } x_n < X_n < x_n + dx_n) \tag{2.167}$$

For the event "X in A" it holds:

$$P(\underline{X} \text{ in } A) = \int_A ... \int f_{\underline{X}}(\underline{x})dx_1 ... dx_n \tag{2.168}$$

If the variables $x_1 .... x_n$ are independent, $f_{\underline{X}}(\underline{x})$ is given by:

$$f_{\underline{X}}(\underline{x}) = f_{X_1}(x_1) f_{X_2}(x_2) .... f_{X_n}(x_n) \tag{2.169}$$

The expected value of a function $g(\underline{x})$ is defined as:

$$E\{g(\underline{x})\} = \int_{-\infty}^{+\infty} ... \int g(\underline{x}) f_{\underline{X}}(\underline{x})dx_1 ... dx_n \tag{2.170}$$

The expected value, variances and covariances are on the basis of the above defined as:

$$\mu(X_i) = E(X_i) \tag{2.171}$$

$$\sigma^2(X_i) = E\left[\{X_i - E(X_i)\}^2\right] \tag{2.172}$$

$$\text{cov}(X_i X_j) = E\left[\{X_i - E(X_i)\}\{X_j - E(X_j)\}\right] \tag{2.173}$$

In the *n*-dimensional case, the correlation coefficient $\rho(X_i X_j) = \text{cov}(X_i X_j) / \sigma(X_i)\sigma(X_j)$ for each pair of variables *i* and *j* constitutes the entries of a matrix. This matrix is symmetric and has on its main diagonal only ones. With independent variables all non-diagonal terms are zero.

The expected value and the standard deviation of an arbitrary function $z = g(\underline{X})$ can in most cases be approximated using:

$$E(Z) = g(\underline{X}_0) + \sum_{i=1}^{n} (E(\mathrm{X}_i - X_{0i}) g_i^{'}(\underline{X}_0) \tag{2.174}$$

$$\sigma^2(Z) = \sum_i \sum_j g_i^{'}(\underline{X}_0) g_j^{'}(\underline{X}_0) \operatorname{cov}(\mathrm{X}_i \mathrm{X}_j) \tag{2.175}$$

If all variables X_i are independent:

$$\sigma^2(Z) = \sum \{ g_i'(\underline{X}_0) \sigma_i \}^2 \tag{2.176}$$

In this notation $\underline{X}_0$ is a chosen approximation point and g_i' is the partial derivative of g to X_i.

References

- Kolmogorov, A.N. (1956) Foundations of the theory of Probability.
- Laplace, Pierre Simon (1814), A Philosophical Essay on Probabilities, translated from the 6th French edition by Frederick Wilson Truscott and Frederick Lincoln Emory. New York: John Wiley & Sons, 1902, p. 19. Dover Publications edition (New York, 1951)
- Keynes J. (1973) A treatise on probability, London.
- Von Mises, R. (1936) Probability, statistics and truth, Vienna
- Savage L. (1956) Foundations of statistics, New York.
- Bayes and Price (1763) An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. Philosophical Transactions (1683-1775) Vol. 53 (1763), pp. 370-418. Published by: The Royal Society. Stable URL: http://www.jstor.org/stable/105741.
- Gumbel, E.J. (1958) Statistics of extremes. Columbia University Press. New York.
- President's Water Resources Policy Commission (1950). Water Policy for the American People. Report. Washington D.C. Vol 1.
- Barlow, R. E., and F. Proschan (1965). Mathematical Theory of Reliability. J. Wiley & Sons. Reprinted (1996) SIAM, Philadelphia, PA
- J.M. van Noortwijk, (2009) A survey of the application of gamma processes in maintenance, *Reliability Engineering & System Safety*, Volume 94, Issue 1, January 2009, Pages 2-21.
- Sklar A (1959) Fonctions de réparation à n dimensions et leurs marges. *Publ. Inst. Statist.* Univ. Paris 8, 229–231
- Paul G. Hoel, Sidney C. Port Charles J. Stone (1971) *Introduction to Probability Theory* University of California, Los Angeles

Appendix 2.1 Solution to Alternative examples & exercises

Example 2.4

About Example 2.4.: The reason is that the events are neither mutually exclusive nor one event includes every other event. Solution is the sum of:

- 3 event in which 6 happens in exactly one throw each with probability, $1/6\cdot(5/6)^2$
- $\binom{3}{2}$ events in which 6 happens in exactly two throws each with probability, $(1/6)^2\cdot(5/6)$
- 1 event in which 6 happens in exactly three throw $(1/6)^3$

Answers Exercise 2.12

$$\begin{aligned}&\mathrm{Var}(X)=E\left[\left(X-E(X)\right)^2\right]=E\left[X^2-2XE(X)+E(X)^2\right]\\&=E(X^2)-2E(X)E(X)+E(E(X)^2)=E(X^2)-2E(X)^2+E(X)^2\\&=E(X^2)-E(X)^2\end{aligned}\qquad(2.177)$$

Looking at the expected value:

$$E(X)=\sum_{x=0}^{n}xf_X(x)=\sum_{x=0}^{n}x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\qquad(2.178)$$

The first term of the summation equals zero, hence

$$E(X)=\sum_{x=1}^{n}x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}=\sum_{x=1}^{n}\frac{n!}{(x-1)!(n-x)!}p^x(1-p)^{n-x}=np\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!(n-x)!}p^{x-1}(1-p)^{n-x}\qquad(2.179)$$

The term with the summation starts to resemble a binomial pdf based on $n-1$ trials. We introduce the variable $Z=X-1$ then:

$$E(X)=np\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!(n-x)!}p^{x-1}(1-p)^{n-x}=np\sum_{z=0}^{n-1}\frac{(n-1)!}{(z)!(n-1-z)!}p^z(1-p)^{n-1-z}\qquad(2.180)$$

Notice that the term inside the summation is the binomial pdf of Z based on $n-1$ trials and hence it must sum to 1 which concludes the first part of the proof $E(X)=np$

For the variance let us start with computing $E\left[X(X-1)\right]=E(X^2)-E(X)$ and then $E(X^2)=E\left[X(X-1)\right]+E(X)$:

$$E\left[X(X-1)\right]=\sum_{x=0}^{n}g(x)f_X(x)=\sum_{x=0}^{n}x(x-1)\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\qquad(2.181)$$

Similarly to the procedure above, the first two terms of the summation ($x=0$ and $x=1$) equal zero, hence:

$$E\left[X(X-1)\right]=\sum_{x=2}^{n}\frac{n!}{(x-2)!(n-x)!}p^x(1-p)^{n-x}\qquad(2.182)$$

We follow next a procedure similar as $E(X)$ and factorise the term $n(n-1)p^2$:

$$E\left[X(X-1)\right] = n(n-1)p^2 \sum_{x=2}^{n} \frac{n-2!}{(x-2)!(n-x)!} p^{x-2}(1-p)^{n-x} \tag{2.183}$$

We introduce the variable $Z = X - 2$ then:

$$E\left[X(X-1)\right] = n(n-1)p^2 \sum_{z=0}^{n-2} \frac{n-2!}{(z)!(n-2-z)!} p^{z}(1-p)^{n-2-z} \tag{2.184}$$

As before, the term inside the summation is the binomial pdf of Z based on $n-2$ trials and hence it must sum to 1. We have thus, $E\left[X(X-1)\right] = n(n-1)p^2$ and consequently $E(X^2) = E\left[X(X-1)\right] + E(X) = n(n-1)p^2 + np$. Finally $\varnothing$, which concludes the proof.

Chapter "Risk analysis and risk evaluation"
Author: S.N. Jonkman

Parts of this chapter are based on the publication Cur 190 "Kansen in de civiele techniek" (CUR, 1997; also the most recent version of 2015)

Contents

Learning objectives of this chapter:

- Understand the basic concepts of risk, risk analysis, decision analysis
- Be able to assess and quantity the risks for a (simplified) system with different risk metrics (individual, societal risk / FN curve, economic risk)
- Be able to apply (simplified) cost benefit analysis, decision analysis and economic optimization for engineering projects
- Be able to quantify understand, apply and derive safety standards for individual and societal risk

3 Risk Analysis and risk evaluation

This section deals with the analysis and evaluation of risks. Firstly, the concepts of risk (section 3.1) and risk analysis (3.2) are introduced. The sections after that introduce approaches for risk evaluation: general decision theory (3.3), cost benefit analysis (3.4) and safety standards that focus on the risk to life (3.5). The relationship between safety standards and engineering design is indicated in section 3.6.

3.1 Risk

3.1.1 General definition of risk

Almost all activities in life are characterized by some level of risk. Especially in the design and management of engineering systems risk and safety are key concepts and need to be taken into account explicitly. In social discussions no unambiguous meaning is assigned to the concept of risk. Two definitions given in the Oxford dictionary are: 1) a situation involving exposure to danger; 2) the possibility that something unpleasant or unwelcome will happen.

The first definition focuses on the consequences, the second on the possibility or probability. Quantifying and evaluating risks based on merely the probabilities or consequences is less realistic. For example, the risk of losing € 100 with a probability of 50% is different than the risk of losing € 1000 with the same probability. Also, the risk associated with losing a given sum of money will depend on the probability of the event.

An often-used definition considers risk as expected value:

Risk is the probability of an undesired event multiplied by the consequences

The unit of risk now depends on the units of probability and consequences. The probability of an event is generally expressed as the probability per unit time, for example per year. The consequences of an undesired event are often multi-dimensional, i.e. they can consist of different types of consequences, such as material, ecological damages, injuries and fatalities (see section 3.2.2 for further details). In many applications in engineering consequences are expressed by means of a monetary value. The unit of the risk (or expected value $E(d)$) then becomes € per year. For a case with one event scenario i with probability p_i it yields:

$$E(d) = p_i d_i \tag{3.1}$$

A more general definition of risk has been given by Kaplan and Garrick (1981):

Risk is a set of scenarios (s_i), each of which has a probability (p_i) and a consequence (d_i)

This definition of Kaplan and Garrick allows the use of various so-called risk metrics (or risk measures) to quantify or depict the risk. The expected value of the damage for a set of multiple discrete scenarios $\{1,\ldots,n\}$, can be expressed as:

$$E(d) = \sum_{s_i=1}^{n} p_i d_i \tag{3.2}$$

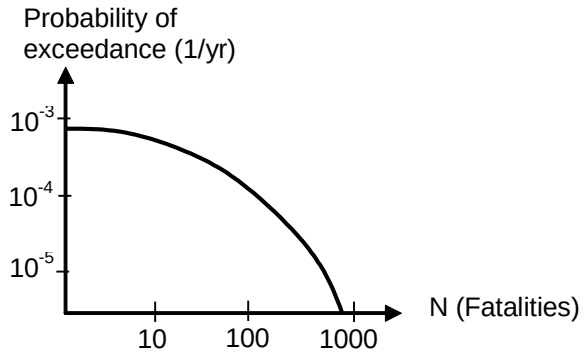


Figure 3.1 FN curve, showing the probability of exceedance of a certain number of fatalities N on Log-Log scale.

The expected value does not give insight in the magnitude of probability and consequences and the contribution of individual scenarios. Therefore, an often-used alternative risk metric is the risk curve. It shows the probability of exceedance of a certain magnitude of consequences. A well-known example of such a risk curve is the FN curve, which displays the probability of exceedance of N fatalities. The values on both axes are generally shown on a logarithmic scale, see Figure 3.1 for an example. The FN curve was originally introduced for the assessment of the risks in the nuclear industry (Farmer, 1967; Kendall *et al.*, 1977) and is now used to display and limit risks in various countries and sectors. Further information on the use of FN curves and a simple example of how to construct such a curve are included in section 3.5.1.

3.1.2 Other risk definitions

In the remainder of this lecture notes and the course the definitions from the previous paragraph will be applied. Since (civil) engineers will work in a broad domain of applications it is useful to highlight some risk concepts used in other domains.

Within economics, risk is generally associated with a deviation from the expected return or the probability of loss. In social sciences risk is often considered as a contextual notion or social construct. Vlek (1996) has summarized 11 formal definitions used in social sciences, see Table 3.1. In some of these definitions (e.g. numbers 2 and 4) the perceived seriousness of the undesired consequences plays an important role. Examples of other, more informal risk definitions used in psychology are "the lack of perceived controllability", "set of possible negative consequences" and "fear of loss" (Vlek, 1996).

Substantial research has focussed on the factors that determine the perception of risk (e.g. Slovic, 1987, Vlek, 1996). Examples of factors that influence risk perception include: the degree of damage, the controllability of and familiarity with the hazards, the extent of benefits from an activity, and voluntariness of exposure.

Table 3.1 Formal definitions of risk used in social sciences (Vlek, 1996)

1	Probability of undesired consequence
2	Seriousness of (maximum) possible undesired consequence
3	Multi-attribute weighted sum of components of possible undesired consequences
4	Probability × seriousness of undesired consequence ('expected loss')
5	Probability-weighted sum of all possible undesired consequences ('average expected loss')
6	Fitted function through graph of points relating probability to extent of undesired consequences
7	Semi variance of possible undesired consequences about their average
8	Variance of all possible consequences about mean expected consequence
9	Weighted sum of expected value and variance of all possible consequences
10	Weighted combination of various parameters of the probability distribution of all possible consequences
11	Weight of possible undesired consequences ('loss') relative to comparable possible desired consequences

The definitions applied in the research on natural hazards, often define risk in terms of more qualitative concepts such as hazard, vulnerability and exposure[1].

- Hazard: A dangerous phenomenon, substance, human activity or condition that may cause loss of life, injury or other health impacts, property damage, loss of livelihoods and services, social and economic disruption, or environmental damage.
- Vulnerability: The characteristics and circumstances of a community, system or asset that make it susceptible to the damaging effects of a hazard.
- People, property, systems, or other elements present in hazard zones that are thereby subject to potential losses.

3.2 Risk Analysis

3.2.1 General

The previous section made it clear that risk is a function of probabilities and consequences. The risk analysis therefore consists of an analysis of probabilities and consequences of undesired events in a given system. Alternative terms used in literature are risk assessment and quantitative risk analysis (QRA).

A risk analysis is carried out because involved parties (e.g. designers, managers, decision makers) want to identify and evaluate the risks and decide on their acceptability. Outcomes of risk analysis can be used in the design process to decide on the required safety levels of new systems (e.g. a new tunnel) or to support decisions on the acceptability of safety levels and the need for measures in existing systems (e.g. a flood defence system). A quantitative measure of some form is needed to transfer decisions on acceptable safety into a technical domain (Voortman, 2004). Examples are choices in the design of civil structures, such as the height of a flood defence or the strength of a building. Also, risk analysis can be used to analyse the effectiveness of risk reduction measures, incl. management and maintenance strategies. Overall, the risk analysis aims to support rational

[1] UN ISDR (United Nations Office of Disaster Risk reduction) Terminology, http://www.unisdr.org/we/inform/terminology accessed 26-1-2015

decision-making regarding risk-bearing activities (Apostolakis, 2004). Moreover, a risk analysis provides insight in mechanisms of system failure and the associated failure probabilities and consequences. As such, it can also serve as a tool of communication and management. Insights of the risk analysis can be used to optimize system design and management and often there is a direct link to quality assurance.

3.2.2 Elements in a risk analysis

In general the following elements can be identified within risk analysis, see the scheme in Figure 3.2[2] (Based on (CUR, 1997; CIB, 2001; Faber and Stewart, 2003; Jongejan, 2008):

- System definition and setting the scope and objectives of the analysis;
- Qualitative analysis of undesired events;
- Quantitative analysis of the risk;
- Risk evaluation (of the acceptability of the risk)

Some more information on these steps is given in the following paragraph. In addition to the steps in risk assessment of a given system, risk management also includes the element 'risk reduction and control'. Dependent on the outcome of the former phase measures can be taken to reduce the risk. This will lead to changes in the system configuration and the risk level. If the risk analysis is used in the design of systems, the steps are often repeated several times with adjusted system specifications to obtain an optimal design. It should also be determined how the risks can be controlled, for example by monitoring, inspection or maintenance.

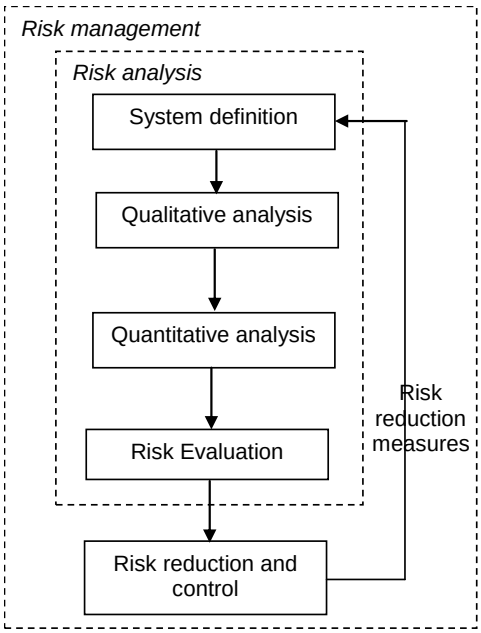


Figure 3.2 Schematic view of steps in risk assessment and risk management.

It is noted that a (probabilistic) risk analysis is different from a (deterministic) scenario analysis. A risk analysis is based on quantitative analysis of all (known) undesired events and their probabilities and consequences. A scenario analysis considers one (or a limited number) of design scenarios, often without considering its failure probability. Both approaches are complementary, as a scenario analysis considers one of the scenarios from the risk analysis.

In the Netherlands, scenario analysis is often used for evaluating disaster preparedness and designing the capabilities of emergency management services. For example, in the design of the Green Heart Tunnel which is a part of the high speed rail line in the Netherlands, a single accident scenario was chosen to design the emergency exits out of the tunnel. Analysing a high impact, low probability scenario in detail may provide helpful clues about the effectiveness of particular safety measures or crisis management actions. While

[2] A more detailed framework of the steps of a risk analysis is included in appendix *.

such analyses may be helpful, it would be wrong to assume that such high-impact, low probability scenarios should always be properly manageable by emergency responders. This is because the probability of harm should be an integral part of any cost-benefit or cost-effectiveness analysis (Jongejan et al., 2011).

3.2.3 Steps in Risk management

This section gives some more information on the steps in the risk analysis and some fundamental concepts in risk analysis.

1. System definition

This step entails the definition and description of the system as well as the scope and objectives of the analysis. The process or system under consideration can usually be described as a so-called input-output element. Here the system is assumed to be failing if no output takes place. Usually a system is divided into components and subsystems, which can all be schematised as an input-output elements. By means of the internal relations the components and subsystems together form a configuration that is representative of the total system. Further information on the decomposition and analysis of system is provided in section 3.5 of these lecture notes.

A system can be represented in terms of physical components, organizations and users, and an external environment. In order to analyse failure and risks, not only physical components, but also organizations and operators and users need to be considered (see e.g. Bea, 1998). Different groups of organizations and persons will be involved in different roles. There are the professionals responsible for the operations and management of the system (e.g. the pilots and crew in an aircraft), potential users of a system (passengers), and external parties (people living near the airport exposed to risk and noise). Each of this group has a different "relationship" and attitude towards the risk and this could affect its acceptability. For example, a higher risk might be acceptable for pilots and crews in an aircraft (who have a direct benefit) than for regular passengers. Finally, the external environment (e.g. wind or waves) will determine the loads on the system and affect the potential failure mechanisms.

2. Qualitative analysis

In this step, potential hazards, undesired events, failure mechanisms and scenarios are identified and described. An important goal of this phase is to gain insight, as complete as possible, into all possible undesired events and their consequences. For most systems, multiple undesired events can be distinguished. For example, two events with different impacts that can both lead to flooding of a polder are 1) the inflow of large amounts of water due to a dike failure; 2) the inflow of smaller amounts of water when a sluice gate is not closed.

When a system or part of it no longer fulfils one or more desired functions, this is known as failure. It means that the state of the system changes from normal operation of failure. The state of failure can be reached through different failure mechanisms (or failure modes). For example, a dike can fail due to overtopping, but also due to geotechnical failure mechanisms such as instability or piping. A **limit state** is a condition of a structure beyond which it no longer fulfils the relevant design criteria (Eurocode, 2001). In practice two types of limit states are

distinguished: the serviceability and the ultimate limit states abbreviated as SLS and ULS respectively.

- In the case of the exceedance of ultimate limit state (ULS) failure or collapse of a system or structure occurs. This, for instance, occurs if the breakwaters of a harbour entrance are destroyed as a result of extreme conditions. An example from structural engineering is the collapse of a roof of a building
- In the case of exceedance of the serviceability limit state (SLS) exceedance leads to temporary and/or partial failure. An example of this state is the non-workability in a harbour, due to waves that are temporarily too high. Another example could be too much vibration of a structure, so that the users experience discomfort.

Further information on the use of the SLS and ULS concepts in civil engineering is included in chapter 10.

In a risk analysis, it is very important to get an overview of the various undesired events and failures before proceeding with a quantitative analysis. In practice, many accidents are caused by failing to identify failure modes.
Finding a list of threats and failing modes that is as complete as possible, is not an easy task. Aids are data banks, literature studies, interviews, experiences with comparable systems, brainstorm sessions et cetera. Techniques for systematically identifying undesired events (e.g. FMEA (Failure Modes and Effects Analysis) are treated in more detail in section 9.4 of the chapter on reliability of systems.

3. Quantitative analysis

The probabilities and consequences of the defined undesired events are determined in this step. The risk is quantified in a risk number or graph as a function of probabilities and consequences (see section 3.5.2 for an example).
The probability of failure can be quantified using the (previously) introduced limit state. A limit state *Z* can be assessed by considering the resistance R and the loads S, i.e.

$$Z = R - S \tag{3.3}$$

Failure occurs when $R < S$, so when $Z < 0$. Techniques for computing the probability of failure, i.e. $P(Z < 0)$, are treated in more detail in later sections of these lecture notes.

After failure has been defined and analysed, the consequences of the event are quantified. First, the physical effects associated with the undesired event have to be considered, e.g. heat and / or smoke from fire, or inflow of water due to dike breach. Depending on whether people or objects are exposed to the physical effects, damages, life loss or other impacts can occur.
As an example the failure of a dike for a set of discrete events is considered:

- The probability that a dike fails, $P(E_1)$
- The conditional probability that water flows into the polder given a dike breach $P(E_2 \mid E_1)$
- The probability of damage given dike breach and inflow into the polder $P(D \mid E_1 \cap E_2)$

The probability of damage can now be computed by combining these terms

$$P(D) = P(E_1)\ P(E_2 \mid E_1)\ P(D \mid E_1 \cap E_2) \tag{3.4}$$

As introduced in section 3.1, multiple types of consequences can be caused by one disaster. Table 3.2 gives an overview of the different types of consequences of the failure of large engineering systems. The damage is divided into tangible and intangible damage, depending on whether or not the losses can be assessed in monetary values. Another distinction is made between the direct damage, caused by physical effects of the event, and damages occurring outside the directly exposed area. The latter occurs when companies outside a flooded area experience damages, due to loss of demand from customers in the flooded area. In a risk analysis it is desired to take into account a complete set of impacts. Since a lot of the items from the table cannot be quantified easily, the quantitative analysis and risk evaluation are often focused on economic damages and life loss.

Table 3.2 General classification of damages, based on (Vrouwenvelder and Vrijling, 1996)

	Tangible	**Intangible**
Direct	• Residences • Structure inventory • Vehicles • Agriculture • Infrastructure and other public facilities • Business interruption • Evacuation and rescue operations • Reconstruction of flood defences • Clean up costs	• Fatalities • Injuries • Animals • Utilities and communication • Historical and cultural losses • Environmental losses
Indirect	• Damage for companies outside the exposed area • Substitution of production outside the area • Temporary housing of evacuees	• Societal disruption • Damage to government

4. Risk evaluation

In the risk evaluation phase the decision is made whether the risk is acceptable or not and whether risk reduction measures need to be implemented. Or in other words, it is attempted to answer the question "how safe is safe enough?" (Starr, 1967). The results of the quantitative analysis provide input for risk evaluation and decision making.

Different quantitative approaches can be used to support risk evaluation, which will be outlined more in detail in the coming sections.

- **Decision making under uncertainties** (Section 3.3): Recording different variants, with associated risks, costs and benefits, in a matrix or decision tree, serves as an aid for making decisions. With this, the optimal selection can be made from a number of alternatives.
- **Cost benefit Analysis** (section 3.4.1): the costs and benefits of risk reduction measures are considered. When a very large number of design choices are possible, an **economic optimization** (section 3.4.2) can be applied to select an optimal system design, based on costs and benefits of risk reduction.
- **Safety standards** (section 3.5): Comparing the risk with predetermined safety standards which often focus on loss of life.

However, given the nature of the key question (how safe is safe enough?) several political, psychological and social processes play a role in the evaluation of risk. This means that risk evaluation is not a purely technical process, but will involve many subjective elements. One of the difficulties facing regulators is that people's preferences and risk attitudes may diverge and that costs and benefits may not be distributed evenly. This means that a single, collective decision, in practice, has to be based on strongly divergent individual preferences. In practice, this implies that devising collective decision making procedures is inevitably political. This ambiguity can also be found in the numerous interpretations of "the" precautionary principle, which is interpreted by some as a decision making criterion that requires proof of harmlessness (a scientific impossibility), whereas it is seen by others as a decision making procedure that puts emphasis on dialogue and stakeholder involvement, e.g. Jongejan (2008).

Risk reduction and risk control

If the risks are considered unacceptable several forms of risk reduction can be implemented. These can be changes to the engineered system, or changes to the organization and management. From analysis of accidents it appears that human and organizational errors are still a major cause of failure in civil engineering. It seems that the only suitable way to reduce human errors is by the incorporation of sufficient control in the different phases of the construction process (Taerwe, 1986) and by a thorough education of all personnel involved. Therefore, an extensive interaction between the safety methodology and the quality management is a necessity in order to guarantee the safety of our structures.

3.3 Decision-making under uncertainties

Decision-making under uncertain conditions is part of everyday life, e.g. when choosing to buy a lottery ticket or choosing to take an umbrella during cloudy weather. In contrast to the rather intuitive decision making in everyday matters, a structured analysis of different alternatives with associated risks, costs and benefits is very useful for decisions in (civil) engineering. This chapter offers a very basic introduction into the decision theory with applications to decision problems in the civil engineering domain. Further reference is made to the work by other scholars for more rigorous and detailed treatment of this topic, see for example Raiffa and Schlaifer, (1961); Benjamin and Cornell, (1970).

3.3.1 Introduction and basics

Making a decision is in fact choosing from alternatives. The decision theory[3] is based on the classic "Homo Economicus" model. The Homo Economicus:

- has complete information about the decision situation;
- knows all the alternatives;
- knows the existing situation;
- knows which advantages and disadvantages each alternative provides, be it in the form of random variables;
- strives to maximise that advantage (formally called utility).

[3] For the basis of decision theory, see for example (Von Neumann and Morgenstern, 1953).

The decision-making concept discussed in these lecture notes assumes this model. Decision-making in practice is often different since the above conditions will not be fulfilled. There can be multiple decision makers and multiple objectives. Also, the decision maker does not know all the alternatives or their outcomes. For many practical cases this has led to an extension of the decision model, but not to a fundamental adjustment of the classical model.
Within a decision problem the following characteristics can be distinguished:

- the set of all possible actions or decisions (a), from which the decision maker can choose
- the set of all (natural) circumstances (θ) that influence the outcomes
- the set of the set of all possible results (ω), which are functions of the actions and circumstances: $\omega = f(a,\theta)$.

The actions, natural circumstances and the outcomes can be shown in a so-called decision tree (Figure 3.3).

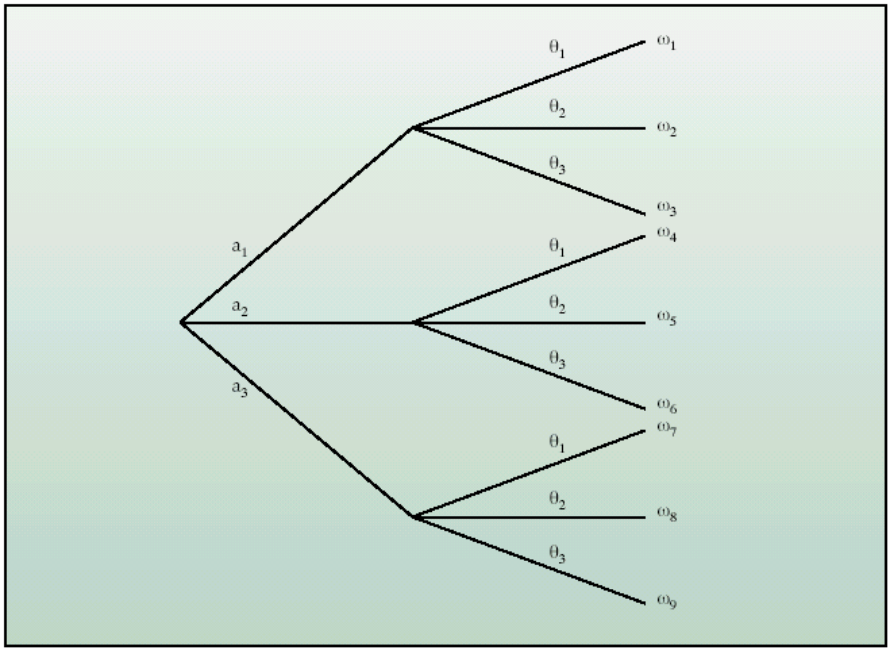


Figure 3.3 Decision tree

Based on the possible results a choice is made for an action. To be able to assess the different results, a numerical value is assigned to each outcome of ω, which can be used to establish the benefit of each outcome. This number can be a monetary value, a number on an arbitrary scale or utility - as long as the decision maker(s) can establish a consistent ranking of the outcomes with it. In the last two cases the benefit has no absolute value, but only gives the relative value of the different outcomes. Utility is a concept used to rank the possible outcomes according to the preferences of the decision maker. Utility (u) values are between $0 \leq u(\omega) \leq 1$. A utility function can be used to characterize the relative utility of various outcomes. The elaborations below are based on the monetary values as a measure for the outcomes and assume a risk neutral decision maker. This is a decision maker who is indifferent between choices with equal expected outcomes, even if one choice is riskier than the other. For example, a risk neutral decision maker would have the same preference for a € 400 pay out, or a 50/50 bet with a coin toss with outcomes of € 0 (head) or € 800 (tail). Utility and risk aversion are further discussed in section 3.3.3

3.3.2 Decision rules

Once a set of actions, circumstances and outcomes is known, various approaches can be used to come to a preferred decision. Various deterministic decision rules are available which do not take into account the probabilities of the possible circumstances and outcomes. One example of such a decision rule is the *minimax* criterion: a decision maker wants to minimize maximum losses. This is in fact a risk-averse criterion. Another example is the maximax criterion: a decision maker chooses the option with the maximum income and is risk seeking.

Although these decision rules are helpful in some cases, the probability of occurrence of certain circumstances is a key feature of the decision problem. Information on the probability of outcomes is needed for an optimal choice of action(s). For example, when making a decision to start a business in soup or ice cream, the decision maker would want to know what the probabilities of rainy or sunny weather are. Selling ice cream in Dutch winter will probably not make a good (expected) profit, but it would be a profitable business in a Mediterranean summer.

Therefore it is necessary to include information on the probabilities of circumstances and outcomes, in order to determine a rational action with the highest expected value of the benefit. This theory is known as the Bayesian decision theory. In a probabilistic or Bayesian decision framework the optimal action a^* is defined as the one maximizing the expected utility. The following formula is found for the case with discrete outcomes.

$$a^*: \quad \max_a E\big(u(a,\theta)\big) = \max_a \sum_\theta u(a,\theta) P(\theta_i) \tag{3.5}$$

In which $u(a,\theta)$ – utility of action a under circumstance θ. $P(\theta)$ is the probability that circumstance θ_i occurs.

A rational decision is choosing the action with the highest expected (utility) value or highest benefit if outcomes are expressed in monetary values. This is illustrated in the example below. Note that other examples in these lecture notes will also be based on monetary values.

Example 3.1: buying shares or bonds?

Suppose a person has EUR 1000 at his or her disposal and is given the choice to invest this money in bonds or in shares of a given company. It is known that, on a yearly basis, 3 % of the current market value is distributed as interest on the bonds. The dividend of the shares depends on the company's profit. Suppose that the board of directors have made the following agreements:

- for a profit smaller than 5 % of the shareholders capital, no dividend is paid;
- for a profit larger than 5 % of the shareholders capital, dividend is paid, the percentage of which corresponds to 3 % of the current market value of the shares;
- for a profit larger than 10 % of the shareholders capital, the dividend corresponds to 6 % of the current market value of the shares.

The set of actions A has two elements: a_1 = investing in shares AND a_2 = investing in bonds

The set or market circumstances N has three elements, namely:

θ_1 = company profit $\leq$5 %

θ_2 = 5 % < company profit ≤ 10 %

θ_3 = company profit > 10 %

Assume the inflation amounts to 2 %. The set of outcomes ω contains three possible outcomes for the shares:

ω_1 = return (0 % - 2 %) = -2 % per annum

ω_2 = return (3 % - 2 %) = 1 % per annum

ω_3 = return (6 % - 2 %) = 4 % per annum

Note that for the bonds the net outcome always yields ω_2=1% (i.e. 3% interest – 2% inflation). The outcomes can be shown in a decision tree (see Figure 3.4) or in a table (see Table 3.3).

Table 3.3 Outcomes given the decisions (a_1,a_2) and market conditions (θ_1,θ_2,θ_3).

	Market circumstances		
Actions	θ_1	θ_2	θ_3
a_1 : buy shares	-2 %	1 %	4 %
a_2 : buy bonds	1 %	1 %	1 %

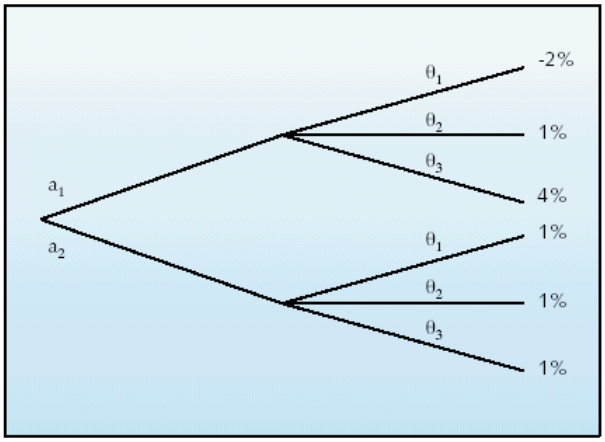


Figure 3.4 Decision tree for the example of buying shares (a_1) or bonds (a_2).

The deterministic decision rules can be applied to this example. Minimax would result in investing in bonds (a_2), maximax would result in buying shares.

The optimal decision can be found by taking into account the probabilities of the market circumstances. These three circumstances are assumed to be exhaustive and mutually exclusive (i.e. outcomes cannot overlap and the sum of probabilities equals 1). The probabilities are estimated at $P(\theta_1) = 0.2$; $P(\theta_2) = 0.3$; $P(\theta_3) = 0.5$. These probabilities can now be included in the decision tree. The expected value of the return of the actions is as follows:

Buying shares: 0.2(-2 %) + 0.3 (1 %) + 0.5 (4 %) = 1.9 %.

Buying bonds: 1%

In this case the expected outcome is larger for buying shares than for buying bonds. So for a risk neutral decision maker buying shares (a_1) would be the preferred action. Note that this action also includes a probability of 0.2 of a loss. This is also expressed by a higher standard deviation of the expected outcomes for buying shares. The above example can also be extended by applying

different utility functions for various outcomes.

In the previous example, the number of circumstances is limited and the probability distribution of the circumstances is discrete. For many decision problems this is not the case. The state of nature, for instance, can assume many values that cannot be made discrete. This, for example, would have been the case if the dividend in example 0 had been a percentage of the profit. In such cases a probability density function can be used to characterize the spectrum of outcomes. Using a continuous form of formula (3.5), the expected value of various actions, and the optimal action / decision can be identified.

In taking decisions with uncertainties, it appears that probabilistic calculation techniques are a valuable aid to reach a rational choice. This is particularly the case if risks are dependent on the possible decisions. In such cases, Bayesian decision theory minimizes the total costs (i.e. investment costs plus risk in terms of potential losses). This can best be illustrated by means of an example from the civil engineering domain.

Example 3.2: drainage of a construction site – decision tree

In a river polder a basement has to be built in an excavated construction site. The construction site is made of sheet piling and the bottom is sealed off with a clay layer with a thickness (d) of, on average, $\mu_d = 2.5$ m. The thickness is not known exactly; it follows from measurements that the thickness has a normal distribution and a standard deviation of $\sigma_d = 0.2$ m.

The river cuts through the clay layer and the underlying sand layer is fed by the river (see Figure 3.5). The groundwater potential in the upper layer equals the potential in the deep sand layer. The upward water pressure under the sealing layer is assumed to be a direct function of the river levels. The fluctuating river levels result in fluctuations of the upward pressure under the sealing layer.

Measurements of the groundwater levels over a long period have given an insight into the extreme groundwater levels. The maximum upward pressure (h) under the sealing layer in the construction period has an a normal distribution with an expected value of $\mu_h = 4$ m water column and a standard deviation of $\sigma_h = 0.75$ m water column.

From these values the probability of flooding due to bursting of the clay layer due to upward water pressure can be calculated for the construction period. We define a limit state function $Z = R - S$.

R is the strength consisting of the weight of the clay layer and S is the water pressure. We find:

$$Z = \rho_c d - \rho_W h \tag{3.6}$$

In which: ρ_c – density of clay (=20kN/m^3); ρ_w – density of water (=10 kN/m^3)

The probability of failure $P(Z<0)$ for this situation can be found by calculating the mean and standard deviation of Z:

$$\mu_Z = \rho_c \mu_d - \rho_W \mu_h = 20\ 2.5 - 10\ 4 = 10 \text{ kN/m}^2$$

$$\sigma_z = (\rho_c^2 \sigma_c^2 + \rho_w^2 \sigma_h^2)^{0.5} = 8.5 \text{ kN/m}^2 \tag{3.7}$$

According to Chapter 2.6.1, we find $P(Z<0) = \Phi(-\mu_Z/\sigma_z) = \Phi(-1.17) = 0.12$. This is indicated as

the initial failure probability $P_{f,0}$.

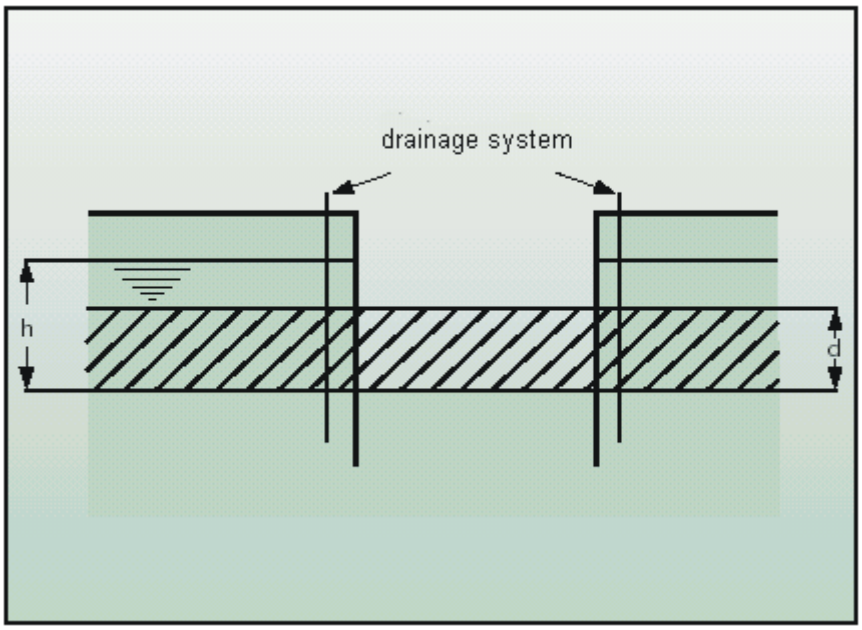


Figure 3.5 Situation: Excavation near a river

The effect of a drainage system in the construction site (see Figure 3.6) on the groundwater levels has been reviewed using groundwater flow calculations. It appears that it reduces the mean value of the maximum water levels to $\mu_h = 3.52$m. and the standard deviation remains the same. In this case the failure probability is reduced to 0.04. Such a drainage system costs EUR 150,000.

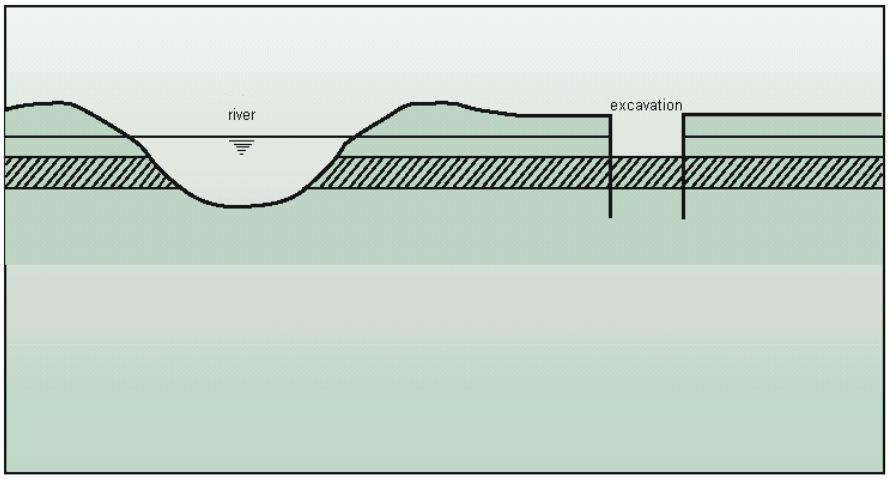


Figure 3.6 Drainage around the excavation.

The flooding of the construction site will result in the buoyancy of the basement resulting in damages are estimated at EUR 5,000,000. The designer of the construction site is faced with the choice whether or not to include drainage facilities in the design of the construction site.

To provide insight the decision problem is illustrated with a decision tree. For this, the sets of actions, circumstances and outcomes have to be defined first:

The set of actions A consists of:

a_1 = excavation without drainage

a_2 = excavation with drainage

The set of circumstances N is formed by:

θ_1 = the sealing layer remains intact

θ_2 = the water pressure exceeds the weight of the sealing layer

The set of outcomes Ω consists of:

ω_1 = nothing happens; loss = € 0

ω_2 = the construction excavation is flooded: loss =€5,000,000

The previous eligibility analysis has shown that the probability of flooding of the excavation equals $P_f = 0.12$ for a situation without drainage and $P_{fd} = 0.04$ with drainage.

Without drainage, the risk, defined as the expected value of the loss, is: $0.12 \cdot$ € 5,000,000 = € 600,000. With drainage the risk is: $0.04 \cdot$ € 5,000,000 = € 200,000.Costs and probabilities can also be shown in the decision tree (see Figure 3.7). The expected values of the costs can be calculated for the different actions by adding the present values of the cost of actions and risk:

a_1 : expected value (additional) costs = risk =€ 600,000

a_2 : expected value (additional) costs = extra costs + risk = €150,000 + € 200,000

= € 350,000

This implies that the construction of the drainage system is rationally speaking the best decision for a risk neutral decision maker.

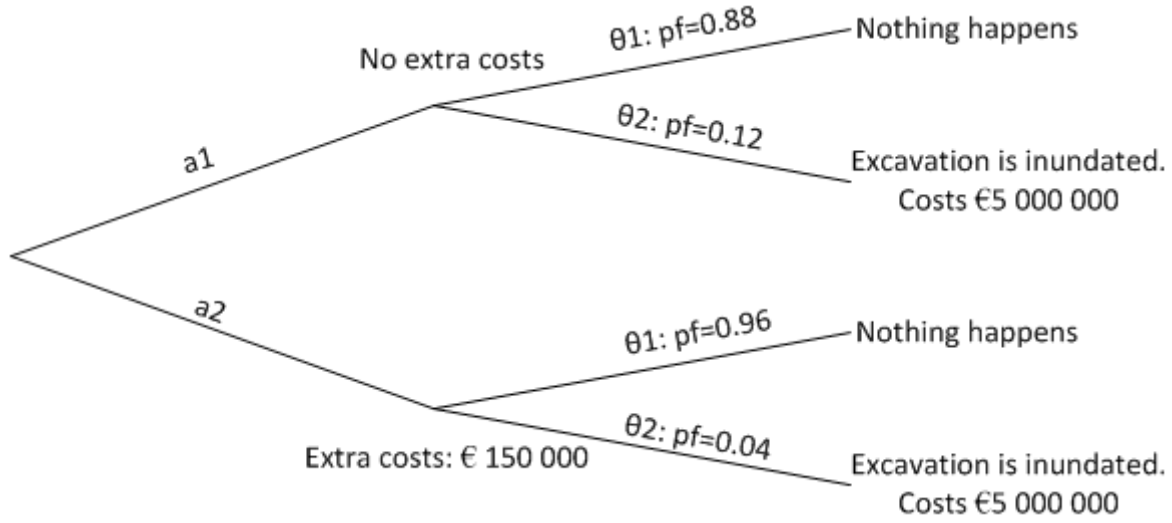


Figure 3.7 Decision tree with probabilities and costs.

3.3.3 Utility functions

The elaborations in the previous sections were based on monetary values as a metric for outcomes and a risk neutral decision maker. This section will provide some basic information on utility functions and various risk attitudes.

Utility is a concept used to rank the possible outcomes according to the preferences of the decision maker. Utility (u) values are between $0 \leq u(\omega) \leq 1$. A utility function can be used to characterize the relative utility of various outcomes. The utility function depends on the preferences and attitude of the decision maker. An example of three different utility functions is given in Figure 3.8 for an example of an activity with potential monetary benefits between € 0 and

€ 1000. In the case of a risk neutral decision maker the utility increases linearly with the benefits. In the case of the risk aversion relatively small benefits are already given a high utility. A risk seeking decision maker gives a relatively low value to smaller benefits, but high utility values to larger incomes.

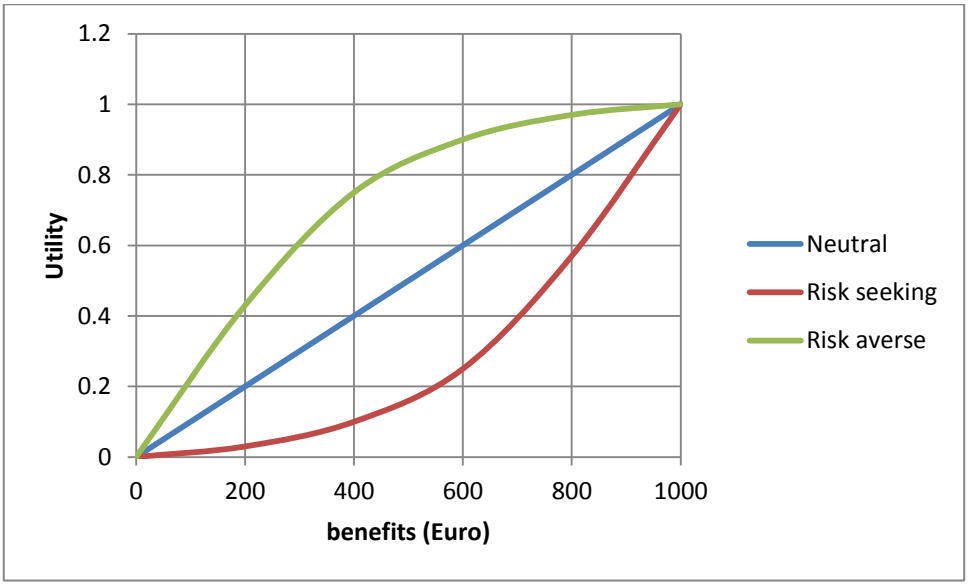


Figure 3.8 Example of a utility function for monetary benefits.

Example 3.3: a bet

The utility functions can be applied to a simple example. Assume that a student has the chance to receive a guranteed payout of €400. Alternatively, a bet can be organized with a $p = 0.5$ chance of a payout of € 800, and a $(1-p) = 0.5$ chance of no payout. A risk neutral decision maker would be indifferent between both choices, as they have the same expected outcome – i.e. €400 gain – and a utility value of $u(€400) = p \cdot u(€800) + (1-p)u(€0) = 0.4$.

We now apply the two other utility functions. For the risk averse function u_{RA} we find the following:

The expected utility of the direct payout equals $u_{RA}(€400) = 0.75$

The expected utility of the bet becomes: $0.5\, u_{RA}(€800) + 0.5\, u_{RA}(€0) = 0.5 \cdot 0.97 + 0.5 \cdot 0 = 0.485$

In this case the expected utility of the direct payout is higher and this is the preferred decision of the risk averse decision maker. For the risk seeking function u_{RS} we find the following:

The expected utility of the direct payout equals $u_{RS}(€400) = 0.1$

The expected utility of the bet becomes: $0.5u_{RS}(€800) + 0.5u_{RS}(€0) = 0.5 \cdot 0.57 + 0.5 \cdot 0 = 0.285$

In this case the expected utility of the bet is higher. The preferred decision of the risk seeking decision maker is the bet.

So for the example of the bet, the risk neutral decision maker is indifferent between the bet and the direct payout. The risk averse decision maker would accept a lower payout, rather than taking the bet. The risk seeking decision maker accepts the bet, even if the guaranteed payout is higher than the expected outcomes of the bet. The same concept can also be applied to losses, a decision-maker can be risk averse against events with large consequences. For example, a civil engineering company could be risk averse in making decisions about bidding for projects with financial risks that could threaten the financial stability of the company. On the other hand, a government with a large portfolio of projects may act more risk neutral. The various risk attitudes can also play a role in making investment decisions (see also 3.4.2). A risk averse investor would prefer investments (such as bonds or guaranteed loans with relatively low expected returns) over investments in stocks that have higher expected returns, but also a higher chance of losing money. The principle of risk aversion is also related to decisions about insurances. For example, most people are risk averse against losing their belongings in a large house fire. They are willing to pay an annual premium that is higher than the expected losses due to the fire. Finally, a further discussion of a related risk aversion concept regarding accidents with large numbers of fatalities is included in section 3.5.

3.4 Cost Benefit Analysis and economic optimization

3.4.1 Cost benefit analysis

This section deals with simplified cost benefit analysis for risk reduction interventions in the engineering domain. An important question in evaluating (engineering) projects is whether the benefits outweigh the costs. Cost benefit analysis (CBA) is generally used for appraisal of a wide range of effects of projects or interventions in order to support decision making. The cost benefit analysis starts with defining the system and existing situation. Then, a broad range of effects of the proposed project and intervention can be identified. Table 3.4 below shows an overview of effects of the Delta works that were built after the 1953 flood disaster in the Netherlands. The main aim of the delta works was to provide flood protection to the Southwest of the Netherlands. However, other effects included the agricultural benefits to the region (benefits) and the effects on environmental quality in the estuaries in which dams were built (costs or negative effects).

A choice has to be made which effects in the CBA are evaluated in monetary terms. For some items, such as construction costs or economic risk reduction (see below) this is straightforward. For other items such as environmental effects or reduction of risk to life, approaches for monetary evaluation exist, but they are not standardized or undisputed.

Table 3.4: Costs and benefits of the delta works (Don en Stolwijk, 2003)

Costs	**Benefits**
Construction and maintenance costs	Reduction of flood risk (damage, loss of life)
Loss of environmental quality	Improvement of infrastructure
	Recreation
	Agricultural benefits
	Economic stimulus for the region
	Economic stimulus for the water engineering sector
	National Pride

When considering investments that primarily focus on risk reduction (e.g. dike reinforcement or reinforcing of buildings for earthquakes), the main benefits will consist of the reduction of expected economic damages. For a measure to be cost effective, the investments should be smaller than the risk reduction.

$$I < \Delta E(D) \tag{3.8}$$

Where: I – investments [€]; $\Delta E(D)$ – present value of risk reduction [€]

This formula can also be used to calculated the benefit/cost ratio, i.e. $\Delta E(D)/I$. This ratio should be greater than 1 for cost effective projects. Note that all costs and the risk reduction are given in terms of present value with a unit of €.

For investments that focus on prevention (i.e. reducing the failure probability of the system) equation (3.9) can be formulated as follows.

$$I < \left(P_{f,0} - P_{f,new}\right)D \tag{3.9}$$

Where: $P_{f,0}$ – initial failure probability [-]; $P_{f,new}$ – failure probability after risk reduction investment

Example 3.4: Drainage of a construction site – costs and benefits

The example from the previous section could also be formulated as a simplified cost-benefit analysis. The same values for the variables are used. The costs of the drainage are € 150,000. The benefits are equal to the risk reduction, i.e.

$$\Delta E(D) = \left(P_{f,0} - P_{f,d}\right)D = (0.12 - 0.04)\ € 5{,}000{,}000 = € 400{,}000 \tag{3.10}$$

This shows that costs are smaller than the risk reduction benefits, i.e. $I < \Delta E(D)$. The benefit/cost ratio equals € 400,000/€ 150,000 = 2.66

One can easily show that the investment would just be acceptable if the benefits and costs are equal, i.e. when the benefit / cost ratio equals 1. This would still be the case if the probability of failure with drainage equals $P_{f,d} = 0.09$.

Other types of interventions do not affect the probability of an accident, but focus on reducing the damages. They are often indicated as mitigation. An example in the field of flood management concerns raising buildings instead of reinforcing the dikes. In this case the criterion becomes:

$$I < P_{f,0}\left(D_0 - D_{new}\right) \tag{3.11}$$

Where: D_0 – initial damage [€]; D_{new} – damage after investments in reducing the consequences [€]

The foregoing assumes that both the costs and benefits are expressed in the same unit, generally in the form of a present value [€].The nett present value represents the sum of the present values of the benefit and cost cash flows over a period of time. In engineering we often deal with situations with larger initial investments, whereas the risk reduction benefits are spread out over a longer

time period. For example, when a government invests this year in flood protection, the costs are made in year 0, but the benefits will be spread over the coming decades.
For such situations the failure probability is generally expressed per unit of time, mostly per year. That means that the risk (reduction) is expressed in terms of € per year, whereas the initial investments[4] have the unit of €. The nett present value of cost or benefit values over a future range of years can be calculated with formula (3.12). To calculate the nett present value NPV [€] a discount rate r should be used. The discount rate represents a required return on an investment.

$$NPV = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t} \tag{3.12}$$

Where: C_t– costs in year t [€]; T – reference period [years]
The contribution of costs in a certain year to the nett present value depend on the discount rate and the reference year. Costs or benefits closest to the present will have the greatest contribution. Figure 3.9 below shows the value of $1/(1+r)^t$ for a given year t for different discount rate. The higher the discount rate, the smaller the contribution to the nett present value of costs or benefits that are far away from the present. For an infinite time horizon it can be shown that $\sum 1/(1+r)^t \approx 1/r$. This approximation can be used in evaluating engineering projects with a long life time.

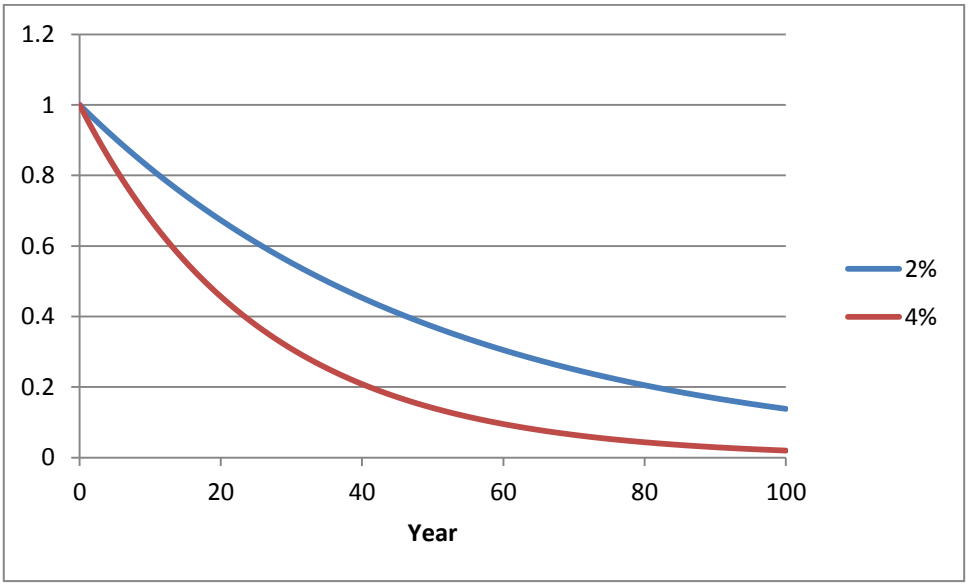


Figure 3.9 Normalized present value of a cost or benefit in year t for different discount rates.

Economic Valuation of human life

The previous paragraphs have focussed on the evaluation of economic costs and benefits of risk reduction interventions. Many of these interventions also directly focus on reducing injuries and fatalities. Examples are regulations and investments in traffic safety which have introduced measures such as airbags and seat belts.

[4] Note that there can also be yearly investments such as management and maintenance.

In literature on risk management the economic valuation of human life is often depicted as a difficult problem as it raises numerous moral questions. Some claim it is unethical to put a price on human life because life is priceless. The actual expenditures on risk reducing prospects show however that the investment in the reduction of risks to humans is always finite. Different approaches are available for evaluating the costs of interventions in relation to the reduction of risk to human life, see Vrijling and van Gelder, (2000); Jongejan et al., (2005) for a further discussion of the various approaches.
One of the options is to add the economic value d of N human fatalities to the economic damages, i.e. $D + N \cdot d$. The value of the number of lives lost can be determined with different approaches.

Several approaches are based on so-called stated preferences. A survey can reveal how much people are willing to pay, e.g. for safety measures. In these cases the value of statistical life (VoSL) is obtained from the willingness to pay expressed by respondents in surveys. For example, in the cost benefit analysis for the flood defences in the Netherlands a value of a statistical life of €6.7 million per fatality is used (Deltares, 2013). The Value of a Statistical life lost in traffic accidents is estimated at €2.6 million (SWOV, 2012).

One alternative approach is based on so-called revealed preferences. The costs of saving and extra (statistical) life (CSX) for actual life-saving interventions that have been taken in the past can be determined.

$$CSX = I / \Delta E(N) \tag{3.13}$$

Where: CSX – costs of saving an extra life [€/(life·year)]; $\Delta E(N)$ – reduction of the expected number of fatalities per year.

An extensive study on CSX values in various sectors, see Tengs et al. (1995) showed that these vary widely across sectors and even within sectors. The highest CSX values are found for risks for small probability – large consequence events, for example in the nuclear domain. For such cases the expected number of fatalities is already small and investments in incremental safety are large.
One other approach is to base the value of a human life on macro-economic indicators. Several metrics have been proposed that relate this value to a person's economic production.
Given the difficulties associated with economic valuation of human life and the associated risk reduction, it is decided in some domains to develop separate criteria for considering the risk to life. This topic is further elaborated in section 3.5.

3.4.2 Economic optimization

The previous sections have focussed on decisions for which the number of actions was limited, e.g. excavation with or without drainage and the associated costs and benefits. However, there are several situations in which the number of actions is unlimited. This occurs when the failure probability level has to be decided for a system that is yet to be designed, with an infinite number of design options. An example of this type of decision problem is the heightening of dikes, as in theory an unlimited amount of values can be chosen for the elevation, e.g. 2m, 5m, 5.1m, 5.11m, 5.1111m 6m etc.

For this situation an economic optimization that takes into account the costs of increasing the safety level and reducing the risks can be applied to derive an optimal level of safety (or the optimal "failure probability"). The economic optimization was developed and applied by van Dantzig (1956), to derive the optimal dike height for South Holland after the 1953 storm surge disaster, as will be further elaborated in the next section.

In the economic optimization the total costs (C_{tot} [€]) are determined, consisting of the investments I [€] in a safer system and the present value of the risk R [€].

$$C_{tot} = I + R \tag{3.14}$$

The annual risk, or expected economic damage is found by:

$$E(D) = P_f D \tag{3.15}$$

Where: $E(D)$ – expected value of the risk [€/yr]; P_f – failure probability of the system per year ([1/year]; D – damage in case of failure [€]

In this approach it is thus assumed that all damages are expressed in monetary terms. Additional criteria for separately considering the loss of human life are included in the next section.

The present value of the risk for an infinite time horizon can be found as follows:

$$R = \frac{P_f D}{r} \tag{3.16}$$

The risk can be reduced by constructing a safer system (a lower P_f), or limiting the damage (smaller D). In this case we assume prevention measures that focus on reducing the failure probability. The investments will become a function of the failure probability of the system, since increasing the safety will lead to an increase of costs.

$$I = I\left(P_f\right) \tag{3.17}$$

Figure 3.10 shows the costs and risks as a function of the accepted failure probability of a system. The economic optimum is found when the total costs are minimal. For this situation the following is valid:

$$\frac{dC_{tot}}{dP_f} = 0 \tag{3.18}$$

This approach can be applied to various decisions problems, such as the optimal dike height (see next section) but also the dimensioning of other interventions such as sprinklers or ventilation in tunnels to reduce fire risks. In cases where no continuous functions are available to create a figure like the one below, the analyst can consider a limited number of design options (e.g. no, small, medium or large sprinklers in a tunnel) and determine investments, risks and total costs for these options.

In addition to the determination of the optimal failure probability, the cost benefit criterion should still be verified. (It is possible that we find an economic optimum with higher total costs than in

the current situation without interventions). It is checked whether the benefits of risk reduction are larger than the costs of the dike heightening, i.e. $\left(P_{f,0} - P_{f,opt}\right)D > I\left(P_{f,opt}\right)$.

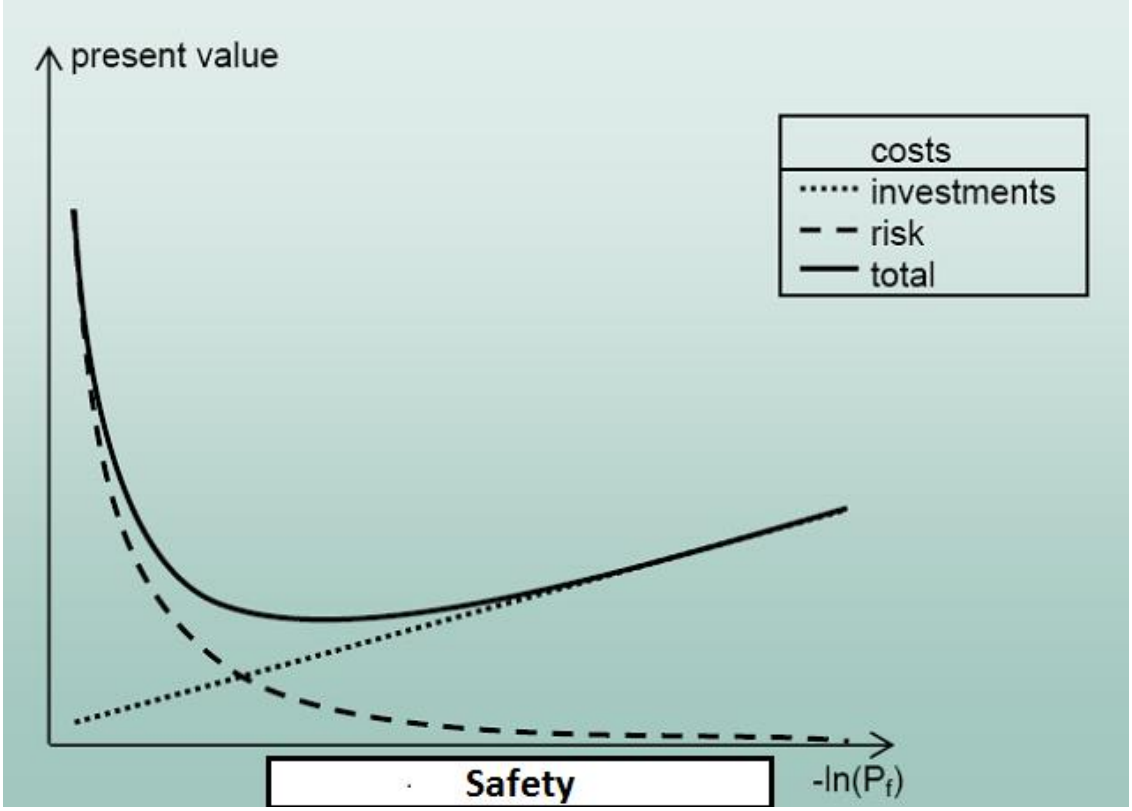


Figure 3.10 Economic optimization, costs, risks and total costs as a function of the failure probability of the system.

3.4.3 Application of the economic optimization: optimal dike heightening

Before the major floods of 1953, dikes in the Netherlands were not designed for a specified safety level but mainly strengthened based on practical experience. One of the main questions after the disaster was the optimal dike height and the "acceptable" probability of flooding. Van Dantzig was a professor in mathematics and a member of the first Delta committee. He developed an econometric approach to determine the optimal probability of flooding (or protection level) and the corresponding dike height (van Dantzig, 1956).

The approach only considers failure of the dikes due to overtopping. The probability distribution of water levels along the Dutch coast can be approximated by an exponential distribution.

$$F\left(h\right) = 1 - e^{\left(-(h-A)/B\right)} \tag{3.19}$$

In which h - water level [m]; A, B – constants of the exponential distribution [m]

Neglecting wave run-up, the probability of failure of the dikes – leading to flooding – can be approximated by the probability of exceedance of the dike height h_d [m], i.e.

$$P_f = P\left(h > h_d\right) = e^{\left(-(h_d-A)/B\right)} \tag{3.20}$$

It is assumed that there is total damage D to all objects and infrastructures in the flooded area if the dikes fail. For the discount rate Van Dantzig applied a reduced interest rate r' = economic

growth – inflation. The net present value of the expected damages, i.e. the risk is found as follows (assuming an infinite time horizon).

$$R = \frac{P_f D}{r'} = \frac{e^{-\frac{h_d - A}{B}} D}{r'} \tag{3.21}$$

Note that the risk is thus dependent on the dike height. To limit the failure probability and thus risks, the dikes can be heightened further. The investments in dike heightening are determined by initial costs for mobilization and variable costs that are dependent on the level of dike heightening (see also Figure 3.11):

$$I = I_0 + I\left(h_d - h_0\right) \tag{3.22}$$

Where: I_0 – fixed costs [€/m]; I – variable costs per m heightening [€/m]; h_d – new dike height after raising [m]; h_0 – actual or initial dike height [m]

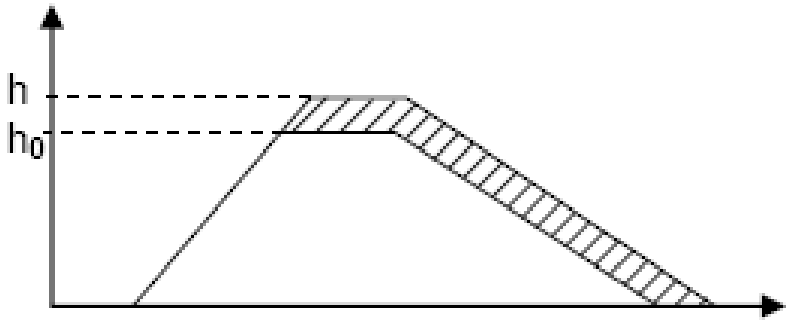


Figure 3.11 Schematic view of dike heightening.

Now, the total costs C_{tot} are the sum of the investment costs and the risk.

$$C_{tot} = I_0 + I\left(h_d - h_0\right) + \frac{e^{-\frac{h_d - A}{B}} D}{r'} \tag{3.23}$$

The optimal dike height is found when the total costs are minimal. This is the case when $dC_{tot} / dh_d = 0$.

$$\frac{dC_{tot}}{dh_d} = I - \frac{e^{-\frac{h_d - A}{B}} D}{Br'} = 0 \tag{3.24}$$

Since the $P_f = e^{\left(-(h_d - A)/B\right)}$, we can find the optimal flooding probability $P_{f,opt}$:

$$P_{f,opt} = \frac{IBr'}{D} \tag{3.25}$$

The larger the damage D, the smaller the optimal failure probability and thus the higher the level of protection. If incremental protection is expensive – i.e. for high values of I – a higher optimal failure probability and thus a lower level of optimal safety will be found. The optimal failure

probability is also dependent on the discount rate r'. Combination with equation (3.20) gives the optimal dike height $h_{d,opt}$:

$$h_{d,opt} = A - B \cdot ln\left(P_{f,opt}\right) \tag{3.26}$$

In addition to the determination of the optimal dike height the cost benefit criterion should still be verified. It is checked whether the benefits of risk reduction are larger than the costs of the dike heightening.

$$I\left(h_{d,opt}\right) < \left(P_{f,0} - P_{f,opt}\right) D / r' \tag{3.27}$$

Where: $P_{f,0}$ – flooding probability in the initial situation

It can be easily shown that this is equivalent to a check of whether the total costs in the optimal situation are smaller than those in the original situation. The first Delta committee used the following values for South Holland:

D = fl 24.2.10^9 [unit is Dutch guilders]; r'=0.015; I = 40.1.10^6 fl/m; A = 1.96m; B = 0.33m; h_0 = 3.25m; I_0 = fl 110.10^6.

Using these values the following optimal dike height and optimal failure probability were derived: $h_{d,opt}$ = 5.83m; $P_{f,opt}$ = 8.10^{-6} per year.

Although the optimal safety level was determined at a failure probability of $P_{f,opt}$ = 1/125,000 per year, in later political derivation a value of 1/10,000 per year was determined for the probability of exceedance of design water levels. This implied that the dikes of South Holland would need to be designed for hydraulic conditions (water levels and waves) with a probability of exceedance of – on average – 1/10,000 per year. In later decision-making safety standards have been derived for other dike rings (see Figure 3.12). For example, flood defenses in the river system are designed for a safety standard of 1/1250 per year.

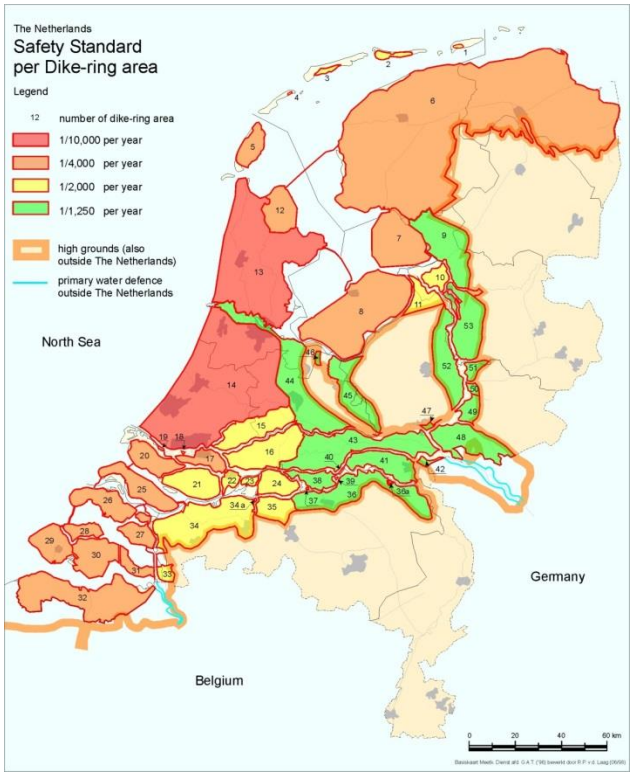


Figure 3.12 Dike Rings in the Netherlands, showing probability of exceedance of the design water levels per dike ring. Note that dike rings along the River Meuse are not shown on this map. Most of these dike rings have a safety standard of 1/250 per year.

It was expected that the actual failure probability for dikes designed for this design load, would be smaller than 1/10,000 per year. Recent risk analysis in the project VNK have shown that this is not the case due to the geotechnical failure mechanisms. For most dike rings the estimated failure probabilities are larger than the probability of exceeding the design loads. For example, for riverine dike rings that have been designed for design levels with a probability of exceedance of 1/1250 per year, failure probabilities are often in the order of magnitude of 1/100 per year (VNK, 2014).

Several extensions of and additions to the model have been proposed. For example, the inclusion of sea level rise (Vrijling and van Beurden, 1980), modelling of the damage as dependent on the water depth in the polder (van Dantzig, 1956), the inclusion of the economic value of loss of life as part of the damage, and inclusion of risk aversion by giving quadratic weight to the damages (van Gelder et al., 1997)

The model by van Dantzig was primarily focused on finding what the optimal safety at that time (when van Dantzig published his model in the 1960's) should have been. However, a single optimal level is not always the best solution. While considering longer timescales and changing conditions, such as economic growth and sea level rise, the model needs to address the possibility of multiple interventions during the period considered. This leads to questions regarding the timing of interventions (when? At which intervals?), as well as the optimal strengthening or raising of the dikes (how much?). Eijgenraam (2006) developed an economic decision model that takes into account both questions.

Figure 3.13 below depicts the "saw-tooth" curve that shows the periodic interventions: both the extent of the intervention (vertical) as well as the timing between interventions (horizontal). In between interventions, the safety will gradually decrease due to sea level rise and / or subsidence of the dike. Over a longer period of time, the dike heightening (or strengthening) should follow sea level rise. Additional dike heightening (lower the failure probability even further) could be considered in case of economic growth, which will lead to an increase of damages and increase of protection standards. The length of the optimal interval between interventions is largely dependent on the initial or mobilization costs. If these are high, for example in the case of structural interventions such as storm surge barriers, a long interval or life time of almost 100 years can be chosen. For regular dike reinforcements this interval will be more in the range of several decades (e.g. 50 years). For interventions with no or very small initial costs, such as nourishments along the coast, it is optimal to intervene more frequently.

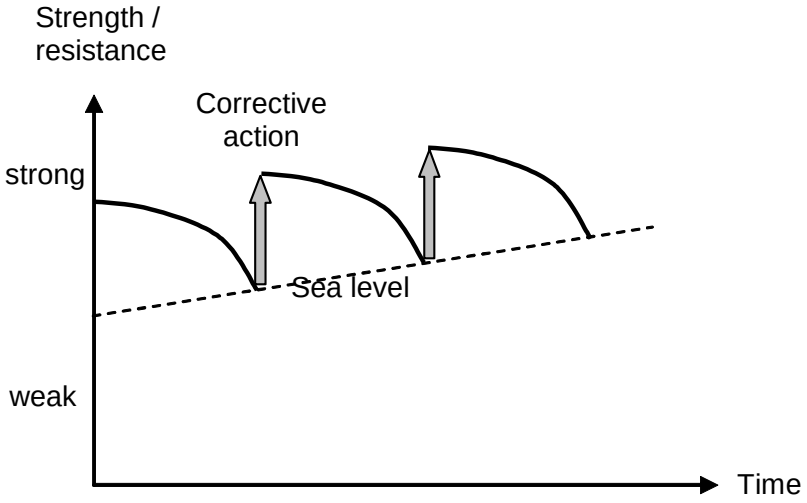


Figure 3.13 Periodic investments in dike reinforcement for a situation with sea level rise.

3.5 Safety standards

When answering the question "how safe is safe enough" a merely economic treatment with cost benefit analysis or economic optimization is often not sufficient for activities with risks to people. Therefore, criteria have been developed that focus on risks to human life. This section focuses on safety standards and criteria for evaluating the risk to life.

3.5.1 Introduction

Two aspects are typically considered when evaluating and regulating risks to the public: the total or population-wide effects, and the distribution of effects within the affected population. Table 3.5 summarizes these perspectives. The societal perspective is concerned with 'total effect' and the effects of large-scale accidents on the society, in terms of economic damages and life loss. The individual perspective is concerned with distributive justice ('equity'), i.e. the distribution of harm over the population.

As risk is often the by-product of an otherwise legitimate and advantageous activity, such as production or transportation, regulating risks is essentially a balancing act between economic and

social activities on the hand and a sufficiently safe society on the other hand (e.g. Jongejan, 2008): just as too lenient regulations are suboptimal, too stringent ones are too.

Table 3.5: Overview of perspectives on risk evaluation.

Aspect	Relevant risk metric	Rationale
Equity ('distributive justice')	Individual risk	Regulation is to prevent that individuals are exposed to disproportionally large risks
Total effect	Societal and economic risk	Regulation is to limit large-scale accidents and optimize the costs and benefits of risk reduction

Based on this general concept, it has been proposed to evaluate risks based on three criteria (TAW, 1985; Vrijling et al, 1995; 1998):

- To limit the **individual risk** to prevent that certain people are exposed to disproportionally large risks
- To limit the **societal risk** to limit the risks of large scale accidents with many fatalities
- **Economic optimization** to balance investments in risk reduction from an economic point of view

The **individual risk** concerns the probability of death of a person due to an accident. Various related definitions exist. The "average individual risk" for a certain activity can be calculated, e.g. the individual risk due to smoking or the risk of an airplane crash for a frequent flyer. Table 3.6 compares the probability of lethal accidents for various types of accidents.

Table 3.6 Average probability of death due to several activities. (statistics based on various sources for the Netherlands and Europe)

Activity	Probability (per year)
Mountain climbing	10^{-2}
Traffic (young men)	10^{-4}
Accidents at home	10^{-4}
Structural failures	10^{-7}

The individual risk due to an accident can be calculated with:

$$IR = P_f P_{d|f} \tag{3.28}$$

Where: IR – individual risk [per year]; P_f – probability of an accident [per year]; $P_{d|f}$ – conditional probability of death given an accident.

Both the individual risk and $P_{d|f}$ can be formulated as a function of a location near a risk source. If it is assumed that a person is permanently present at that location, the individual risk becomes a property of that location and it can be used for zoning and land use planning. This is applied in the industrial safety policy in the Netherlands (see next section).

Societal risk refers to the probability of an accident with multiple fatalities. It is often graphically represented by an FN-curve that shows the exceedance probabilities of the potential numbers of

fatalities $(P(N > n))$ on double log scale. A simplified example of the calculation of an FN curve is given below.

The **economic optimization** has already been discussed in detail in section 3.4.2. Possible criteria for acceptable individual and societal risk are discussed in the coming paragraphs.

Example 3.5: difference between individual and societal risk

To illustrate the difference between the individual and societal perspective, consider a fairly safe car that has a probability of 10^{-5} per year of causing a fatality due to technical failure. This may well be acceptable to an individual. Such a probability is in the same range as the average death rate in traffic (650 fatalities / 17 million people which is about 4.10^{-5} per year). If only 1000 cars are sold in the Netherlands, the number of fatal accidents with this car is rare, i.e. 0.01 fatalities per year. However, when the car becomes popular and 10 million cars are sold, the average number of fatalities due to technical failure becomes 100 per year. This may well lead to public concerns and indignation. From a societal point of view this may no longer be acceptable, whereas the individual risk has not changed.

Example 3.6: calculation of an FN curve

The following example shows the composition of an FN curve for a system with two mutually exclusive event scenarios.

Accident 1 with N_1=10 fatalities and a probability of P_1=10^{-2} per year

Accident 2 with N_2 =100 fatalities and a probability of P_2 =10^{-3} per year

Based on this information the probability mass function can be formed (first graph). Consequently, the cumulative distribution function can be made (second graph). Finally, the probability of exceedance or the FN curve is made (third graph).

Finally, we note that the expected value of the number of fatalities equals:

$E(N) = P_1N_1 + P_2N_2 = 0.2$ fatalities per year

This value will also be found when the surface of the area under the FN curve is computed (Vrijling and van Gelder, 1997).

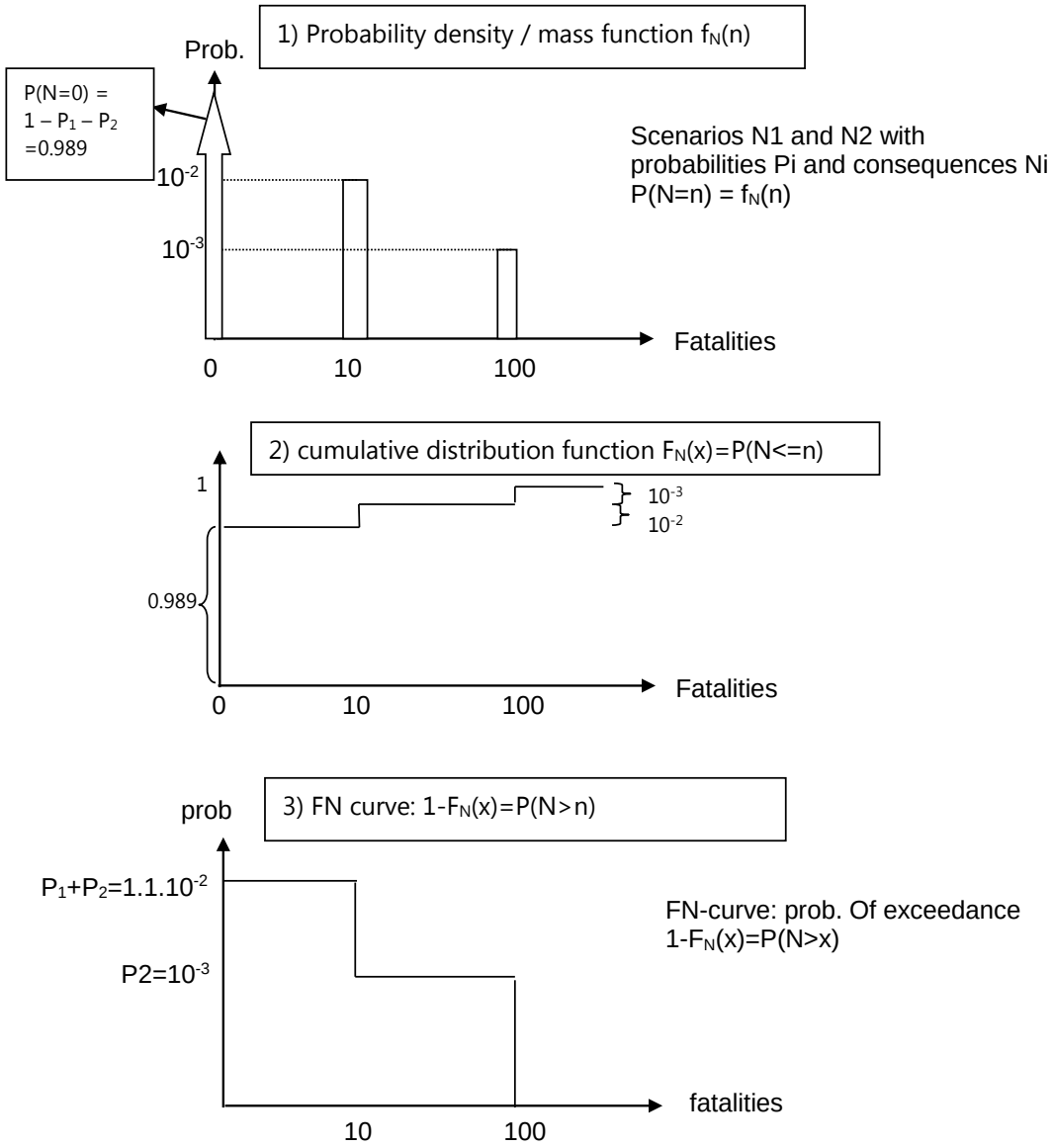


Figure 3.14 Composition of an FN curve for a simplified example.

3.5.2 Limits for individual and societal risk

Safety standards can be used to set limits to the individual and societal risk. A **limit** value can be set for **individual risks**. Such a limit value is to avoid disproportionate exposures by laying down a minimum safety level for all individuals. In various fields of applications limit values in the range of 10^{-4} to 10^{-6} per year are used (see below for more information). To put the stringency of the individual risk criteria into perspective, one could consider the effect of the probability of an accident on life expectancy. When an average person would be constantly exposed to a maximum allowable level of risk of 10^{-6} per year, the decrease of his or her life expectancy would be only 1 day as shown in Table 3.7.

Table 3.7: Relationship between the additional individual risk and life expectancy

Additional probability of death in a year	**Life expectancy in years**	**Decrease of life expectancy in years**	**Decrease of life expectancy in days**
10^{-3}	74.97	3.16	1153
10^{-4}	77.81	0.32	117
10^{-5}	78.10	0.03	11
10^{-6}	78.13	0.00	1
10^{-7}	78.13	0.00	0

Societal risk can be evaluated by means of an FN limit line. The calculated FN curve of the system should, in principle, not exceed the limit line. An FN-criterion is defined by three variables: (1) its base point (the exceedance probability of 1 fatality, i.e. *C*), (2) its slope (α), and (3) its probability and/or consequence cut-off (A and B). Figure 3.15 shows the different constraints that could make up an FN limit line.

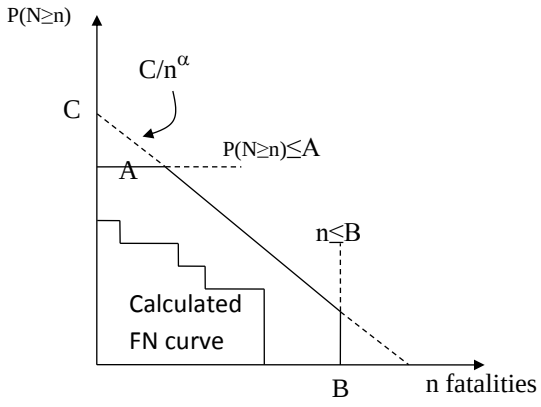


Figure 3.15 A fictitious FN-curve (grey) and different FN limit lines.

The general formulation for such a limit line without horizontal or vertical cut-off equals:

$$1 - F_N(n) \le C / n^{\alpha} \tag{3.29}$$

Where: C – constant that determines the vertical position of the limit line; α - coefficient that determines the steepness of the limit line

The limit line is called risk neutral[5] if $\alpha = 1$, since it places equal weight on exceedance probabilities and numbers of fatalities. If $\alpha = 2$, the limit is risk averse. This means that that the exceedance probability of 10 times as many fatalities should be 100 times lower. This has been motivated by public aversion to large numbers of fatalities. For example, the loss of 1000 people

[5] It should be noted that the usage of the term "risk neutral" to describe the FN-curve limit line for $\alpha = 1$ is widespread but, strictly speaking, incorrect in the context of decision theory (see section 3.3). This is because the cost of risk bearing to a risk neutral decision maker equals expected loss (i.e. the product of probabilities and damages). The FN-curve shows cumulative probabilities, however. Also, an individual crossing of the limit line would not necessarily disturb a risk neutral decision maker, provided the other accident scenarios have relatively small probabilities.

in one accident (e.g. a major explosion) could be valued differently than 1000 losses of 1 person in separate accidents (e.g. in traffic).

For different applications limit lines have been developed with varying constants and steepness. Examples of application areas include industrial risks in the Netherlands (next section), dams in the United States and Canada, and chemical risks in Hongkong and the UK (Jonkman et al., 2003).

Example 3.7 Risk matrix

Risk matrices are often used in various industries for risk evaluation and decision support, for example for the evaluation of health and safety risks within a facility. For a given undesired event the extent of probability and consequences are estimated on a qualitative or semi-quantitative scale, see Figure 3.16 for an example. Ranges of failure probabilities or consequences can be assigned to the qualitative terms on the axes in the example. The combination of probability and consequence estimates determines whether the risk is acceptable or whether it requires more attention and risk reduction efforts. However, unlike the FN curve, the cumulative effects of multiple possible events are generally not considered in a risk matrix.

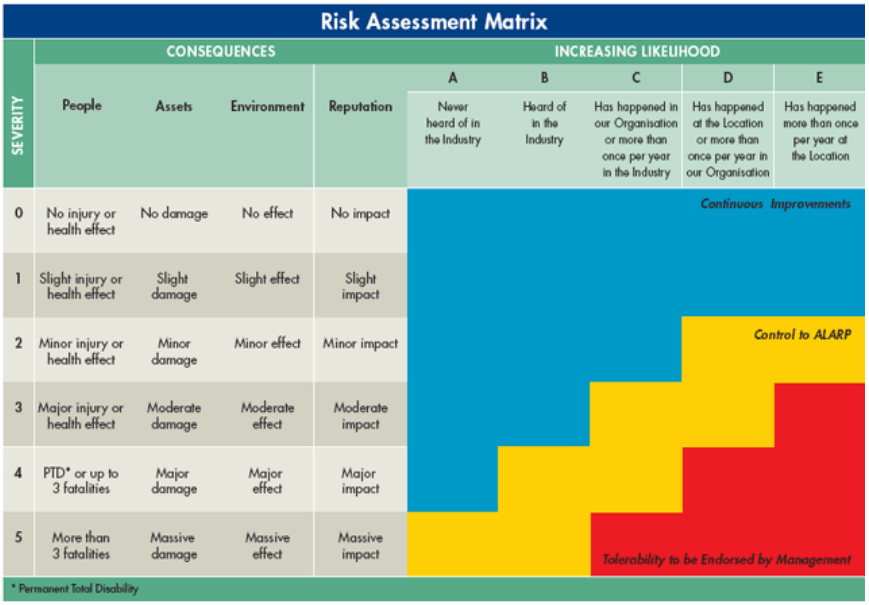


Figure 3.16 Example of a risk matrix (Source: http://www.eimicrosites.org/heartsandminds/ram.php)

3.5.3 The Dutch safety standards for industrial major hazards

The Dutch major hazards policy deals with the risks to those living in the vicinity of major industrial hazards such as chemical plants and LPG-fuelling stations. The development of the Dutch major hazards policy was strongly incident driven, as were European efforts aimed at the prevention of major industrial accidents. After a number of severe industrial accidents, including the Bhopal accident in 1984 which killed an estimated 3000 people and severely injured over 200.000, a European directive was drafted concerning the prevention of major accidents: the 1982 Seveso Directive. This was later replaced by the Seveso II Directive.

The cornerstones of the Dutch major hazards policy are a) the use of quantitative risk analysis (QRA); b) comparison of QRA outcomes with limits to individual and societal risks (Bottelberghs, 2000).

Within the Dutch major hazards policy, individual risk is defined as the probability of death of an average, unprotected person that is constantly present at a certain location. It is thereby a property of location and iso-risk contours can be plotted on a map (see Figure 3.17). Individual risk is therefore also named local risk ("plaatsgebonden risico") in the Netherlands. The shape of the risk contours for other applications will look different. For airports the contours will follow the shape of the runway and flight paths, for polders and flooding the risk contours will be highest in the deepest part of the polder.

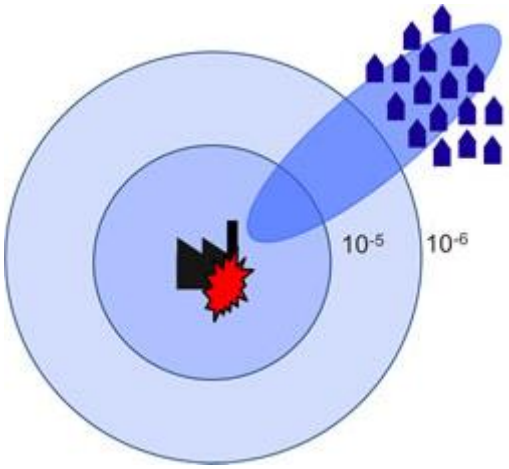


Figure 3.17 Example of a schematic individual risk contour (circles) for a chemical installation.

Table 3.8 Individual risk criteria used in the Netherlands.

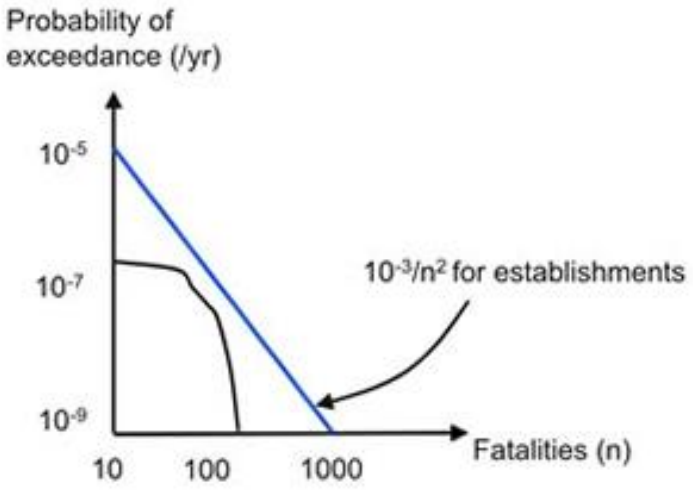
		Individual (or local) risk criterion
Existing situations	Vulnerable	10^{-6} per year
	Limitedly vulnerable	10^{-5} per year, strive for 10^{-6} per year
New situations	Vulnerable	10^{-6} per year
	Limitedly vulnerable	10^{-6} per year

A distinction is made between vulnerable objects such as schools and houses and limitedly vulnerable objects such as small offices. The following criteria apply. For new situations a limit of 10^{-6} per year applies. A comparison with Table 3.8 shows that this individual risk value has a negligible effect on life expectancy.

The criterion for societal risk that is used in the Netherlands for evaluating the third party risks posed by major industrial hazards is shown in Figure 3.18 below. It serves as a reference in the broader assessment of third party risks by competent authorities. Exceedances of the criterion line also have to be motivated by competent authorities. When the criterion line is not exceeded, the acceptability of the third party risk still has to be motivated.

The limit line is characterized by $C = 10^{-3}$ and a steepness of $\alpha = 2$, making it a risk averse criterion. The criterion is used to assess the acceptability of the risks of individual facilities.

Figure 3.18 FN limit line used for installations in the Netherlands.



3.5.4 The TAW model: a general model for deriving safety standards

Engineers are sometimes faced with the question 'how safe is safe enough'. For systems for which no regulations are available, this can be a difficult question to resolve. A model has therefore been developed by the Technical Advice Committee for Water (TAW)[6] defences for deriving safety standards on the basis of accident statistics (TAW, 1985; Vrijling et al 1995; 1998). The assumption underlying the model is that accident statistics are the result of a process of societal optimisation and that they thereby reflect what is apparently considered acceptable by society at large. Such an approach is commonly referred to as a 'revealed preference' approach.

Individual risk

Accident statistics reveal that the extent to which participation in the activity is voluntary strongly influences the level of risk that is accepted by individuals. Relatively high individual risks are accepted for activities that are voluntary and have a (personal) benefit, such as mountain climbing. Much smaller individual risk values are accepted for involuntary activities for which the risks are imposed by others, e.g. for chemical and nuclear industry. A policy factor (β) is therefore introduced to account for voluntariness of exposure. This factor is set at $\beta = 1$ for an individual risk value of 10^{-4} per year. This represents the “baseline” individual risk for the group of young men who are most at risk from traffic.

[6] TAW is nowadays called ENW: Expertise Network on Flood Protection.

Table 3.9 Accident statistics and proposed policy factor and characteristics of the activity (Sources: CUR, 2015; Vrijling, 2001; Vrijling et al., 1998).

Prob. Of death (per year)	Example / application	β	Voluntariness	Benefit
10^{-2}	Mountain climbing	100	Voluntary	Direct benefit
10^{-3}	Driving a motor cycle	10	Voluntary	Direct benefit
10^{-4}	Driving a car	1	Neutral	Direct benefit
10^{-5}	Flooding	0.1	involuntary	Some benefit
10^{-6}	Factory / LPG station	0.01	involuntary	No benefit

The proposed individual risk limit becomes:

$$IR < \beta \cdot 10^{-4} \tag{3.30}$$

An appropriate value for the policy factor β can be chosen depending on the characteristics of the activity. If the conditional probability of death due to an accident $P(d|f)$ is known, the acceptable failure probability can be computed:

$$P_f \leq \beta \cdot 10^{-4} \,/\, P_{d|f} \tag{3.31}$$

Note that the Dutch individual risk criterion for hazardous installations would be obtained for $\beta =$ 0.01 and a conditional probability of death of 1 ($P_{d|f} = 1$).

Societal risk

The societal risk criterion proposed by the TAW is based on the thought that societal risk should be evaluated primarily at a national level as local developments may lead to a situation that is considered unacceptable by society as a whole (Vrijling et al., 1995). The societal risk criterion at a national scale proposed by the TAW is:

$$E(N) + k\sigma(N) \leq \beta \cdot 100 \tag{3.32}$$

Where:

- $E(N)$ Expected number of fatalities per year
- k Risk aversion index (proposed value, k = 3)
- $\sigma(N)$ Standard deviation of the number of fatalities per year
- β Policy factor

A risk aversion index k has been introduced to account for risk aversion. For accidents with small probabilities and large consequences the standard deviation $\sigma(N)$ is large relative to $E(N)$, see example below. The total risk takes a risk aversion index k [-] into account. For $k > 1$, the "cost of risk bearing" exceed expected loss, implying a risk averse attitude.

Example 3.8: Expected value and standards deviation for two systems

We consider two systems

1. This system has a high failure probability of 0.01 per year and 1 fatality
2. The second system has a smaller failure probability of 0.0001 per year but a higher number of 100 fatalities.

For both systems a Bernoulli distribution of the number of fatalities is applied meaning that the number of fatalities in case of failure is exactly known. The expected value and standard deviation of the number of fatalities are found as follows

$$E(N) = P_f N$$
$$\sigma^2(N) = P_f(1-P_f)N^2 \tag{3.33}$$

The resulting expected value and standard deviation are shown in Table 3.10 below. Although both systems have the same expected value, the standard deviation for the "small probability – large consequence" event for system 2 is higher. Taking into account the standard deviation in the TAW criterion thus accounts for risk aversion against accidents with large numbers of fatalities.

Table 3.10 Calculation of expected value and standard deviation for two activities.

	P_f	N	$E(N)$	$\sigma(N)$
1	10^{-2}	1	10^{-2}	0.099
2	10^{-4}	100	10^{-2}	0.99

The next step would be to distribute this maximum allowable level of societal risk over individual installations. After all, locally imposed societal risk criteria are necessary for achieving the desired national level of societal risk. The translation of the nationally acceptable level of risk to a criterion for a single local installation depends on the type of probability distribution of the number of fatalities. In Vrijling et al (1998) a formulation of the risk acceptance at a local level is presented conform equation (3.29):

$$1 - F_N(n) \le C / n^{\alpha} \tag{3.34}$$

For a binomial distribution this yields:

$$C = \left(\frac{\beta \cdot 100}{k\sqrt{N_A}} \right)^2 \tag{3.35}$$

in which: N_a is the number of independent locations where the activity takes place.

This requirement corresponds to the requirement set for chemical installations if $\beta = 0.03$, $N_a = 1000$ and $k = 3$.

Combination of individual and societal risk

According to the approach by TAW the three approaches (individual, societal and economic risk) lead to three acceptable failure probabilities. The most stringent of the three criteria can be chosen to determine the acceptable probability of failure of the system and to make sure that all three conditions are fulfilled. This can best be illustrated with an example (see below). The principles of

this approach have been applied to derive the proposed new safety standards for flood defences in the Netherland by the Delta Program (2014).

Example 3.9: Combination of individual, societal and economic risk for a dike ring area

We consider application of the three criteria to the case of dike rings in the Netherlands. There are about 100 dike rings in the Netherlands of different sizes. A dike ring is a flood prone area protected from flooding by flood defences and high grounds. The conditional probability of death given flooding depends on the depth of the polder and flood characteristics. Research on loss of life due to floods shows that a conservative estimate would lead to $P_{d|f} = 0.1$

First, we consider the individual risk. A value of $\beta = 0.1$ is proposed as being exposed is considered as an involuntary activity with some benefit (i.e. living in a prosperous delta). This leads to an acceptable individual risk value of 10^{-5} per year. This limit has also been proposed by the Dutch government in the year 2014 ("basisveiligheid").

The acceptable flooding probability according to the individual risk becomes:

$$P_f \leq \beta \cdot 10^{-4} / P_{d|f} = 0.1 \cdot 10^{-4} / 0.1 = 10^{-4} \text{ per year} \tag{3.36}$$

The societal risk criterion can be determined according to equation (3.29). Assuming a risk averse criterion $\alpha = 2$. We can determine the constant C of the limit line for $N_A =$ 100 installations and $\beta = 0.1$.

$$C = \left(\frac{\beta \cdot 100}{k\sqrt{N_A}} \right)^2 = \left(\frac{0.1 \cdot 100}{3\sqrt{100}} \right)^2 = 0.11 \tag{3.37}$$

The limit line for societal risk becomes $1 - F_N(n) \leq 0.11 / n^2$. Both the individual and societal risk criteria are plotted in Figure 3.19 below. As a third criterion the economic optimization can be added. The optimal or acceptable probability of failure depends on the damage and investment costs. A relationship with the graph below can be established by assuming that the number of fatalities is related to the economic damages. A dike ring with many inhabitants and potential fatalities will generally also represent a large economic value. For the sake of the example we assume that every fatality corresponds to an economic damage of € $5 \cdot 10^7$ (note: this is not equal to the value of a human life). To calculate the economic optimum for the example we assume arbitrary values of $r = 0.025$ and $I =$ € $5 \cdot 10^6$; $B = 0.33$. Figure 3.19 shows the combination for the three criteria.

For a given number of fatalities in a dike ring the acceptable failure probability according to the three criteria can be derived. The individual risk criterion is independent on the number of fatalities. The economic criterion shows a linear relation between the failure probability and damage or number of fatalities. The societal criterion is risk averse so shows a decreasing quadratic relationship between acceptable failure probability and consequences.

For a given number of inhabitants and potential fatalities for a dike ring, the acceptable failure probability can be determined. For dike rings with between 1 and 10 fatalities – generally small

areas-the individual risk criterion is the most stringent. For dike rings with large numbers of fatalities, the risk averse societal risk criterion becomes dominant.

Several extensions of this model are possible. One can consider to add the economic value of life loss or consider a different distribution of the nationally acceptable societal risk over dike rings with different sizes.

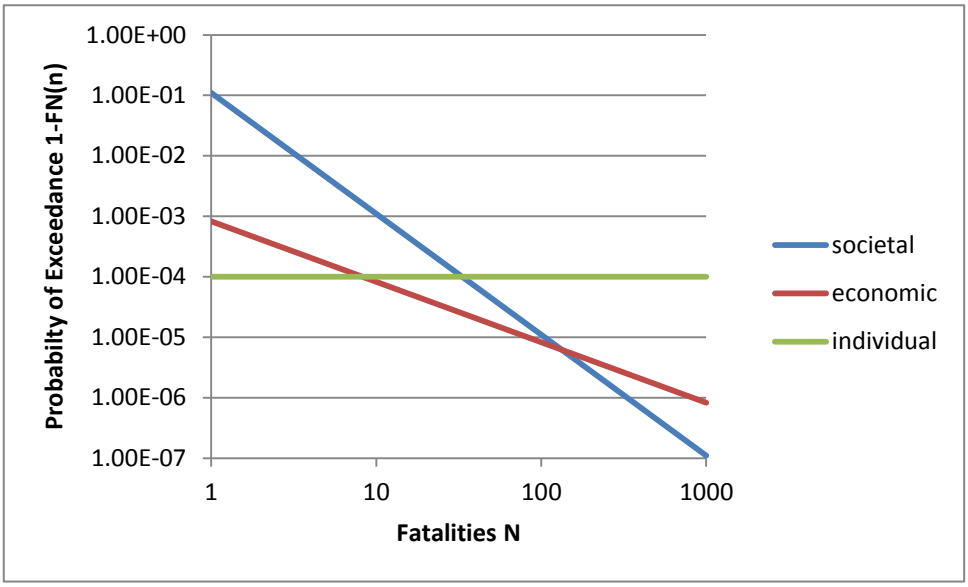


Figure 3.19 Combination of the individual, societal and economic risk criteria for a hypothetical example.

One field of application where these concepts have been applied concerns the flood protection standards in the Netherlands. In the year 2014 new safety standards have been proposed for primary flood defences in the Netherlands. These have been derived based on a risk assessment, and standards have been formulated as a (tolerable) probability of failure of the flood defences. More information on this field of application will be given during the lectures.

3.6 Probabilistic design: the relationship between safety standards and engineering design

The overall objective of **probabilistic design** is to design (and maintain) systems with an acceptable risk level in an optimal way. The first part of this description entails the consideration of acceptable risk in deriving the safety level (or failure probability) of the system. This topic has been the focus of the previous sections on safety standards – see for example the section on the economic optimization.

The first step is to derive an accepted failure probability (also called target value for the reliability). These values are generally expressed by means of failure probability or reliability index β. These values also depend on the reference period to which they are applied. Further details regarding typical target reliability values are given in section 10.3.

For complex systems, consisting of multiple elements with multiple failure mechanisms, the acceptable failure probability needs to be distributed over failure mechanisms and elements. A "failure probability budget" can be established for this distribution. Once the target reliability values for individual elements and failure mechanisms have been identified, partial factors for load and strength, the so-called γ's, can be derived as a basis for practical engineering design (see section 8.4 for further details). Examples of these partial factors are the material and load factors used in structural design.

Preferably, the distribution over failure modes and elements is chosen in such a way that an economical design results. In this respect, it is noted that the design choices will affect the costs of the system. Since the failure probability that follows from the economic optimization is a function of the costs, there can be interactions between the design choices and the optimal safety standard. The following sections will focus on more technical aspects of analysis of failure probabilities of individual elements and systems, as well as the derivation of design values for load and resistance.

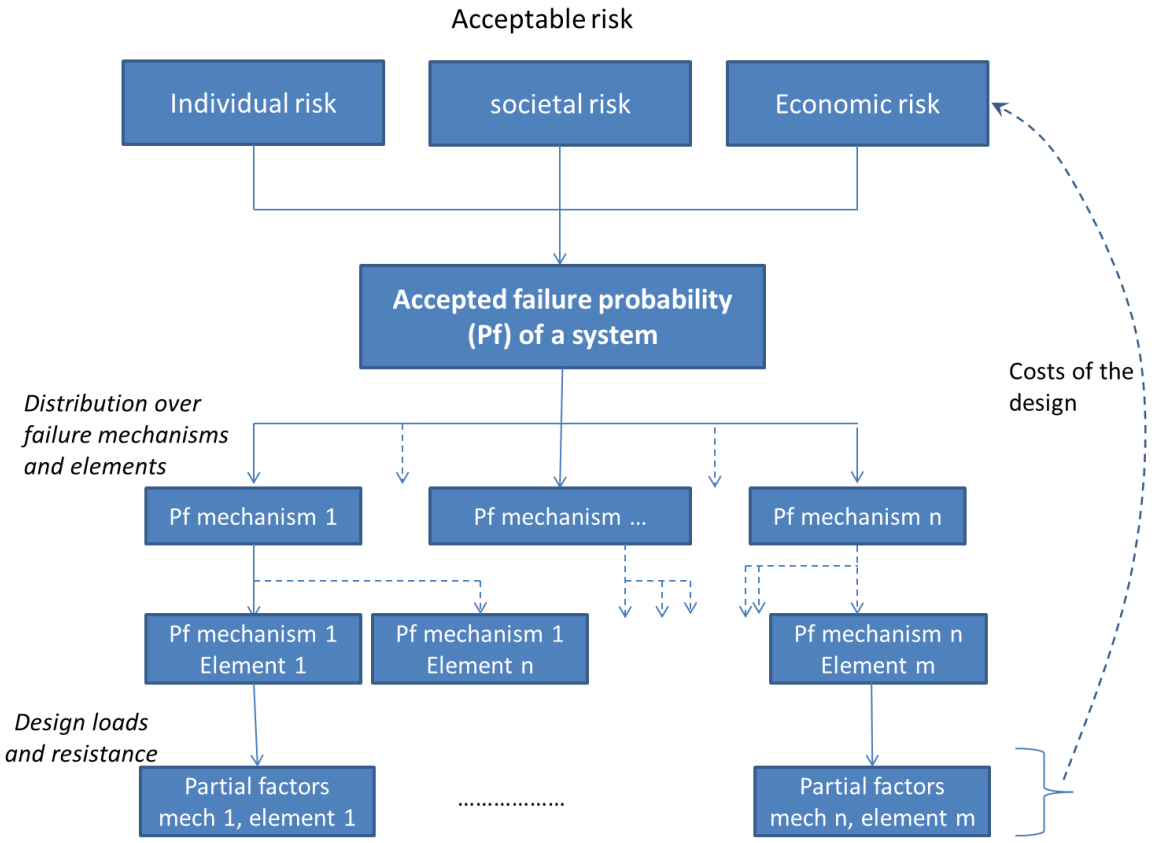


Figure 3.20: Relationship between acceptable risk, failure probability of a system and system design. Adopted from Schweckendiek et al. (2013).

References

- Apostolakis G.E. (2004) How useful is quantitative risk assessment? Risk Analysis Vol. 24 No. 3, pp. 515-520
- Benjamin J.R., Cornell C.A. (1970) Probability, statistics and decision for civil engineers. McGraw Hill
- Bottelberghs, P.H. (2000). Risk analysis and safety policy developments in the Netherlands. Journal of Hazardous Materials, 71: 117-123.
- CIB Workgroup 32 (2001) A framework for risk management and risk communications, CIB report 259
- CUR (Civieltechnisch Centrum Uitvoering research en Regelgeving) (1997) Kansen in de Civiele Techniek, Deel 1: Probabilistisch ontwerpen in theorie. CUR rapport 190
- Deltaprogramma (2014) Deltaprogramma Veiligheid. Synthesedocument veiligheid. Achtergronddocument B1. 25 juli 2014
- Deltares (2014) Technisch-inhoudelijke uitwerking DPV 2.1. Werkdocument Deelprogramma Veiligheid Concept van 26 februari 2014
- Don F.H., Stolwijk H.J.J. (2003) Investeren in veiligheid: de watersnoodramp van 1953 en de kosten en baten van het deltaplan. P. 79-98. In: CUR (2003) 50 jaar na stormvloed 1953, terugblik en toekomst. CUR212.
- Eijgenraam CJJ (2006) Optimal safety standards for dike ring areas. CPB discussion paper 62. March 2006.
- EN 1990:2002 E, Eurocode - Basis of Structural Design, CEN, November 29, 2001
- Faber M.H., Stewart M.G. (2003) Risk assessment for civil engineering facilities: critical overview and discussion, Reliability Engineering and System Safety Vol. 80 pp. 173 - 184
- Farmer F.R. (1967) Siting criteria – a new approach. Atom 128 pp.152-179
- Jongejan RB. How safe is safe enough? The government's response to industrial and flood risks. Doctoral thesis. Delft University of Technology. Delft, the Netherlands, 2008.
- Taerwe L., "Mastering global quality by an interactive concept", Proc. IABSE Symposium Tokyo 1986: Safety and Quality Assurance of Civil Engineering Structures, pp. 317-325.
- Jongejan RB, Jonkman SN, Vrijling JK (***) Methods for the valuation of human life. ESREL 2005
- Jongejan RB, Helsloot I, Beerens RJ, Vrijling JK (2011) How prepared is prepared enough? Disasters 2011 Jan;35(1):130-42. doi: 10.1111/j.1467-7717.2010.01196.x.
- Jonkman S.N., van Gelder P.H.A.J.M., Vrijling J.K. (2003) An overview of quantitative risk measures for loss of life and economic damage, Journal of Hazardous Materials. A99, pp.1-30
- Kaplan S., Garrick B.J. (1991) On the quantitative definition of risk, Risk Analysis Vol. 1 pp. 11-27
- Kendall H.W., Hubbard R.B., Minor G.C., Bryan W.M. (Union of Concerned Scientists) (1997) The Risks of Nuclear Power Reactors: a Review of the NRC Reactor Safety Study WASH-1400, Cambridge, 1977
- Raiffa H., Schlaifer R. (1961) Applied statistical decision theory, Harvard University press, Cambridge, Massachusetts, United States.

- Schweckendiek, T., Vrouwenvelder, A.C.W.M., Calle, E.O.F., Jongejan, R.B., Kanning, W.: Partial Factors for Flood Defenses in the Netherlands. Modern Geotechnical Codes of Practice – Development and Calibration, Fenton, G.A. et al. (eds.). Special Geotechnical Publication. Taylor & Francis. 2013.
- Slijkhuis
- Slovic, P. (1987) Perception of Risk. Science Vol. 236 pp.280–285
- Starr C. (1967) Social benefit versus technological risk, Science Vol. 165 pp.1232-1238
- SWOV (2012) Waardering van immateriële kosten van verkeersdoden. Factsheet. http://www.swov.nl/rapport/Factsheets/NL/Factsheet_Immateriele_kosten.pdf
- TAW (Technical Advisory Committee on Water Defences) (1985) Some considerations of an acceptable level of risk in the Netherlands, Report by TAW workgroup 10 'Probabilistic methods'
- Tengs, T.O., Adams, M.E., Pliskin J.S., Gelb Safran D., Siegel J.E., Weinstein M.C., Graham J.D. (1995) Five-hundred live saving interventions and their cost effectiveness, Risk Analysis Vol. 15 No.3 pp.369-390
- Van Dantzig D. (1956) Economic decision problems for flood prevention, Econometrica Vol. 24 pp. 276-287
- Van Gelder, P.H.A.J.M., Vrijling, J.K., and Slijkhuis, K.A.H., 1997a. Coping with uncertainty in the economical optimization of a dike design. In: Water for a Changing Global Community, pp. 554-559.
- Vlek C.A.J. (1996) A multi-level, multi-stage and multi-attribute perspective on risk assessment, decision-making and risk control, Risk Decision and Policy Vol. 1 No. 1 pp.9-31
- Von Neumann, J. Morgenstern O. (1953) Theory of Games and Economic Behavior, Princeton, NJ, Princeton University Press, 1944, second ed. 1947, third ed. 1953.
- Voortman H.G. (2004) Risk-based design of large-scale flood defence systems, PhD thesis, Delft University
- Vrijling, J.K., Hengel, W. van, Houben, R.J. (1995) A framework for risk evaluation, Journal of hazardous materials Vol. 43 No. 3 pp.245-261
- Vrijling, J.K., Hengel, W. van, Houben, R.J. (1998) Acceptable risk as a basis for design, Reliability Engineering and System Safety Vol. 59 pp. 141-150
- Vrijling, J.K., van Gelder, P.H.A.J.M van (1997) Societal risk and the concept of risk aversion, In: C.Guedes Soares (ed.), Advances in Safety and Reliability, Proc. Of ESREL 1997 Lissabon, Vol. 1, pp. 45-52, 1997
- Vrijling, J.K.; Gelder, P.H.A.J.M. van (2000) An analysis of the valuation of a human life In: M.P. Cottam, D.W. Harvey, R.P. Pape, & J.Tait (eds.) ESREL 2000 - "Foresight and precaution", Volume 1, pp.197-200, Edinburgh, Scotland, UK
- J.K. Vrijling, J.A. van Beurden, Sealevel rise: a probabilistic design problem, Proceedings of the International Conference on Coastal Engineering Delft. p.1160-1171, 1990
- VNK (Veiligheid Nederland in Kaart) (2014) De veiligheid vanNederland in kaart. Eindrapportage. Report HB 2540621
- Vrouwenvelder A.C.W.M., Vrijling J.K. (1996) Collegedictaat Probabilistisch ontwerpen b3, TU Delft

Appendix 3.1 More detailed scheme for risk assessment (CUR 190)

The scheme below was developed in the book "probability in civil engineering" (CUR 190). It uses roughly the same main groups of steps as introduced in the main text of these lectures notes: qualitative and quantitative analysis and risk evaluation.

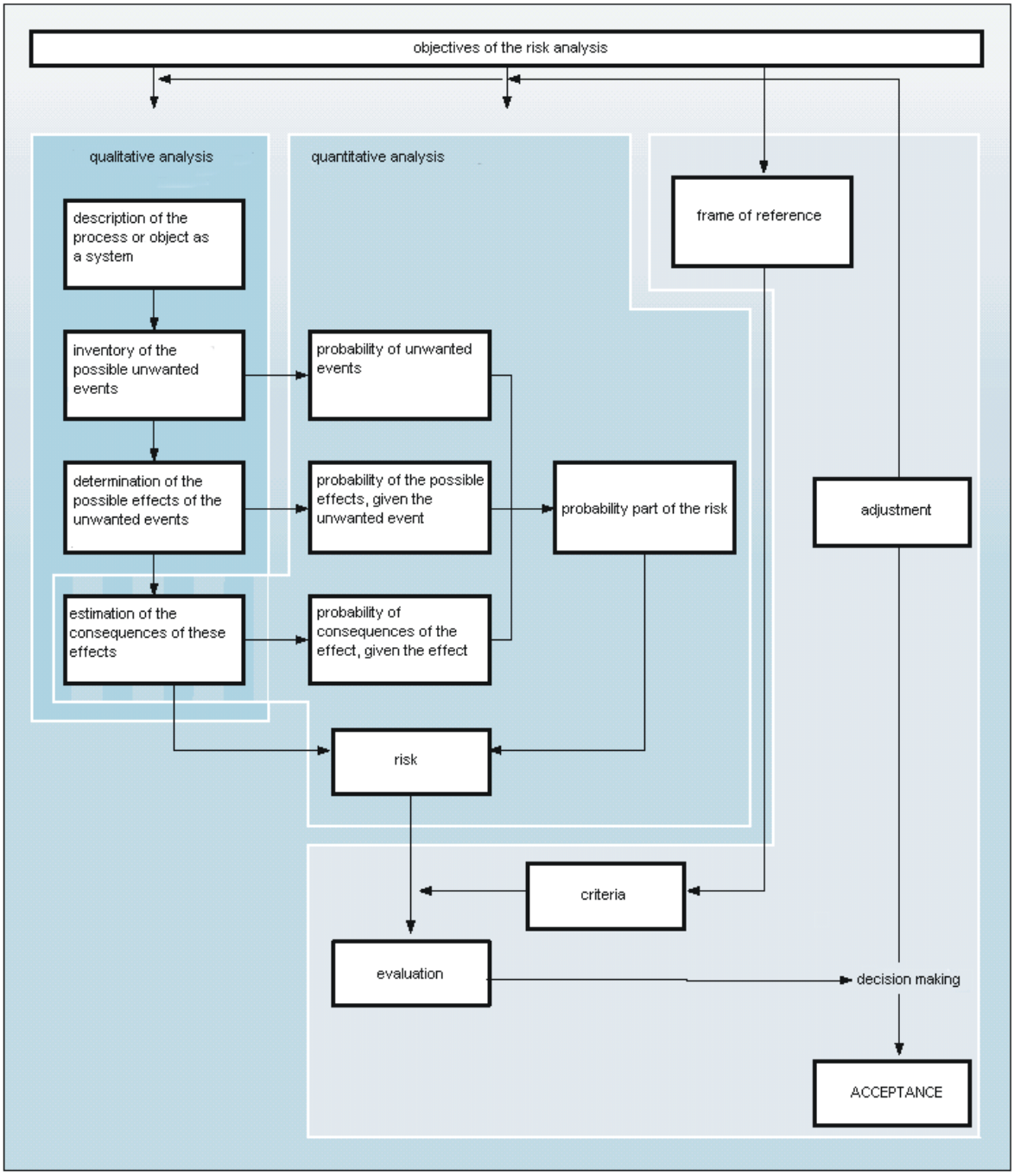


Figure 3.21 Detailed scheme for risk assessment

Appendix 3.2 FN curves for different distribution types

From (CUR, 2015)

Different distribution types can be used for the number of fatalities and some options are shown below.

Note that every activity has a non-zero probability of zero deaths (no accident) and a non-zero probability of deaths. The differences in the probability density functions lay in the assumed distributions of the number of fatalities given an accident. The first function in Figure 3.22 is the Bernoulli-distribution which has only two possibilities: N deaths, with a probability p, or zero deaths, with a probability $1-p$. This distribution is suitable as a boundary condition for activities where failure leads to a loss that is known exactly. The second function in Figure 3.22 is the exponential distribution which is continuously and exponentially distributed over the interval between zero and infinity. This distribution is suitable for modelling economic losses or the number of dead as a result of an activity. The third is the less well-known inverse square Pareto-distribution. This distribution corresponds to the Dutch standards using. The Pareto distributed variable can assume the value zero or every value between one and a specified maximum.

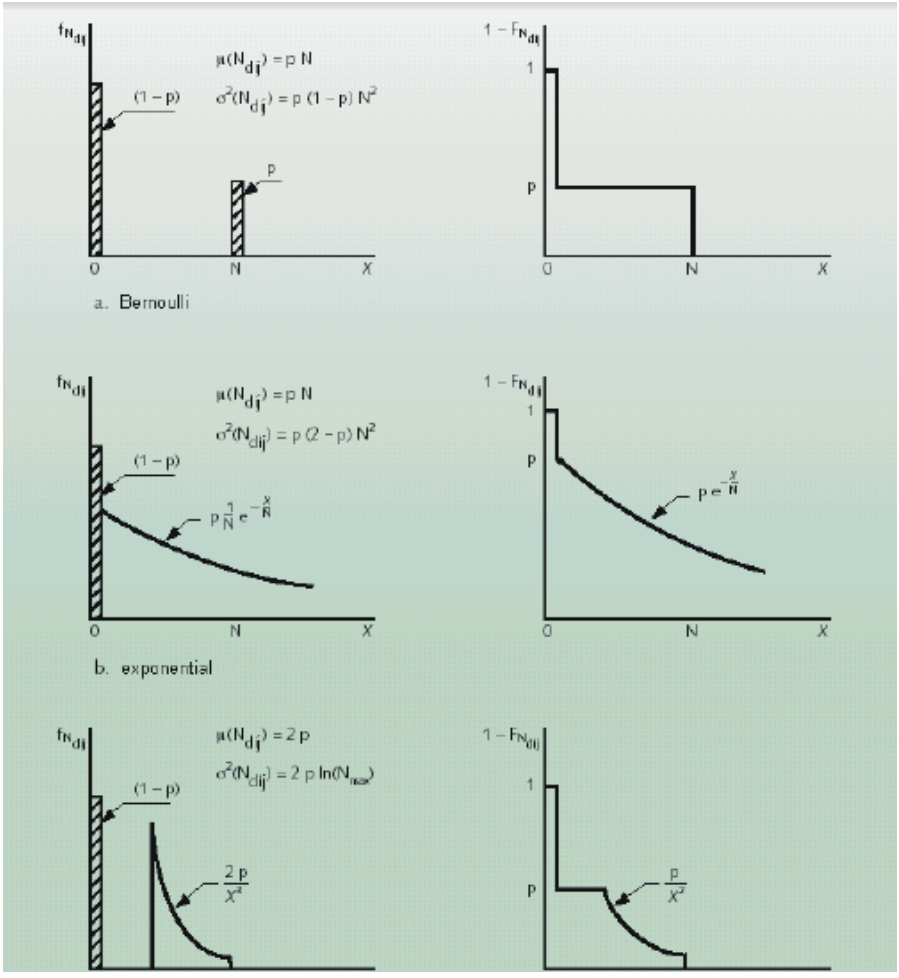


Figure 3.22 Multiple distributions that can be used for the number of fatalities

Part II – Reliability Analysis

Chapter "Reliability Calculations, Basics"
Author: R.D.J.M. Steenbergen and A. Kosters

This chapter is based on 'Risk Analysis of Construction Processes' by L. Taerwe and R. Caspeele (UGent lecture notes) and 'Introduction to Safety and Reliability of Structures' by J. Schneider and T. Vrouwenvelder.

Contents

4 Reliability Calculations, Basics

4.1 The problem Z=R-S

Often, the reliability of a system can be assessed by comparing two stochastic quantities: the resistance R of the system on the one hand and the load (or solicitation) S on the other[1]. Some examples of these quantities are given below:

R	**S**
Flow capacity of a river bed	Discharge of the river
Flow capacity of a sewage pipe	Discharge of waste water
Bending resistance	Existing bending moment
Permissible deflection of beam	Existing deflection of beam
Soil cohesion and shear strength	Stresses in soil due to external loads
Traffic capacity of a road junction	Intensity of traffic

In these lecture notes the focus will be on the reliability analysis of structures, but the examples above illustrate that this reliability analysis can also be applied in other fields of civil engineering, such as hydraulic, geotechnical and traffic engineering. In its simplest format, the safety of the system can be assessed by verifying if the resistance is larger than the load, such that no failure occurs. In terms of the examples: the river does not overflow its banks, the beam does not fail, the slope does not become unstable, the traffic does not come to a standstill, there is no electrical power failure, etc. Thus, a reliable system complies with the following statement:

$$R > S \tag{4.1}$$

Traditionally, the verification of structural safety follows deterministic patterns. A certain value r_d of the resistance of a structural component is derived from a number of characteristic values. In a similar manner a defined value s_d representing the load effects is derived from a number of characteristic values of loads. In order to check for safety or failure these two single values r_d and s_d are then compared:

$$r_d > s_d \tag{4.2}$$

However, in the probabilistic approach advocated here, resistance and load effects are not regarded as deterministic quantities, but as random variables, which can each be described by a certain distribution type and accompanying parameters. Indeed, the strength of the applied materials is – within certain boundaries– different from element to element and the loads acting on a structure not only show a certain spatial variation, but also a variation in time. The ultimate aim is therefore to keep the probability low that a rather low resistance and a very high load occur at the same time. In case the probability density function of R and S are known, the failure probability P_f can be calculated as the probability that S is larger than R:

[1] It is noted that in some other sources alternative terms are used for resistance and load. For example capacity or strength for resistance, and solicitation or demand for loads. In these lecture notes mainly load and resistance will be used.

$$P_f = P[S > R] \tag{4.3}$$

The same problem can be formulated by means of a limit state. This is a condition beyond which the structure or part of the structure does no longer fulfil one of its performance requirements. The limit state *Z* can be assessed by considering the resistance *R* and the loads *S*, i.e.

$$Z = R - S \tag{4.4}$$

Failure occurs when $R < S$, so when $Z < 0$. The probability of failure equals P_f = *P[Z<0]* = *P[S>R]*.

In case of one-dimensional problems and simple distribution functions (normal, lognormal, extreme value distributions…) P_f can be easily calculated – often using analytical methods. However, in practice multiple basic variables influence the limit state, making it difficult and most often impossible to evaluate the multidimensional integrals exactly. Therefore, several methods are available for reliability analysis (see section 4.5 for an overview).

Different standards mention target values for the failure probability as a function of the reference period and the consequences of failure with respect to human lives and economic considerations.

In the following sections, we will first take a look at the resistance and load parameters, R and S. Subsequently, a mathematical formulation for the limit state Z will be derived.

4.2 Resistance of structural elements

The abbreviation *R* denotes the resistance of structural elements in a given cross-section of a structure, or it can stand generally for some other capacity of the system under consideration (see section 4.1).

The reasoning will be illustrated for the structural resistance *R*. The model for *R* has, as a rule, the following typical form:

$$R = M \cdot F \cdot D \tag{4.5}$$

In which:
M model uncertainty variable
F material properties (strength, elastic modulus, …)
D dimensions and the derived quantities

Because the resistance appears as a product of variables and because negative resistances are hardly possible, *R* tends to a lognormal distribution (see section 2.6.5).

The following sections will elaborate upon the abovementioned variables.

4.2.1 Model uncertainties

Since in developing a resistance model certain influences are either consciously or unconsciously neglected, deviations between analysis and tests are to be expected. This fact is considered by

introducing a model variable M that may be determined from tests. The test results r_{Exp} are divided by the corresponding results r_{Mod} obtained using the resistance model:

$$m = \frac{r_{Exp}}{r_{Mod}} \tag{4.6}$$

From a large number of tests, the mean value m_M and the standard deviation s_M and a histogram for M is obtained. These experimental results are then replaced by a suitable distribution.

For good models, $\mu_M \approx 1$ is obtained. Since conservative models are frequently used, it follows that often $\mu_M > 1$. The value of σ_M differs from a few percent for good models up to values in the region of 10% to 20% for poor models.

4.2.2 Material properties

The values for strengths F and other material properties are obtained mostly from tension and compression tests. Usually the results cannot be used directly because of the following problems:

- Conditions in a laboratory test are often quite different from those in the structure;
- The scatter in the material properties of the structure is usually greater than the scatter in results from laboratory tests;
- Material properties may vary in time.

Therefore the laboratory result is often multiplied by a so-called transfer variable, that takes into account the ratios of the properties of a structural component and the respective quantity measured in the test.

4.2.3 Geometrical properties

Geometrical properties D may be measured directly. The dimensions may be checked and compared with the corresponding tolerances. The mean values are usually close to the nominal values, although occasionally systematic influences arise. For example, a formwork may deform when concrete is poured into it, such that the actual dimensions exceed the planned dimensions.

4.3 Load effects in structural elements

When talking about loads, three terms have to be distinguished, illustrated here for wind load:

- Influence: v_{wind} [m/s]
- Load: w [kN/m^2]
- Load effect: M, V, N … due to wind

By way of example, a number of loads normally taken into account in the design of structures is shown in Figure 4.1 plotted as concurrent stochastic processes. Load values are plotted along the horizontal axis while the time axis is vertical.

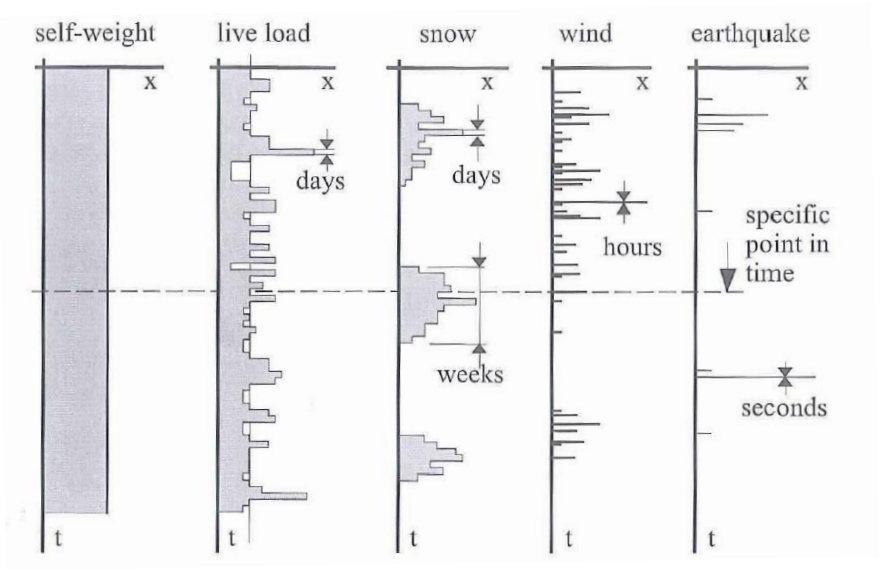


Figure 4.1 Various actions on a vertical time scale

The processes illustrated in Figure 4.1 are briefly discussed below:

- The *self-weight* may to a good approximation be regarded as constant over time.
- Each structure supports in addition to the self-weight *quasi-permanent* live loads. These are superimposed by *short term* live loads. In buildings the latter have a duration in the order of hours or days. On a bridge, high values of loading occur at known times (rush hours) and reach peak values during traffic jams and accidents. These maximum values last from a few minutes to several hours.
- In lower laying areas of many countries there is *snow* for only a short period in the winter months. Maximum values are recorded over periods of days, average values over periods of weeks and months. In mountainous regions the snow may lie on the ground for longer periods and, for snow-prone countries, may be considered to exhibit the character of permanent influence. In some countries snow may be only exceptional or completely non-existent.
- *Wind* only occurs for short periods, say, several minutes to a few hours. Strong gusts are seldom. Maximum gusts last for only a few seconds to minutes.
- Finally, *earthquakes* occur very rarely. Their period of strong motion is in the region of several seconds. The intensity is highly variable.

4.3.1 Modelling of loads

Loads are as a rule stochastic processes in time. In the evaluation of these processes, both the extreme values of a certain load and the so-called arbitrary-point-in-time (APT) values are of interest. To represent these values, two kinds of variables are defined from the observed data: the leading and the accompanying loads.

The leading load is determined essentially by analysing the stochastic process with respect to its extreme values e. Usually these exhibit an extreme value distribution $E_i(e)$, which is defined by its type (e.g. a Gumbel distribution) together with the respective parameters (see section 2.6.6).

Accompanying loads are derived from the APT values of the stochastic process and exhibit a more or less symmetrical distribution $A_i(x)$ with respect to the mean and are usually modelled by normal or lognormal distributions.

Section 7.7 explains how the different loads on a structure can be combined, using the concept of leading and accompanying loads.

4.3.2 Model uncertainties

Uncertainties in the modelling of loads can be taken account by introducing a model uncertainty variable *M*, as was done in section 4.2 for the resistance. The model variable takes into account the uncertainties introduced by simplifications of e.g. the static system, of the load pattern or shape and of the influences of stiffness and cracking of structural parts.

A distinction can be made between a general model variable *M* for all loads and specific model variables M_{Ei} and M_{Ai} that are used only for the leading and accompanying loads respectively.

4.4 General formulation for limit state design

We return to the limit state *Z* that was defined in equation (4.4). All variables that are used to model the resistance and load effects, that together make up the limit state, can be assimilated in one structural model:

$$g(\underline{X}) = Z = 0 \tag{4.7}$$

where the vector $\underline{X}$ consists of n basic variables such as:

- material properties
- actions (loads)
- geometrical properties
- model uncertainties.

For all basic variables one has to consider an appropriate probabilistic model. In case a basic variable has a negligible variation in time or space, one can consider that variable as **deterministic**.

Also the function $g(\underline{X})$ can be considered as a random variable, which we will denote as *Z* and use as condensed notation for the limit state equation. The function $g(\underline{X})$ is defined so that $g(\underline{X}) > 0$ (i.e. $Z > 0$) corresponds to safe conditions, while $g(\underline{X}) < 0$ (i.e. $Z < 0$) corresponds to failure. With $f_{\underline{X}}(\underline{x})$ the n-dimensional probability density function of the *n* basic variables X_i, the **failure probability P_f** becomes

$$P_f = \int_{g(\underline{X})<0} f_{\underline{X}}(\underline{x})\, d\underline{x} \tag{4.8}$$

In case of $n = 2$, the failure probability P_f is the volume under the joint probability density function corresponding to the domain where $g(X) < 0$ (see also
Figure 4.4). The elaboration or application of (4.8) can be performed using several methods with a different level of accuracy and complexity. This is further elaborated in the following sections.

P_f is calculated for a certain **reference period** t_{ref} which corresponds most often, but not necessarily, to the design working lifetime. Chapter 10 gives some examples of reference periods used in structural engineering.

The probability of survival (or the reliability) is defined as:

$$P_s = 1 - P_f \tag{4.9}$$

To illustrate the limit state formulation as explained above, consider the simple structural system depicted in figure Figure 4.2. This system may fail in multiple ways, for example at the midspan due to too large bending moments, or at the supports due to too large shear forces. In this example the limit state of bending in the midspan will be considered, which is represented by

$$g(\underline{X}) = Z = M_R - M_S = 0 \tag{4.10}$$

where M_R is the resisting moment of the cross-section and M_S is the bending moment as a result of the applied load, both calculated with respect to the midspan.

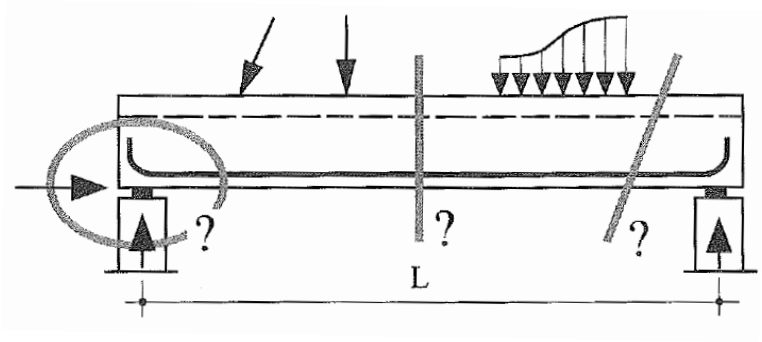


Figure 4.2 Reinforced concrete beam and three failure possibilities

In general, resistance effects are represented by R and load effects by S. Considering this notation, the limit state equation becomes:

$$g(\underline{X}) = Z = R - S = 0 \tag{4.11}$$

In the *(R, S)* plane (Figure 4.3) this linear relationship represents the boundary between the "safe domain" D_s *(Z > 0)* and the "unsafe domain" D_f *(Z < 0)*.

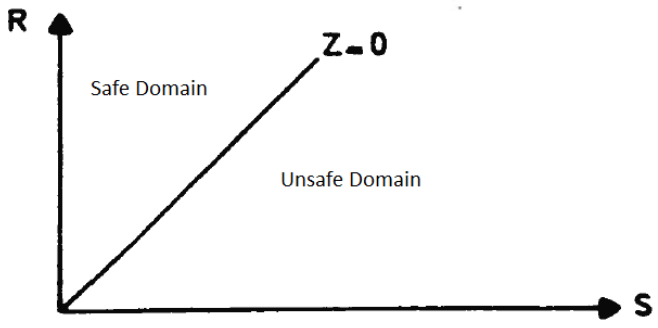


Figure 4.3 Safe and Unsafe Domain in case of a linear limit state equation

As *R* and *S* are both stochastic variables, their distribution functions are given by:

$$f_R(r) \tag{4.12}$$

$$f_S(s) \tag{4.13}$$

If *R* and *S* are independent the joint probability density function $f_{RS}(r,s)$ follows by multiplication:

$$f_{RS}(r,s) = f_R(r) \cdot f_S(s) \tag{4.14}$$

This joint probability density function can be drawn using altitude lines in the *R*, *S* space. It is now easy to see that the probability of failure of the structure is equal to the volume of the joint probability density function in the unsafe region; see
Figure 4.4.

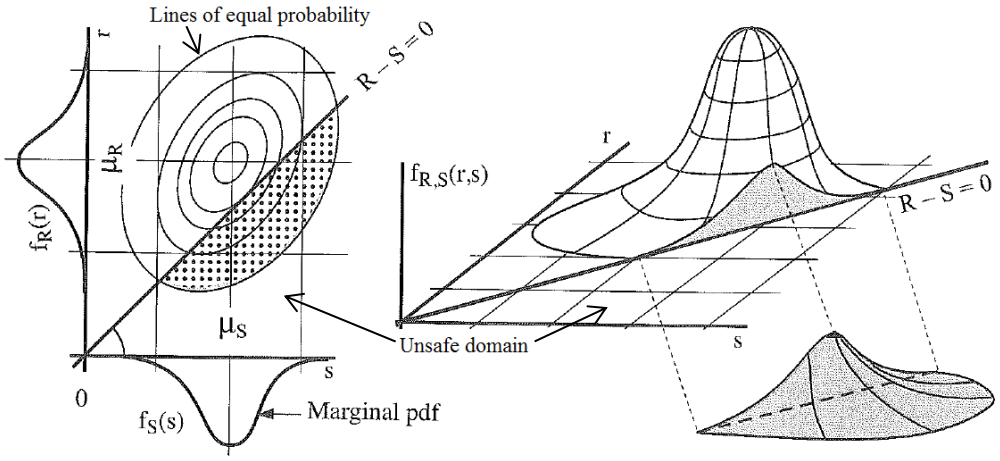


Figure 4.4 Probability of Failure

Mathematically this can be written as:

$$P_f = \iint_{Z<0} f_R(r)\, f_S(s)\, dr\, ds \tag{4.15}$$

However, as we have seen in sections 4.2 and 4.3, the resistance effect *R*, here represented by the resisting moment M_R, and the load effect *S*, here represented by the bending moment M_S, depend on a number of quantities. Thus, *R* and *S* are functions of variables that can be both stochastic and deterministic and can be represented as follows:

$$R = R(X_1, X_2 \ldots\ldots X_m) \tag{4.16}$$

$$S = S(X_{m+1}, X_{m+2} \ldots\ldots X_n) \tag{4.17}$$

The distribution functions of *R* and *S* therefore depend on the distribution functions of these variables.
The limit state function is now:

$$Z = R - S = Z(X_1, X_2 \ldots\ldots X_n) \tag{4.18}$$

And the failure probability (*R* and *S* being independent) becomes:

$$P_f = \iiint\limits_{Z<0} \ldots \int f_R(x_1, x_2 \ldots\ldots x_m)\, f_S(x_{m+1}, x_{m+2} \ldots\ldots x_n)\, dx_1 dx_2 \ldots\ldots dx_n \tag{4.19}$$

There are several ways of determining the failure probabilities. They are explained in the next section.

4.5 Reliability Methods

4.5.1 General

Generally, methods to calculate the reliability of a structure can be divided into five groups:

- **Level IV methods (risk-based)**: In these methods the consequences (costs) of failure are also taken into account and the risk (consequence multiplied by the probability of failure) is used as a measure of the reliability. In this way different designs can be compared on an economic basis taking into account uncertainty, costs and benefits.
- **Level III methods (numerical)**: The uncertain quantities are modelled by their joint distribution functions. The probability of failure is calculated exactly, e.g. by numerical integration.
- **Level II methods (approximation)**: The uncertain parameters are modelled by the mean values and the standard deviations, and by the correlation coefficients between the stochastic variables. The stochastic variables are implicitly assumed to be normally distributed.
- **Level I methods (semi-probabilistic design)**: The uncertain parameters are modelled by one characteristic value for load and resistance as for example in codes based on the partial coefficients (γ's) concept.
- **Level 0 methods**: Deterministic calculations.

If the reliability methods are used in design they have to be calibrated so that consistent reliability levels are obtained. Level I methods can e.g. be calibrated using level II methods, level II methods can be calibrated using level III methods, etc.

4.5.2 Level III methods

When a reliability method of level III is applied, the probabilistic formulation for P_f is calculated exactly, using analytical formulations, numerical integration or Monte Carlo simulations. Solving the problem using analytical expressions is only possible in a limited number of simple cases; numerical integration is only practical when the number of basic variables n is small.

4.5.3 Level II methods

In case of level II methods only the mean values of the basic variables and the moments of first and second order (covariance matrix) are used in most cases. The joint probability density function is simplified and the computational effort is reduced by linearization of the limit state function, usually with a technique called the First Order Reliability Method. In this method, the limit state function is linearized in the so-called design point, i.e. the point on $g(\underline{X}) = 0$ with the highest probability density, thus the point where failure is most probable.

4.5.4 Level I methods

When applying semi-probabilistic methods of level I, the variables whose probabilistic distributions have to be taken into account are represented by a characteristic value that corresponds to a low percentile in case of strength distributions or a high percentile in case of distributions related to actions. Further, partial factors are introduced with values that are based on level II calculations.

The basic verification format mentioned in EN 1990 consists of verifying whether the limit state is not exceeded when all basic variables in the limit state equation are replaced by so-called design values (designated with subscript "*d*"). In case of a simple limit state function as given in $Z=R-S$ one has to verify whether the design resistance R_d is at least equal to the design value of the load effect S_d, i.e.:

$$S_d \leq R_d \tag{4.20}$$

where

$$S_d = E(F_{d1}, F_{d2}, \ldots, a_{d1}, a_{d2}, \ldots, \theta_{d1}, \theta_{d2}, \ldots) \tag{4.21}$$

$$R_d = E(X_{d1}, X_{d2}, \ldots, a_{d1}, a_{d2}, \ldots, \theta_{d1}, \theta_{d2}, \ldots) \tag{4.22}$$

with

- F an action
- X a material property
- a a geometrical property
- θ a variable representing the model uncertainty

4.5.5 Level 0 methods

These are deterministic methods which use deterministic or nominal values of the basic variables and one (empirical) global safety factor. The verification is performed according to an equation with the following format:

$$R_{nom} \geq \gamma \, S_{nom} \tag{4.23}$$

4.5.6 Closing remark

The introduction of probability-based calculation methods was among others due to the observation that deterministic methods resulted in scattered safety levels and that no coherent safety methodology is available in case of a new technology/material.

The following chapters in these lecture notes will further elaborate on the level III, II and I probabilistic methods.

Chapter "Level III Methods"
Author: R.D.J.M. Steenbergen and A. Kosters

A few parts of this chapter are based on 'Risk Analysis of Construction Processes' by L. Taerwe and R. Caspeele (UGent lecture notes)

Contents

5 Level III Methods

5.1 Introduction

Methods of level III evaluate the following integral explicitly:

$$P_f = \int_{g(\underline{X})<0} f_{\underline{X}}(\underline{x})\, d\,\underline{x} \tag{5.1}$$

In the following sections a number of basic cases are explained for which exact calculation is possible.

5.2 General Case with Independent Normally Distributed Variables

In the case of independent normally distributed random variables and a linear limit state function, a level III probabilistic calculation can be easily done by hand.
As R and S are independent normally distributed random variables, the variable $Z = R - S$ is also normally distributed with parameters:

$$\mu_Z = \mu_R - \mu_S, \qquad \sigma_Z = \sqrt{\sigma_R^2 + \sigma_S^2} \tag{5.2}$$

Hence one obtains:

$$P_f = P\left[Z < 0\right] = \Phi\left[\frac{0 - \mu_Z}{\sigma_Z}\right] = \Phi(-\beta) \tag{5.3}$$

The **reliability index** β gives the distance between μ_Z (the mean value of Z) and $Z = 0$ in σ_Z (standard deviation of Z) units as indicated in Figure 5.1. This figure depicts the distribution of:

$$u_Z = \frac{Z - \mu_Z}{\sigma_Z} \tag{5.4}$$

Such that $\mu_{uz} = \frac{\mu_Z - \mu_Z}{\sigma_Z} = 0$.

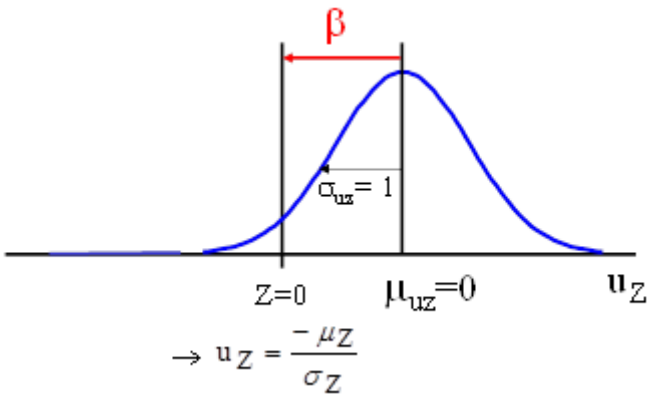


Figure 5.1 Distribution of $Z = R - S$ and reliability index

The relation between β and P_f is illustrated in Figure 5.2 and Table 5.1. It is noted that Table 2.2 gives a more comprehensive overview of the values for the cumulative normal distribution.

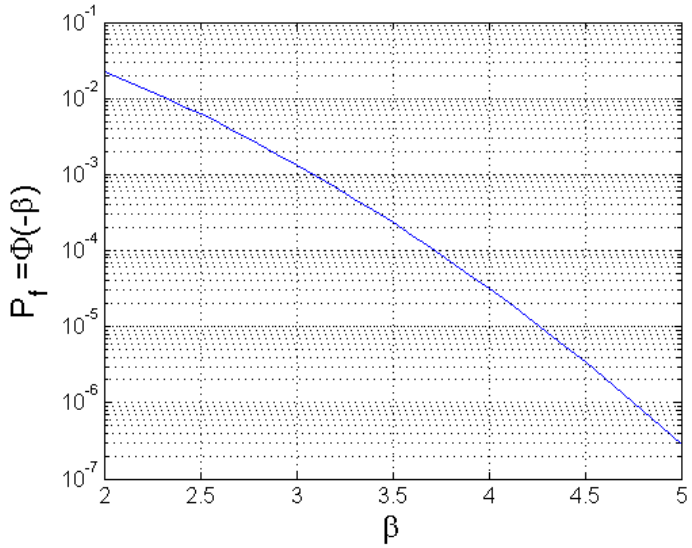


Figure 5.2 Probability of failure P_f against reliability index β on 10Log scale

Table 5.1 Relationship between P_f and β .

P_f	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
β	1.28	2.32	3.09	3.72	4.27	4.75	5.20

As $\Phi(-\beta) = 1-\Phi(\beta)$, the following holds for the probability of survival according to (5.3):

$$P_s = 1 - P_f = \Phi(\beta) \tag{5.5}$$

Example 5.1

Consider a system with in which the load S has a normal distribution $\mu(S)$=20kN and $\sigma(S)$=3kN. The strength of the system is normally distributed with $\mu(R)$=25kN and $\sigma(R)$=1kN. Loads and strength are independent. The probability of failure can be assessed by considering the limit state $Z = R - S$; and $P(Z<0)$.

It follows that $\mu(Z) = \mu(R) - \mu(S)$= 25 kN – 20 kN = 5 kN.

The standard deviation equals: $\sigma(Z) = (\sigma(R)^2 + \sigma(S)^2)^{0.5}$ = 3.16 kN .

The probability that the system fails is $P(Z<0)$. It follows from the normal distribution (see section 2.6.1) that $P(Z<0) = \Phi(- \mu(Z)/\sigma(Z))$ = 0.057.

Example 5.2

A reinforced concrete column with cross-section 500 x 300 mm consists of 14 longitudinal reinforcement bars Ø 20 mm. The strength R of the column can be calculated according to the following formula:

$$R = A_C \cdot R_C + A_S \cdot R_S \tag{5.6}$$

with

- A_c : the cross-sectional concrete area
- A_s : the cross-sectional steel area
- R_c : the concrete compressive strength
- R_s : the yield strength of the steel

The strength of the materials are considered to be normally distributed according to:

$$\begin{aligned} &R_c : N\ (35\ \text{N/mm}^2; 5\ \text{N/mm}^2) \\ &R_s : N\ (450\ \text{N/mm}^2; 30\ \text{N/mm}^2) \end{aligned} \tag{5.7}$$

The total load S on the column consists of a permanent load G and a variable load Q according to:

$$S = G + Q \tag{5.8}$$

The distribution of both types of loads is considered to be normal with the following mean and standard deviation:

$$\begin{aligned} &G : N\ (2000\ \text{kN}; 150\ \text{kN}) \\ &Q : N\ (1500\ \text{kN}; 500\ \text{kN}) \qquad \text{(distribution annual maxima)} \end{aligned} \tag{5.9}$$

Determine the failure probability P_f of the column. The strength R is normally distributed with parameters:

$$\begin{aligned} &\mu_R = A_c\, \mu_{R_c} + A_s\, \mu_{R_s} = 145602 \cdot 35 + 4398 \cdot 450 = 7075\ \text{kN} \\ &\sigma_R^2 = A_c^2\, Var[R_c] + A_s^2\, Var[R_s] = (150000 - 4398)^2 \cdot 25 + 4398^2 \cdot 900 \\ &\sigma_R = 740\ \text{kN} \end{aligned} \tag{5.10}$$

Hence, R : N(7075 kN, 740 kN) and V_R = 10.5%.

The total load *S* is also normally distributed with parameters:

$$\mu_S = 2000 + 1500 = 3500 \text{ kN}$$
$$\sigma_S = \sqrt{150^2 + 500^2} = 522 \text{ kN} \quad (5.11)$$

Hence, *S : N* (3500 kN, 522 kN) and V_S = 14.9 %. It follows that $\mu_Z = \mu_R - \mu_S = 7075 - 3500 = 3575$ kN and $\sigma_Z = (\sigma_R^2+\sigma_S^2)^{0.5} = (740^2+522^2)^{0.5} = 905.6$ kN.

This results in: $P_f = \Phi(-3.9459) = 0.3976 \cdot 10^{-4} = 4 \cdot 10^{-5}$.

5.3 (Numerical) Integration

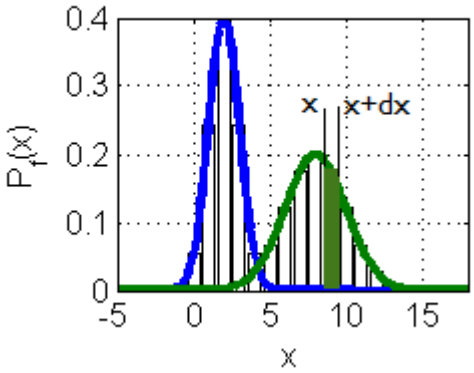
Consider the resistance variable *R* and the load effect *S* so that the limit state function can be written as $g(\underline{X}) = Z = R - S = 0$ and $P_f = P[R < S]$. Figure 5.3A shows the probability density functions of R and S and Figure 5.3B shows the joint probability density function. The domain *X* is divided up into bins $x_i \in X$ with a spacing of dx. When the limit of the spacing $dx \to 0$ is taken, the joined probability density function can be formulated in two ways, namely as:

$$P_f = \lim_{dx\to 0} \sum_{x\in X} P\left[(R < x) \cap (x \le S \le x + dx)\right] = \int_{-\infty}^{+\infty} F_R(x)\, f_S(x)\, dx \quad (5.12)$$

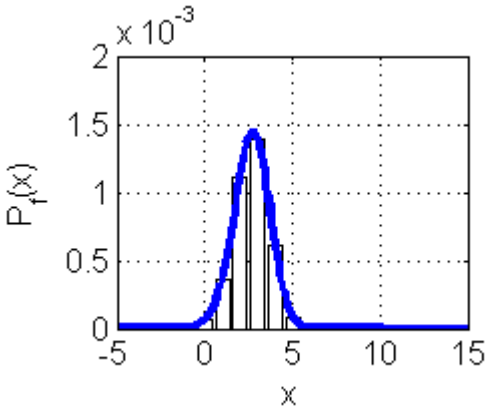
As depicted in Figure 5.3C, or alternatively as:

$$P_f = \lim_{dx\to 0} \sum_{x} P\left[(x \le R \le x + dx) \cap (S > x)\right] = \int_{-\infty}^{+\infty} f_R(x)\left[1 - F_S(x)\right] dx \quad (5.13)$$

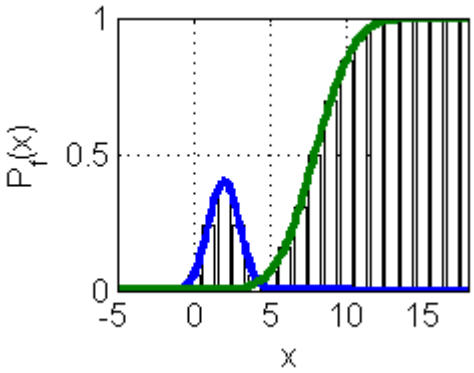
As depicted in Figure 5.3D. The indicated integration boundaries should be considered as formal designations.



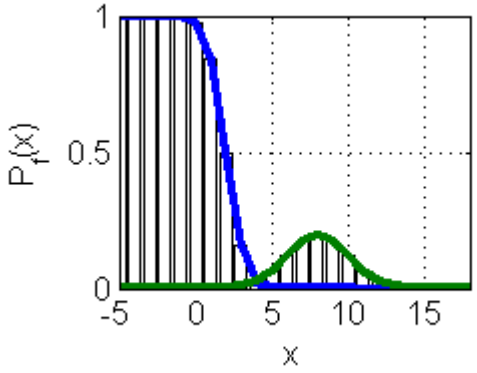
A. Probability density functions of load S (blue) and resistance R (green).



B. Joint probability density function.



C. First integration method: probability density function of load S (blue) and cumulative distribution function of resistance R (green).



D. Second integration method: 1 minus the cumulative distribution function of load S (blue) and probability density function of resistance R (green).

Figure 5.3 Distribution functions associated with the calculation of P_f

Writing this in another way we obtain the same result:

$$P_f = \iint_{r<s} f_R(r)\, f_S(s)\, dr\, ds = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{R=s} f_R(r)\, dr \right] f_S(s)\, ds = \int_{-\infty}^{+\infty} F_R(s)\, f_S(s)\, ds \tag{5.14}$$

Or:

$$P_f = \int_{-\infty}^{+\infty} \left[\int_{S=r}^{+\infty} f_S(s)\, ds \right] f_R(r)\, dr = \int_{-\infty}^{+\infty} \left[1 - F_S(r)\right] f_R(r)\, dr \tag{5.15}$$

This integral is in literature often indicated as the “convolution integral”.

This procedure is illustrated in Figure 5.4 for $Z=R-S$ with R normally distributed with mean 60 kN and standard deviation of 5 kN and S normally distributed with mean 40 kN and standard

deviation of 10 kN and *R* and *S* being independent. It can be seen that the volume under the joint probability density function in the region $Z < 0$ can be written as:

$$P_f = \int_{-\infty}^{+\infty} \left[f_R(r) \int_{S=r}^{+\infty} f_S(s)\, ds \right] dr \tag{5.16}$$

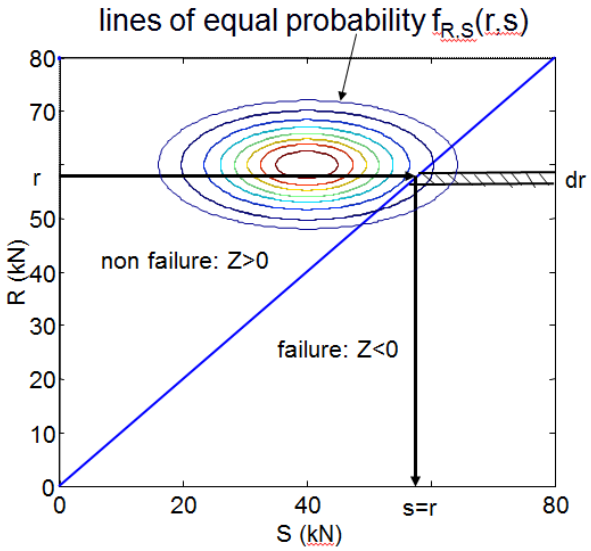


Figure 5.4 Analytical solving of integral

The above mentioned integrals can also be solved numerically, i.e. by means of numerical integration. An example is given below:

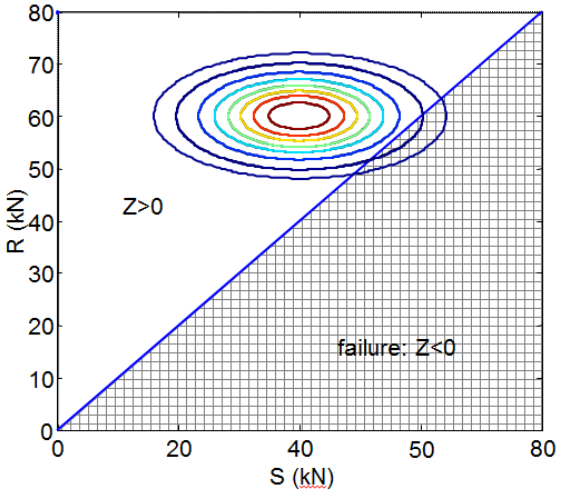


Figure 5.5 Numerical integration

The failure probability is now calculated splitting the volume $Z<0$ in small volumes

$$P_f = \sum_i \sum_j f_{R,S}\left(r_i, s_j\right) \Delta r \Delta s \tag{5.17}$$

For non-linear limit state functions the same procedure can be used. In general *R* as well as *S* will also be functions of a number of random variables, hence:

$$g(\underline{X}) = g_R(R_1, ..., R_i) - g_S(S_1, ..., S_j) \tag{5.18}$$

The single integrals can thus become multiple integrals which makes the calculation difficult. Standard numerical integration techniques are computationally intensive since the number of integration steps increases exponentially with the number of stochastic variables.

5.4 Monte Carlo Simulations

5.4.1 General

As previously mentioned, the calculation of P_f through numerical integration is rather difficult in case $n > 2$. In those cases, the only practical solution in order to calculate P_f is by Monte-Carlo simulations, generating random samples. One of the possible ways to generate random samples from a certain distribution function is through the cumulative distribution function.

The Monte Carlo method uses the possibility of drawing random numbers from a uniform probability density function F_U between zero and one. Practically all programming languages include a standard procedure for this, for example the RAND program in Excel. The idea behind the method is to generate a random number x from an arbitrary distribution $F_X(x)$ by drawing a number x_u from the uniform distribution between zero and one.

The cumulative probability $P(X \leq x)$of a uniform distributed random variable on the interval $[a, b]$ is as given in equation 2.77. In case of a uniform distribution on [0,1] the CDF reads:

$$F_U(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases} \tag{5.19}$$

Let $F_X(x)$be an arbitrary cumulative distribution function of a random variable $X$of interest, and $F_X^{-1}(x)$ its inverse. Let x_U denote a random realization of $F_U(x)$ which is then treated as a cumulative probability for $F_X(x)$. The realization x corresponding to the cumulative probability x_U is then

$$x = F_X^{-1}(x_u) \tag{5.20}$$

See Figure 5.6 for a graphical interpretation of the procedure.

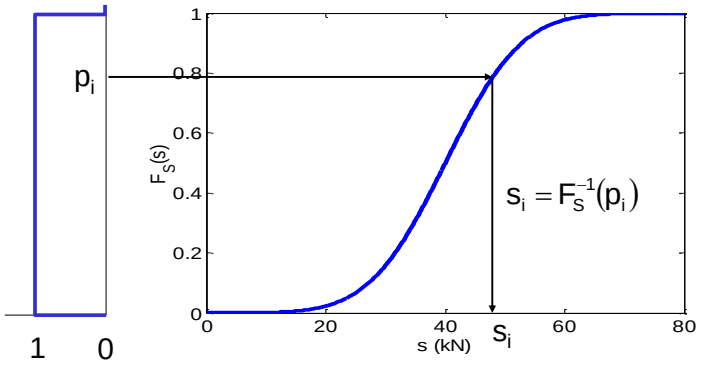


Figure 5.6 Generation of random samples

This way of drawing random numbers is generally applicable. However, for distributions, for which the inverse probability distribution function $F_X^{-1}(x_u)$ is not known analytically, this method can lead to a lot of iterative calculations. Hence, other less computationally intensive methods for drawing from (for example) a normal distribution exist.

In more or less the same way, base variables of a statistical vector can be drawn from a known joint probability distribution function. However, the joint probability distribution function must then be formulated as the product of the conditional probability distributions of the base variables of the vector. In formula this is:

$$F_{\vec{X}}(\vec{x}) = F_{X_1}(\mathrm{x}_1)F_{X_2|X_1}(\mathrm{x}_2 \mid \mathrm{x}_1) \ldots F_{X_m|X_1,X_2,\ldots,X_{m-1}}(\mathrm{x}_m \mid \mathrm{x}_1,\mathrm{x}_2,\ldots,\mathrm{x}_{m-1}) \tag{5.21}$$

By taking *m* realizations of the uniform probability distribution between zero and one, a value can be determined for every x_i:

$$\begin{aligned} x_1 &= F_{X_1}^{-1}\left(x_{u_1}\right) \\ x_2 &= F_{X_2|X_1}^{-1}\left(x_{u_2}\,|x_1\right) \\ &\cdot \\ &\cdot \\ &\cdot \\ x_m &= F_{X_m|X_1,X_2,\ldots,X_{m-1}}^{-1}\left(x_{u_m}\,|x_1,\mathrm{x}_2,\ldots,\mathrm{x}_{m-1}\right) \end{aligned} \tag{5.22}$$

If the base variables are statistically independent, this can be simplified to:

$$x_i = F_{X_i}^{-1}(\mathrm{x}_{u_i}) \tag{5.23}$$

This corresponds to equation (5.20). By inserting the values for the reliability function(s) one can check whether the obtained vector $(x_1, x_2, \ldots, x_m)$ is located in the safe area.

The Monte Carlo procedure now works as follows: For each basic variable X_i $(i = 1, \ldots, n)$ one simulates *N* realizations $x_{i1}, x_{i2}, \ldots, x_{iN}$.

For each set j $(j = 1, \ldots, N)$ one calculates $g(x_{1j}, x_{2j}, \ldots, x_{nj})$. In case $g(.) < 0$ a counter N_f is increased by one. After *N* simulations one calculates:

$$\hat{P}_f = \frac{N_f}{N} \tag{5.24}$$

In case $N \rightarrow \infty$ one obtains the failure probability P_f. Equation (5.24) can also be rewritten as:

$$\hat{P}_f = \frac{1}{N}\sum_{j=1}^{N} I\left[g(\underline{x}) < 0\right] \tag{5.25}$$

with *I[.]*, the indicator function which is equal to 1 in case the argument of the operator is true and otherwise 0 in case the argument is false. An example of the outcomes of a Monte Carlo simulation is shown in Figure 5.7.

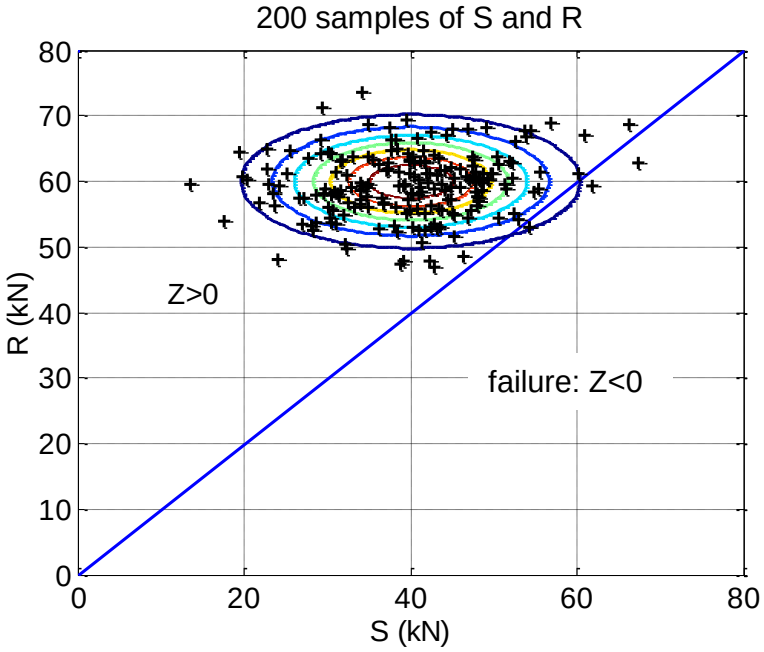


Figure 5.7 Example of the output of a Monte Carlo simulation with 200 samples. The line shows the limit state.

The number of simulations N is determined in function of the number of significant digits with which one wants to calculate P_f or based on the relative accuracy.

Considering that N_f follows a binomial distribution (see Table 2.1), the variance or standard deviation of the relative frequency $\hat{P}_f$ can be found as follows:

$$Var[N_f] = NP_f(1-P_f) \tag{5.26}$$

$$\hat{P}_f = \frac{N_f}{N} \tag{5.27}$$

$$Var\left[\hat{P}_f\right] = Var\left[\frac{N_f}{N}\right] = \frac{1}{N^2}Var\left[N_f\right] = \frac{1}{N}P_f(1-P_f) \tag{5.28}$$

$$\sigma_{\hat{P}_f} = \sqrt{\frac{P_f(1-P_f)}{N}} \tag{5.29}$$

As $P_f << 1$, the standard deviation can be approximated with:

$$\sigma_{\hat{P}_f} \cong \sqrt{\frac{P_f}{N}} \tag{5.30}$$

And the following coefficient of variation can be obtained:

$$V_{\hat{P}_f} = \frac{\sigma_{\hat{P}_f}}{P_f} \cong \sqrt{\frac{P_f}{NP_f{}^2}} = \frac{1}{\sqrt{NP_f}} \tag{5.31}$$

In case of a target coefficient of variation (relative error) the number of required simulations N increases as P_f decreases. For $V = 0.01$ and $P_f = 1 \cdot 10^{-5}$ one has to execute 109 simulations. As for this case on average only one of 105 simulations leads to an increase of N_f, the frequency of obtaining a "success" (realization in the unsafe domain D_f) is very low. Hence, different "variance reducing" techniques have been developed. These techniques allow the simulations to be performed in a more effective way. In the following the so-called "importance sampling" technique is explained.

5.4.2 Importance sampling

The aim of this technique is to obtain more realizations of the random vector $\underline{X}$ which are located in the unsafe domain D_f in order to increase the frequency of "success". In order to achieve this, an appropriate "sampling function" $f_S(\underline{x})$ is chosen so that its maximum is located in the domain that contributes most to P_f. For example, one can choose the mean value $\underline{\mu}$ of $f_s(\underline{x})$ so that it coincides with the point on $g(\underline{x}) = 0$ which has the highest value according to $f_{\underline{X}}(\underline{x})$. Hence, the following equations can be obtained.

$$P_f = \int_{D_f} f_{\underline{X}}(\underline{x})\, d\underline{x} = \int_{D_{\underline{X}}} I\left[g(\underline{x}) < 0\right] f_{\underline{X}}(\underline{x})\, d\underline{x} \tag{5.32}$$

with D_f the unsafe domain and $D_{\underline{X}}$ the entire domain over which $\underline{X}$ is defined. The evaluation of the right-hand side of equation (5.32) is in practice performed according to (5.25) in which $\underline{x}$ follows a distribution according to $f_{\underline{X}}(\underline{x})$. Equation (5.32) can also be rewritten as follows:

$$P_f = \int_{D_{\underline{X}}} I\left[g(\underline{x}) < 0\right] \frac{f_{\underline{X}}(\underline{x})}{f_S(\underline{x})} f_S(\underline{x})\, d\underline{x} \tag{5.33}$$

With $f_S(\underline{x})$ the n-dimensional "importance sampling" PDF, which is also defined over the same domain $D_{\underline{X}}$. In this case the random numbers x follow a distribution according to $f_S(\underline{x})$. One can show that according to (5.33) the following holds:

$$P_f = \frac{1}{N} \sum_{j=1}^{N} I\left[g(\underline{x}) < 0\right] \tag{5.34}$$

with $\underline{x}$ distributed according to $f_{\underline{X}}(\underline{x})$. One can similarly prove that:

$$P_f = \frac{1}{N} \sum_{j=1}^{N} I\left[(g(\underline{x}) < 0\right] \frac{f_{\underline{X}}(\underline{x})}{f_S(\underline{x})} \tag{5.35}$$

with $\underline{x}$ distributed according to $f_S(\underline{x})$. This formulation can be compared to (5.25) holding for crude Monte Carlo simulations. The efficiency of this importance sampling technique is strongly dependent on the choice of $f_S(\underline{x})$.

In case one locates $f_S(\underline{x})$ at the design point $\underline{x}^*$ (see Chapter 6), i.e. the point on $g(\underline{x}) = 0$ with the highest probability of occurrence, the "success" rate will approximate 50% in case the limit state

function $g(\underline{x}) = 0$ is not too strongly curved in the neighbourhood of $\underline{x}^*$. One can for example obtain $f_S(\underline{x})$ by shifting $f_{\underline{X}}(\underline{x})$ to $\underline{x}^*$ or to the point with coordinates according to the design values of a method of level I.

Figure 5.8 gives an example of the importance sampling technique. In this example, the distribution functions of R and S are equal to those in Figure 5.4: R is normally distributed with μ_R= 60 kN and σ_R= 5 kN, whereas S is normally distributed with μ_S= 40 kN and σ_S= 10 kN. For this problem, the design point is located at S=R=56 kN (see Chapter 6).

Both for R and S, the same sampling distribution is chosen, with the mean located close to the design point, at S=R=60 kN, whereas the standard deviation is set at 5 kN. From this distribution function, 200 samples are generated that represent values for S, and another 200 samples that represent values for R. For each sample pair, Z=R-S is calculated. Because the values for R and S are taken from the same distribution function, failure will be observed in about 50% of the cases. Subsequently, the generated sample pairs (r,s) are substituted in the joint sampling distribution function $f_S(\underline{x})$, which is in this case described by:

$$f_S(\underline{x}) = f_S(r) \cdot f_S(s) \tag{5.36}$$

And in the real joint probability density function:

$$f_{\underline{x}}(\underline{x}) = f_R(r) \cdot f_S(s) \tag{5.37}$$

Taking the ratio $f_{\underline{x}}(\underline{x}) / f_S(\underline{x})$ for each sample pair, the failure probability can now be calculated according to equation (5.35).

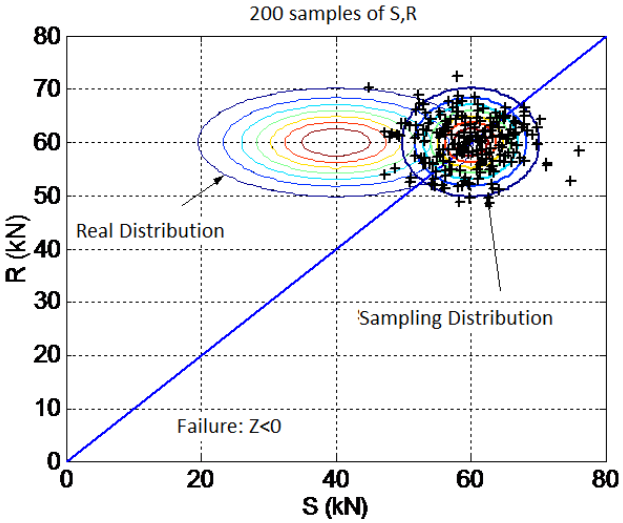


Figure 5.8 Importance sampling

Chapter "Level II Methods"
Author: R.D.J.M. Steenbergen and A. Kosters

Parts of this chapter are based on 'Introduction to Safety and Reliability of Structures' by J. Schneider and T. Vrouwenvelder.

Contents

6 Level II Methods

In level II methods the mean of the base variables and their covariance matrix are taken into account to determine the failure probability of the limit-state function. In section 6.1 the notion of design point is introduced. In section 6.2.1 a geometric argument is made, identifying the reliability index as the shortest distance from the origin to the boundary of the failure space in phase space. In section 6.2.2 the reliability index is then connected to the design point. Thereafter, two iterative methods are introduced to find the design point and from there the reliability index in the case of independent base variables. The first method makes use of transformation to standard normal variables (section 6.2.3). It is derived that one can also directly iterate on the original limit state function (section 6.2.4). In section 6.3 it is discussed how the design point should be interpreted in the case of non-normally distributed variables. At last, in section 6.4 it is explained how the design point can be found in the case of dependent base variables, taking the dependence between the variables into account through conditional probabilities.

6.1 General

If the reliability function is linear, the expected value and the standard deviation of this function can be determined with:

$$
\begin{aligned}
Z &= a_1 X_1 + a_2 X_2 + ... + a_n X_n + b \\
\mu_Z &= a_1 \mu_{X_1} + a_2 \mu_{X_2} + ... + a_n \mu_{X_n} + b \\
\sigma_Z &= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)}
\end{aligned}
\tag{6.1}
$$

With σ_Z as defined in equation (2.155) if it is recognized that $g_i' = \frac{dZ}{dX_i} = a_i$ and $Cov(X_i, X_j)$ as defined in equation (2.116). If the base variables $X_1, X_2, ..., X_n$ are normally distributed, Z is also normally distributed. The probability that $Z < 0$, can then be determined using the standard normal distribution:

$$
P(Z<0) = \Phi\left(\frac{0 - \mu_Z}{\sigma_Z}\right) = \Phi\left(-\frac{\mu_Z}{\sigma_Z}\right) = \Phi(-\beta) \tag{6.2}
$$

Hence, for a linear reliability function with normally distributed base variables it is relatively simple to calculate the probability of failure. The reliability index β was defined by Cornell (1969) as:

$$
\beta = \frac{0 - \mu_Z}{\sigma_Z} = \frac{\mu_Z}{\sigma_Z} = \frac{1}{V_Z} \tag{6.3}
$$

Here V_Z is the coefficient of variation of Z. β was also already introduced in section 5.3.1 for normally distributed variables. As β increases, the failure probability decreases and the reliability or safety increases.

In case of a non-linear limit state equation, the limit state equation can be linearized using a Taylor expansion around the point $\underline{X} = \underline{\mu}$. This is also called the Mean-value approach, as was described in section 2.5.8. Hence:

$$
g(\underline{X}) = Z \cong g(\mu_1, ..., \mu_n) + \sum_{i=1}^{n} \frac{\partial g(\mu_i)}{\partial X_i} (X_i - \mu_i) \tag{6.4}
$$

which is again a linear function of the basic variables X_i . Further, one can show that:

$$
\mu_Z \cong g(\mu_1, ..., \mu_n) \tag{6.5}
$$

$$
\sigma_Z^2 \cong \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial g(\underline{\mu})}{\partial X_i} \frac{\partial g(\underline{\mu})}{\partial X_j} Cov\left[X_i, X_j\right] \tag{6.6}
$$

The reliability index β can be calculated using (6.3). The disadvantage of this calculation method (which considers a linearization in the centre of gravity of the probability mass), is the fact that the reliability index is not invariant with respect to the formulation of the limit state equation.

It is of course not beneficial that the numerical value of β depends on the way in which the problem is mathematically formulated. This problem can be overcome by executing the linearization at the so-called design point.

An illustration with an example is given below. We consider a ball which is connected to the ceiling, see Figure 6.1. The strength of the wire is depending on the ultimate stress f of the material and the diameter d of the wire; these are assumed to be independent normally distributed variables. The load S is deterministic.

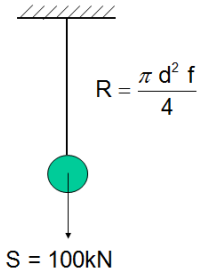


Figure 6.1 Example: ball connected to the ceiling

We take S=100 kN, f = N(290 N/mm^2, 25 N/mm^2) and d = N(30 mm, 3 mm). The limit state function becomes:

$$Z = \frac{\pi d^2 f}{4} - S \tag{6.7}$$

In the picture below the joint probability density function of f and d is drawn together with the marginal distributions and the limit state function Z. The failure probability is the volume under the joint probability density function in the region $Z<0$. In the left hand picture below the Z-function is linearized in the mean values of f (written as σ in the figure) and d. In the right hand picture the Z-function is linearized in the design point. The exact definition of the design point will be given in the next sections, but already here it becomes clear that the linearization of the Z-function in the design point gives a much better approximation of the failure probability.

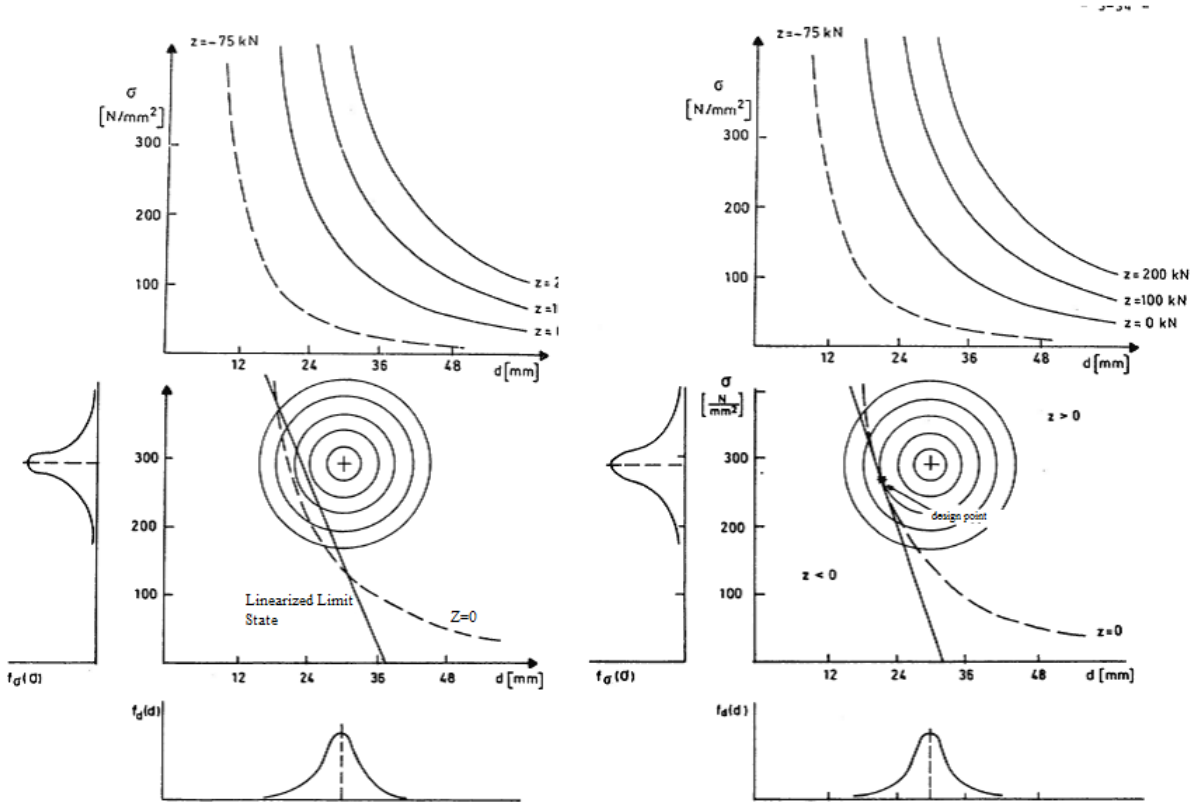


Figure 6.2 Linearization in mean values and in design point

6.2 Reliability Index According to Hasofer en Lind

6.2.1 Basic formulation

Hasofer and Lind (1974) introduced a generalized reliability index which is invariant with respect to the formulation of the limit state equation and which is currently commonly applied in structural reliability analysis.

Let us first consider **uncorrelated normally distributed variables**. First, the basic variables X_i have to be normalized according to:

$$U_i = \frac{X_i - \mu_i}{\sigma_i} \tag{6.8}$$

with $\mu_i = E[X_i]$ and $\sigma_i^2 = Var[X_i]$. In case of normalized basic variables U_i it holds that $E[U_i] = 0$ and $Var[U_i] = 1$. The limit state equation becomes $g(\underline{U}) = 0$ in the n-dimensional U-space and the surface $\partial\omega$ that is described by this equation divides the U-space in an unsafe domain D_f (*failure region*) and a safe domain D_s (*safe region*). The origin of the U-space, which coincides with the maximum of the joint probability density function, is normally located in the domain D_s. The joint density function can be written as:

$$f_{\underline{U}}(\underline{u}) = \frac{1}{(2\pi)^{n/2}} . e^{-\frac{u_1^2}{2}} . e^{-\frac{u_2^2}{2}} ... e^{-\frac{u_n^2}{2}}$$
$$f_{\underline{U}}(\underline{u}) = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\left(u_1^2 + u_2^2 + ... + u_n^2\right)\right] \tag{6.9}$$

which is rotation symmetrical and hence the contour lines of equal probability are concentric circles.

The basic equation $g(\underline{X}) = R - S = 0$ can – based on (6.8) – be rewritten in function of the normalized basic variables U_R and U_S as:

$$g(\underline{U}) = \sigma_R U_R - \sigma_S U_S + (\mu_R - \mu_S) = 0 \tag{6.10}$$

In Figure 6.3 the equations $g(\underline{X}) = 0$ and $g(\underline{U}) = 0$ are drawn for this simple case. Note that in the U-coordinate system, the straight line $g(\underline{U}) = 0$ no longer passes through the origin. From the intersections with the axes, which can be found by substituting $U_R=0$ and $U_S=0$ in (6.10), it follows that:

$$OC = \frac{\mu_R - \mu_S}{\sigma_S}$$
$$OB = \frac{\mu_R - \mu_S}{\sigma_R} \tag{6.11}$$

So, by using Pythagoras, that:

$$BC = (\mu_R - \mu_S)\sqrt{\frac{1}{\sigma_R^2} + \frac{1}{\sigma_S^2}} \tag{6.12}$$

From the proportionality $OA/OB = OC/BC$ it follows that:

$$OA = \frac{\mu_R - \mu_S}{\sigma_S} . \frac{\mu_R - \mu_S}{\sigma_R} . \frac{1}{(\mu_R - \mu_S)\sqrt{\frac{1}{\sigma_R^2} + \frac{1}{\sigma_S^2}}} = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}} \tag{6.13}$$

and hence ***OA* is equal to the reliability index β**.

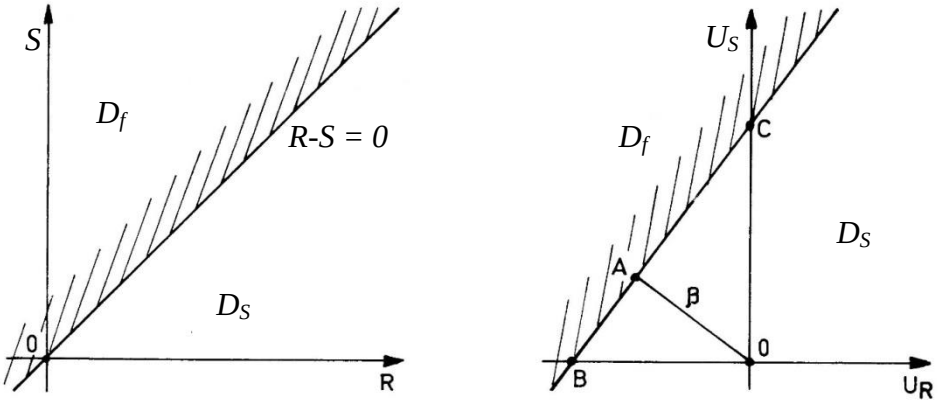


Figure 6.3 Linear limit state equation in the *X*- and *U*-space

This property also follows from the following derivation. From analytical geometry it is known that when an equation of a linear relationship:

$$a_0 + a_1 X_1 + a_2 X_2 = 0 \tag{6.14}$$

is rewritten in the shape:

$$\frac{a_1 X_1 + a_2 X_2 + a_0}{\sqrt{a_1^2 + a_2^2}} = 0 \tag{6.15}$$

which is called the Hesse normal form – the factor:

$$\left| \frac{a_0}{\sqrt{a_1^2 + a_2^2}} \right| \tag{6.16}$$

is the distance between the origin and the linear relationship. In case one rewrites (6.10) in the shape (6.15), one obtains that (6.13), i.e. β, is the distance from the origin to $g(\underline{U})$.

Hasofer and Lind generalized this property by postulating:

> *The reliability index β is equal to the shortest distance from the origin to the surface described by g(U) = 0 in the space of the normalized basic variables.*

This statement is illustrated with Figure 6.4, which shows the concentric circles of the joint probability density function, and the reliability index β as the shortest distance from the origin to the limit state function $g(\underline{U}) = 0$. In this figure it can easily be seen that a larger value of β gives a more reliable system: The further the straight line passes from the origin, the greater is β and the smaller is the cut-away volume representing P_f.

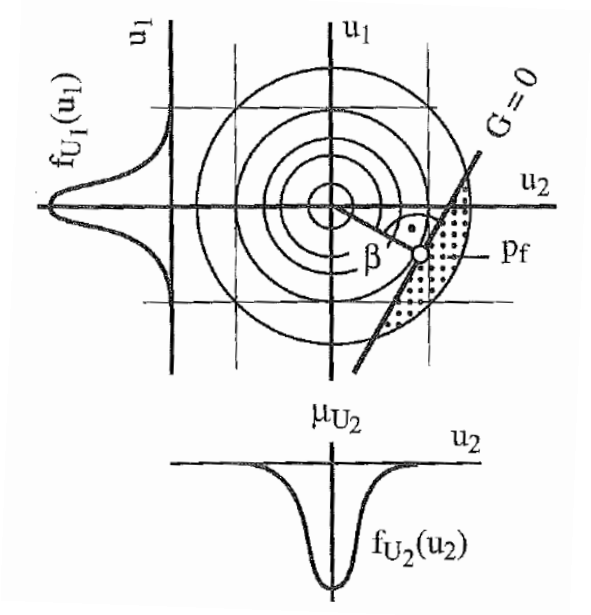


Figure 6.4 Standardised normal space

6.2.2 Non Linear Limit State Functions

The failure area is indicated in the U_1,U_2-plane in Figure 6.5. The figure shows that a linearization of the Z-function in different points leads to different values for the approximation of the reliability index. The formulation of the reliability index according to equation (6.13) may therefore not be used indiscriminately.

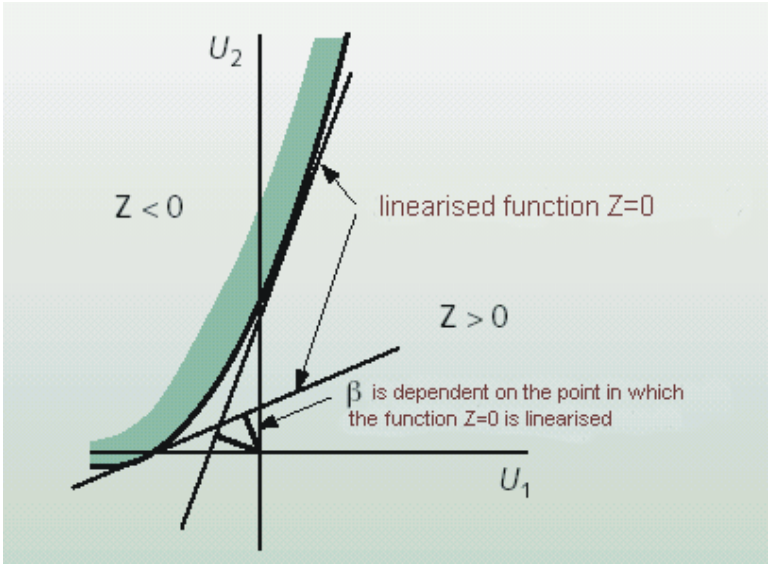


Figure 6.5 Linearization of the reliability function.

The definition of the reliability index according to Hasofer and Lind (1974) does not depend on whether or not the reliability function is linear. The distance from the edge of the failure area ($Z = 0$) to the origin of the transformed coordinate system is:

$$\beta = \min_{Z=0}\left(\sqrt{U_1^2 + U_2^2}\right) \tag{6.17}$$

The point A on the edge of the failure area, with the smallest distance to the origin is the design point.

From the geometrical representation it follows that in the general situation of more random variables the vector β is perpendicular to the hyperplane which is tangent to the failure surface $\partial\omega$ in the point closest to the origin. This point is called the design point $\underline{u}^*$. The corresponding point in the original space of the basic variables is designated as $\underline{x}^*$. The design point is the point of the limit state equation with the highest probability density, hence in literature one often mentions this as the "***most probable failure point***". This property becomes immediately clear when considering that the contour lines of equal probability are hyperspheres in the normalized space (circles in case $n = 2$, see Figure 6.4). This follows also analytically from the joint density function:

$$\begin{aligned} f_{\underline{U}}(\underline{u}) &= \frac{1}{(2\pi)^{n/2}} . e^{-\frac{u_1^2}{2}} . e^{-\frac{u_2^2}{2}} \ldots e^{-\frac{u_n^2}{2}} \\ f_{\underline{U}}(\underline{u}) &= \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\left(u_1^2 + u_2^2 + \ldots + u_n^2\right)\right] \end{aligned} \tag{6.18}$$

When $g(\underline{U}) = 0$, this expression becomes maximal for the point where Σu_i^2 becomes minimal.

Finding the design point is an iterative process, for which several methods are available. The numerical calculation of β can be executed according to a general calculation algorithm for the determination of a constraint minimum of a function.

Two methods will be explained here. The core of the methods is actually the same, but the formulation of the reliability function differs.

6.2.3 Method 1: transformation to normal variables

The first method is based on the standardized reliability function, i.e. all base variables are transformed to the standard normally distributed variables.

We apply a first order Taylor expansion (a linearization) of $g(\underline{U})$ around the design point $\underline{u}^*$:

$$g(\underline{U}) = Z = g(\underline{u}^*) + \sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} (U_i - u_i^*) \tag{6.19}$$

and as $\underline{u}^*$ lies on $g(\underline{U}) = 0$ it follows that $g(\underline{u}^*) = 0$ so that:

$$Z = 0 + \sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} (U_i - u_i^*) = \sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} U_i - \sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} u_i^* \tag{6.20}$$

Based on equations (6.14), (6.15) if we identify $a_i = \frac{\partial g(\underline{u}^*)}{\partial U_i}$ and $a_0 = -\sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} u_i^*$

we can use (6.16) to write β as follows:

$$\beta = \frac{a_0}{\sqrt{\sum_{i=1}^{n} a_i^2}} = \frac{-\sum_{i=1}^{n} \frac{\partial g(\underline{u}^*)}{\partial U_i} . u_i^*}{\sqrt{\sum_{i=1}^{n} \left[\frac{\partial g(\underline{u}^*)}{\partial U_i} \right]^2}} = \frac{-\sum_{i=1}^{n} \frac{\partial g(\underline{x}^*)}{\partial X_i} (x_i^* - \mu_i)}{\sqrt{\sum_{i=1}^{n} \left[\frac{\partial g(\underline{x}^*)}{\partial X_i} \sigma_i \right]^2}} \tag{6.21}$$

Considering the notations in equation (6.14), one finds for the 2-dimensional case:

$$\alpha_1 = \frac{a_1}{\sqrt{a_1^2 + a_2^2}} \quad \text{and} \quad \alpha_2 = \frac{a_2}{\sqrt{a_1^2 + a_2^2}} \tag{6.22}$$

These are the direction cosines of the perpendicular or vector $\vec{\beta}$ under consideration. In case of n dimensions and application of (6.20) one can expand these equations towards:

$$\alpha_i = \frac{a_i}{\sqrt{\sum_{j=1}^{n} \left[a_j \right]^2}} = \frac{\frac{\partial g(\underline{u}^*)}{\partial U_i}}{\sqrt{\sum_{i=1}^{n} \left[\frac{\partial g(\underline{u}^*)}{\partial U_i} \right]^2}} = \frac{\frac{\partial g(\underline{x}^*)}{\partial X_i} \sigma_i}{\sqrt{\sum_{i=1}^{n} \left[\frac{\partial g(\underline{x}^*)}{\partial X_i} \sigma_i \right]^2}} \tag{6.23}$$

These values are called the **weight or sensitivity factors** as they are a measure for the relative importance of the standard deviation of a basic variable to the reliability index. The α_i values are the components of the unit vector according to $\overrightarrow{OA}$ (see equation (6.13)) that is:

$$\overrightarrow{OA} = \beta \vec{\alpha} = \vec{\beta} \tag{6.24}$$

From (6.21) it follows that $\beta = -\Sigma\, \alpha_i\, u_i^*$. This is fulfilled in case:

$$u_i^* = -\alpha_i\, \beta \text{ or } x_i^* = \mu_i - \alpha_i\, \beta\, \sigma_i \tag{6.25}$$

as $\Sigma\alpha_i^2 = 1$.

The equations in (6.25) determine the **coordinates u_i^* of the design point** in case of normally distributed variables. One obtains $\alpha_i > 0$ for resistance variables and the coordinate of the design point corresponds to a low percentile of the resistance distribution. In case of loads $\alpha_i < 0$ and according to (6.25) the corresponding design value is found to be a high value of the load distribution.

In order to determine the $n+1$ unknown parameters, i.e. α_1, α_2, ..., α_n and β, one can make use of the n equations (6.23) and $g(\underline{U}) = 0$ or:

$$\alpha_i = \frac{\dfrac{\partial\, g(-\underline{\alpha}\,\beta)}{\partial U_i}}{\sqrt{\displaystyle\sum_{i=1}^{n}\left[\frac{\partial\, g(-\underline{\alpha}\,\beta)}{\partial U_i}\right]^2}} \tag{6.26}$$

$$g(-\alpha_1\,\beta, -\alpha_2\,\beta, ..., -\alpha_n\,\beta) = 0 \tag{6.27}$$

These equations can be solved iteratively, preferably considering some carefully chosen starting solutions. In case one assigns a priori an equal weight to all variables, one can consider:

$$\alpha_i = \pm\, n^{-1/2} \tag{6.28}$$

In which the + should be applied for resistance variables and the – sign for load variables (see application examples). The method's procedure can best be illustrated with an example.

Example 6.1: Method 1, Transformation to normal variables

Consider the following limit state equation:

$$g(\underline{X}) = Z = X_1 X_2 - X_3 \tag{6.29}$$

The variables X_1, X_2 and X_3 are independent normally distributed random variables. Also given:

$$\begin{aligned} \mu_1 &= 8, \quad & \sigma_1 &= 2 \\ \mu_2 &= 3, \quad & \sigma_2 &= 1 \\ \mu_3 &= 4, \quad & \sigma_3 &= 2 \end{aligned} \tag{6.30}$$

Wanted is the determination of the design point with the corresponding reliability index. First, the base variables are transformed to standard normally distributed variables:

$$U_1 = \frac{X_1 - 8}{2}, \qquad U_2 = \frac{X_2 - 3}{1}, \qquad U_3 = \frac{X_3 - 4}{2} \tag{6.31}$$

The reliability function of the transformed variables is:

$$g(\underline{U}) = 6 \cdot U_1 + 2 \cdot U_1 U_2 + 8 \cdot U_2 - 2 \cdot U_3 + 20 \tag{6.32}$$

In the design point the following is valid (see equation (6.25)):

$$\begin{aligned} &g(-\alpha_1\beta, -\alpha_2\beta, -\alpha_3\beta) = 0 \\ &-6\alpha_1\beta + 2\alpha_1\alpha_2\beta^2 - 8\alpha_2\beta + 2\alpha_3\beta = -20 \end{aligned} \tag{6.33}$$

From this a formulation for β can be derived:

$$\beta = \frac{-20}{-6\,\alpha_1 + 2\,\alpha_1\alpha_2\,\beta - 8\alpha_2 + 2\alpha_3} \tag{6.34}$$

With equation (6.23)), the formulations for α_1, α_2 and α_3 can be derived as follows:

$$\frac{\partial g(\underline{U})}{\partial U_1} = 6 + 2U_2, \quad \frac{\partial g(\underline{U})}{\partial U_2} = 2U_1 + 8, \quad \frac{\partial g(\underline{U})}{\partial U_3} = -2 \tag{6.35}$$

$$\begin{aligned} \alpha_1 &= \frac{6 - 2\alpha_2\beta}{\sqrt{(6 - 2\alpha_2\beta)^2 + (8 - 2\alpha_1\beta)^2 + 2^2}} \\ \alpha_2 &= \frac{8 - 2\alpha_1\beta}{\sqrt{(6 - 2\alpha_2\beta)^2 + (8 - 2\alpha_1\beta)^2 + 2^2}} \\ \alpha_3 &= \frac{-2}{\sqrt{(6 - 2\alpha_2\beta)^2 + (8 - 2\alpha_1\beta)^2 + 2^2}} \end{aligned} \tag{6.36}$$

This system of equations formed by (6.34) and (6.36) can be solved by means of successive substitution. However, this is complex and therefore an iterative approach is chosen.

In this case, there is the problem of choosing realistic initial values for β, α_1, α_2 and α_3. The initial value of β can be determined with a Mean Value approximation:

$$\beta = \frac{\mu_Z}{\sigma_Z} \approx \frac{g(\mu_1, \mu_2, \mu_3)}{\sqrt{\left(\frac{\partial g}{\partial X_1}(\mu_1, \mu_2, \mu_3)\,\sigma_1\right)^2 + \left(\frac{\partial g}{\partial X_2}(\mu_1, \mu_2, \mu_3)\,\sigma_2\right)^2 + \left(\frac{\partial g}{\partial X_3}(\mu_1, \mu_2, \mu_3)\,\sigma_3\right)^2}} \quad (6.37)$$

$$\beta = \frac{8 \cdot 3 - 4}{\sqrt{(3 \cdot 2)^2 + (8 \cdot 1)^2 + (1 \cdot 2)^2}} = 1.96$$

The chosen initial values of α_1, α_2 and α_3 are equal absolute values, however, with a different sign (load parameters with a minus sign and resistance parameters with a plus sign). With the values for β, α_1, α_2 and α_3 new values are calculated until these remain stable. Table 6.1 gives the result for 6 iterations.

With the values found for β, α_1, α_2 and α_3 the design point and the probability of failure can be calculated. The eventual design point is:

$$\begin{aligned} X_1^* &= \mu_1 - \alpha_1 \beta \sigma_1 = 8 - 0.20 \cdot 2.39 \cdot 2 = 7.04 \\ X_2^* &= \mu_2 - \alpha_2 \beta \sigma_2 = 3 - 0.94 \cdot 2.39 \cdot 1 = 0.75 \\ X_3^* &= \mu_3 - \alpha_3 \beta \sigma_3 = 4 - -0.27 \cdot 2.39 \cdot 2 = 5.29 \end{aligned} \quad (6.38)$$

And the probability of failure is:

$$P_f = \Phi(-\beta) = \Phi(-2.39) = 0.0084 \quad (6.39)$$

Table 6.1 Iterations Example 6.1: Method 1

	initial value	**iteration step number**					
		1	2	3	4	5	6
β	1.96	2.51	2.49	2.42	2.39	2.39	2.39
α_1	0.58	0.52	0.32	0.23	0.21	0.20	0.20
α_2	0.58	0.80	0.89	0.93	0.94	0.94	0.94
α_3	-0.58	-0.28	-0.33	-0.29	-0.27	-0.27	-0.27

6.2.4 Method 2: direct iteration based on the limit state function

The second method is actually derived from the previously discussed method and follows similar steps. It does have the advantage that the limit state function does not need to be transformed to a function of standard normally distributed variables. In this case, the β-value is calculated with equation (6.3) for the reliability function linearized in a chosen point. This β-value is subsequently used to determine a new point, in which the reliability function is linearized. Finally we iterate towards the design point $\underline{X}^*$. In this case, the α_i-values are calculated with:

$$\alpha_i = \frac{\dfrac{\partial}{\partial X_i} g(\underline{X}^*)\sigma_{X_i}}{\sqrt{\sum_{i=1}^{n}\left(\dfrac{\partial}{\partial X_i} g(\underline{X}^*)\sigma_{X_i}\right)^2}} = \frac{\left\{\dfrac{\partial}{\partial X_i} g(\underline{X}^*)\right\}\sigma_{X_i}}{\sigma_Z} \tag{6.40}$$

The new point, in which β- and α_i-values are calculated anew, is determined by:

$$X_i^* = \mu - \alpha_i \beta \sigma_{X_i} \tag{6.41}$$

The method can best be illustrated by means of an example.

Example 6.2: Method 2, Direct iteration based on the limit state function

To illustrate, the same problem as in Example 6.1 is used. Consider the following limit state equation:

$$g(\underline{X}) = Z = X_1 X_2 - X_3 \tag{6.42}$$

The variables X_1, X_2 and X_3 are independent normally distributed random variables. Also given:

$$\begin{aligned} &\mu_1 = 8, \quad && \sigma_1 = 2 \\ &\mu_2 = 3, \quad && \sigma_2 = 1 \\ &\mu_3 = 4, \quad && \sigma_3 = 2 \end{aligned} \tag{6.43}$$

Once again the aim is to determine the design point with the corresponding reliability index. The partial derivatives in the design points are:

$$\frac{\partial g}{\partial X_1}(X_1^*, X_2^*, X_3^*) = X_2^*, \quad \frac{\partial g}{\partial X_2}(X_1^*, X_2^*, X_3^*) = X_1^*, \quad \frac{\partial g}{\partial X_3}(X_1^*, X_2^*, X_3^*) = -1 \tag{6.44}$$

This leads to the calculation of the mean and standard deviation of Z (see equations (2.138) and (2.139)) and to the reliability index β and sensitivity factors α following from equation (6.40).

$$\begin{aligned} &\sigma_Z = \sqrt{(X_2^*\sigma_1)^2 + (X_1^*\sigma_2)^2 + \sigma_3^2} \\ &\mu_Z = (X_1^* \cdot X_2^* - X_3^*) + X_2^*(8 - X_1^*) + X_1^*(3 - X_2^*) - (4 - X_3^*) \\ &\beta = \frac{\mu_Z}{\sigma_Z}, \qquad \alpha_1 = \frac{X_2^*\sigma_1}{\sigma_Z}, \qquad \alpha_2 = \frac{X_1^*\sigma_2}{\sigma_Z}, \qquad \alpha_3 = \frac{-1\cdot\sigma_3}{\sigma_Z} \end{aligned} \tag{6.45}$$

With the preceding formulas a new estimate of the design point can be calculated for a reliability function linearized in a point. Table 6.2 displays the results of six iterations.

As a starting point of the iteration we choose μ_Z and σ_Z calculated on the basis of the assumption

that the mean values are the design values. In the following step in which β- and α_i-values are calculated anew, the design values are:

$$X_i^* = \mu_i - \alpha_i \beta \sigma_{X_i} \tag{6.46}$$

A comparison of the results in Table 6.2 and the results in Table 6.1 reveals that both methods approximately converge to the design point equally quickly. However, the number of calculations per iteration is greater for the second method. On the other hand, it is not necessary to transform the reliability function. On grounds of the latter argument, the second method is easier to apply in a computer program.

Table 6.2 Iterations Example 6.2: Method 2

	Initial value	**Iteration step number**					
		1	2	3	4	5	6
σ_z		10.20	6.70	6.46	7.12	7.35	7.43
μ_z		20.00	16.45	15.54	17.02	17.56	17.75
β		1.96	2.45	2.41	2.39	2.39	2.39
α_1		0.59	0.44	0.28	0.23	0.21	0.20
α_2		0.78	0.85	0.91	0.93	0.94	0.94
α_3		-0.20	-0.30	-0.31	-0.28	-0.27	-0.27
X_1^*	8	5.69	5.86	6.63	6.90	7.00	7.03
X_2^*	3	1.46	0.92	0.82	0.77	0.76	0.75
X_3^*	4	4.77	5.46	5.49	5.34	5.30	5.29

The procedure above can be operationalized using a systematic approach in a table.

1. The first step is to choose a first estimate for the design point to linearize the limit state function (e.g. the mean values).
2. Then the following table is made in the case of independent variables:

Table 6.3 Fill in Table for Level II calculation

X_i	μ	σ	X_i*	$\frac{\partial Z}{\partial X_i}\left(\vec{X}^*\right)$	$\left\{\frac{\partial Z}{\partial X_i}\left(\vec{X}^*\right)\times\sigma_{X_i}\right\}^2$	$\frac{\partial Z}{\partial X_i}\left(\vec{X}^*\right)\times\left(\mu_{X_i}-X_i^*\right)$	α_i	α_i^2
..								
..								
				Σ				1

We then determine the mean value and standard deviation of $Z = g\left(\vec{X}\right)$:

$$\mu(Z) = g\left(X_1^*, X_2^* .... X_n^*\right) + \left(\mu_{X_1} - X_1^*\right)\frac{\partial g}{\partial X_1}\left(X_1^*, X_2^* .... X_n^*\right) + ..$$
$$.. + \left(\mu_{X_2} - X_2^*\right)\frac{\partial g}{\partial X_2}\left(X_1^*, X_2^* .... X_n^*\right) + \left(\mu_{X_n} - X_n^*\right)\frac{\partial g}{\partial X_n}\left(X_1^*, X_2^* .... X_n^*\right) \tag{6.47}$$

$$\sigma^2(Z)=\left(\frac{\partial g}{\partial X_1}\left(X_1^*,X_2^*....X_n^*\right)\right)^2\sigma_{x_1}{}^2+\left(\frac{\partial g}{\partial X_2}\left(X_1^*,X_2^*....X_n^*\right)\right)^2\sigma_{x_2}{}^2+..$$
$$..+\left(\frac{\partial g}{\partial X_n}\left(X_1^*,X_2^*....X_n^*\right)\right)^2\sigma_{x_n}{}^2 \qquad (6.48)$$

3. We then calculate $\beta=\mu(Z)/\sigma(Z)$ and the α-values according to equation (6.40).
4. We determine the new design point according to equation (6.46).

Steps 2 till 5 are repeated until the values for β, α and the design point have reached convergence.

Example 6.3 gives another example of the first method: transformation to normal variables.

Example 6.3: Beam in reinforced concrete

Consider a rectangular bottom reinforced concrete cross-section with characteristics:

- *B (width)* = 200 [mm]
- A_s *(rebar area)* = 800 [mm^2]
- f_c *(concrete compressive strength)* = 30 [N/mm^2]
- f_y *(yield stress)* = N(420, 40) [N/mm^2]
- *d (effective depth)* = N(400, 10) [mm]
- *M (external moment)* = N(40, 18) [kNm]; V = 45 %

The first three variables are considered as deterministic variables in this example, the last three as random variables. For the modelling of the resisting moment of the rectangular cross-section we make use of a simplified calculation method in which a rectangular distribution of the concrete compressive stresses is used over a height of *h=0.8x*. The formula for the resistance (see e.g. lecture notes 'Reinforced Concrete') is given by equation (6.50). We calculate:

$$\omega=\frac{A_s f_y}{Bdf_c}=\frac{800\cdot 420}{200\cdot 400\cdot 30}=0.140 \quad \text{(based on the mean values)} \qquad (6.49)$$

The resistance of the beam becomes:

$$R=A_s\ f_y\ d\left(1-\frac{0,4}{0.8\,.\,0.85}\omega\right)=0.917\ A_s\ f_y\ d \qquad (6.50)$$

This gives the following limit state function:

$$g(\underline{X})=0.917\cdot 800\cdot f_y\cdot d-M=0.734\cdot 10^{-3}\ f_y\ d-M \qquad [kNm] \qquad (6.51)$$

We write $g(\underline{X})$ in function of the normalized variables:

$$U=(X-\mu)/\sigma \quad or \quad X=U\,\sigma+\mu \qquad (6.52)$$

$$X_1=M,\quad X_2=f_y,\quad X_3=d \qquad (6.53)$$

$$g(\underline{U}) = -18U_1 - 40 + 0.734 \cdot 10^{-3}\,(40U_2 + 420)\,(10U_3 + 400)$$
$$g(\underline{U}) = -18U_1 + 11.74 \cdot U_2 + 3.08 \cdot U_3 + 0.29 \cdot U_2 U_3 + 83.31 \qquad (6.54)$$

The following set of equations has to be solved iteratively:

$$\beta = \frac{83.31}{-18\alpha_1 + 11.74\alpha_2 + 3.08\alpha_3 - 0.29\alpha_2\alpha_3\beta} \qquad (6.55)$$

$$\alpha_1 = -\frac{1}{k} \cdot 18 \qquad (6.56)$$

$$\alpha_2 = +\frac{1}{k}(11.74 - 0.29\,\beta\alpha_3) \qquad (6.57)$$

$$\alpha_3 = +\frac{1}{k}(3.08 - 0.29\,\beta\alpha_2) \qquad (6.58)$$

In which k follows from the condition $\sum \alpha_i^2 = 1$. Considering the starting values β = 3, $\alpha_1 = -1/\sqrt{3}$, $\alpha_2 = \alpha_3 = 1/\sqrt{3}$ we obtain the following values in the subsequent iteration steps.

Table 6.4 Iterations Example 6.3

Step	0	1	2	3
β	3	4.465	3.853	3.851
α_1	-0.577	-0.842	-0.836	-0.835
α_2	0.577	0.526	0.538	0.539
α_3	0.577	0.120	0.111	0.115

From the found values it follows that: $\Phi(-\beta)$ = 5.88 $\cdot 10^{-5}$. The coordinates of the design point are $X_i^* = \mu_i - \beta\,\alpha_i\,\sigma_i$:

$$M^* = 40 + 3.851 \cdot 0.835 \cdot 18 = 97.88 \text{ kNm}$$
$$f_y^* = 420 - 3.851 \cdot 0.539 \cdot 40 = 337.0 \text{ N/mm}^2 \qquad (6.59)$$
$$d^* = 400 - 3.851 \cdot 0.115 \cdot 10 = 395.6 \text{ mm}$$

6.3 Non Normally Distributed Variables

If independent non-normally distributed random base variables are involved, it seems plausible that the reliability function is not normally distributed either. To be able to apply level II approximation methods in these cases, the non-normally distributed base variables have to be transformed to normally distributed base variables.

In case the **variables are not normally distributed** one often determines a design value on the basis of the following formula (See also Annex C of EN 1990):

$$F_{X_i}(x_i^*) = \Phi(-\alpha_i\,\beta) \qquad (6.60)$$

Or:

$$x_i^* = F_{X_i}^{-1}\left[\Phi(-\alpha_i\,\beta)\right] \tag{6.61}$$

In which it is expressed that x_i^* should correspond with the same percentile in $F_{X_i}(x)$ as $-\alpha_i\beta$ in $\Phi(u)$. We elaborate this further for variables that follow a lognormal or a Gumbel distribution.

a) Lognormal distribution

It follows that:

$$\frac{\ln(x_i^* / \breve{\mu}_{X_i})}{\sigma_{\ln X_i}} = -\alpha_i\,\beta \tag{6.62}$$

$$\ln x_i^* = \ln \breve{\mu}_{X_i} - \alpha_i\,\beta\,\sigma_{\ln X_i} = \ln\left[\mu_{X_i}\,\exp\left(-\frac{1}{2}\ln(1+V_{X_i}^2)\right)\right] - \alpha_i\,\beta\sqrt{\ln(1+V_{X_i}^2)} \tag{6.63}$$

$$x_i^* = \frac{\mu_{X_i}}{\sqrt{1+V_{X_i}^2}}\exp\left(-\alpha_i\,\beta\sqrt{\ln(1+V_{X_i}^2)}\right) \tag{6.64}$$

In case $V < 0.2$ the following holds with good approximation:

$$x_i^* = \mu_{X_i}\,\exp\left(-\alpha_i\,\beta\,V_{X_i}\right) \tag{6.65}$$

b) Gumbel distribution (EXI_L)

It follows that:

$$\exp\left[-e^{-\alpha(x_i^*-u)}\right] = \Phi(-\alpha_i\,\beta) \tag{6.66}$$

$$\ln\left[-\ln\Phi(-\alpha_i\,\beta)\right] = -\alpha\,(x_i^* - u) \tag{6.67}$$

$$x_i^* = u - \frac{1}{\alpha}\ln\left[-\ln\Phi(-\alpha_i\,\beta)\right] \tag{6.68}$$

In many cases, particularly in case of skewed distributions, the tails of the distribution will be considerably different from the normal distribution. When only considering the first and second order moments (i.e. the mean and variance), one does not obtain a good approximation of the problem. A further refinement of the calculation method can be obtained by applying the Rackwitz-Fiessler algorithm.

Rackwitz and Fiessler (1977) introduced a transformation, to transform a variable with an arbitrary distribution to a normally distributed variable. This transformation assumes that the values of the real and the approximated probability density function and probability distribution function are equal in the design point. This is shown in Figure 6.6.

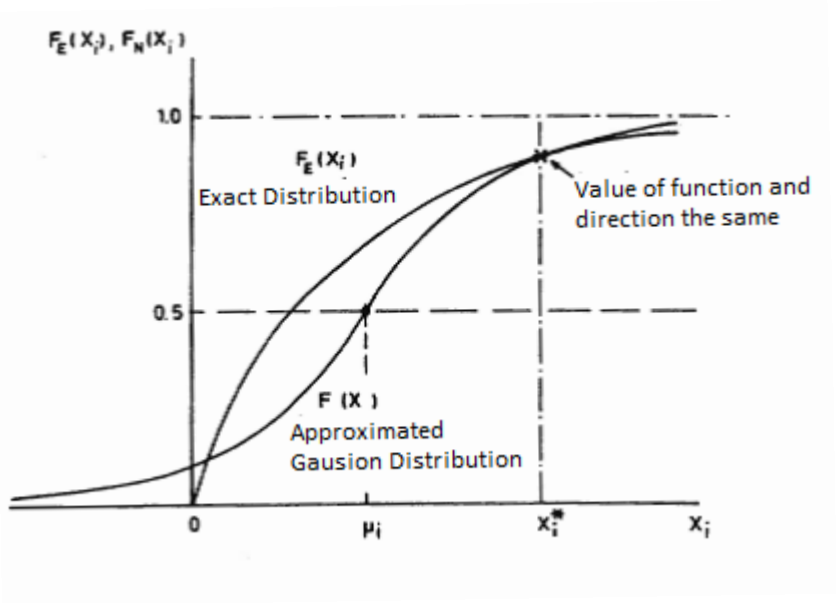


Figure 6.6 Transformation to normal distribution in the design point

In Figure 6.7 it is explained what it means that the values of the real and the approximated probability density function and probability distribution function are equal in the design point.

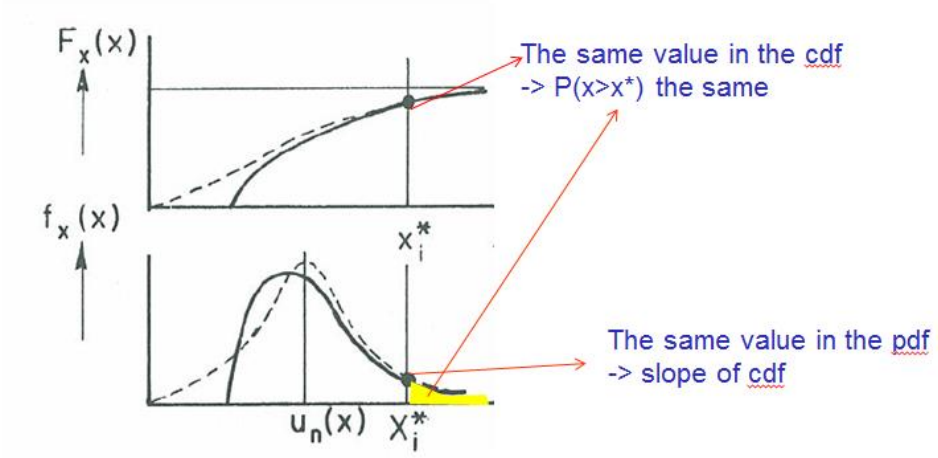


Figure 6.7 Transformation to normal distribution in the design point, the same values in cdf and pdf

More specifically, this means that the parameters μ_i^N and σ_i^N (the superscript indicating a Normal distribution) can be calculated from the equations:

$$F_{X_i}(x_i^*) = \Phi\left(\frac{x_i^* - \mu_i^N}{\sigma_i^N}\right) \tag{6.69}$$

$$f_{X_i}(x_i^*) = \frac{1}{\sigma_i^N}\varphi\left(\frac{x_i^* - \mu_i^N}{\sigma_i^N}\right) \tag{6.70}$$

From equation (6.69) it follows that:

$$\frac{x_i^* - \mu_i^N}{\sigma_i^N} = \Phi^{-1}\left[F_{X_i}(x_i^*)\right] \tag{6.71}$$

Rewriting this expression and substituting it into equation (6.70) yields:

$$\mu_i^N = x_i^* - \sigma_i^N \; \Phi^{-1}\left[F_{X_i}(x_i^*)\right] \tag{6.72}$$

$$\sigma_i^N = \frac{\varphi[\Phi^{-1}[\mathrm{F}_{\mathrm{X_i}}(\mathrm{x}_\mathrm{i}^*)]]}{\mathrm{f}_{\mathrm{X_i}}(x_i^*)} \tag{6.73}$$

φ() is the standard normal probability density function. When further also considering equation (6.60), i.e.:

$$F_{X_i}(x_i^*) = \left[\Phi(-\alpha_i \; \beta)\right] \tag{6.74}$$

equations (6.72) and (6.73) can be further simplified towards:

$$\mu_i^N = x_i^* + \alpha_i \; \beta \; \sigma_i^N \tag{6.75}$$

$$\sigma_i^N = \frac{\varphi(\alpha_i \; \beta)}{f_{X_i}(x_i^*)} \tag{6.76}$$

From the equations above it appears that the standard deviation and the average of the approximating normal distribution depend on the value of X in the design point. Therefore, in the iterative calculation of the design point and of the reliability index, new values for $\sigma_x^{'}$ and $\mu_x^{'}$ must be calculated for every step. This is illustrated in Example 6.4.

Example 6.4

Once again the same problem is considered as in Example 6.1. Consider the following limit state equation:

$$g(\underline{X}) = Z = X_1 X_2 - X_3 \tag{6.77}$$

All variables are independent. The variables X_1, X_2 are normally distributed random variables:

$$\begin{aligned} &\mu_1 = 8, \qquad \sigma_1 = 2 \\ &\mu_2 = 3, \qquad \sigma_2 = 1 \end{aligned} \tag{6.78}$$

However, this time it is assumed that the base variable X_3 is uniformly distributed over the interval (-20, 28). The mean and the standard deviation of X_3 are thus the same as in Example 6.1, i.e. $\mu_3 = 4$ and $\sigma_3 = 2$. The probability density function and the probability distribution of X_3 read:

$$\begin{aligned} &f_{X_3}(x_3) = \frac{1}{48} \\ &-20 \le x_3 \le 28 \\ &F_{X_3}(x_3) = \frac{x_3 + 20}{48} \end{aligned} \tag{6.79}$$

In this case the transformed reliability function is:

$$Z = 6 \cdot U_1 + 2 \cdot U_1 U_2 + 8 \cdot U_2 + 24 - \mu_{X_3}{}' - \sigma_{X_3}{}' U_3 \tag{6.80}$$

Substitution of U_i =- $\alpha_i \beta$ in the design point gives:

$$Z = -6\alpha_1\beta + 2\alpha_1\alpha_2\beta^2 - 8\alpha_2\beta + 24 - \mu_{X_3}{}' + \sigma_{X_3}{}'\alpha_3\beta \tag{6.81}$$

Subsequently, the iteration formulas can be drawn up as was done in Example 6.1. The system of equations to be solved is now:

$$\begin{aligned}
\beta &= \frac{-24 + \mu_{X3}^{'}}{-6\alpha_1 + 2\alpha_1\alpha_2\beta - 8\alpha_2 + \sigma_{X3}^{'}\alpha_3} \\
\alpha_1 &= \frac{6 - 2\alpha_2\beta}{\sqrt{(6-2\alpha_2\beta)^2 + (8-2\alpha_1\beta)^2 + \sigma_{X3}^{'\,2}}} \\
\alpha_2 &= \frac{8 - 2\alpha_1\beta}{\sqrt{(6-2\alpha_2\beta)^2 + (8-2\alpha_1\beta)^2 + \sigma_{X3}^{'\,2}}} \\
\alpha_3 &= \frac{-\sigma_{X3}^{'}}{\sqrt{(6-2\alpha_2\beta)^2 + (8-+2\alpha_1\beta)^2 + \sigma_{X3}^{'\,2}}} \\
X_3^{*} &= F_{X3}^{-1}(\Phi(-\alpha_3\beta)) = 48 \cdot \Phi(-\alpha_3\beta) - 20 \\
\sigma_{X3}^{'} &= \frac{\varphi(-\alpha_3\beta)}{f_c(X_3^{*})} = \varphi\left(-\alpha_3\beta\right) \cdot 48 \\
\mu_{X3}^{'} &= X_3^{*} + \alpha_3\left(\beta\,\sigma_{X3}^{'}\right)
\end{aligned} \tag{6.82}$$

Table 6.5 gives the results of a number of iterations. The calculation shows that the reliability index is considerably lower than in Example 6.1.

To illustrate the transformation of the uniform to the normal distribution, the real and the approximating probability distribution in the design point X_3* are indicated in Figure 6.8.

Table 6.5 Iterations Example 6.4

	Initial value	**iteration step number**		
		1	2	3
β	1.96	1.07	1.03	1.03
α_1	0.58	0.31	0.32	0.31
α_2	0.58	0.47	0.47	0.46
α_3	-0.58	-0.83	-0.82	-0.83
X_3^{*}	21.87	19.00	18.50	18.60
σ	10.04	12.92	13.36	13.27
μ	10.46	7.54	7.16	7.23

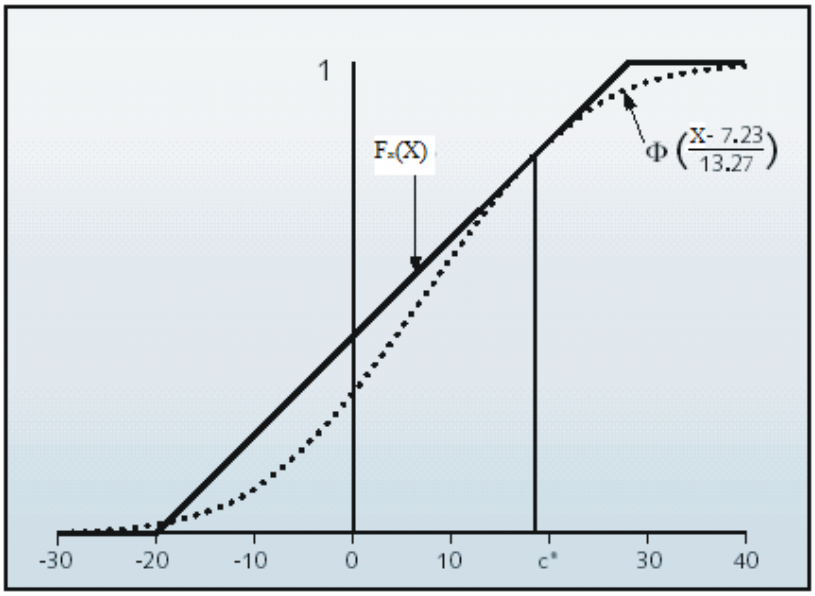


Figure 6.8 Approximation of the real probability distribution by a normal distribution.

6.4 Dependent random base variables

(informative, will not be part of the exam)

The probability density function with multiple base variables was introduced in section 2.9. In section 4.2 it was explained how the probability of failure is defined:

$$P_f = \int_{g(\underline{X})<0} f_{\underline{X}}(\underline{x}) d\underline{x} \tag{6.83}$$

In the case of independent base variables, the expression can be simplified to

$$P_f = \int_{g(X_1, X_2, \ldots, X_N)<0} \prod_{i=1}^{N} f_{X_i}(x_i) dx_1 dx_2 \ldots dx_N \tag{6.84}$$

This allows to approximate for each variable individually the real distribution function by a normal distribution in the design point. If the case of dependent base variables, this approximation can't be done independently for each variable. Therefore, a transformation can be performed to a new set of variables in which independence is 'forced'. In this section, it well be demonstrated how continuous dependent variables $X_1, X_2, \ldots, X_N$ can be transformed to independent normal distributed variables, the Rosenblatt transformations, from where the design point with its associated failure probability can be computed iteratively.

Recall from section 2.4.1 that the total probability density function can be determined as the product of the marginal conditional probability density functions of the respective base variables:

$$f_{(X_1, X_2, \ldots, X_N)}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2|X_1=x_1}(x_2 \mid x_1) \cdot \ldots \cdot f_{X_N|\{X_j\}_{j=1}^{N-1}}(x_N \mid \{x_j\}_{j=1}^{N-1}) \tag{6.85}$$

The marginal probability density function of any group of base variables $\{X_j\}_{j=1}^{i}$ can be found as:

$$f_{X_1 \cdots X_i}(x_1, x_2, \cdots, x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}}(\underline{x}) \, dX_{i+1} \, dX_{i+2} \cdots dX_n \tag{6.86}$$

When the distributions of the individual base variables is known, the conditional cumulative probability distributions can be determined as:

$$F_{X_1}(x_1) = \int_{-\infty}^{X_1} f_{X_1}(\zeta) d\zeta \tag{6.87}$$

$$F_{X_2|X_1=x_1}(x_2 \mid x_1) = \frac{1}{N_2} \int_{-\infty}^{X_2} f_{X_1, X_2}(x_1, \zeta) d\zeta \tag{6.88}$$

$$F_{X_i|\{X_j=x_j\}_{j=1}^{i-1}}(x_i \mid \{x_j\}_{j=1}^{i-1}) = \frac{1}{N_i} \int_{-\infty}^{X_i} f_{X_1, X_2, \cdots, X_i}(x_1, x_2, \cdots x_{i-1}, \zeta) d\zeta \tag{6.89}$$

In which N_i is function of $\{x_j\}_{j=1}^{i-1}$ which normalizes the distribution function such that $\lim_{x_i \to \infty} F_i = 1$.

From this the conditional probability density function follows:

$$f_i(X_i \mid \mathrm{X}_1,...,\mathrm{X}_{i-1}) = \frac{\partial}{\partial X_i} F_1(X_i \mid X_1,..., X_{i-1}) = \frac{1}{N_i} f_i(X_1,..., X_{i-1}, X_i) \tag{6.90}$$

Such that the failure region is defined by:

$$P_f = \int_{g(X_1,X_2,...,X_N)<0} \prod_{i=1}^{N} f_{X_1|\{X_j=x_j\}_{j=1}^{i-1}}(x_i \mid \{x_j\}_{j=1}^{i-1}) dx_1 dx_2 ... dx_N \tag{6.91}$$

The Rosenblatt-transformation transforms the variables X_i with the value X_i into standard normally distributed variables with value Z_i. In the design point, the values of the probability distribution of the normally distributed variables are equal to the conditional probability distribution of the original variables, the same approximation procedure as explained in Section 6.3. To that end:

$$\begin{aligned} Z_1 &= \Phi^{-1}(F_1(\mathrm{X}_1))) \\ Z_2 &= \Phi^{-1}(F_2(\mathrm{X}_2 \mid \mathrm{X}_1)) \\ &\cdot \\ &\cdot \\ &\cdot \\ Z_n &= \Phi^{-1}(F_n(\mathrm{X}_n \mid \mathrm{X}_1,..., \mathrm{X}_{n-1})) \end{aligned} \tag{6.92}$$

To transform the reliability function the inverse of the preceding transformation is of importance. This is given by:

$$\begin{aligned} X_1 &= F_1^{-1}(\Phi(Z_1)) \\ X_2 &= F_2^{-1}(\Phi(Z_2) \mid X_1) \\ &\cdot \\ &\cdot \\ &\cdot \\ X_n &= F_n^{-1}(\Phi(Z_n) \mid X_1,..., X_{n-1}) \end{aligned} \tag{6.93}$$

A comparison of the foregoing transformation with the equations (6.93) shows that the Rosenblatt-transformation is a generally applicable transformation for non-normally distributed and/or independent random base variables.

The standard deviation and the mean of the approximating normal distribution are, as given in (6.75) and (6.76):

$$\begin{aligned} \sigma_{X_i}' &= \frac{\varphi(Z_i)}{f_X(x_i^*)} \\ \mu_{X_i}' &= x_i^* + Z_i \sigma_{X_i}' \end{aligned} \tag{6.94}$$

The way a Rosenblatt-transformation takes place, can best be illustrated with a calculation example.

Example 6.5

Once again the same reliability function of Example 6.1 is considered. The limit state function is

$$g(X_1, X_2, X_3) = X_1 \cdot X_2 - X_3 \tag{6.95}$$

The base variables X_1 and X_2 are statistically dependent and their joint probability density function is:

$$f_{X_1, X_2}(x_1, x_2) = \frac{2}{175}\left(\frac{x_1}{14} + \frac{x_2}{5} + \frac{x_1 x_2}{14 \cdot 5}\right) \tag{6.96}$$

with $0 \le x_1 \le 14$ and $0 \le b \le 5$

The base variable X_3 is normally distributed with mean $\mu_3 = 4$ and $\sigma_3 = 2$.

We search for the probability of failure of (6.95). In short, the steps to be carried out:

a) Because the variables X_1 and X_2 are dependent, we need to find the joint distribution functions in the variables $X_2 \mid X_1 = x_1$ and X_1, such that the design point can be calculated sequentially, see section 6.4.
b) In the design point, the joint distribution in terms of X_1 and $X_2 \mid X_1 = x_1^*$ is approximated by two normal distributions with parameters μ_1^N, σ_1^N and μ_2^N, σ_2^N, see section 6.2.3 .
c) The design point in standard normal space is given by (z_1^*, z_2^*, z_3^*). The transformation function to express this design point in the real variables (x_1^*, x_2^*, x_3^*) has to found.
d) The system of equations obtained has to be solved, which can be done in an iterative manner with starting assumptions for $(\alpha_1, \alpha_2, \alpha_3)$ and β.

a)

De distribution function for X_1:

$$f_{X_1}(x_1) = \int_0^5 f_{X_1, X_2}(x_1, x_2) dx_2 = \frac{3x_1}{490} + \frac{1}{35} \tag{6.97}$$

The cumulative distribution function transfers variables in physical space X_1 to the probability space $U_1 = [0\,1]$. This is denoted as $F : X_1 \to U_1$. It can be found from the density function:

$$F_{X_1}(x_1) = \int_0^{x_1} f_{X_1}(\zeta) d\zeta = \frac{x_1}{980} \cdot (3x_1 + 28) \tag{6.98}$$

With its inverse $F^{-1} : U_1 \to X_1$, that transfers points in probability space U_1 to physical space X_1:

$$F_{U_1}^{-1}(u_1) = \frac{490}{3} \cdot \sqrt{\frac{3x_1}{245} + \frac{1}{1225}} - \frac{14}{3} \tag{6.99}$$

The distribution function for $X_2 \mid X_1 = x_1$:

$$F_{(X_2|X_1=x_1)}(x_1,x_2)=\frac{1}{N(x_1)}\int_0^{x_2} f_{X_1,X_2}(x_1,\zeta)d\zeta=\frac{1}{N(x_1)}\left(\frac{x_1}{12250}+\frac{1}{875}\right)\cdot x_2^2+\frac{x_1\cdot x_2}{1225}\right) \tag{6.100}$$

With $N(x_1)$ normalizing the distribution function:

$$N(x_1)=\int_0^5 f_{X_1,X_2}(x_1,\zeta)d\zeta=\frac{3x_1}{490}+\frac{1}{35} \tag{6.101}$$

$F_{(X_2|X_1=x_1)}(x_1,x_2)$ is a function of one variable (X_2) and one parameter (x_1^*). To find its inverse, the parameter (x_1^*) should be known, to make sure the inverse is taken over the right domain.

To clarify this requirement, imagine the function $y=x^2$. The inverse function over the domain $x\in[0\,10]$ would be given by $x=\sqrt{y}$, over the domain $x\in[-10\,0]$ would be given by $x=-\sqrt{y}$, but over the domain $x\in[-5\,5]$ there are two possible solutions x and $-x$, so the inverse is not properly defined.

Therefore, the inverse of (6.100) has to be found every time again after updating x_1^*. We can than substitute to reduce the amount of variables:

$$F_{X_2|X_1=x_1}(x_2)=\frac{1}{N(x_1^*)}\int_0^{x_2} f_{X_1,X_2}(x_1^*,\zeta)d\zeta \tag{6.102}$$

The inverse function $F_{U_2|X_1}=x_1^{*-1)}:U_2\rightarrow X_2\mid X_1=x_1^*$ can then be found as:

$$F^{-1}_{U_2|X_1=x_1^*}(u_2)=(F_{X_2|X_1=x_1^*}(x_2))^{-1} \tag{6.103}$$

b)

We assume a normal approximation in the design point. The mean and standard deviation of the approximate normal distributions can be found with equations (6.124) and (6.125):

$$\sigma_1^N=\frac{\phi(z_1^*)}{(f_{X_1}(x_1^*)} \tag{6.104}$$

$$\sigma_2^N=\frac{\phi(z_1^*)}{(f_{X_2|X_1=x_1^*}(x_2^*)} \tag{6.105}$$

Where we also need the distribution function of $F_{X_2|X_1=x_1}$:

$$f_{X_2|X_1=x_1^*}(x_2)=\frac{d}{dx_2}(F_{X_2|X_1=x_1}(x_2)) \tag{6.106}$$

Subsequently the means:

$$\mu_1^N = x_1^* - z_1^* \cdot \sigma_1^N \tag{6.107}$$

$$\mu_2^N = x_2^* - z_2^* \cdot \sigma_2^N \tag{6.108}$$

c)

The design point in standard normal variables is given by:

$$\begin{aligned} z_1^* &= -\alpha_1 \beta \\ z_2^* &= -\alpha_2 \beta \\ z_3^* &= -\alpha_3 \beta \end{aligned} \tag{6.109}$$

The transformation of the normal variables to the physical parameters is:

$$x_1^* = F_{Z_1}^{-1}(F(z_1^*)) \tag{6.110}$$

$$x_2^* = F_{Z_2|X_1=x_1^*}^{-1}(F(z_2^*)) \tag{6.111}$$

$$x_3^* = \mu_3 + z_3^* \cdot \sigma_3 \tag{6.112}$$

The transformation for the independent variable x_1^* is done with (6.99). For x_2^*, equation (6.103) has to be updated in every iteration step. For x_3^* a transformation from standard normal space Z to X_3 is necessary, which is simpler and given directly by (6.112).

d)

This gives all the ingredients to make an update of the α_i:

$$\alpha_1 = \frac{\sigma_1^N \cdot (\mu_2^N + \sigma_2^N \cdot z_2^*)}{X} \tag{6.113}$$

$$\alpha_2 = \frac{\sigma_2^N \cdot (\mu_1^N + \sigma_1^N \cdot z_1^*)}{X} \tag{6.114}$$

$$\alpha_3 = -\frac{\sigma_3}{X} \tag{6.115}$$

With

$$X = \sqrt{(\sigma_1^N \cdot (\mu_2^N + z_2^* \sigma_2^N))^2 + (\sigma_1^N \cdot (\mu_2^N + z_2^* \sigma_2^N))^2 + \sigma_3^2} \tag{6.116}$$

The reliability index, as defined in (6.3), can be found from the reliability function expressed in the approximate normally distributed variables:

$$\beta = \frac{\mu_{g_N}}{\sigma_{g_N}} \tag{6.117}$$

Because the normal variables approximate the variables $X_2 \mid X_1 = x_1$ and X_1 which behave independently, the expectation of the limit state function g_N and its variance are given by:

$$\mu_{g_N} = \mu_1^N \cdot \mu_2^N - \mu_3 \tag{6.118}$$

$$\sigma_{g_N} = \sqrt{(\sigma_1^N)^2 \cdot (\sigma_2^N)^2 + (\sigma_1^N)^2 \cdot (\mu_2^N)^2 + (\sigma_2^N)^2 \cdot (\mu_1^N)^2 + \sigma_3^2} \tag{6.119}$$

The system of equations is a closed system for $\alpha_1, \alpha_2, \alpha_3, \beta, \mu_1^N, \mu_2^N, \sigma_1^N, \sigma_2^N, x_1^*, x_2^*$ and x_3^*.
To solve the system of equations, we could iterate towards a convergent answer. Starting values need to be chosen, for example: starting values $\alpha_1(1) = 0.6; \alpha_2(1) = 0.6; \alpha_3(1) = -0.6; \beta(1) = 2$.

We can iterate through the system of equations until the solution for β converges to a constant value. This is best done with the help of a computer, for example in a while-loop in Matlab, since in every iteration step the inverse of $F^{-1}_{U_2|X_1=x_1^*}(u_2)$ and the density function $f_{X_2|X_1=x_1^*}(x_2)$ need to be found for the updated x_1^*. A stopping criterion for the while-loop could for example be $\beta(i) - \beta(i-1) < 0.0001$, where i denotes the iteration step.

In short: from starting values α_i and β:

1. Calculate the design point in standard normal variables with (6.109).
2. Transform these variables to the physical design point with (6.110), (6.111) and (6.112).
3. Find the parameters of the approximative normal distribution with (6.104), (6.105), (6.107) and (6.108)
4. Update the influence factors and the reliability index with (6.113), (6.114), (6.115) and (6.117)

The iteration steps are summarized in the table below:

	initial value	**iteration step number**					
		1	2	3	4	5	6
β	1.5	1.0337	1.0079	1.0058	1.0058	1.0058	1.0059
α_1	0.7181	0.7258	0.7277	0.7286	0.7291	0.7294	0.7295
α_2	0.6743	0.6725	0.6709	0.67	0.6695	0.6692	0.669
α_3	-0.172	-0.144	-0.142	-0.142	-0.142	-0.142	-0.142
x_1^*	4.5283	5.1603	5.2143	5.2124	5.2082	5.2056	5.204
x_2^*	1.7183	1.9935	2.0258	2.0313	2.0332	2.0342	2.0348
x_3^*	5.74	4.356	4.2916	4.2864	4.286	4.286	4.286
μ_{x1}'	5.74	4.356	4.2916	4.2864	4.286	4.286	4.286
σ_{x1}'	8.751	8.897	8.906	8.9057	8.905	8.9046	8.9043
μ_{x2}'	3.2229	3.2494	3.251	3.2516	3.2519	3.2521	3.2522
σ_{x2}'	4.8538	5.034	5.0462	5.0458	5.0449	5.0443	5.0439

Therefore, the failure probability of (6.95) is $\Phi(-\beta) = 0.157$

References

- Cornell C.A., A probability-based structural code, ACI Journal, Vol. 66, 1969, pp. 974-985.
- Hasofer A., Lind N., An exact and invariant first order reliability format, Proc. ASCE, J. Eng. Mech. Div., 1974, pp. 111-121.
- Rackwitz R., Fiessler B., An algorithm for calculation of structural reliability under combined loading, Berichte zur Sicherheitstheorie der Bauwerke, Lab. für den Konstruktiven Ingenieurbau, München, 1977.
- DIN, General principles on the specification of safety requirements for structures, Beuth, 1981.
- Guidelines for loading and safety regulations for structural design, NKB Report No. 55E, June 1987.

Chapter "Time Dependence"
Author: R.D.J.M. Steenbergen

This chapter has been compiled using materials from the CUR book (1997; also version 2015).

Contents

7 Time dependence

7.1 Probability of failure and life span

In general, both the loads and the strength are functions of time. It is therefore typically of little use to speak of a probability of failure without mentioning the period, to which this refers. In case only the load varies in time, it is wise to define the representative value of the load for the considered period. If the representative value of the load equals the maximum load during the considered period, the probability distribution can be determined by using the theory for extreme values. If fatigue problems are concerned, the representative value of the load must be established by addition of the loads over the period. The calculation of the probability of failure can be carried out with level III or level II methods, after defining the representative value of the load.

If both the resistance strength and the load are time dependent, the characteristic strength and load have to be defined carefully. After all, it is possible that the maximum value of the load does not coincide with the minimum value of the strength (see Figure 7.1). In such a case the instantaneous distributions of the strength and the loads have to be assumed.

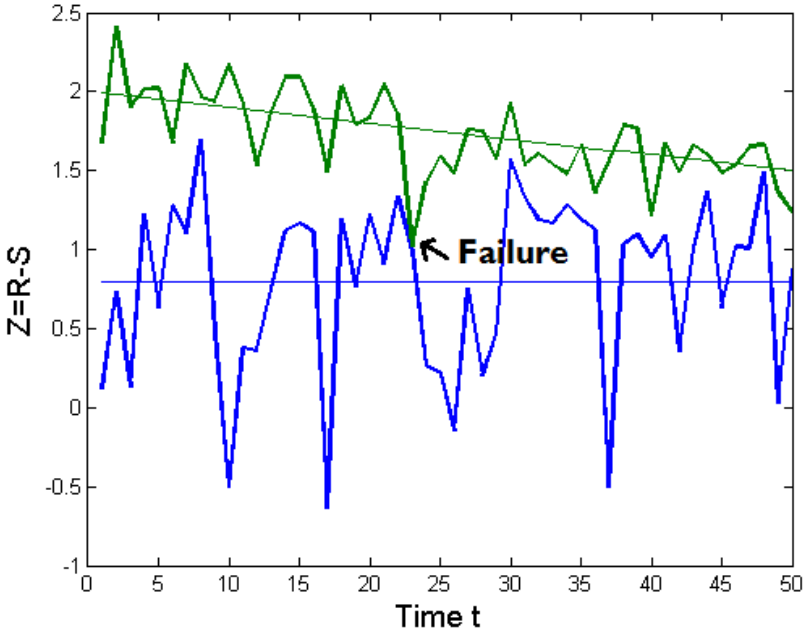


Figure 7.1 Green: Strength (R) normally distributed with mean decreasing over time. Blue: load (S) having an extreme value distribution in blue, plotted along its mean. At the point where $Z = R - S < 0$, failure occurs.

The probability of failure in the time span (0, t) is equal to the complement of the probability that no failure occurs in the interval. In the form of a formula this is:

$$P_f(t) = 1 - P\left((R(\tau) > S(\tau)) \text{ for all } \tau \in (0,t)\right) \tag{7.1}$$

in which:

- $R(\tau)$ is the strength at time τ;
- $S(\tau)$ is the load at time τ.

The strength $R(\tau)$ can be a function of the load over the period (0, τ):

$$R(\tau) = R(R_0, S(t) \text{ with } 0 < t < \tau) \tag{7.2}$$

For example, this is the case with material fatigue. Failure can be seen as the termination of the life span of the element under consideration. The life span is a random variable. The probability distribution of the life span is:

$$F_L(t) = P(L < t) = P_f(t) \tag{7.3}$$

The probability density function of the life span is found by differentiating the probability distribution:

$$f_L(t) = \frac{d\,F_L(t)}{d\,t} \tag{7.4}$$

The probability that the life span ends between the times t and $t + dt$, can be written as:

$$P(t < L < t + dt) = F_L(t + dt) - F_L(t) = f_L(t)\,dt \tag{7.5}$$

Implicitly, the ending of the life span (*t, t + dt*) means that no failure occurred in the interval (0, t), so:

$$f_L(t)\,dt = P\left((R(\tau) < S(\tau)) \text{ for } \tau \in (t, t+dt) \cap (R(\tau) > S(\tau)) \text{ for } \tau \in (0, t)\right) \tag{7.6}$$

The probability density function is also viewed as the probability of failure per unit of time at a certain point in time, or as the rate of failure. The literature often refers to this as the **unconditional rate of failure**. Another important concept is **the conditional rate of failure** r(t), also known as the "hazard function". This is defined as:

$$r(t)\,dt = P\left((R(\tau) < S(\tau)) \text{ for } \tau \in (t, t+dt) \,|\, (R(\tau) > S(\tau)) \text{ for } \tau \in (0, t)\right) \tag{7.7}$$

According to the calculation rule $\left(P(A|B) = P(A \cap B)P(B)\right)$:

$$\begin{aligned} f_L(t)\,dt &= r(t)\,dt\,(1 - F_L(t)) \\ r(t) &= \frac{f_L(t)}{1 - F_L(t)} \end{aligned} \tag{7.8}$$

For small values of the probability distribution function of the life span, this formulation shows that the values of the unconditional rate of failure and those of the conditional rate of failure are equal. Substitution of:

$$f_L(t) = \frac{d\,F_L(t)}{dt} \tag{7.9}$$

in equation (7.8) gives the following differential equation:

$$\frac{d\,F_L(t)}{dt} = (1 - F_L(t))\,r(t) \tag{7.10}$$

The solution to this differential equation is:

$$F_L(t) = 1 - exp\int_0^t -r(\tau)\,d\tau \tag{7.11}$$

The functions $F_L(t)$, $f_L(t)$ and $r(t)$ are plotted for a number of cases in Figure 7.2. In situation a, there is a constant $r(t)$. This means that the fact that the element did not fail during a period (0, *t*) has no influence on the probability of failure in the interval (*t, t + dt*). This situation arises if the knowledge concerning the strength and the load does not vary in the course of time.

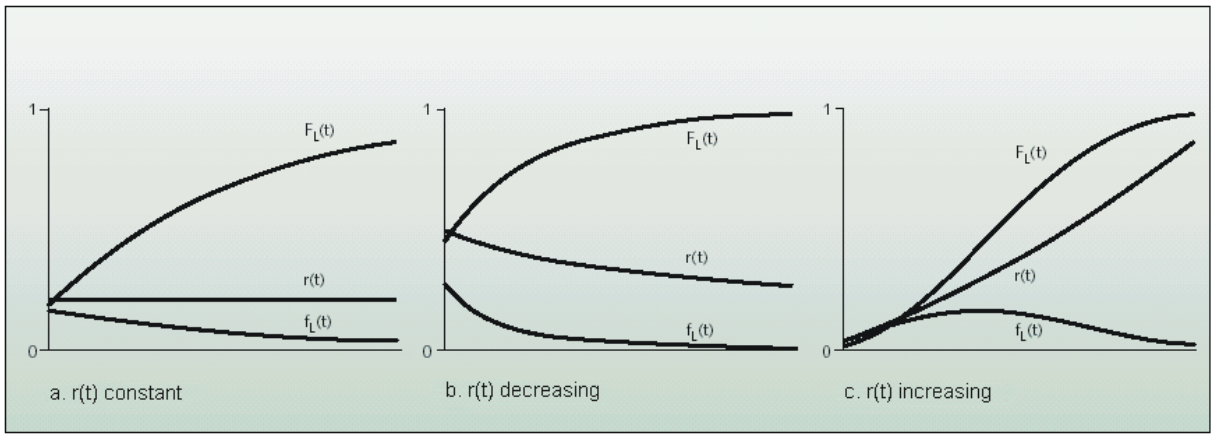


Figure 7.2 Conditional failure rate, unconditional failure rate and probability distribution of the life span.

In situation b the conditional rate of failure decreases. This can be a correct consideration for elements for which the expected value of the strength increases or for which the expected value of the load decrease in time. It could also be the case that the standard deviation of the strength and/or of the load decreases. Situation c involves an increasing conditional rate of failure. Such is the case if deterioration, wear or fatigue is concerned.

In practice, often combinations of the different situations occur. In general this results in a function r(t), as plotted in Figure 7.3, which is referred to as the bathtub curve. Three phases can be distinguished:

I. the begin phase, in which failure can occur as a result of structural and design flaws;
II. the middle phase, in which calamities and extreme circumstances play a part;
III. the end phase with an increased rate of failure as a result of deterioration or wear.

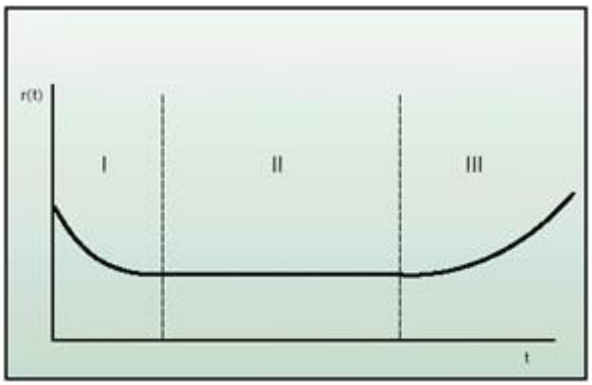


Figure 7.3 Bath tub curve.

The start phase, during which the probability of failure is often greater than in the middle phase, is generally unavoidable. The end phase, with an increased rate of failure as a result of wear and deterioration, however, can be prevented or delayed by carrying out maintenance or by replacement. The question as to when one should resort to maintenance or replacement is discussed further in chapter 11.

In engineering, and particularly in process engineering, the rate of failure in the middle phase is important. Often a constant rate of failure can be assumed. The probability distribution of the life span can then be denoted as:

$$F_L(t) = 1 - e^{-\lambda t} \tag{7.12}$$

in which $\lambda = r(t)$ = constant. In this case, the probability density of the life span is:

$$f_L(t) = \lambda e^{-\lambda t} \tag{7.13}$$

The expected value of the life span is known as the Mean Time to Failure and is determined by integration:

$$\mu_L = E(t) = \int_0^\infty t\, \lambda e^{-\lambda t}\, dt = \frac{1}{\lambda} \tag{7.14}$$

The Mean Time to Failure plays an important part in the determination of non-availability of elements with a constant rate of failure and in planning of maintenance and inspections. If the rate of failure is not a constant in time, this simplification is not possible. For these elements, the determination of the reliability and of the expected value of the life span is described in the following section.

7.2 Application to Different Cases

As the first example we take a structure that is subsequently loaded with independent loads $S_1, S_2, S_3 \cdots$. We assume the resistance R is time-independent. The question that is addressed is as follows: What is the probability that failure occurs if the construction is loaded n-times (see Figure 7.4).

Essential for the failure of the structure is the maximum $\hat{S}$ of the instantaneous loads S_i. If R is larger than $\hat{S}$ the structure will not fail. If R is smaller than $\hat{S}$ the construction fails. So:

$$P\{\text{failure by } S_1, ..., S_n\} = P\{R < \hat{S}\} \tag{7.15}$$

The probability density function of $\hat{S} = \max\{S_i\}$ follows from:

$$F_{\hat{S}}(s) = P\{\hat{S} < s\} = P\{S_1 < s \text{ and } S_2 < s \text{ and } ... S_n < s\} \tag{7.16}$$

Because of the assumed independency:

$$F_{\hat{S}}(s) = P\{S_i < s\}^n = \{F_{S_i}(s)\}^n \tag{7.17}$$

The failure probability is given by the convolution integral, see chapter 3:

$$P(\text{failure by } S_1..S_n) = \int_0^\infty f_R(r)\{1 - F_{\hat{S}}(r)\}dr$$

$$P(\text{failure by } S_1..S_n) = \int_0^\infty f_R(r)\{1 - F_{S_1}^n(r)\}dr \tag{7.18}$$

In the previous equation S_i is the instantaneous load: the leading load is $\hat{S} = \max\{S_i\}$. Determining the distribution of $\hat{S}$ is from a numerical point of view not difficult. In the special case that S_i is an extreme-value distribution for maxima even a analytical solution exists: $\hat{S}$ then also has an extreme-value distribution with parameters that can be determined in a simple way.
The problem in this example has great similarities with a series system characterized by brittle failure. It is indeed possible to see a number of sequential loads as a series system in time. We can therefore use this in our advantage to use the lower and upper boundaries derived for those systems (see chapter 9):

$$\max P\{R < S_i\} \le P\{\text{failure by } S_1..S_n\} \le \sum P\{R < S_i\} \tag{7.19}$$

The probability of failure for the period in time (1..*n*) is bounded by the maximum instantaneous probability of failure and the sum of all the instantaneous probabilities of failure. In Figure 7.4 this is shown. While refering to chapter 9 the following can be said:

- The upper boundary occurs when the events $R<S_i$ are mutually exclusive. Physically speaking, this occurs seldom but also in this case it holds that the upper boundary is of great importance because it is a very accurate approximation of events that do occur often.

If we take the derivative of the upper boundary to find the probability density function of the lifetime, a constant *f*-function appears (we neglect the fact that the upper boundary is a discrete function): the conditional probability of failure *r* is an ascending function. If $F_L(t)$ reaches the value 1, the function f is set back to 0 and *r* will take the value of ∞.

- If the events $R<S_i$ are independent, which is the case when $\sigma(R) << \sigma(S_i)$ and S_i independent, for $F_L(t)$ an exponential function can be obtained in the same manner as for a series system. The lifetime then has an exponential distribution and the unconditional failure rate is constant.

- Finally, the lower boundary produces a corrupted density function: It holds that for $t \to \infty$ the lim $F_L(t) \neq 1$. This case represents the situation when $\sigma(R) >> \sigma(S_i)$. The survival of the first load mainly gives information about the strength of the construction and gives some kind of certainty that following loads are also survived.

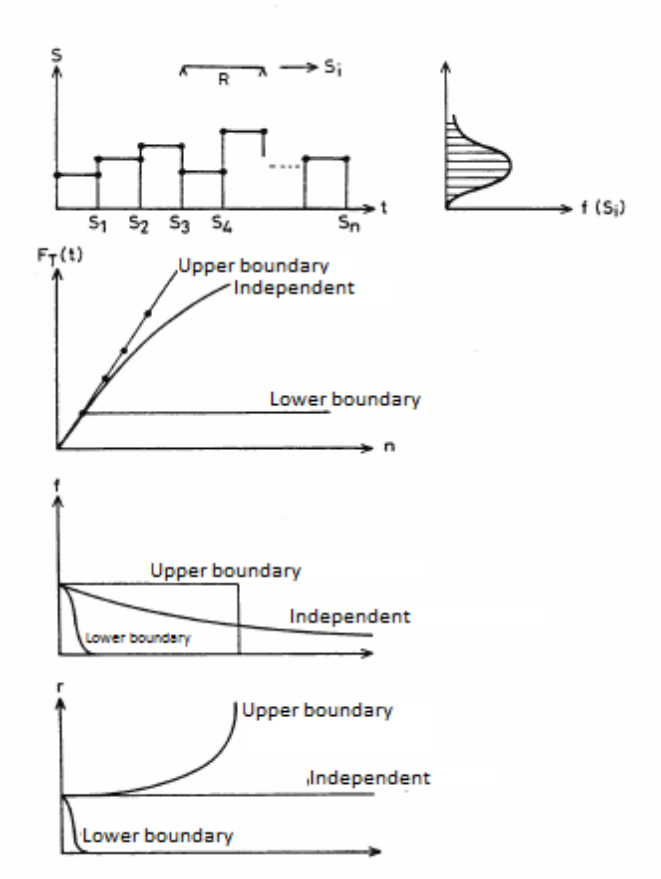


Figure 7.4 Distribution Function, Probability Density Function and the conditional failure probability with independent load changes in time and time independent loads.

7.3 Deterioration processes

The fact that the failure rate is not constant can be caused by a change of the strength or the load in time. The variations of the strength and the loads in time can be referred to as deterioration processes.

These deterioration processes may include, for example, metal fatigue, corrosion, wear, chemical action et cetera. Such processes can best be described as statistical processes. These processes are of great importance for the determination of the reliability of systems and for planning maintenance.

The variation processes of the loads are of great importance for the design and modification of systems. Consider the relative sea level rise, climatic changes et cetera. The strength of a component in a system at a certain time is determined by the deterioration process. In reference to the strength, the deterioration process is understood to mean: the variation of the strength in time. This change of the strength can be a result of external loads and internal processes.

A striking example of this is the course of the strength of a reinforced concrete element. Directly after pouring the concrete the strength of the concrete is virtually none. During the hardening and maturation a strength development takes place. Usually concrete reaches its characteristic strength after 28 days, after that the strength gradually continues to increase. If the reinforced concrete

structure is exposed to mechanical or chemical loads, the strength of the structure can change. As a result of penetration of moisture and chemicals into the reinforcement zone, the reinforcement can be harmed. The strength of the element then decreases. Processes in the pressure zone can also initiate strength reductions of the concrete.

Figure 7.5 shows that the deterioration process of the strength occupies a prominent position in the determination of the probability of failure. However, the required knowledge concerning the exact deterioration behaviour of a system is seldom if never available. Attempts are often made to find a mathematical description of the course of the strength in time, by means of physical research and observations of systems. Because the model research and/or observations of systems merely cover a limited period of time, curve fitting and extrapolation are used to obtain a deterioration model. This generally means the introduction of a model uncertainty. The input parameters for the model usually also contain uncertainties. These may originate from soil analysis or they may sometimes even be based on experience and intuition. The input parameters are thus best described as random variables, each with its own probability density function.

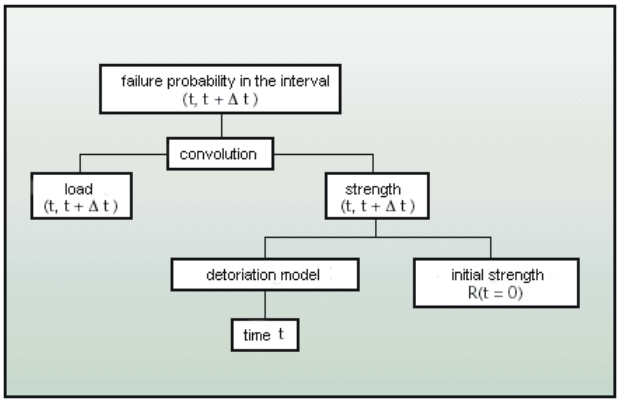


Figure 7.5 Position of the deterioration process in the calculation diagram for the risk

The deterioration process can therefore best be modelled as a statistical process. The result is that a mean value and a variance for the strength of the system are found for each moment in time. The variance increases as the uncertainties in the mathematical model and in the input parameters are larger (see chapter 11 for more information).

It is possible to model the deterioration process in two essentially different ways. The first method uses observations of the strength in time. The change of the strength per unit of time can be modelled with a mean and a variance. This model does not include an analysis of the cause of the deterioration. Example 7.1 elaborates on this method.

Example 7.1
Since 1964 annual registrations have been carried out of the profile of a narrow strip of the Dutch coast line. For that purpose a set of over 2500 measurement sections were defined along the entire North Sea coast, spaced at 200 to 250 metres. The amount of sand volume in a cross-section (m3/m) can be considered a measure for the strength of the section.

The change of the sand volume in a section in a year can be used as a deterioration model.

A section in Noord-Holland is used as an example. The measured sand volumes are given in Table 7.1. It is assumed that the changes of the sand volumes are independent every year. The mean change of the sand volume during a year over the measured period is $\overline{\Delta V}$ = -47.1 m^3/m per year and the standard deviation amounts to $\sigma_{\Delta V}$ = 110 m^3/m per year.

Table 7.1 Yearly measured sand volume in a cross section

year	volume (m3/m)	ΔV	Year	volume (m3/m)	ΔV
1964	8967		1975	8434	-118
1965	8915	-52	1976	8404	-30
1966	8750	-165	1977	8287	-117
1967	8860	110	1978	8209	-78
1968	8872	12	1979	8211	2
1969	8655	-217	1980	8395	184
1970	8746	91	1981	8265	-130
1971	8807	61	1982	8226	-39
1972	8705	-102	1983	8018	-208
1973	8748	43	1984	8050	32
1974	8552	-196	1985	7977	-73

$\Sigma(\Delta V)$ = -990 m^3/m = -47.1 m^3/(m year)
$\Sigma(\Delta V^2)$ = 288428
VAR = {288428 - 21(-47.1)^2}/20 = 12088 $s_{\Delta V}$ = 110 m^3/(m year)

With a lack of more information one assumes that the change of sand volume in one year in the section is normally distributed with the parameters $\mu_{\Delta V}$ = -47.1 m^3/m and $\sigma_{\Delta V}$ = 110 m^3/m.
The deterioration model is:

$$\Delta R(t) = \sum_{i=1}^{t} \Delta V_i \qquad (7.20)$$

The second method is based on the external or internal loads that cause the deterioration and on the resistance of an element counteracting deterioration. In this case, deterioration is defined as a function of the load and the resistance. Both this function and the load and resistance parameters contain uncertainties. This method can best be illustrated with an example.

Example 7.2
The example considers an apron on a geotextile, behind a weir in a waterway (see Figure 7.6).
The function of the apron is to ballast the geotextile, which has to prevent the washing out of bed material. The mass of the rock fill per m^2 apron is defined as strength.

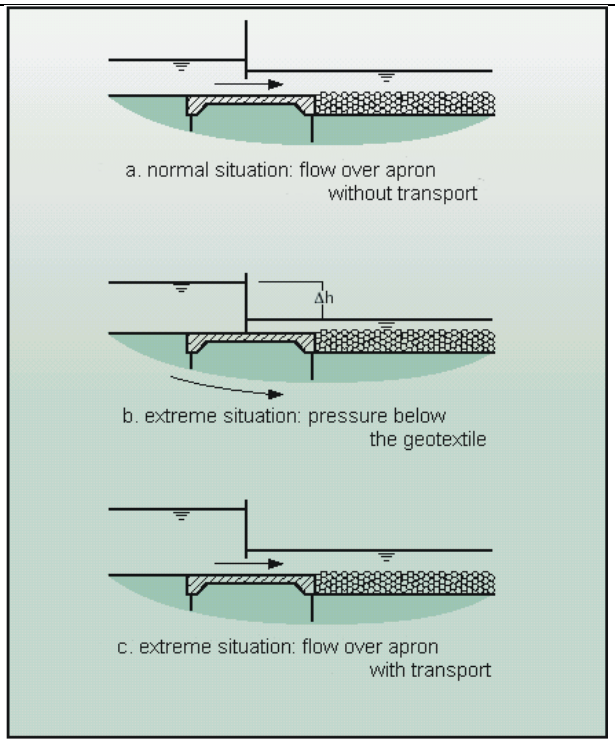


Figure 7.6 Cross section of the apron

Two types of loads are distinguished, namely:

- extreme pressure under the filter;
- flow over the filter.

The decrease of the strength (mass per area) is caused by the transport of material in water flow. The load in the deterioration model is therefore the current. During 11 months of the year the current is negligible and no transport of rock fill takes place. During one month every year a larger amount of water is discharged. It is assumed that during this one-month period the flow is stationary (the current does not vary in time) and that the flow velocities are independent each year. However, the flow velocity is not known in advance and is schematised as a random variable. The flow velocity is normally distributed with a mean $\mu_U = 4$ m/s and a standard deviation $\sigma_U = 0.5$ m/s. The probability density function of the flow velocity is thus:

$$f_U(U) = \frac{1}{0.5\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{4-U}{0.5} \right)^2 \right) \tag{7.21}$$

According to this distribution negative and very large flow velocities are also possible. This is incorrect if the flow only takes place in one direction and if the flow velocity is limited. The latter is almost always the case. However, the probability of these values is negligibly small for the given probability density function.

Amongst other factors, the dimensions (the mass) of the rock fill determine the resistance of the structure against transport. Below a certain critical flow velocity no transport of rock fill occurs. When the critical flow velocity is exceeded the transport increases with a certain exponent of the flow velocity.

In this example the relation that indicates the decline in strength for a given load, is assumed as follows:

$$U \geq U_{kr} \quad -> \quad T = A(U - U_{kr})^n$$
$$U < U_{kr} \quad -> \quad T = 0$$

in which:

- T is the decline of the strength in a year in kg/m^2;
- A is a constant, which depends on the geometry of the structure and on the resistance of the rock fill against transport, in kg A s^n/m^{n+2} per year;
- n is a constant, which depends on the geometry of the structure and on the resistance of the rock fill against transport;
- U is the flow velocity behind the weir in m/s;
- U_{kr} is the flow velocity, for which transport starts in m/s.

Usually the values for A, U_{kr} and n are determined with model research, they are then random variables. The relation that describes the transport (= reduction of the strength during a year), can therefore better be written as:

$$U \geq U_{kr} \quad -> \quad T = A(U - U_{kr})^n \varepsilon \tag{7.22}$$

in which ε is the model uncertainty.

The probability distribution of the transport can be approximated using simulation techniques. Figure 7.7 shows the results of 5000 simulations. For these simulations all variables were kept normally distributed with the parameters:

$$\begin{aligned} \mu_A &= 1260, & \sigma_A &= 126 \\ \mu_U &= 4, & \sigma_U &= 0.5 \\ \mu_{U_{kr}} &= 6, & \sigma_{U_{kr}} &= 1.2 \\ \mu_n &= 2, & \sigma_n &= 0.01 \\ \mu_\varepsilon &= 1, & \sigma_\varepsilon &= 0.02 \end{aligned} \tag{7.23}$$

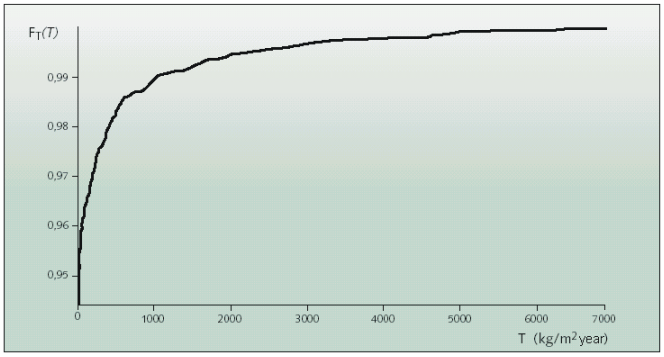


Figure 7.7 Probability distribution of the transport in a year

According to the probability distribution, there is a probability of approximately 95% that no transport of material takes place, because in those cases the critical flow velocity is not

exceeded. The expected value of the transport during a year is therefore close to T=0.

The expected value of the transport during a year is μ_T = 40 kg/m2 and the standard deviation σ_T = 336 kg/m2.

The deterioration model is:

$$\Delta R(t) = \sum_{i=1}^{t} T_i \tag{7.24}$$

in which T_i is the transport during year *i*.

With the help of the deterioration model *ΔR(t)* the strength is determined statistically at every point in time:

$$R(t) = R(0) - \Delta R(t) \tag{7.25}$$

in which:

- *R(0)* is the initial strength
- *ΔR(t)* is the deterioration after period (0,*t*)

The rate of failure of the element can be determined with the strength at time t. If maintenance is defined as restoring the strength to the initial level, the amount of maintenance required at time *t* is *ΔR(t)*.

For a load that varies in time, the same applies as for the deterioration concerning the strength of a system, bearing in mind that the definition of the limit state as regards failure, reads:

$$Z = \big(R(0) + \Delta R(t)\big) - \big(S(0) + \Delta S(t)\big) \tag{7.26}$$

An example of a deterioration process of the load can be found in a power plant. Suppose that the capacity of a plant is schematised as the strength *R(0)*. The capacity is tuned at a peak load at time *t* = 0. This load is based on the number of connections to the plant. Possibly planned expansions of residential areas and industrial areas will increase the number of connections.
By considering the number of new connections as a function of time, the increase of the peak load in time *Δ S(t)* can be estimated.

The definition of the limit state helps determine the probability of a shortage of capacity during a peak load at an arbitrary time *t*. This way, it can be judged beforehand whether or not modification of the power plant is necessary at a given point in time. Other examples of variation processes of the loads are: the relative sea level rise, increasing traffic intensity, increasing river discharge, increasing environmental loading et cetera.

7.4 Risk Calculation for Systems with a Variable Rate of Failure

(not part of the examination)

The previous section discussed the probability of failure of an element as a function of time. The relation between the unconditional rate of failure, the probability distribution of the life span and the conditional rate of failure is given as (see also section 7.1):

$$r(t) = \frac{f_L(t)}{1 - F_L(t)} \tag{7.27}$$

For small values of the probability distribution of the life span the values of $r(t)$ and $f_L(t)$ are virtually equal and the following approximately is valid:

$$F_L(t) \approx \int_0^t -r(\tau)\, d\tau \tag{7.28}$$

Figure 7.8 shows the plot of an example of both: $F_L(t)$ and $\int_0^t -r(\tau)\, d\tau$.

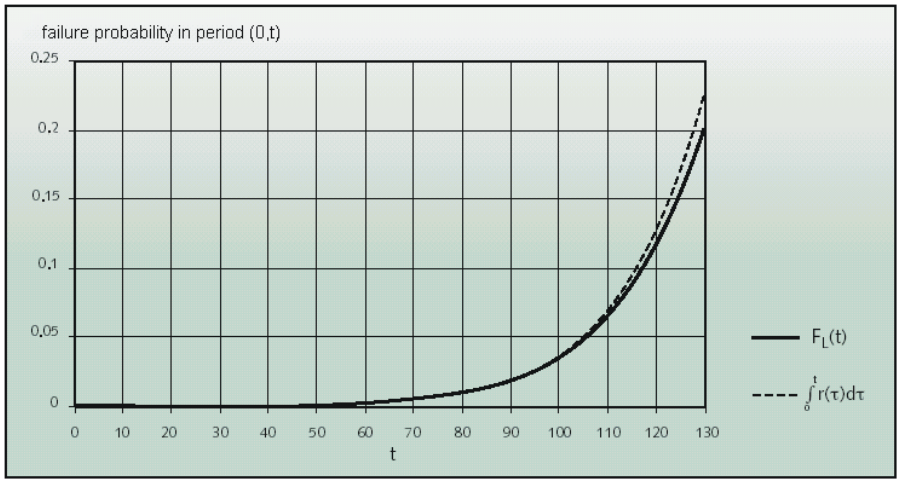


Figure 7.8 Comparison of the exact failure probability with the approximation in equation (7.28).

It seems that the approximation is very good for probabilities of failure smaller than 0.1. If the consequence of failure is independent of time, the risk over a period (0,*t*) can be defined as follows :

$$\text{Risk} = F_L(t)\, g(x) \tag{7.29}$$

in which:

- x is the consequence;
- $g(x)$ is the function of the consequence;
- $F_L(t)$ is the probability of failure over the period (0, t).

Often the consequence is also time dependent. Consider matters such as inflation, economic growth and developments which cause a change of value. An economic analysis of the risk always involves time dependence of the consequence. Usually the consequence is defined as the discounted value of the damage. In this case the economic risk can be approximated by:

$$C_R = \int_0^t g\left(x(\tau)\right) r(\tau) d\tau \tag{7.30}$$

in which:

- C_R is a discounted value of the risk;
- $g(x(\tau))$ is a discounted value of the consequence of failure at time τ;
- $r(\tau)$ is the conditional rate of failure.

The conditional rate of failure depends on the initial strength, the deterioration process and the load. For one element the following applies:

$$r(t)\,dt = P\left(R(0) + \Delta R(t) < S(\tau), \tau \in (t, t+dt) \middle| R(0) + \Delta R(\tau) > S(\tau), \tau \in (0,t)\right) \tag{7.31}$$

With the help of level II failure probability calculations this can be approximated, using the reliability function:

$$Z = R(0) + \Delta R(t) - S(t) \tag{7.32}$$

This approximation is reasonably accurate if the probability of failure is statistically independent during the different time steps. Of course, there is statistical independence which follows from the initial strength at $t = 0$ and the deterioration model. Kuijper (1992) showed that such an estimate of the rate of failure is an upper limit approximation.

Using level III simulation techniques and level II transformations the statistical dependence is taken into account entirely. Figure 7.9 gives a flowchart for the estimation of the risk over a period (0, T) with the help of Monte Carlo simulation.

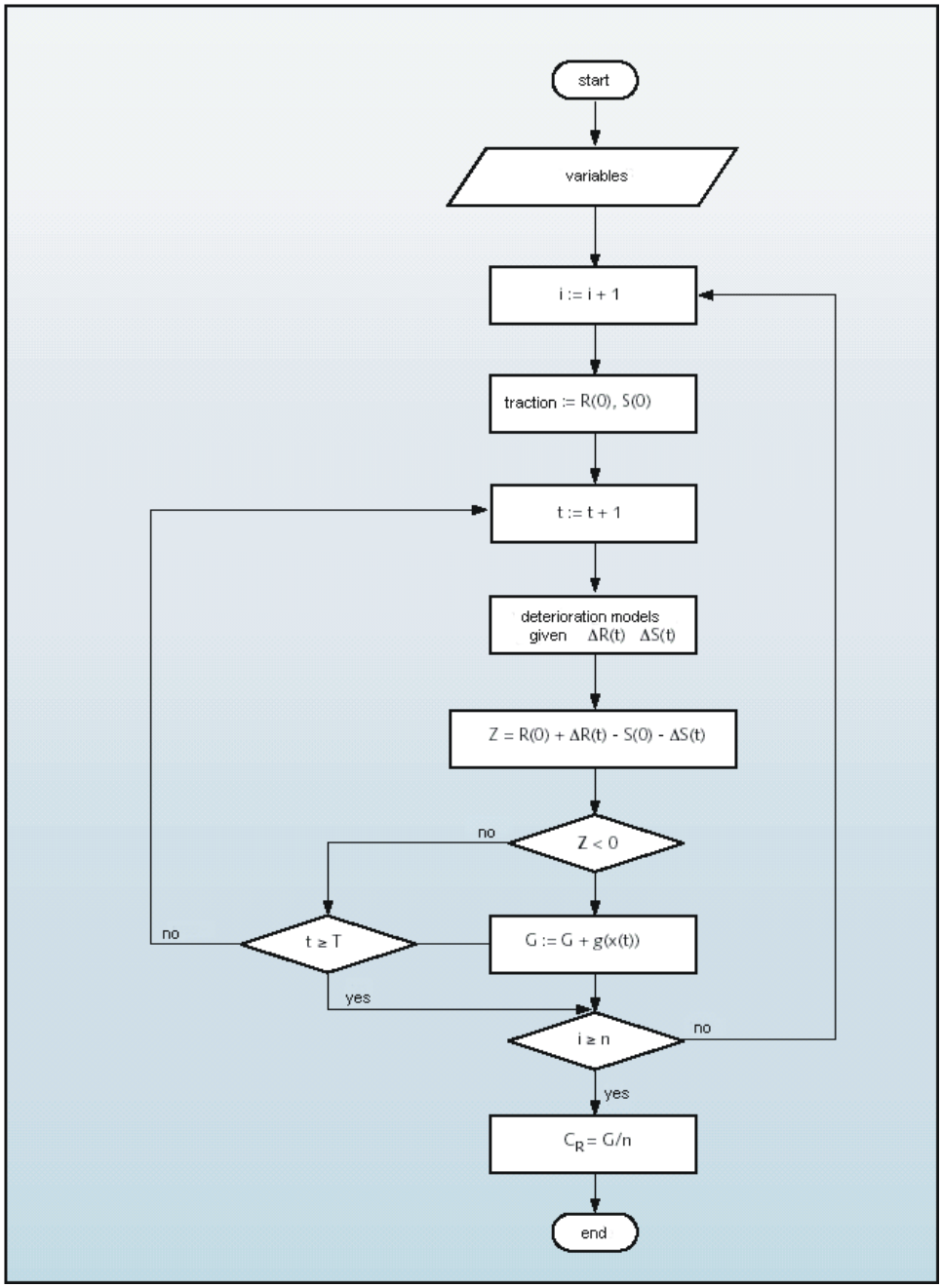


Figure 7.9 Flow chart using the Monte Carlo simulation.

7.5 Non Availability

Non-availability refers to the fraction of time that an element cannot fulfil its function. According to the reliability theory the non-availability corresponds with the probability that an element cannot perform its function satisfactorily at an arbitrary moment in time. In considering non-availability of an element or a system, in which the element is located, it must be taken into account whether or not the failure of the element can be directly observed or perceived.

Generally, if an element can fail without being directly noticed, it is tested periodically to see whether or not the function can still be fulfilled. A telephone is a good example of this. If a

telephone is broken, the owner of the phone cannot be called. The owner only notices the problem when he wants to make a call himself.

The non-availability can be divided into three types according to the cause. Non availability due to:

- Non-noticeable failure. The element does not function between the time of failure and testing, this time span is known as the down time.
- Testing or periodical maintenance.
- Repair or replacement after noticing the failure (e.g. during testing or periodical maintenance).

The down time is a random variable, for which the probability density function can be derived from the probability density function of the life span by means of a simple linear transformation. The function reads:

$$f_{T_D}(T_D) = f_L(T - T_D) \tag{7.33}$$

in which:

- T_D is the down time;
- T is the test interval.

For a constant rate of failure $r(t) = \lambda$ during period T the following applies:

$$f_{T_D}(T_D) = \lambda e^{-\lambda(T-T_D)} \tag{7.34}$$

In this case, the mean down time is defined as:

$$\mu_{T_D} = E(T_D) = \int_0^T \lambda e^{-\lambda(T-T_D)} T_D dT_D = T - \frac{1-e^{-\lambda T}}{\lambda} \approx \frac{1}{2}\lambda T^2 \tag{7.35}$$

The non-availability as a result of non-noticeable failure is then:

$$U_{nmf} = \frac{\mu_{T_D}}{T} = 1 - \frac{1-e^{-\lambda T}}{\lambda T} \approx \frac{1}{2}\lambda T \tag{7.36}$$

After each test interval there is a period during which the element is unavailable as a result of testing. This period is called the test duration τ. The non-availability as a result of testing or periodical maintenance is:

$$U_{test} = \frac{\tau}{T + \tau} \tag{7.37}$$

If testing highlights the failure of an element, repair or replacement will have to take place. A certain amount of time is needed for this, called the repair time θ. The expected non-availability caused by the repair of an element with a constant rate of failure equals:

$$U_{rep} = \frac{\theta}{\mu_L} = \frac{\theta}{1/\lambda} = \lambda\theta \tag{7.38}$$

In the case of failures that are directly noticed, the element will generally be repaired or replaced directly after failure. The non-availability is then only caused by the duration of the repair.

7.6 Markov Processes

not part of the examination

A statistical process is a process that can be in different states at different points in time, for which the course of the state in time is determined by a stochastic process.

An example of this is the description of the state of a binary element. Two states can be distinguished for the element: functioning and non-functioning. The probability that an element does not function at a certain time is described by the probability distribution of the life span.
The expected value of the fraction of the time that a repairable element does not function, is the non-availability. In the foregoing both the probability distributions of the life span and the non-availability have been discussed. This section describes a method for determining the non-availability of elements, for which the state at an arbitrary time is described by a Markov-process.

A statistical process is a Markov-process if the probability distribution of the progress of the process from an arbitrary point in time, is determined exclusively by the state at that point in time and not by manner in which the state was reached. A Markov-process is therefore a process without a memory. In the reliability analysis of a component this means that a repaired component is equivalent to a new component.

It is customary to present a Markov-process in a diagram. In the diagram the states in which the process can be are indicated with a number.

The transition probabilities between the different states are designated with connection lines. The transition speed is shown on these lines. The transition speed is understood to be, for example, the rate of failure or the repair speed within a given time interval. Figure 7.10 gives a Markov-diagram for a binary element, for which the rate of failure and the repair speed are constant and for which failure immediately leads to repair. This means that the failure can be directly noticed.

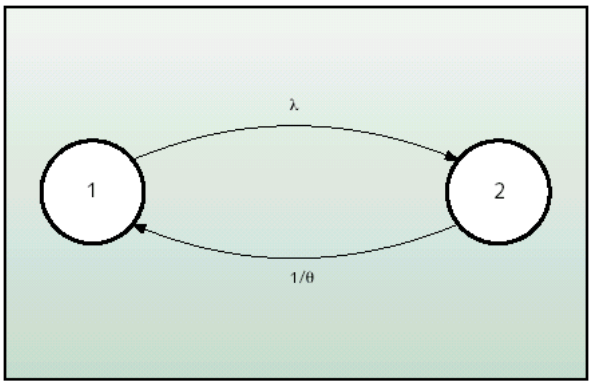


Figure 7.10 Markov diagram for a binary element

The probability that the state of an element moves from functioning to failure during a time interval (t, $t + \Delta t$) is:

$$P_{1\to 2}(t, \Delta t) = P_1(t)\,\lambda\,\Delta t \tag{7.39}$$

in which:

- $P_1(t)$ is the probability that the element is in state 1 at time t;
- $\lambda\,\Delta t$ is the probability that the element moves from state 1 to state 2 during the interval $(t, t + \Delta t)$, provided the element is in state 1 at time t.

The change of the probability of state 1 or state 2 during an interval with time span dt is:

$$\begin{aligned} dP_1(t) &= -\lambda\,dt\,P_1(t) + \frac{1}{\theta}\,dt\,P_2(t) \\ dP_2(t) &= +\lambda\,dt\,P_1(t) - \frac{1}{\theta}\,dt\,P_2(t) \end{aligned} \tag{7.40}$$

Dividing by dt gives two linked first order linear differential equations:

$$\begin{aligned} \frac{dP_1(t)}{dt} &= -\lambda P_1(t) + \frac{1}{\theta}P_2(t) \\ \frac{dP_2(t)}{dt} &= +\lambda P_1(t) - \frac{1}{\theta}P_2(t) \end{aligned} \tag{7.41}$$

The boundary condition for solving these differential equations is:

$$P_1(t) + P_2(t) = 1 \tag{7.42}$$

Substitution of the boundary condition in equation (7.41) gives:

$$\frac{dP_2(t)}{dt} + \left(\lambda + \frac{1}{\theta}\right)P_2(t) = \lambda \tag{7.43}$$

Supposing that it is certain that at point in time $t = 0$ the element is in state 1, the initial condition is $P_2(0) = 0$. With this initial condition the solution to the differential equation reads:

$$P_2(t) = \frac{\lambda}{\lambda + 1/\theta}\left(1 - e^{-(\lambda + 1/\theta)}\right) \tag{7.44}$$

By inserting $t = \infty$ the so-called stationary state is found. This reads:

$$P_2(\infty) = \frac{\lambda}{\lambda + 1/\theta} \approx \lambda\,\theta \tag{7.45}$$

Figure 7.11 gives the course of $P_2(t)$ and the stationary situation.

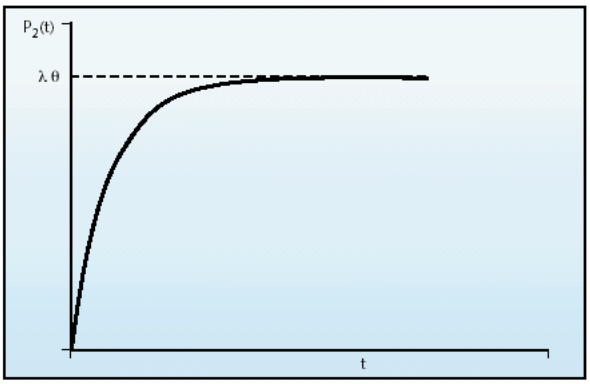


Figure 7.11 Course of the probability $P_2(t)$ in time.

For a binary element the stationary situation of a Markov-process, in which perceptible failure occurs, is reached after approximately three times the average duration of repair.

The stationary situation can also be found quickly by assuming the conditions for this situation:

$$\frac{dP_1(\infty)}{dt} = 0$$
$$\frac{dP_2(\infty)}{dt} = 0 \tag{7.46}$$

This leads to a simple set of linear equations:

$$-\lambda P_1(\infty) + \frac{1}{\theta} P_2(\infty) = 0$$
$$+\lambda P_1(\infty) - \frac{1}{\theta} P_2(\infty) = 0 \tag{7.47}$$

The solution to these equations reads:

$$P_1(\infty) = \frac{1/\theta}{\lambda + 1/\theta} = \frac{1}{\lambda\theta + 1}$$
$$P_2(\infty) = \frac{\lambda}{\lambda + 1/\theta} = \frac{\lambda\theta}{\lambda\theta + 1} \tag{7.48}$$

$P_2(\infty)$ is the probability that the element is in state 2 in the stationary situation. This can also be interpreted as the fraction of the time that the element is not available.

7.7 Combinations of Time Dependent Loads

In many cases loads are not stochastic variables but stochastic processes, i.e. fluctuations in time occur. In Figure 7.12 some typical fluctuations have been drawn. The load effects caused by self-weight has very few fluctuations in time and can be considered to be constant in time. With respect to the variable loads we subdivide the load types as shown in Figure 7.12. Examples are respectively floor loads by the inventory, floor loads during receptions and wind loads. In all these case it is necessary to make a clear distinction between the 'instantaneous' or 'arbitrary point in time (APT)' distribution and the extreme value distribution. Both have been drawn in Figure 7.12.

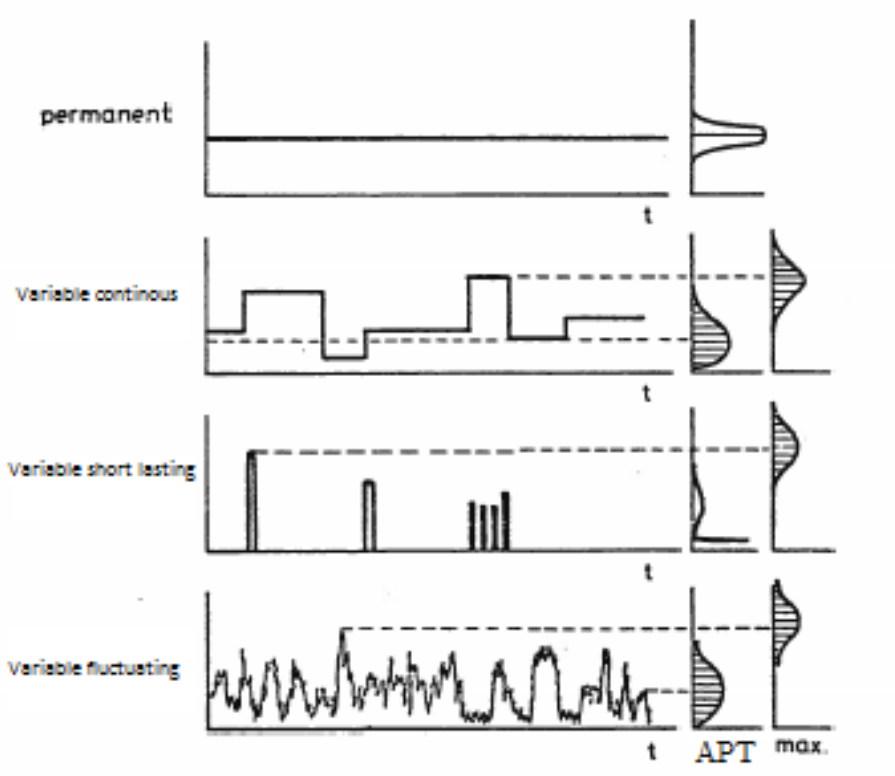


Figure 7.12 Different loadtypes as a function of the time with their 'instantaneous' or 'APT'distribution and extreme value distribution of the maxima functions

In combining loads one has to solve the point that the maxima of the different loads do not have to occur simultaneously. For a specific structure, first the loads should be summed up and then the maximum total load (effect) can be determined.

A practical solution for the determination of the failure probability of an element, which is subject to loads varying differently in time, is given by the Turkstra-rule (Turkstra, 1980) According to this rule the extreme value of one of the loads is combined with the instantaneous values of the other loads. This way the probability of failure is calculated for every combination and the combination with the greatest probability of failure is considered normative.
This method slightly underestimates the probability of failure.

Ferry Borges and Castanheta's load model (1972) permits a slightly more balanced handling of load combinations. Both models are elaborated below.

7.7.1 Turkstra's combination rule

We will now generalize the representation in Figure 7.12 towards k variable loads $X_1(t)$ to $X_k(t)$. The maximum load $X_{max}(t_{ref})$ over the reference period t_{ref} is equal to:

$$X_{\max}(t_{ref}) = \max_{t_{ref}}\left[X_1(t) + X_2(t) + ... + X_k(t)\right] \tag{7.49}$$

Turkstra suggested the following approximation:

$$
\begin{aligned}
Y_1 &= \max_{t_{ref}} \left[X_1(t) \right] + X_2(t) + \ldots + X_k(t) \\
Y_2 &= X_1(t) + \max_{t_{ref}} \left[X_2(t) \right] + \ldots + X_k(t) \\
&\vdots \\
Y_k &= X_1(t) + X_2(t) + \ldots + \max_{t_{ref}} \left[X_k(t) \right]
\end{aligned}
\tag{7.50}
$$

For these variables, the maximum of each variable load is combined with the instantaneous (APT) value of the other variable loads. Hence, the following holds:

$$
X_{\max}(t_{ref}) \cong \max_i \{ Y_i \}
\tag{7.51}
$$

A more refined approximation is obtained when applying the Borges-Castanheta model for variable loads.

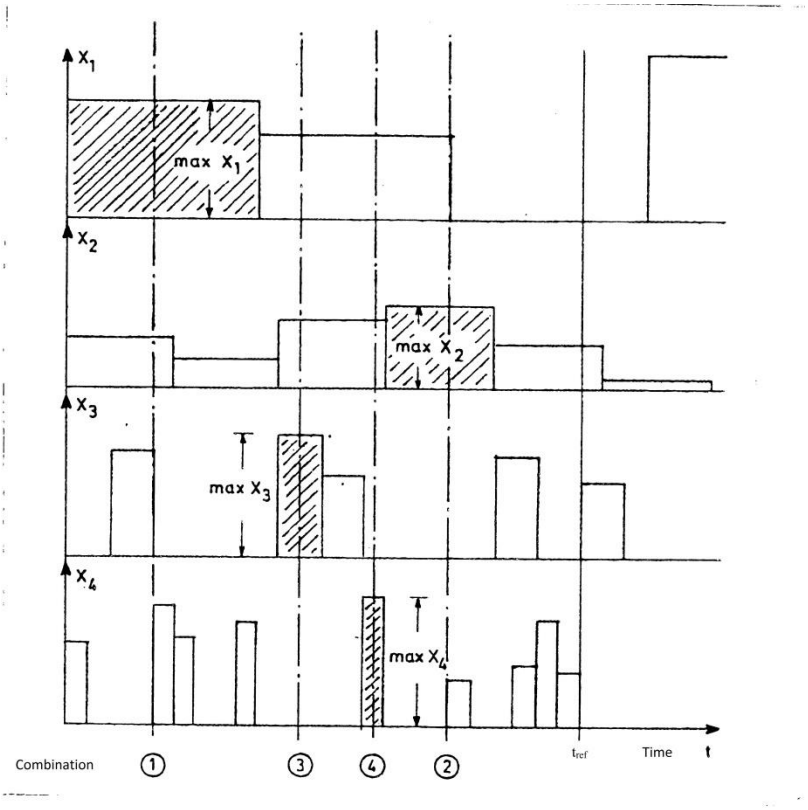


Figure 7.13 Turkstra's combination rule

7.7.2 Ferry Borges-Castanheta load model (FBC model)

This load model assumes that every load S_i is constant during a certain unit interval τ_i. During the reference period t_{ref}, for which the probability of failure is calculated, n_i load repetitions occur. This number is:

$$
n_i = \frac{t_{ref}}{\tau_i}
\tag{7.52}
$$

Further, it is also assumed that the load level in successive intervals is independent and identically distributed according to the distribution function $f_{X_i}(x_i)$. For this kind of rectangular pulse process one simply obtains:

$$F_{\max}(x_i) = \left[F_{X_i}(x_i) \right]^{n_i} \tag{7.53}$$

The number of load recurrences is not equal for all loads. The loads are arranged such that $n_1 \leq n_2 \leq ... \leq n_i \leq ... \leq n_m$. Moreover n_i/n_{i-1} must be a positive natural number. Figure 7.14 gives an example of loads according to this model.

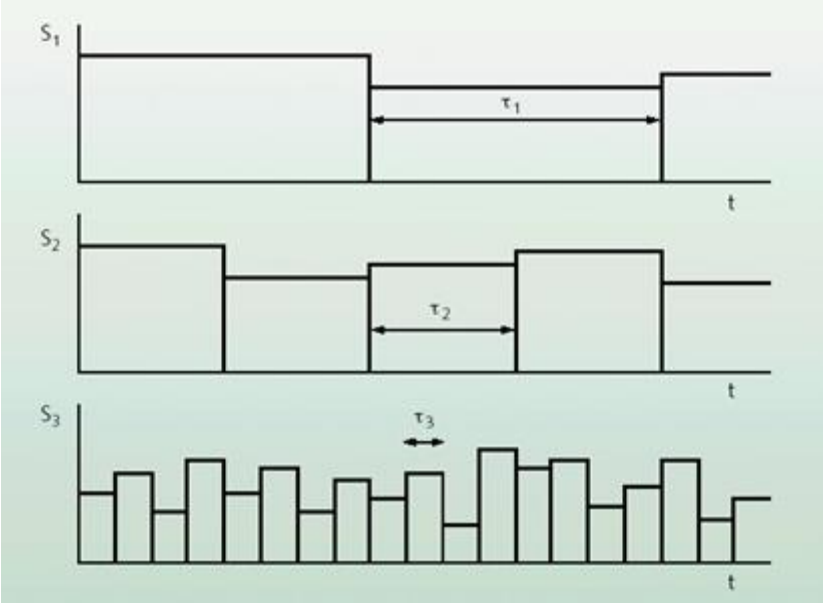


Figure 7.14 Schematisation of load varying in time according to the FBC model.

The determination of the combinations takes place as follows:

Make all possible combinations of the loads. For *m* loads *m!* combinations are possible. For the first load in the combination the extreme value of the number of load repetitions during period t_{ref} are taken into account. For every subsequent load in the combination the number of possible load repetitions in the unit interval of the preceding load has to be ascertained and for that the extreme value distribution has to be determined.

Example 7.3

Determine the load combinations of the three loads that vary independently in time, given in Figure 7.14. In total 6 combinations are possible:

- combination 1: S_1 S_2 S_3
- combination 2: S_1 S_3 S_2
- combination 3: S_2 S_1 S_3
- combination 4: S_3 S_1 S_2
- combination 5: S_2 S_3 S_1
- combination 6: S_3 S_2 S_1

Because $\tau_1 \geq \tau_2$ the combinations 3 and 5 are equivalent and because $\tau_2 \geq \tau_3$ the combinations 4 and 5 are also equivalent. Thus four combinations remain. For these combinations the following reasoning applies:

- Combination 1: During the reference period t_{ref}, n_1 load repetitions of load S_1 are possible. During the unit interval τ_1 another n_2/n_1 load repetitions of load S_2 can occur. Subsequently, it is possible that another n_3/n_2 load repetitions occur of load S_3 during unit interval τ_2.
- Combination 2: For this combination n_1 load repetitions of load S_1 are also possible. During the unit interval τ_1 another n_3/n_1 load repetitions can occur of load S_3. In τ_3 no more load repetitions are possible for S_2.
- Combination 3: During the reference period t_{ref}, n_2 load repetitions of load S_2 are possible. During the unit interval τ_2 another n_3/n_2 load repetitions can occur of load S_3, but no more load repetitions are possible for S_1.
- Combination 4: In this combination n_3 load repetitions of load S_3 are possible. During the unit interval τ_3 no load repetitions can occur for loads S_1 and S_2.

The load combinations and the number of load repetitions that have to be taken into account are given in Table 7.2.

Table 7.2 Possible Load combinations.

combination	**number of load repetitions**		
	S_1	S_2	S_3
1	n_1	n_2/n_1	n_3/n_2
2	n_1	1	n_3/n_1
3	1	n_2	n_3/n_2
4	1	1	n_3

In chapter 9 an application will be shown for combinations of loads in a Level I probabilistic analysis.

References

- Ferry Borges, J. en M. Castanheta, Structural safety - 2nd edition. Laboratorio Nacional de Engenharia Civil, Lissabon, 1972.
- Kuijper, H.K.T., Maintenance in hydraulic engineering, economically sound planning of maintenance (in Dutch: "Onderhoud in de waterbouw, economisch verantwoord plannen van onderhoud"). Delft University of Technology, Delft, 1992.
- Turkstra, C.J. and Madsen, H.O., Load combinations in codified structural design. Journal of Engineering Structural Division., ASCE, Volume 106, nr. St. 12, December 1980.

Chapter "Level I Methods"
Author: R.D.J.M. Steenbergen

Large parts of this chapter are based on 'Risk Analysis of Construction Processes' by L. Taerwe and R. Caspeele (UGent lecture notes)

Contents

8 Level I Methods

8.1 Introduction

The most common way of creating a design is by means of regulations and guidelines. The essence of the standards is that a certain characteristic value of the strength is divided by a factor and that the characteristic value of the load is multiplied by a factor, for which the following must apply:

$$\frac{R_k}{\gamma_R} > \gamma_S S_k \tag{8.1}$$

The factors γ_R and γ_S are known as partial safety factors. The representative values of the strength and the load are generally calculated with:

$$\begin{aligned} R_k &= \mu_R + k_R \sigma_R \\ S_k &= \mu_S + k_S \sigma_S \end{aligned} \tag{8.2}$$

In which k_R will be negative and k_S can be positive or negative depending if the action is favourable or non-favourable. For definitions of the characteristic values, see section 10.4. For material properties often R_k is defined as that value that has a probability of non-exceedance of 5%. For loads on structures often the characteristic value is also called the representative value F_{rep}.

The procedure above is shown in Figure 8.1.

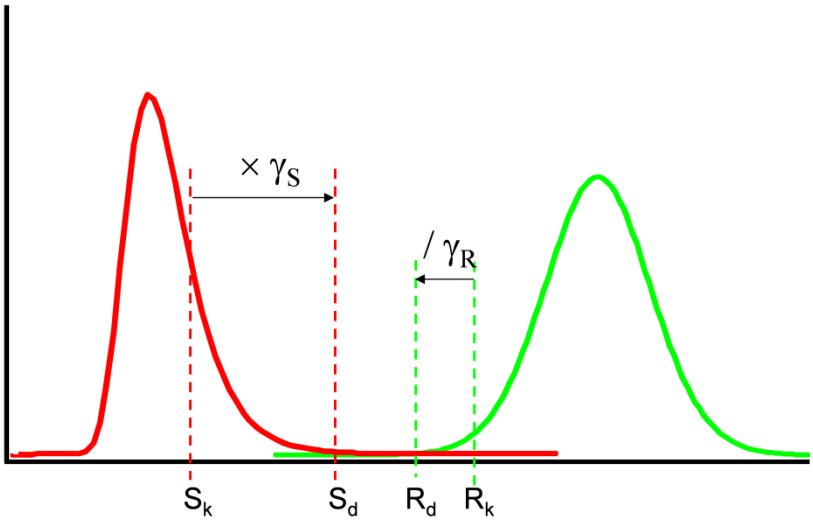


Figure 8.1: Probability density functions showing the variations in load (red, left) and resistance (green, right). The design load and resistance are chosen in such a way that a structure with a sufficiently low probability of failure can be designed. Codes and guidelines provide information on the load and resistance factors (γ's) that can be used to derive these values

Safety factors should be applicable is many (different) situations; therefore some standardisation is needed.

In order to get insight in the influence of variables on the reliability, the values of sensitivity factors α should be determined using a probabilistic calculation. The most appropriate method for this is the level II method because the α values directly result from the analysis. Using level III methods it is also possible to calculate the α-values, but this is less straightforward.
According to the Eurocode, the core of the level I design method is that the α-values are standardized and that they are considered independent of an arbitrary specific case. The standardized α-values can be obtained by carrying out failure probability calculations for a number of reference cases. Subsequently a weighted average of the calculated α-values can be determined, provided that the error in the resulting probabilities of failure is minimized. This leads to the α-values in Table 8.1. The derivation is given in the next section.

Table 8.1 Standardized α values for structures according to the Eurocode

Variable	α
Dominant strength parameter	0.80
Remaining strength parameter	0.32
Dominant load parameter	-0.70
Remaining load parameter	-0.28

Usually it is not possible to point to the dominant parameter amongst a number of load parameters on beforehand. Therefore, alternately one of the parameters is considered dominant. This results in a number of load scenarios, each with another dominant load parameter. The different load scenarios are mutually exclusive. The design point follows from the normative load scenario. The ratio of 40% (i.e. 0.32/0.8 or 0.28/0.7) between the dominant parameter and the remaining parameters is explained in the next section.

In this chapter the symbol E is introduced for the load effect as also done in Eurocode NEN-EN 1990. The basic limit state equation becomes $Z = R - E$. For direct loads in general, such as wind loads or wave heights, the symbol S is used. In the verification of the structural reliability in most of the cases we use resulting load effects such as bending moments in a structure. For the load effects in many cases models are used that are functions of the loads. For these models in a probabilistic calculation also model uncertainties should be taken into consideration. Therefore a different symbol E is used in the case of resulting or derived load effects.

8.2 Simplified Level II Method

8.2.1 General Formulation

As a very common extension of the basic limit state equation

$$g(\underline{X}) = R - E = 0 \tag{8.3}$$

we now consider that R and E are themselves very often a function of multiple variables and that most often the contribution of the resistance and the load to the limit state function can be separated, i.e.:

$$g(\underline{X}) = g_R(R_1, ..., R_i, ..., R_m) - g_E(E_1, ..., E_j, ..., E_n) \tag{8.4}$$

Due to dependence on several load factors, the load effect will be denoted by E rather than S. Hence, the sensitivity factors can in general be calculated as (expression (6.23)):

$$\alpha_{R_i} = \frac{\frac{\partial g_R(r^*)}{\partial R_i}\sigma_{R_i}}{\sqrt{\sum_{i=1}^{m}\left(\frac{\partial g_R(r^*)}{\partial R_i}\cdot\sigma_{R_i}\right)^2 + \sum_{j=1}^{n}\left(\frac{\partial g_E(e^*)}{\partial E_j}\cdot\sigma_{E_j}\right)^2}} \tag{8.5}$$

In which r^* and e^* are the design points values.

And similarly in case of α_{E_j}. Designating the summations in the denominator as σ_R^{*2} and σ_E^{*2} respectively, leads to:

$$\alpha_{R_i} = \frac{\frac{\partial g_R(r^*)}{\partial R_i}\sigma_{R_i}}{\sqrt{\sigma_R^{*2} + \sigma_E^{*2}}} \tag{8.6}$$

This expression can be split up as follows:

$$\alpha_{R_i} = \frac{\sigma_R^*}{\sqrt{\sigma_R^{*2} + \sigma_E^{*2}}} \cdot \frac{\frac{\partial g_R(r^*)}{\partial R_i}\sigma_{R_i}}{\sigma_R^*} = \alpha_R \cdot \tilde{\alpha}_{R_i} \tag{8.7}$$

and similarly $\alpha_{E_j} = \alpha_E \cdot \tilde{\alpha}_{E_j}$. The factor α_R (resp. α_E) represents the contribution of the general / total resistance (resp. the load) in the total variability at the design point. The factor $\tilde{\alpha}_{R_i}$ (resp. $\tilde{\alpha}_{E_j}$) represents the contribution of the variable R_i (resp. E_j) in the variability of the resistance function (resp. the load function) at the design point. In case one allows a deviation $\Delta\beta = \pm 0.5$ compared to the values for β_{target} that are mentioned in Table 10.6 and in accordance to König and Hosser:

$$0.15 \le \frac{\sigma_E}{\sigma_R} \le 3.48 \tag{8.8}$$

holds, then the global sensitivity factors can be derived as:

$$\alpha_R = 0.8, \qquad \alpha_E = -0.7 \tag{8.9}$$

In Annex C of EN 1990 the application boundaries $0.16 < \sigma_E/\sigma_R < 7.6$ are mentioned. In case σ_E/σ_R falls outside these boundaries, one should consider $\alpha = \pm 1$ in case of the variable with the largest standard deviation and $\alpha = \pm 0.4$ in case of the variable with the smallest standard deviation.

When R_i and E_j are rearranged according to decreasing σ_i or respectively σ_j values (standard deviations), one can derive the following expressions for the partial sensitivity factors:

$$\tilde{\alpha}_{R_i} = \sqrt{i} - \sqrt{i-1}, \qquad i = 1, ..., m \tag{8.10}$$

$$\tilde{\alpha}_{E_j} = \sqrt{j} - \sqrt{j-1}, \qquad j = 1, ..., n \tag{8.11}$$

so that:

$$\tilde{\alpha}_{R_1} = \tilde{\alpha}_{E_1} = 1.0 \tag{8.12}$$

$$\tilde{\alpha}_{R_i} = \tilde{\alpha}_{E_j} = 0.4 \qquad \text{for } i \text{ and } j \geq 2 \text{ (approximative for } i, j > 2) \tag{8.13}$$

Hence, R_1 and E_1 are the basic variables with the most important influence on g_R and g_E (largest standard deviation). When more than 3 variables occur in g_R or $g_E$the approximative procedure leads to a more conservative design and one can reasonably consider $\alpha_i = 0$ for $i > 3$ in order to avoid exceeding $\beta_{target} + 0.5$.

8.2.2 Background of sensitivity factors

(not part of the examination)

In order to explain the origin of equation (8.11) we first consider the linear combination:

$$E = E_1 + E_2 \tag{8.14}$$

with $E_1 : N(\mu_1, \sigma_1)$ and $E_2 : N(\mu_2, \sigma_2)$. In the plane of the normalized variables U_1 and U_2 the expression $E_1 + E_2 = E^*$ can be written as:

$$U_1 \sigma_1 + U_2 \sigma_2 + \mu_1 + \mu_2 = E^* \tag{8.15}$$

The design value E^* has to be determined so that there exists a fraction $p = \Phi(\alpha_E \beta)$ that is larger than E^*. This means that the straight line (8.15) has to be located at a distance $|\alpha_E \beta|$ from the origin (Figure 8.2). The slope of the line (8.15) is determined by σ_1/σ_2. The envelop of all straight lines which are located at a distance $|\alpha_E \beta|$ from the origin is given by a circle with radius $|\alpha_E \beta|$. Hence, the design point with coordinates u_1^* and u_2^* will correspond to the tangent point of the line under consideration (i.e. determined by σ_1/σ_2) to the circle. The line which is drawn in Figure 8.2 corresponds to $\sigma_1/\sigma_2 = 1$. In order to be able to obtain fixed design values, we now choose the points *A* and *B* on this line corresponding with the following coordinates:

$$\begin{aligned} u_1^* &= -\alpha_E \beta \qquad u_2^* = -(\sqrt{2}-1)\,\alpha_E \beta \qquad \text{in case } \sigma_1 > \sigma_2 \\ u_2^* &= -\alpha_E \beta \qquad u_1^* = -(\sqrt{2}-1)\,\alpha_E \beta \qquad \text{in case } \sigma_2 > \sigma_1 \end{aligned} \tag{8.16}$$

in which the second coordinate follows from the simple geometrical considerations illustrated in Figure 8.2.

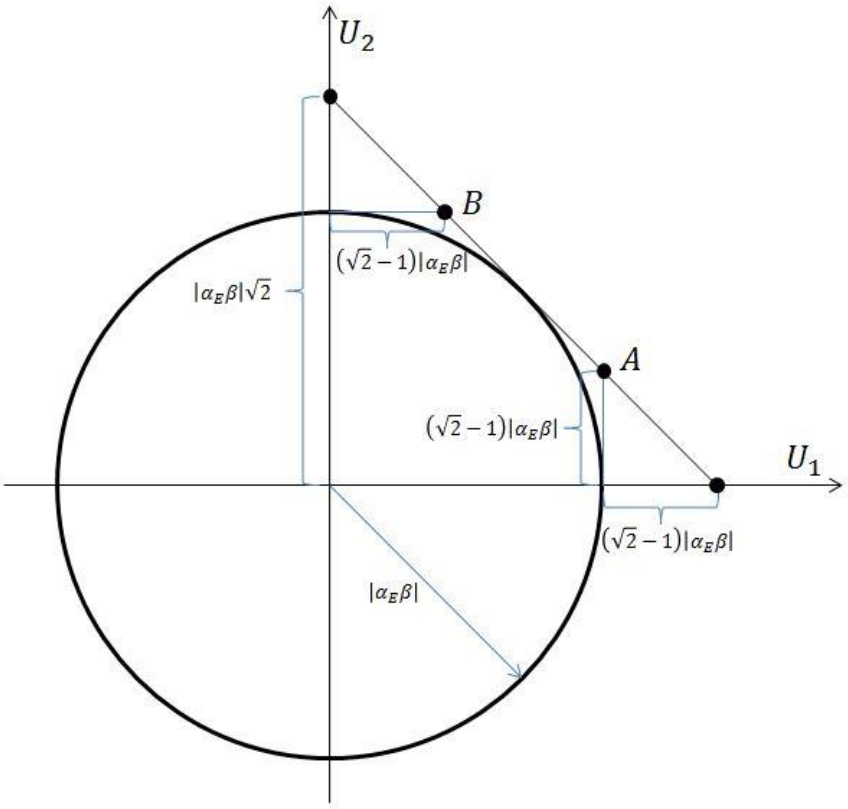


Figure 8.2 Simplified combination of two basic variables E_1 and E_2

Lines which are going through the first point, will for all values of the ratio σ_1/σ_2 be located between the illustrated line ($\sigma_1 = \sigma_2$) and the vertical line through $U_1 = u_1^* = -\alpha_E\ \beta$ (in case $\sigma_2 = 0$ or $\sigma_1/\sigma_2 \rightarrow \infty$). The lines corresponding to these two limit cases are tangent to the circle. For intermediate values of σ_1/σ_2 the line through *A* will not be tangent to the circle with radius $|\alpha_E\ \beta|$ but will be tangent to a somewhat larger circle. Hence, the probability of exceedance of the combination will be smaller than the anticipated value $\Phi(\alpha_E\ \beta)$. This larger circle goes through *A* and thus has a radius:

$$\sqrt{(\alpha_E\ \beta)^2 + (\sqrt{2}-1)^2\ (\alpha_E\ \beta)^2} = 2\sqrt{1-\frac{\sqrt{2}}{2}}\,|\alpha_E\ \beta| = 1.082\,|\alpha_E\ \beta| \tag{8.17}$$

In case of α_E = 0.7 the coefficient of β becomes 0.76, which is only a small deviation compared to the target value 0.7 β.

In case of a summation of *n* variables:

$$E = E_1 + ... + E_n \tag{8.18}$$

one can use the following analytical derivation in order to show that the simplified values of $\tilde{\alpha}_{E_j}$ in:

$$E^* = \sum_{j=1}^{n} \mu_{E_j} - \alpha_E\ \beta \sum_{j=1}^{n} \tilde{\alpha}_{E_j}\ \sigma_{E_j} \tag{8.19}$$

can be calculated according to (8.11). We remind that the indices are designated in such a way that:

$$\sigma_{E_1} \geq \sigma_{E_2} \geq ... \geq \sigma_{E_n} \tag{8.20}$$

When considering (8.18), the equivalent expression of (8.6) in case of load variables becomes:

$$\tilde{\alpha}_{E_j} = \frac{\frac{\partial g_E(e^*)}{\partial E_j} \cdot \sigma_{E_j}}{\sigma_E^*} = \frac{\sigma_{E_j}}{\sqrt{\sigma_{E_1}^2 + ... + \sigma_{E_n}^2}} \tag{8.21}$$

Hence, the last summation in (8.19) can be rewritten as:

$$\sum_{j=1}^{n} \tilde{\alpha}_{E_j} \cdot \sigma_{E_j} = \frac{\sigma_{E_1}^2 + ... + \sigma_{E_n}^2}{\sqrt{\sigma_{E_1}^2 + ... \sigma_{E_n}^2}} \tag{8.22}$$

Or:

$$\sum_{j=1}^{n} \tilde{\alpha}_{E_j} \cdot \sigma_{E_j} = \sqrt{\sigma_{E_1}^2 + ... + \sigma_{E_n}^2} \tag{8.23}$$

As a safe approximation for the determination of the value $\tilde{\alpha}_{E_j}$ for the j^{th} variable load E_j one can consider in the previous expression:

$$\sigma_{E_1} = \sigma_{E_2} = ... = \sigma_{E_j} \qquad \sigma_{E_{j+1}} = ... = \sigma_{E_n} = 0 \tag{8.24}$$

In case $j = 1$ one obtains $\tilde{\alpha}_{E_j} = 1$. In case $j = 2$ the following holds, when considering that $\tilde{\alpha}_{E_1} = 1$:

$$\tilde{\alpha}_{E_1} \sigma_{E_1} + \tilde{\alpha}_{E_2} \sigma_{E_2} = \sigma_{E_2} (1 + \tilde{\alpha}_{E_2}) = \sqrt{\sigma_{E_2}^2 + \sigma_{E_2}^2} \tag{8.25}$$

Or:

$$\tilde{\alpha}_{E_2} = \sqrt{2} - 1 \tag{8.26}$$

In case $j = 3$ one obtains the following as an approximation for (8.23):

$$\sigma_{E_3} (1 + \sqrt{2} - 1 + \tilde{\alpha}_{E_3}) = \sqrt{\sigma_{E_3}^2 + \sigma_{E_3}^2 + \sigma_{E_3}^2} \tag{8.27}$$

$$\tilde{\alpha}_{E_3} = \sqrt{3} - \sqrt{2} \tag{8.28}$$

or in general:

$$\tilde{\alpha}_{E_j} = \sqrt{j} - \sqrt{j-1}\,. \tag{8.29}$$

8.3 Verification Procedure According to NEN-EN 1990

In this section the implementation of the Level I calculations in Eurocode NEN-EN 1990 is discussed. It is used to derive design values for load actions and resistance.

8.3.1 Design values of actions

The design value F_d of an action F is defined in general as:

$$F_d = \gamma_f \cdot F_{rep} \qquad \text{with } F_{rep} = \psi\, F_k \tag{8.30}$$

with γ_f a partial factor that accounts for unfavourable fluctuations of F compared to F_{rep} and ψ a combination factor. The value of ψ is equal to 1, ψ_0, ψ_1 or ψ_2 (see section 10.5.4) for an explanation of ψ-factors).

8.3.2 Design value of load effects

In general the design value of the load effect E_d (internal forces, stresses, deformations, displacements, etc.) can be written as:

$$E_d = \gamma_{Sd}\; E\,(\gamma_{f,i}\; F_{rep,i}, \underline{X}_d, \underline{a}_d) \qquad i \geq 1 \tag{8.31}$$

With:

- $\underline{a}_d$ the vector with the design values of the geometrical variables;
- $\underline{X}_d$ the vector with the design values of the material characteristics (optional);
- γ_{Sd} a partial factor that accounts for uncertainties related to:
  - the calculation model for the determination of the load effects, designated as *E(...)*;
  - the probabilistic models for the loads.

Most often (8.31) can be simplified as:

$$E_d = E\,(\gamma_{F,i}\; F_{rep,i}, \underline{X}_d, \underline{a}_d) \qquad (i \geq 1) \tag{8.32}$$

With:

$$\gamma_{F,i} = \gamma_{f,i}\; \gamma_{Sd} \tag{8.33}$$

In case of a linear elastic analysis, (8.31) and (8.32) both yield the same result. In case of a nonlinear analysis the following simplifications are safe approximations (in case of a single dominating variable):

a) E increases more than proportional to F:

$$E_d = E\,(\gamma_F\; F_{rep}, ...) \tag{8.34}$$

b) E increases less than proportional to F:

$$E_d = \gamma_F\; E\,(F_{rep}, ...) \tag{8.35}$$

In both cases one obtains larger (more conservative) values compared to the use of (8.31).

8.3.3 Design values of material and product properties

The design value X_d of a material or product characteristic is in general defined as:

$$X_d = \eta \frac{X_k}{\gamma_m} \tag{8.36}$$

with:

- η the mean value of a conversion factor considering:
  - volume and scale effects;
  - influences of temperature and humidity;

  - other relevant parameters;
- γ_m a partial factor considering:
  - unfavourable fluctuations of X compared to X_k;
  - the stochastic nature of η.

8.3.4 Design values of geometrical characteristics

a) Most often the design value a_d of a geometrical characteristic is considered equal to a nominal value or:

$$a_d = a_{nom} \tag{8.37}$$

b) In case fluctuations of geometrical characteristics (e.g. the location of loads) significantly influence the structural safety, the design value is defined as:

$$a_d = a_{nom} \pm \Delta a \tag{8.38}$$

where Δa takes into account:

- unfavourable fluctuations compared to the characteristic or nominal value;
- the cumulative effect of the fluctuations of different geometrical characteristics.

The influence of other deviations is covered by γ_F and γ_M.

8.3.5 Design value of resistance effects

The design value R_d of a resistance effect is in general defined as:

$$R_d = \frac{1}{\gamma_{Rd}} R\left(X_{d,i}\,;\underline{a}_d\right) = \frac{1}{\gamma_{Rd}} R\left(\eta_i \frac{X_{k,i}}{\gamma_{m,i}}\,;\underline{a}_d\right) \qquad i \geq 1 \tag{8.39}$$

with γ_{Rd} a partial factor accounting for the uncertainties related to the calculation model $R(...)$ and possible fluctuations of geometrical properties not explicitly accounted for through (8.38).

As a simplification of (8.39) one can write:

$$R_d = R\left(\eta_i \frac{X_{k,i}}{\gamma_{M,i}}\,;\underline{a}_d\right) \qquad i \geq 1 \tag{8.40}$$

With:

$$\gamma_{M,i} = \gamma_{Rd}\ \gamma_{m,i} \tag{8.41}$$

Alternative to the use of (8.39), one can also determine the design value R_d of a material or product resistance by direct assessment of the characteristic value of the strength, without explicit determination of the design values of the different basic variables or:

$$R_d = \frac{R_k}{\gamma_M} \tag{8.42}$$

Equation (8.42) is applicable on elements consisting of one type of material (e.g. a steel profile, pile foundations, etc.) and is also used in relation to *design assisted by testing*.

8.4 Derivation of Partial Factors

8.4.1 General

The standards offer values of partial safety factors for the most common strength and load parameters. The link with the probabilistic level II and III method is found in the definition of the design point. The design point is the point in the failure space with the greatest joint probability density of the strength and the load. It is therefore plausible that for failure the values of the strength and the load are close to the values for the design point. These values are (for normally distributed variables):

$$\begin{aligned} r^* &= \mu_R - \alpha_R \beta \sigma_R = \mu_R (1 - \alpha_R \beta V_R) \\ s^* &= \mu_S - \alpha_S \beta \sigma_S = \mu_S (1 - \alpha_S \beta V_S) \end{aligned} \tag{8.43}$$

As a design criterion it is safe to abide by:

$$r^* > s^* \tag{8.44}$$

Equaling the equations (8.44) and (8.1) results in a number of equations for the partial safety factors:

$$\gamma_m = \frac{R_k}{r^*} \qquad \text{for resistances} \tag{8.45}$$

$$\gamma_E = \frac{e^*}{E_k} \qquad \text{for load effects} \tag{8.46}$$

In which R_k and E_k are the characteristic values of resistance and load effects

In case one considers that the design value can be obtained through a Level II method, i.e. $x_i^* = X_{d,i}$, the partial factor γ_i is introduced in order to establish the relationship with $X_{k,i}$.

A partial factor is larger if:

a) the absolute value of the influence coefficient α is larger;
b) the target reliability index β is higher;
c) the coefficient of variation V_i is larger.

Example 8.1

Suppose that both the strength and the load are normally distributed, with:

$$\begin{aligned} &S:\mu_S = 10 \quad \text{and} \quad V_S = 0.5 \quad \text{so} \quad \sigma_S = 5 \\ &R:V_R = 0.2 \end{aligned} \tag{8.47}$$

The question is to determine the partial safety factor for the strength with a reliability index β = 3.6 and k_R = -1.64 (5% non-exceedance) and what mean value of the strength is needed.

The influence coefficients are (using equation (8.6)):

$$\alpha_R = \frac{0.2\,\mu_R}{\sqrt{0.04\,\mu_R^2 + 25}} \quad \wedge \quad \alpha_S = -\frac{5}{\sqrt{0.04\,\mu_R^2 + 25}} \tag{8.48}$$

Using equation (8.43), The design point is given by:

$$r^* = \mu_R - \frac{0.04\,\mu_R^2}{\sqrt{0.04\,\mu_R^2 + 25}} 3.6 \quad \text{and} \quad s^* = 10 + \frac{25}{\sqrt{0.04\,\mu_R^2 + 25}} 3.6 \tag{8.49}$$

From $r^* - s^* = 0$ follows:

$$(\mu_R - 10) - 3.6 \cdot \sqrt{0.04\,\mu_R^2 + 25} = 0 \tag{8.50}$$

Solving this equation gives: μ_R = 51. The design point value for the strength is r^* = 18.0 and the characteristic value is R_k = 34.2. The partial safety factor for the strength therefore is equal to γ_m = 34.3/18.0 = 1.9. If the spread in the load is greater, the required strength will change. Assume σ_R = 2, then the following must apply:

$$\mu_R - 10 - \sqrt{0.04 - \mu_R^2 + 4} \cdot 3.6 = 0 \quad \rightarrow \quad \mu_R = 39.0 \tag{8.51}$$

In this case, the partial safety factor for the strength comes to:

$$\gamma_m = \frac{26.2}{11.8} = 2.2 \tag{8.52}$$

8.4.2 Partial factors for permanent loads

In this section we derive typical values for the partial factors for permanent loads.
Consider $G_k = \mu_G$, α_E = -0.7 and $\tilde{\alpha}_{E_1}$ = 1.0 so that for the unfavourable effect of G:

$$\gamma_g = \frac{g^*}{G_k} = \frac{\mu_G\,(1 + 0.7 \cdot 3.8 \cdot V_G)}{\mu_G} = 1 + 2.66\,V_G \tag{8.53}$$

In case the permanent load acts as a favourable effect, it can be considered as a resistance variable and hence:

$$\gamma_g = \frac{g^*}{G_k} = 1 - 0.8 \cdot 3.8 \cdot V_G \tag{8.54}$$

In section 10.4.3 we mention that V_G varies between 5% and 10%. These values actually relate to load effects (including model uncertainties) rather than on the loads themselves (without model uncertainties).

8.4.3 Partial factors for variable loads

In this section we derive typical values for the partial factors for Gumbel distributed variable loads.

Using (6.68) and $\alpha_E = -0.7$ and $\beta = 3.8$ yields:

$$q^* = u - \frac{1}{\alpha}\ln\left[-\ln\Phi\,(0.7 \cdot 3.8)\right] = u + \frac{1}{\alpha} \cdot 5.543 \tag{8.55}$$

According to table (2.3) the following relationships for the parameters are obtained:

$$\frac{1}{\alpha} = \frac{\sigma_R}{1.282} \qquad u = \mu - \frac{0.5772}{1.282} \cdot \sigma \tag{8.56}$$

or hence:

$$q^* = \mu - \frac{0.5772}{1.282}\,\sigma + \frac{5.543}{1.282}\,\sigma = \mu + 3.873\,\sigma \tag{8.57}$$

For the partial factor we obtain:

$$\gamma_q = \frac{q^*}{Q_k} = \frac{\mu + 3.873\,\sigma}{\mu - 0.4584\,\sigma} = \frac{1 + 3.873\,V}{1 - 0.4584\,V} \tag{8.58}$$

In case V_{tref} = 0.08 for the extreme value distribution of the extremes in t_{ref} a value γ_Q = 1.36 is obtained. Considering the model uncertainty γ_{Sd} = 1.1 this leads to:

$$\gamma_Q = 1.36 \cdot 1.1 = 1.50 \tag{8.59}$$

8.4.4 Partial factors for concrete

When defining the characteristic strength as the 5-quantile of the strength distribution (as usually considered) and considering α_R = 0.8 it follows from equation (8.45) and (8.46) that in case of a normal distribution:

$$\gamma_c = \gamma_m = \frac{R_k}{r^*} = \frac{1 - 1.645\,V_c}{1 - 0.8 \cdot 3.8 \cdot V_c} \tag{8.60}$$

For a lognormal distribution the following holds according to (6.115) in case $V_c < 0.2$:

$$R_k = \mu_R \exp(-1.645\,V_c) \qquad r^* = \mu_R \exp(-\alpha_R\,\beta\,V_c) \tag{8.61}$$

so that:

$$\gamma_c = \gamma_m = \frac{R_k}{r^*} = \exp(-1.645\,V_c + 0.8 \cdot 3.8 \cdot V_c) = \exp(1.395\,V_c) \tag{8.62}$$

In Table 8.2 the values for γ_c according to (8.60) and (8.62) are mentioned for different values of V_c.

Table 8.2 Partial factors for Concrete

V_c	γ_c (γ_m)	
	N	LN
0.05	1.082	1.072
0.10	1.200	1.150
0.15	1.385	1.233
0.20	1.712	1.322

Further, we consider that the uncertainties due to the conversion from the conventional concrete strength determined using control specimens to the actual concrete strength in the concrete structure are included in V_c. Further, a value γ_{Rd} = 1.1 is considered in order to take into account uncertainties in the resistance model and execution errors. Hence, for V_c = 0.15 one obtains:

$$\gamma_c = \gamma_M = 1.1 \cdot 1.385 = 1.52 \qquad \text{(normal strength distribution)} \tag{8.63}$$

$$\gamma_c = \gamma_M = 1.1 \cdot 1.233 = 1.36 \qquad \text{(lognormal strength distribution)} \tag{8.64}$$

This corresponds to the basic case γ_c = 1.5. Smaller values of V_c correspond to smaller values of γ_c.

8.4.5 Partial factors for reinforcement and pre-stressing steel

A similar derivation as the one considered for (8.60) yields:

$$\gamma_s = \gamma_m = \frac{1 - 1.645\,V_s}{1 - 0.8 \cdot 3.8 \cdot V_s} \tag{8.65}$$

When applying V_s = 0.05 one obtains $\gamma_s = \gamma_m$ = 1.082. When considering further that γ_{Rd} = 1.1 for taking into account execution errors, one obtains $\gamma_s = \gamma_M$ = 1.082 · 1.1 = 1.19. This result can be compared to the commonly applied value γ_s = 1.15.

8.5 Numerical values for the combination coefficient ψ_0

This chapter will introduce approaches for deriving the factor ψ that is used for a safety verification in which several loads are combined.

8.5.1 General

The variation in time of the value of a variable action can in general be modelled by a stochastic process (see also Chapter 7). In Figure 8.3 two load processes $Q_1(t)$ are $Q_2(t)$ illustrated. From this figure it becomes clear that:

$$\max_{t_{ref}}[Q_1(t)+Q_2(t)] < \max_{t_{ref}}[Q_1(t)] + \max_{t_{ref}}[Q_2(t)] \tag{8.66}$$

Because of this reason (and considering two independent variable loads), the alternative load combination $Q_{1k} + \psi_0 Q_{2k}$ is introduced instead of $Q_{1k} + Q_{2k}$.

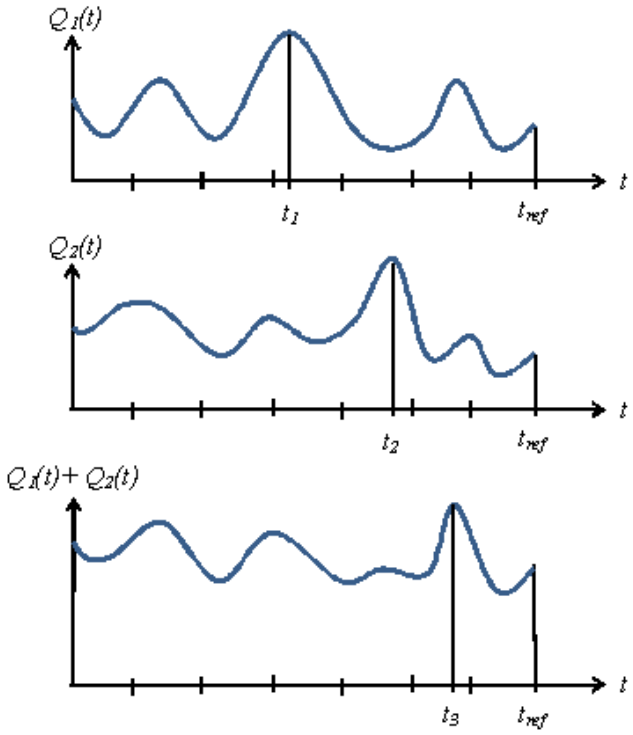


Figure 8.3 Two Load Processes showing the points in time at which the maximum values of the individual loads (Q_1 and Q_2) and combined load (Q_3) occur.

8.5.2 Application: normal distributions

First, we will assume normal distributions for Q_1 and Q_2. Assume that the characteristic values of the two variable loads are defined as the 95-quantiles of a normal distribution. This yields:

$$Q_{1k} = \mu_1 + 1.645\,\sigma_1 \tag{8.67}$$

$$Q_{2k} = \mu_2 + 1.645\,\sigma_2 \tag{8.68}$$

For the characteristic value of the sum $Q_{12} = Q_1 + Q_2$ the following holds:

$$Q_{12k} = (\mu_1 + \mu_2) + 1.645\sqrt{\sigma_1^2 + \sigma_2^2} \tag{8.69}$$

Transforming (8.69) yields:

$$Q_{12k} = (\mu_1 + 1.645\,\sigma_1) + (\mu_2 + 1.645\,\sigma_2) + 1.645\left[\sqrt{\sigma_1^2 + \sigma_2^2} - (\sigma_1 + \sigma_2)\right]$$
$$Q_{12k} = Q_{1k} + Q_{2k} + 1.645\left[\sqrt{\sigma_1^2 + \sigma_2^2} - (\sigma_1 + \sigma_2)\right] \tag{8.70}$$

The value between square brackets is negative from which follows that $Q_{12k} < Q_{1k} + Q_{2k}$ which is consistent with (8.66).

Assume that the subscripts are chosen so that $Q_{1k} > Q_{2k}$. Hence we can write:

$$Q_{12k} = Q_{1k} + \psi_0\, Q_{2k} \qquad \text{with } \psi_0 < 1 \tag{8.71}$$

From the previous it follows that:

$$\psi_0 = \frac{\mu_2 - 1.645\cdot\sigma_1 + 1.645\cdot\sqrt{\sigma_1^2 + \sigma_2^2}}{\mu_2 + 1.645\cdot\sigma_2} = \frac{1 - 1.645\cdot V_1\cdot\xi_{12} + 1.645\cdot\sqrt{V_1^2\,\xi_{12}^2 + V_2^2}}{1 + 1.645\cdot V_2} \tag{8.72}$$

with $\xi_{12} = \mu_1/\mu_2$.

We calculate some numerical values for ψ_0 in case $V_1 = V_2 = V$, with:

- V = 0.1 ; 0.3 ; 0.5 ; 0.7
- ξ_{12} = 1 ; 2
- k_q = 0.8416 ; 1.645 ; 2.576

The three values of k_q relate to the definition of Q_k as the 80, 95 or 99.5-percentile, respectively. The resulting ψ_0 values are summarized in Table 8.3.

Table 8.3 Results Normal Distribution

V	$\xi_{12} = 1$			$\xi_{12} = 2$
	k_q = 0.8416	k_q = 1.645	k_q = 2.576	k_q = 1.645
0.1	0.955	0.917	0.880	0.892
0.3	0.882	0.806	0.744	0.748
0.5	0.827	0.736	0.670	0.655
0.7	0.783	0.686	0.623	0.591

8.5.3 Borges-Castanheta load model (application in EN 1990)

We here apply the FBC model (see section 7.7.2) for the combination of loads. We discuss the case in which the loads are assumed to be Gaussian distributed.

Two loads, normal distributions

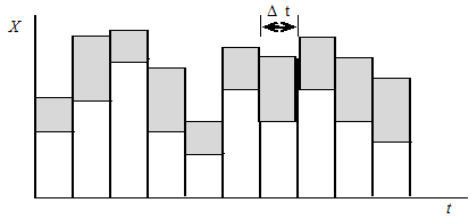


Figure 8.4 Two loads, normal distributions

We consider a combination of two loads S_1 and S_2 with extreme value distributions F_{S1max} and F_{S2max}. Both loads have the same number of load changes in the reference period t_{ref}, so both loads have an equal Δt. For the first load in the combination the FBC model leads to the use of the extreme value distribution for T, we assume m $(=t_{ref} / \Delta t)$ load changes in t_{ref}. For the second load because of the equal Δt the APT (arbitrary point in time or instantaneous) distribution should be used. As an example we take the loads to be normally distributed with mean value of 10 and standard deviation of 2.

The load S has a normal distribution with μ_S = 10 and σ_S = 2. We take α_S = -0.7 and β = 4. The design value is:

$$S^* = \mu_S - \alpha_S \cdot \beta \cdot \sigma_S = 15.2 \tag{8.73}$$

The time independent combination value is:

$$S_{com} = \mu_S - 0.4 \cdot \alpha_S \cdot \beta \cdot \sigma_S = 12.2 \tag{8.74}$$

So for the time independent case we find ψ_o = 12.2/15.2 = 0.8.

For the time dependent case with t_{ref} = 100yr, Δt = 1yr (so m = 100), we use the following derivation. In the 100 year distribution function the probability that the design value is exceeded is:

$$P\{S > S^*\} = \Phi\left(-\alpha_S \cdot ß\right) = 0.0026 \tag{8.75}$$

This means for the exceedance probability of the design value in the APT distribution:

$$P\{S_{apt} > S^*\} = 1 - \left[1 - \Phi\left(-\alpha_S \cdot ß\right)\right]^{1/m} = 0.000026 = \Phi\left(4.05\right) \tag{8.76}$$

The second load in the combination is considered with its APT distribution as mentioned above. In the APT distribution we apply the reduced α-factor for the combination value S_{com}:

$$P\left\{S_{apt} > S_{com}\right\} = \Phi\left(0.4 \cdot 4.05\right) = \Phi\left(1.63\right) = 0.05 \tag{8.77}$$

In the extreme value distribution this means:

$$P\{S > S_{com}\} = 1-(1-0.05)^m = 0.995 = \Phi(-2.5) \tag{8.78}$$

The combination value becomes:

$$S_{com} = \mu_S - 2.5 \cdot \sigma_S = 8.3 \tag{8.79}$$

We find: $\psi_o = 5/15.2 = 0.3$.

Two Gumbel distributions

We consider now the case of two Gumbel distributed loads q. We consider the second load and determine q^* based on the maximum over t_{ref} (with n_2 repetitions, designated in the following as n). The combination value q_{com} is determined based on $n' = n_2/n_1$ repetitions.

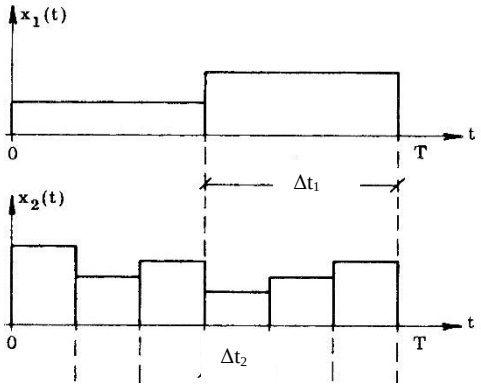


Figure 8.5 FBC model with two loads

From equation (6.117) we obtain the design value:

$$\begin{aligned} q^* &= u_{t_{ref}} - \frac{1}{\alpha}\ln\left[-\ln\Phi(0.7\,\beta)\right] \\ &= \mu_{t_{ref}} - 0.5772 \cdot \frac{\sqrt{6}}{\pi}\sigma - \frac{\sqrt{6}}{\pi}\sigma\ln(-\ln\Phi(0.7\,\beta)) \\ &= \mu_{t_{ref}}\left[1 - \frac{\sqrt{6}}{\pi}V_{t_{ref}}\,(0.5772 + \ln(-\ln\Phi(0.7\,\beta)))\right] \end{aligned} \tag{8.80}$$

with μ_{tref} determined for n elementary time intervals. Consequently, the following holds according to table (2.3):

$$\mu_{t_{ref}} = \mu_1 + \frac{\ln n}{\alpha}$$

We now consider the maximum over a shorter period that consists only of n' intervals ($n' < n$) and we consider the action as non-dominant multiplying the α-factor with 0.4. The corresponding design value q_{com} then becomes:

$$q_{com} = u_{n'} - \frac{1}{\alpha}\ln\left[-\ln\Phi(0.28\,\beta)\right] \tag{8.81}$$

With:

$$u_{n'} = u_1 + \frac{\ln n'}{\alpha} = u_{t_{ref}} - \frac{\ln n}{\alpha} + \frac{\ln n'}{\alpha} = u_{t_{ref}} - \frac{1}{\alpha}\ln\left(\frac{n}{n'}\right) \tag{8.82}$$

so that:

$$\begin{aligned} q_{com} &= \mu_{t_{ref}} - \frac{1}{\alpha}\ln\left(\frac{n}{n'}\right) - \frac{1}{\alpha}\ln\left[-\ln\Phi\left(0.28\,\beta\right)\right] \\ &= \mu_{t_{ref}}\left[1 - \frac{\sqrt{6}}{\pi}V_{t_{ref}}\left(0.5772 + \ln\left(-\ln\Phi\left(0.28\,\beta\right)\right) + \ln\left(\frac{n}{n'}\right)\right)\right] \end{aligned} \tag{8.83}$$

Consequently, the combination value is obtained as $\psi_0 = q_0^*/q^*$. In Table C4 of EN 1990 the following equation is mentioned:

$$\psi_0 = \frac{1 - 0.78\,V\left[0.58 + \ln(-\ln\Phi\left(0.28\,\beta\right)) + \ln N_1\right]}{1 - 0.78\,V\left[0.58 + \ln(-\ln\Phi\left(0.7\,\beta\right))\right]} \tag{8.84}$$

with $N_1 = T/T_1$ (rounded off to integers) in which T is the reference period and T_1 is the largest of the basic periods of the actions which have to be combined. This corresponds to equation (8.83) since:

$$\frac{n}{n'} = \frac{n_2}{n_2/n_1} = n_1 \tag{8.85}$$

References

- KÖNIG G., HOSSER D., The simplified level II method and its application on the derivation of safety elements for level I, CEB Bulletin no. 147, February 1982.

Chapter "Reliability of Systems"

Authors: S.N. Jonkman, A.C.W.M. Vrouwenvelder

This chapter has been compiled using materials from the lecture notes b3 (Vrouwenvelder and Vrijling, 1987) and the CUR book (1997; also version 2015).

Contents

Learning objectives of this chapter:

- To be able to perform a failure probability calculation for series, parallel and combined systems for various situations :(in)dependence and for various correlation values.
- To be able to create a fault or event tree for a (simplified) engineering system.

9 Reliability of systems

9.1 General introduction

A structure or system in civil engineering generally consists of a set of elements. Whereas the previous chapter has focused on the assessment of the reliability of a single limit state, this chapter addresses the reliability analysis for systems that consist of multiple elements.

The consequences of the failure of an element will depend on the type of system that is considered. In some cases, the failure of an element leads to subsequent failure of the whole system (progressive collapse). In other cases and system configurations, there could be no failure if a single element fails and other elements take over the function of the failed element.

In reliability analysis two types of basic systems can be distinguished, i.e. the series and the parallel (see Figure 9.1). Within a series system, failure of a single element will always lead to the failure of the entire system. An example (Figure 9.2- left) would be a bridge (statically determinate) for which failure of one the deck or girder elements leads to failure of the bridge. Within a parallel system, failure of one element can be compensated by the performance of another element. An example could be a pile foundation with many piles. In practice, a complex system can generally be represented by means of a combination of parallel and series subsystems.

The left side of Figure 9.2 shows the representation of the series and parallel system as an input-output system. This is similar to the possible systems configurations of light bulbs and other electronic applications. When one Christmas tree light fails in the series system, the entire system ceases to function. If the lights connected are in parallel, the other lights will remain functioning when one light fails.

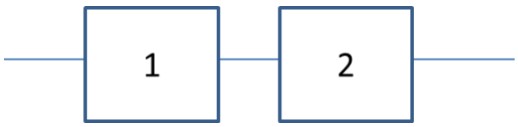
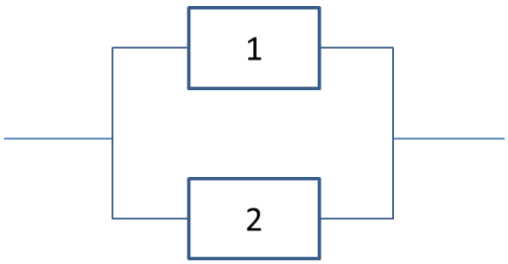
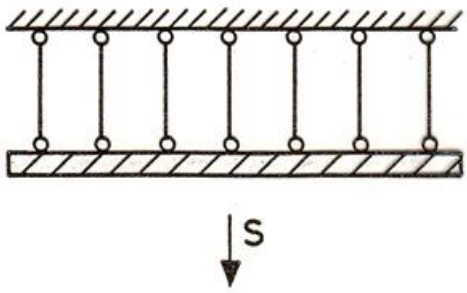
Type	System representation	Simple example (structural engineering)
Series		
Parallel		

Figure 9.1: Series and parallel system: schematic representation (left) and example (right) (Vrouwenvelder and Vrijling, 1987).

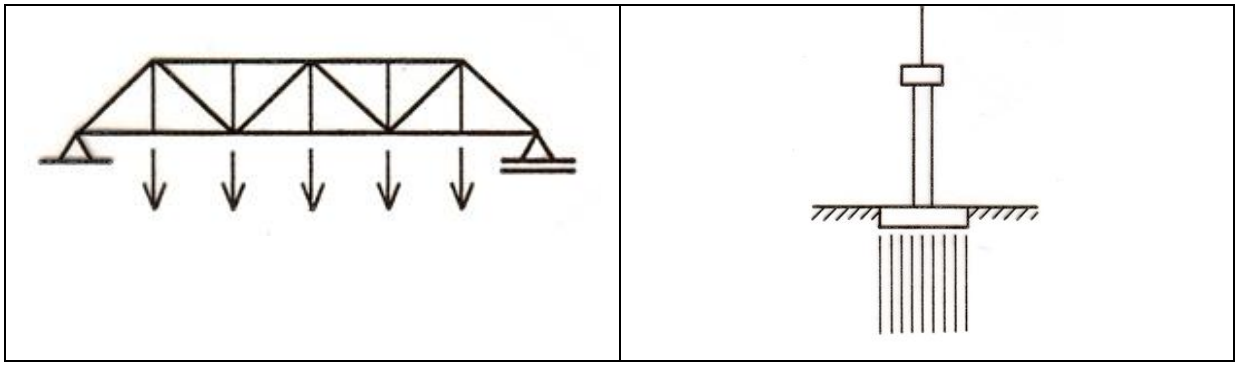


Figure 9.2: Examples of a series and parallel system: a bridge (series system; left); and a pile foundation (parallel system, right)

The remainder of this chapter will now focus on the analysis and quantification of reliability for series (section 9.2) and parallel systems (section 9.3). Techniques for systems analysis, such fault and event trees, will be discussed in chapter 9.4. The relationship between system reliability and (system) design considerations is addressed in chapter 9.5.

9.2 Series systems

9.2.1 General

A simple series system with two elements is considered. Element 1 has a strength of R_i and element 2 a strength R_2. The load on the system is represented by S. For every element a limit state function Z_i can be formulated:

$$\begin{aligned} Z_1 &= R_1 - S \\ Z_2 &= R_2 - S \end{aligned} \tag{9.1}$$

If the average, standard deviations and distributions of R_i and S are known, the reliability indices β_i and corresponding failure probabilities $P(F_i)$ can be determined.

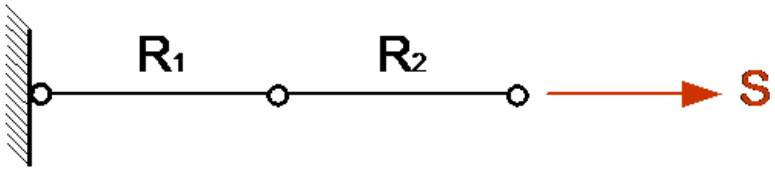


Figure 9.3: A simple series system.

The general formulation is as follows:

$$P(F) = P(F_1 \cup F_2) = P(F_1) + P(F_2) - P(F_1 \cap F_2) \tag{9.2}$$

The upper and lower boundaries of the system failure probability $P(F)$ can be assessed by taking into account the different possible cases of dependence, see Figure 9.4. This gives the lower and upper bound for the system failure probability for the system with two elements:

$$Max(P(F_1), P(F_2)) \le P(F) \le P(F_1) + P(F_2) \tag{9.3}$$

The lower bound is found for a situation with full dependence of both elements. Failure of one element, implies failure of the other and the two events in the Venn diagram overlap. The upper

bound is found if the failures of the elements are mutually exclusive. If the failures are independent ($\rho = 0$) the following solution is found:

$$P(F) = P(F_1) + P(F_2) - P(F_1) \cdot P(F_2) \tag{9.4}$$

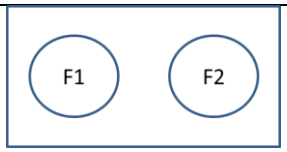
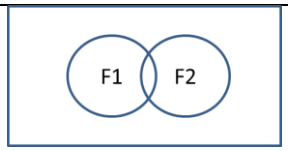
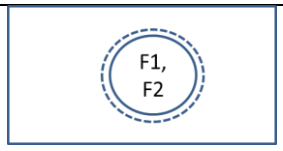
Case	Mutually exclusive	Independent	Dependent
Correlation coefficient. $\rho_{Z1,Z2}$ =	-1	0	1
Venn diagram			
System failure probability $P(F)$	$P(F_1) + P(F_2)$	$P(F_1) + P(F_2) - P(F_1) \cdot P(F_2)$	$Max(P(F_1), P(F_2))$

Figure 9.4 Typical cases and outcomes for three levels of dependence. In sketching the Venn diagrams it is assumed that $P(F_1) = P(F_2)$.

9.2.2 Systems reliability as a function of the correlation coefficient

The previous section presented the failure probability for a simple system consisting of two elements for cases in which the system failure probability can be determined exactly with analytical formulations. This is the case for mutually exclusive, independent and dependent failures, i.e. for $\rho = 0$ and $\rho = \pm 1$. For other cases and values of the correlation coefficient the failure probability of the series system, will be between the lower and upper bound, see Figure 9.5. The relation in this figure is typical for variables whose joint distribution is described by the Gaussian copula. In Appendix 9.2, The failure probability as a function of the correlation coefficient in the case of other copulae is displayed.

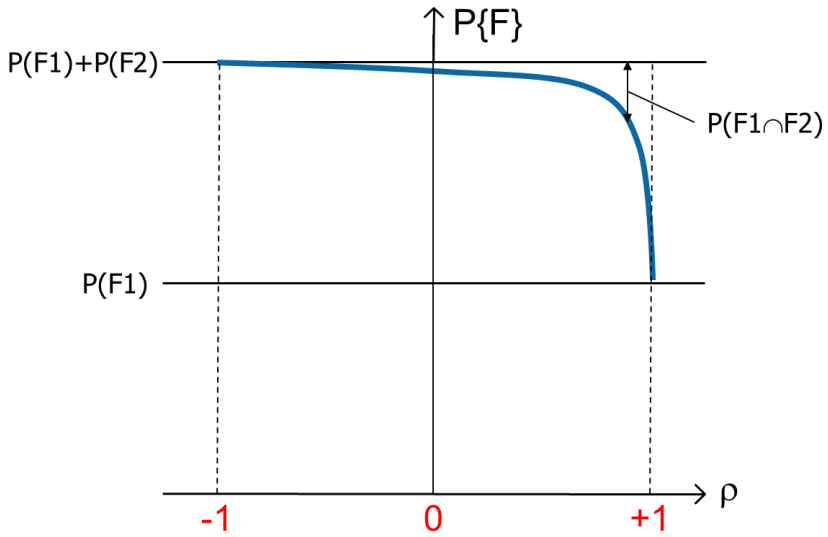


Figure 9.5: Failure probability of a series system of two identical elements as a function of the value of the correlation coefficient for joint distribution of variables described by the Gaussian copula (schematic representation for small failure probabilities).

An exact calculation of the failure probability may be challenging. However, for variables joint by a Gaussian copula, good approximation methods are available, such as the method of Ditlevsen (1977) that is elaborated below. Ditlevsen developed an approximation method which can be used to calculate more narrow boundaries for the system failure probability than the upper and lower bound. This method assumes a known correlation between the failure modes, expressed by a correlation coefficient ρ and normally distributed reliability functions. Ditlevsen deduced that:

$$Max\left(\Phi(-\beta_1)\Phi(-\beta_2{}^*);\Phi(-\beta_1{}^*)\Phi(-\beta_2)\right) \le P(F_1 \cap F_2) \le \Phi(-\beta_1)\Phi(-\beta_2{}^*) + \Phi(-\beta_1{}^*)\Phi(-\beta_2) \tag{9.5}$$

in which:

$$\beta_1 = -\Phi^{-1}(P(F_1)),\ \text{so}\ P(F_1) = \Phi(-\beta_1) \tag{9.6}$$

$$\beta_2 = -\Phi^{-1}(P(F_2)) \tag{9.7}$$

$$\beta_1{}^* = \frac{\beta_1 - \rho\beta_2}{\sqrt{1-\rho^2}} \tag{9.8}$$

$$\beta_2{}^* = \frac{\beta_2 - \rho\beta_1}{\sqrt{1-\rho^2}} \tag{9.9}$$

ρ is the correlation coefficient between Z_1 and Z_2, the limit states corresponding to F_1 and F_2.

Thus, for the probability of failure of the series system $P(F_1)$ with two elements the following applies:

$$\begin{aligned}&\Phi(-\beta_1) + \Phi(-\beta_2) - \Phi(-\beta_1)\Phi(-\beta_2^*) - \Phi(-\beta_2)\Phi(-\beta_1^*) \le P(F) \le \\ &\Phi(-\beta_1) + \Phi(-\beta_2) - Max\left(\Phi(-\beta_1)\Phi(-\beta_2^*);\Phi(-\beta_1^*)\Phi(-\beta_2)\right)\end{aligned} \tag{9.10}$$

For situations with two identical elements $\beta_1{}^* = \beta_2{}^* = \beta^*$ and this simplifies to:

$$\begin{aligned}&2\Phi(-\beta_i) - 2\Phi(-\beta_i)\Phi(-\beta_i^*) \le P(F) \le 2\Phi(-\beta_i) - \Phi(-\beta_i)\Phi(-\beta_i^*) \\ &2\Phi(-\beta_i)\left(1-\Phi(-\beta_i^*)\right) \le P(F) \le \Phi(-\beta_i)\left(2-\Phi(-\beta_i^*)\right)\end{aligned} \tag{9.11}$$

Example 9.1: Ditlevsen bounds for a system with two elements

A given series system consists of two elements with resistances R_1 and R_2, and a load S (see Figure 9.3). Both resistances are independent, and the load and resistance are also independent. For both load and resistance normally distributed variables are assumed with $\mu(R_i) = 10\text{kN}$ and $\sigma(R_i) = 1\text{kN}$, $\mu(S) = 4\text{kN}$ and $\sigma(S) = 1\text{kN}$. For both elements the following applies:

$$\begin{aligned}&\mu(Z_i) = \mu(R_i) - \mu(S) = 6kN \\ &\sigma^2(Z_i) = \sigma^2(R_i) + \sigma^2(S) = 2kN^2 \\ &\beta_i = \frac{\mu(Z_i)}{\sigma(Z_i)} = \frac{6}{1.4} = 4.29\end{aligned} \tag{9.12}$$

It follows that $P(Z_i < 0) = \Phi(-\beta_i) = \Phi(-4.29) = 8.94 \cdot 10^{-6}$.

The lower and upper bound for the system can now be determined:

$$8.94 \cdot 10^{-6} \le P(F) \le 1.79 \cdot 10^{-5} \tag{9.13}$$

To determine the bounds of the system failure probability more accurately the correlation coefficient ρ needs to be determined for the mechanisms Z_1 and Z_2. First, the co-variance is determined (see also section 9.2.2):

$$\mathrm{Cov}(Z_1, Z_2) = \mathrm{Cov}(R_1, R_2) - \mathrm{Cov}(R_1, S) - \mathrm{Cov}(R_2, S) + \mathrm{Cov}(S, S) \tag{9.14}$$

If load and resistance are independent, it follows that $\mathrm{Cov}(R_i, S) = 0$. It is also assumed that both resistance are independent, so $\mathrm{Cov}(R_1, R_2) = 0$. It follows that:

$$\mathrm{Cov}(Z_1, Z_2) = \sigma^2(S) \tag{9.15}$$

Consequently, the correlation coefficient can be determined:

$$\rho(Z_1, Z_2) = \frac{\mathrm{Cov}(Z_1, Z_2)}{\sigma(Z_1)\sigma(Z_2)} = \frac{\sigma^2(S)}{\sigma^2(R_i) + \sigma^2(S)} = \frac{1}{1+1} = 0.5 \tag{9.16}$$

Based on Figure 9.4 it can already be concluded that for this value of the correlation coefficient, the upper bound is a good approximation. The upper bound gives an estimate of the system failure probability of $P(F) = 1.79 \cdot 10^{-5}$.

The Ditlevsen bounds are found according to equation (9.11). To use these formulas the following need to be calculated first:

$$\Phi(-\beta_i) = \Phi(-4.29) = 8.94 \cdot 10^{-6} \tag{9.17}$$

$$\beta^* = \frac{\beta - \rho\beta}{\sqrt{1-\rho^2}} = \frac{4.29 - 0.5 \cdot 4.29}{\sqrt{1-0.5^2}} = \frac{2.15}{\sqrt{0.75}} = 2.48 \tag{9.18}$$

Substitution gives:

$$\begin{aligned} &2\Phi(-\beta)\left(1-\Phi\left(-\beta^*\right)\right) \le P(F) \le \Phi(-\beta)\left(2-\Phi\left(-\beta^*\right)\right) \\ &2\Phi(-4.29)\left(1-\Phi(-2.48)\right) \le P(F) \le \Phi(-4.29)\left(2-\Phi(-2.48)\right) \\ &2 \cdot 8.94 \cdot 10^{-6} \cdot \left(1 - 6.57 \cdot 10^{-3}\right) \le P(F) \le 8.94 \cdot 10^{-6} \cdot \left(2 - 6.57 \cdot 10^{-3}\right) \\ &1.716 \cdot 10^{-5} \le P(F) \le 1.722 \cdot 10^{-5} \end{aligned} \tag{9.19}$$

Example 9.2

The same assumptions as in Example 9.1 are used, only the standard deviations of the load and resistance are changed into: $\sigma(R_i) = 0.45$ kN, $\sigma(S) = 1.34$ kN. $\sigma^2(\mathrm{Z}_i)$ does not change so the

failure probabilities of individual elements do not change. However, the correlation coefficient $\rho(Z_1, Z_2)$ does change:

$$\rho(Z_1, Z_2) = \frac{\sigma^2(S)}{\sigma^2(R_i) + \sigma^2(S)} = \frac{1.8}{1+1} = 0.9 \tag{9.20}$$

The approach is applied for the Ditlevsen bounds for a system with two identical elements. It yields:

$$\beta_i^* = \frac{\beta_i - \rho\beta_i}{\sqrt{1-\rho^2}} = \frac{4.2 - 0.9 \cdot 4.2}{\sqrt{1-0.9^2}} = 0.96 \tag{9.21}$$

The corresponding failure probability equals $\Phi(-\beta_i^*) = 0.17$. For the system failure probability the following is found using equation (9.11):

$$\begin{aligned} &2\Phi(-\beta_i) - 2\Phi(-\beta_i)\Phi(-\beta_i^*) \le P(F) \le 2\Phi(-\beta_i) - \Phi(-\beta_i)\Phi(-\beta_i^*) \\ &2.6 \cdot 10^{-5}(1 - 0.17) \le P(F) \le 1.3 \cdot 10^{-5}(2 - 0.17) \\ &2.2 \cdot 10^{-5} \le P(F) \le 2.4 \cdot 10^{-5} \end{aligned} \tag{9.22}$$

Note that the upper bound in the above example is fairly close to fundamental upper bound ($2.6 \cdot 10^{-5}$). Apparently, the correlation values need to be very high to have a significant effect on the system failure probability.

9.2.3 Series system with multiple elements

In this section we consider the general case of a series system with multiple elements (see Figure 9.6). For a series system with n elements, the failure space is defined by:

$$R_1 < S_1 \cup R_2 < S_2 \cup R_3 < S_3 \cup \ldots \cup R_n < S_n \tag{9.23}$$

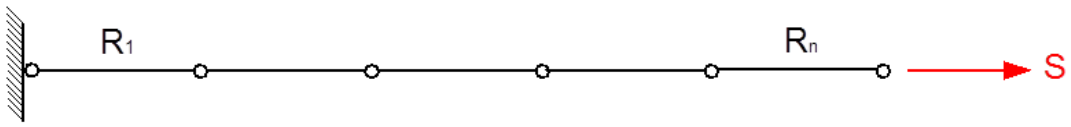


Figure 9.6: Series system with n elements

Following the principles from section 9.2.1 the fundamental boundaries of the probability of failure of the series system with n elements is given by:

$$\max\left(P\left(R_i < S_i\right)\right) \le P_f \le \sum_{i=1}^{n} P\left(R_i < S_i\right) \tag{9.24}$$

The lower bound is valid for the fully dependent case and the upper bound for a mutually exclusive case. For a series system with n independent elements the failure probability can be determined as follows. The failure probability of an element can be written as $P(F_i) = p_i$. It can also be formulated as the complement of the reliability, (i.e. probability of non-failure) $P(F_i) = 1 - P(\overline{F_i})$. Consequently, the system failure probability $P(F)$ for the case of independent elements can be determined as follows:

$$P(F) = 1 - P(\overline{F_1})P(\overline{F_2}).....P(\overline{F_n}) = 1 - \prod_{i=1}^{n}(1 - p_i) \tag{9.25}$$

If the failure probabilities of all n elements are identical ($p_i = p$) this simplifies to $P(F) = 1 - (1-p)^n$. For small values of the probability p approximations are available. If $np << 1$ the following approximation is often applied for the series system (this is similar to the upper bound formulation):

$$P(F) = np \tag{9.26}$$

If $p << 1$ and $n >> 1$ the following approximation can be used.

$$P(F) = 1 - e^{(-np)} \tag{9.27}$$

The following example shows the differences between the various approaches.

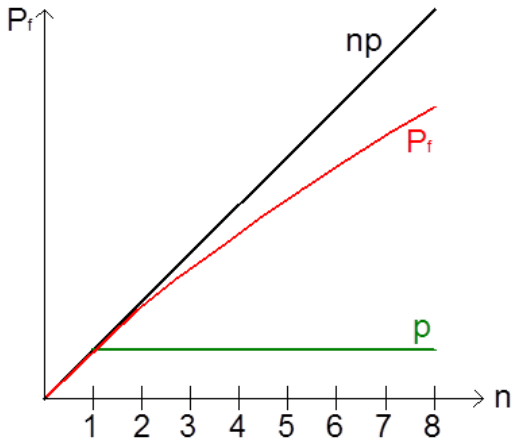
Example 9.3

A series system with $n = 10$ elements is considered. Two cases are considered. In case 1, a relatively large value of the probability of an element is consider, i.e. $p = 0.1$ and in the second case a smaller probability, $p = 0.01$. Table 9.1 below shows the outcomes for the various cases. For the small failure probability (the second case) the two approximations are close to the exact solution. For the first case with the larger p value there are significant differences.

Table 9.1: System failure probabilities for a series system with n=10 elements for varying values of the element failure probability and various approximations.

Case	p	Exact solution: $1-(1-p)^n$	Approximation 1: np	Approximation 2: $1-e^{(-np)}$
1	0.1	0.65	1	0.63
2	0.01	0.0956	0.1	0.0952

Figure 9.7 below summarizes the outcomes and bounds for the system failure probability for the various cases. The lower bound is found for the case of full dependence. The upper bound is found in case of mutually exclusive failures. The outcomes for the independent case will be close to the upper bound in case of small failure probabilities.



$$p \le P\{F\} \le np$$

$$\max p_i \le P\{F\} \le \sum p_i$$

Figure 9.7: Failure probability of the series system for various cases, lower (green) and upper bound (black) and independence (red).

Bounds for a series system with more than two elements

As can be seen from Figure 9.7, the fundamental upper and lower boundaries are often rather wide and provide only information for specific boundary cases (mutually exclusive and dependent). It is therefore useful to have the Ditlevsen boundaries. This has been introduced in section 9.2.2 for a system with $n=2$ elements. It is relatively easy to extend this approach to 3 or more elements. Figure 9.8 shows a Venn diagram for a system with three elements.

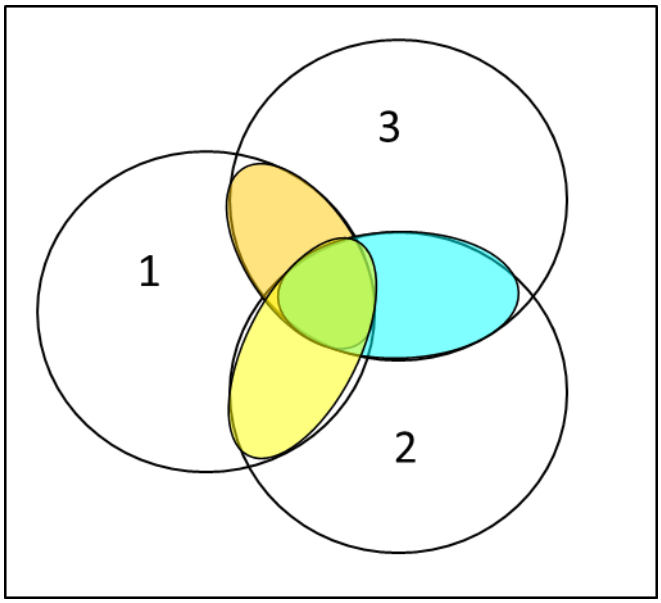


Figure 9.8: Venn diagram for a system with 3 elements

We know that for a system with two elements, the exact failure probability can be determined as follows:

$$P(F) = P(F_1) + P(F_2) - P(F_1 \cap F_2) \tag{9.28}$$

This is used to derive an upper bound for a 3 element system. If a third element is added, the total failure probability increases with $\left(F_3 \cap \overline{F_1} \cap \overline{F_2}\right)$. However, to approximate the exact value we

have two options. In a lower bound approximation we subtract all the additional "overlapping" areas in the Venn diagram, i.e.:

Lower bound:

$$P(F) = P(F_1) + P(F_2) - P(F_1 \cap F_2) + P(F_3) - P(F_1 \cap F_3) - P(F_2 \cap F_3) \quad (9.28)$$

In this case we may subtract a bit too much and underestimate the system failure probability. For the upper bound approximation we subtract the largest of the two additional overlapping areas.

Upper bound:

$$P(F) = P(F_1) + P(F_2) - P(F_1 \cap F_2) + P(F_3) - Max\{ P(F_1 \cap F_3), P(F_2 \cap F_3) \} \quad (9.28)$$

Ditlevsen has used this approach to formulate more general bounds for the same problem with n elements:

$$\begin{aligned} P(F) &\le \sum_i P_i - \sum_{i \ge 2} \max_{j<i} P_{ij,LOW} \\ P(F) &\ge \sum_i \max(P_i - \sum_{j<i} P_{ij,UP}, 0) \end{aligned} \quad (9.29)$$

In which in accordance to equation (9.5):

$$\begin{aligned} P_i &= P\{Z_i < 0\} = \Phi(-\beta_i) \\ P_{ij} &= P\{Z_i < 0 \text{ and } Z_j < 0\} \le \Phi(-\beta_i)\Phi(-\beta_j^*) + \Phi(-\beta_i^*)\Phi(-\beta_j) = P_{ij,UP} \\ P_{ij} &= P\{Z_i < 0 \text{ and } Z_j < 0\} \ge \max(\Phi(-\beta_i)\Phi(-\beta_j^*); \Phi(-\beta_i^*)\Phi(-\beta_j)) = P_{ij,LOW} \end{aligned} \quad (9.30)$$

The application of this approach is demonstrated by means of an example:

Example 9.4

A series system with n elements is given, according to Figure 9.6. The load on all elements is exactly the same and the load is assumed normally distributed. The strength of the elements is statistical and normally distributed. The strength of the various elements is assumed to be correlated.

An equal correlation between the strength of all elements is assumed with a correlation coefficient $\rho = 0.7$. The parameters of the distribution of the strength and the load are:

$$\begin{aligned} &\mu_{R_i} = 280 \quad \sigma_{R_i} = 20 \quad \rho_{R_i} = 0.7 \\ &\mu_{S_i} = 160 \quad \sigma_{S_i} = 20 \quad \rho_{S_i} = 1.0 \end{aligned} \quad (9.30)$$

One is asked to determine the lower and the upper boundaries according to Ditlevsen as well as the elementary lower and upper boundaries for the probability of failure for a system of 2 to 10 elements.

The failure of a single element follows from:

$$\begin{aligned}
&\mu_{Z_i} = \mu_{R_i} - \mu_{S_i} = 280 - 160 = 120 \\
&\sigma_{Z_i} = \sqrt{\sigma_{R_i}^2 + \sigma_{S_i}^2} = 28.28 \\
&\beta_i = \frac{\mu_{Z_i}}{\sigma_{Z_i}} = 4.24 \\
&P_{f_i} = \Phi(-\beta_i) = 1.1 \cdot 10^{-5}
\end{aligned} \tag{9.31}$$

The elementary boundaries are $1.1 \cdot 10^{-5} \le P_f \le n \cdot 1.1 \cdot 10^{-5}$

To calculate the Ditlevsen-boundaries the correlation coefficient between Z_i and Z_j must be determined first. The general expression for the covariance is:

$$\mathrm{Cov}(Z_i Z_j) = \mathrm{Cov}(R_i R_j) - \mathrm{Cov}(R_i S_i) - \mathrm{Cov}(R_j S_j) + \mathrm{Cov}(S_i S_j) \tag{9.32}$$

Because the load and the strength are independent the following now applies:

$$\begin{aligned}
&\mathrm{Cov}(Z_i Z_j) = \mathrm{Cov}(R_i R_j) + \mathrm{Cov}(S_i S_j) \\
&\mathrm{Cov}(Z_i Z_j) = \rho_{R_{ij}} \sigma_{R_i} \sigma_{R_j} + \rho_{S_{ij}} \sigma_{S_i} \sigma_{S_j} \\
&\mathrm{Cov}(Z_i Z_j) = 0.7 \cdot 20^2 + 1.0 \cdot 20^2 = 680
\end{aligned} \tag{9.33}$$

The correlation coefficient $\rho_{Z_i Z_j}$ then follows from:

$$\rho_{Z_i Z_j} = \frac{\mathrm{Cov}(Z_i Z_j)}{\sigma_{Z_i} \sigma_{Z_j}} = \frac{680}{(28.28)^2} = 0.85 \tag{9.34}$$

As the correlation coefficient is now known the Ditlevsen boundaries can be calculated. According to Ditlevsen, if $\beta_i = \beta_j = \beta$ then:

$$P_{f_{ij}} = P(Z_i < 0 \cap Z_j < 0) \le 2\,\Phi(-\beta)\,\Phi(-\beta^*) \tag{9.35}$$

with:

$$\begin{aligned}
&\beta^* = \frac{\beta - \rho_{Z_i Z_j} \beta}{\sqrt{1 - \rho_{Z_i Z_j}^2}} = \frac{4.24 - 0.85 \cdot 4.24}{\sqrt{1 - 0.85^2}} = 1.21 \\
&1.25 \cdot 10^{-6} = \Phi(-\beta)\,\Phi(-\beta^*) \le P_{ij} \le 2\,\Phi(-\beta)\,\Phi(-\beta^*) = 2.5 \cdot 10^{-6}
\end{aligned} \tag{9.36}$$

With these values for $P_{f_{ij}}$ the lower and upper boundaries according to Ditlevsen are calculated in Table 9.2, using the lower bound $P_{ij_{LOW}} = 1.25 \cdot 10^{-6}$ from (9.36) for the upper bound of the system and the upper bound $P_{ij_{UP}} = 2.5 \cdot 10^{-6}$ from (9.36) for the calculation of the lower bound of the system with equations (9.29). For n=2 the calculation corresponds to the elaboration presented earlier in section 9.2.2. For example, for n=3, for the columns A, B and C:

$A = P_i + P_i + P_i = 3 \cdot 1.1 \cdot 10^{-5} = 3.3 \cdot 10^{-5}$,

$B = \min(P_{ij,UP}, P_i) + \min\left(2 \cdot P_{ij,UP}, P_i\right) = 2.5 \cdot 10^{-6} + 2 \cdot 2.5 \cdot 10^{-6} = 7.5 \cdot 10^{-6}$,

$C = P_{ij,LOW} + P_{ij,LOW} = 2 \cdot 1.25 \cdot 10^{-5} = 2.5 \cdot 10^{-6}$.

Table 9.2: Lower and upper boundaries of the probability of failure according to Ditlevsen.

N	$A=\sum_{i=1}^{n} P_i$	$B=\sum_{i=2}^{n}\min(\sum_{j=1}^{i-1} P_{ij,UP}, P_i)$	$C=\sum_{i=2}^{n} \max_{j<i} P_{ij,LOW}$	**lower limit =A-B**	**upper limit =A-C**
1	$1.1\cdot10^{-5}$	0	0	$1.1\cdot10^{-5}$	$1.1\cdot10^{-5}$
2	$2.2\cdot10^{-5}$	$2.5\cdot10^{-6}$	$1.3\cdot10^{-6}$	$2.0\cdot10^{-5}$	$2.1\cdot10^{-5}$
3	$3.3\cdot10^{-5}$	$7.5\cdot10^{-6}$	$2.5\cdot10^{-6}$	$2.6\cdot10^{-5}$	$3.0\cdot10^{-5}$
4	$4.4\cdot10^{-5}$	$1.5\cdot10^{-5}$	$3.8\cdot10^{-6}$	$2.9\cdot10^{-5}$	$4.0\cdot10^{-5}$
5	$5.5\cdot10^{-5}$	$2.5\cdot10^{-5}$	$5.0\cdot10^{-6}$	$3.0\cdot10^{-5}$	$5.0\cdot10^{-5}$
6	$6.6\cdot10^{-5}$	$3.6\cdot10^{-5}$	$6.3\cdot10^{-6}$	$3.0\cdot10^{-5}$	$6.0\cdot10^{-5}$
7	$7.7\cdot10^{-5}$	$4.7\cdot10^{-5}$	$7.5\cdot10^{-6}$	$3.0\cdot10^{-5}$	$6.9\cdot10^{-5}$
8	$8.8\cdot10^{-5}$	$5.8\cdot10^{-5}$	$8.8\cdot10^{-6}$	$3.0\cdot10^{-5}$	$7.9\cdot10^{-5}$
9	$9.9\cdot10^{-5}$	$6.9\cdot10^{-5}$	$1.0\cdot10^{-5}$	$3.0\cdot10^{-5}$	$8.9\cdot10^{-5}$
10	$1.1\cdot10^{-4}$	$8.0\cdot10^{-5}$	$1.3\cdot10^{-5}$	$3.0\cdot10^{-5}$	$1.0\cdot10^{-4}$

Figure 9.9 shows the exact probability of failure, the elementary lower and upper boundaries and the narrower boundaries according to Ditlevsen.

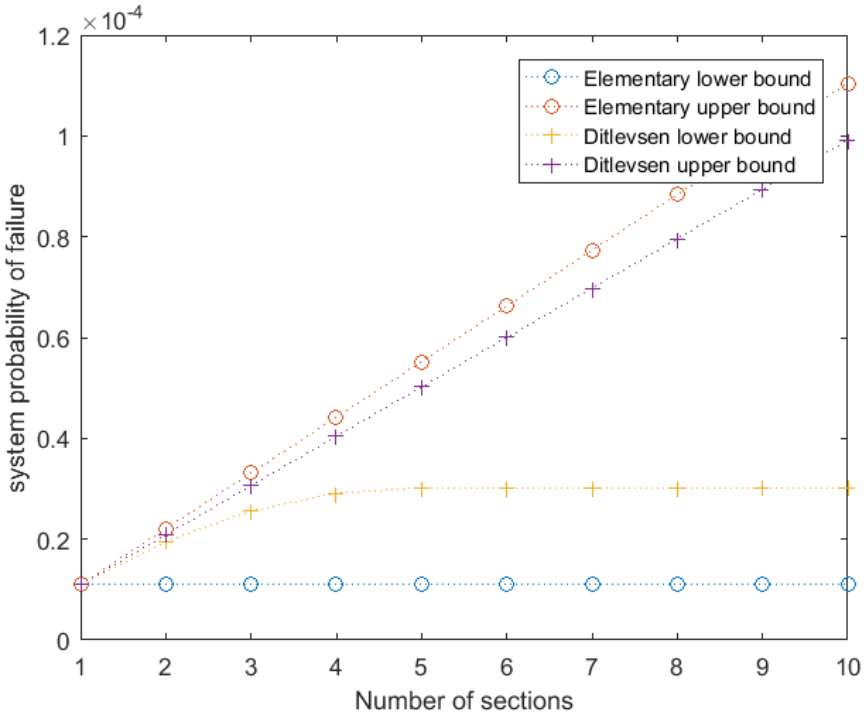


Figure 9.9: Probability of failure of the series system of Example 9.4.

An even better estimate of the probability of failure is attained with the upper and lower boundaries approximation according to Hohenbichler and Rackwitz (1983). This approximation uses a transformation of non-normally distributed statistically dependent basic variables to standard normally distributed independent basic variables. This method is less simple and requires a greater calculation capacity than the Ditlevsen method. Finally, of course also numerical integration and Monte Carlo analysis may be used, but that is generally even more time

consuming.

9.2.4 Continuous series systems *(not part of the examination)*

Besides the discrete series systems with n elements, continuous systems also exist, e.g. a dike body or a concrete bar. In every point, the strength of the continuous series system can be described as a random variable with an expected value and a standard deviation.

The strength in two different points will usually be correlated. The degree of correlation depends, amongst other factors, on the distance Δx between the two points considered. The relation between the correlation and the distance is described with the so-called correlation function. A commonly used expression for the correlation function is:

$$\rho\left(R_x, R_{x+\Delta x}\right) = \exp\left(-\left(\frac{\Delta x}{d}\right)^2\right) \tag{9.37}$$

in which:

- $\rho(R_x, R_{x+\Delta x})$ is the correlation between the strengths at locations x and $x + \Delta x$;
- d is the so-called correlation distance, which depends on the problem.

The figure below shows the observed values of some parameter (e.g. cone resistance in a CPT) over a distance x. In case A the correlation distance is larger than for situations B. This means that in situation A there will be a larger correlation between two observations at two sites than for situation B. The more variation there is in the observations over the distance, the smaller the predictive value of an observation for another location.

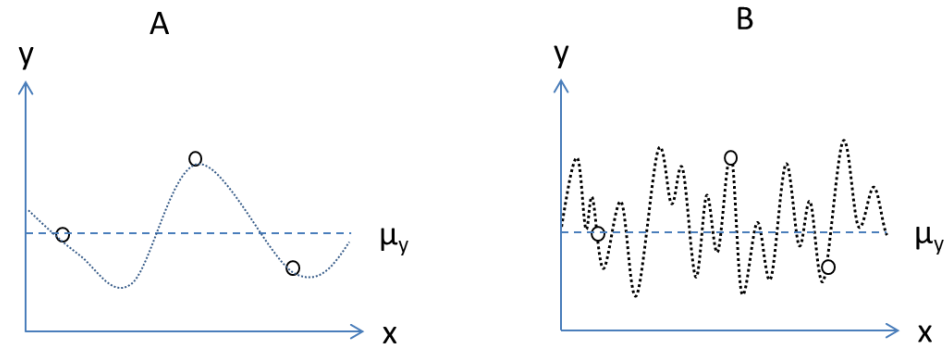


Figure 9.10: Two observation signals over distance x

If every location has the same reliability index β, (and assuming that the spatial correlation of the load is very high) the probability of failure of the system can be approximated by:

$$P_f = \Phi(-\beta)\left(1 + \alpha_R \frac{\beta L}{\sqrt{\pi}\, d}\right) \tag{9.38}$$

in which:

- $\alpha_R = \frac{\sigma_R}{\sqrt{\sigma_R^2 + \sigma_S^2}}$ (follows from the level II calculation);
- L is the length of the system.

This formulation provides a simple way to determine length effects. An example of an application concerns the fluctuation of soil properties for dikes. Large fluctuations may lead to a larger probability of weak soil layers and a higher probability of failure due to geotechnical mechanisms.

9.3 Parallel systems

In general, a parallel system is characterized by the fact that elements can compensate for each other: failure of one element does not automatically lead to failure of the entire system.
Thereby, the parallel system is the opposite of the series system. A series system fails if one of the elements fails. The pure parallel system functions if one of the elements functions. In more popular terms it is often said that the strength of a series system is determined by its weakest link, while a parallel system is as strong as its strongest link.
The pure parallel system fails if all elements of the system fail. The failure space is defined by:

$$F_1 \cap F_2 \cap ... \cap F_n \tag{9.39}$$

in which F_i is the failure of element i $(i=1,2, ...,n)$. The probability of failure of the system is:

$$P_f = P(F_1)\, P(F_2 \mid F_1)\, P(F_3 \mid F_1, F_2) ... P(F_n \mid F_1 ... F_{n-1}) \tag{9.40}$$

If the events F_i are statistically independent:

$$P_f = P(F_1)\, P(F_2) ... P(F_n) \tag{9.41}$$

The bounds for a system with multiple identical elements with each failure probability P_i are as follows:

$$0 \le P_f \le Min(P_i) \tag{9.42}$$

The lower bound is found for a situation in which failures are mutually exclusive. The upper bound is valid for a case in which the failure are fully dependent. These characteristic situations can also be recognized in the Venn diagrams introduced in Figure 9.4. Failure of the parallel system occurs in the surface where the two events overlap.

Two examples of parallel systems are given in Figure 9.11 and Figure 9.12. Another example (not shown) of a system that could function as a parallel system is a foundation with many piles that jointly support a foundation plate.
The first example (Figure 9.11) concerns a city protected by two dikes. If the first dike fails, no immediate flooding of the city occurs as it is still protected by the second dike. The strength of the system is determined by the strongest of the two dikes.
The failures are dependent when both dikes have the same design and elevation and the load on the second line of defence would be the same as on the first line after failure. In that case the upper bound from equation (9.30) can be applied. In case of dependence of elements, the parallel configuration does not lead to an increase of reliability.

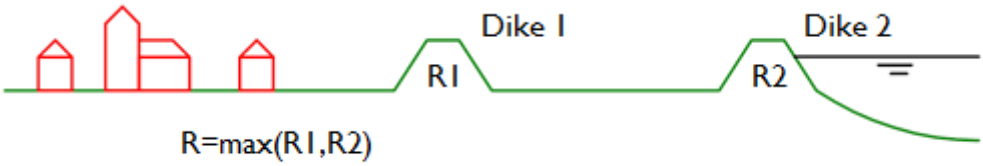


Figure 9.11: Example of a parallel system

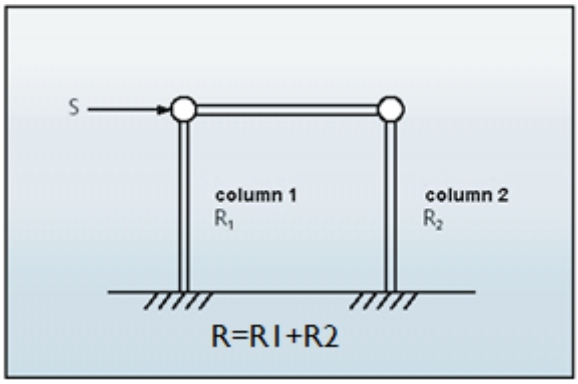


Figure 9.12: Portal structure.

The other case shows a different behaviour. Figure 9.12 shows a portal structure that consists of two columns with a beam on top. The beam is loaded by a horizontal load S. The resistance against collapse of the portal structure is made up of the combined strength of the columns.

The strength of the system against the horizontal force is determined by the sum of the strength of the two columns (note that the system strength is larger than for the previous case with the dikes, since both columns add to the strength):

$$R = R_1 + R_2 \tag{9.43}$$

in which:

- R_1, R_2 is the maximum reaction force, which can be exerted by columns 1 and column 2, respectively;
- R is the strength of the system.

However, the validity of this formulation of the strength greatly depends on the behaviour of the elements after reaching the maximum reaction force. The resistances of the elements can indeed be added if the columns show ductile behaviour. However, if columns show brittle behaviour the total strength will be smaller than the sum of the resistances of the two columns. Figure 9.13 shows the difference between ductile and brittle material behaviour. The ductile material can sustain significant plastic deformation before failure, but this is not the case for the brittle material.

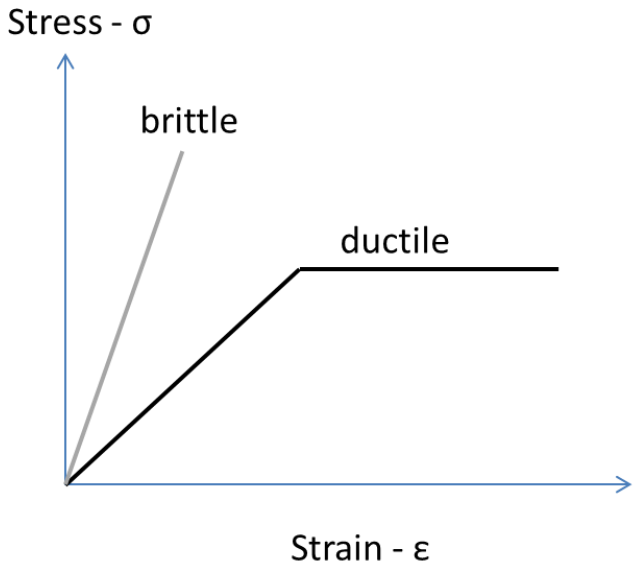


Figure 9.13: Stress-strain diagrams for ductile and brittle material behaviour.

We continue the case assuming ductile behaviour, so that the resistances can be added. It follows that:

$$\begin{aligned}\mu(R) &= \mu(R_1) + \mu(R_2) = 2\mu \\ \sigma^2(R) &= \sigma^2(R_1) + \sigma^2(R_2) + 2\rho\sigma(R_1)\sigma(R_1)\end{aligned} \tag{9.44}$$

The general expression for the variation coefficient for two elements becomes:

$$V(R) = \frac{\sigma(R)}{\mu(R)} = \frac{\sigma(R_i)\sqrt{2+2\rho}}{2\mu(R_i)} = V(R_i)\sqrt{\tfrac{1}{2}(1+\rho)} \tag{9.45}$$

If the two elements are independent ($\rho = 0$) and when both standard deviations are equal to σ it follows that:

$$\sigma^2(R) = \sigma^2(R_1) + \sigma^2(R_2) = 2\sigma^2 \tag{9.46}$$

The variation coefficient can be expressed as follows:

$$V(R) = \sigma(R) \big/ \mu(R) = \sigma\sqrt{2} \big/ 2\mu = V \big/ \sqrt{2} \tag{9.47}$$

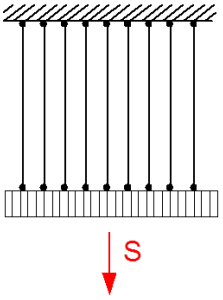
If the two elements are dependent ($\rho = 1$) it follows that:

$$\begin{aligned}\sigma^2(R) &= \sigma^2(R_1) + \sigma^2(R_2) + 2\rho\sigma(R_1)\sigma(R_1) = 4\sigma^2 \\ V(R) &= \sigma(R) \big/ \mu(R) = 2\sigma \big/ 2\mu = V\end{aligned} \tag{9.48}$$

For the more general case of the parallel system, it can be shown that:

$$\mu(R) = n\mu(R_i) \text{ and } \sigma^2(R) = \sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma(R_i)\sigma(R_j) = \sigma^2(R_i)\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij} \tag{9.49}$$

With f an equation which differs with dependency. For cases with multiple elements, this is summarized in the table below:



n elements R_i

$$\mu(R_i) = \mu$$
$$\sigma(R_i) = \sigma$$
$$V(R_i) = V$$

Figure 9.14: Parallel system with multiple elements.

Table 9.3: Outcomes for the parallel system with n elements for different levels of dependence.

	Independent	Dependent
R_{sys}	$\sum R_i$	nR_i
$\mu(R_{sys})$	$n\mu$	$n\mu$
$\sigma^2(R_{sys})$	$n\sigma^2$	$n^2\sigma^2$
$V(R_{sys})$	$V/\sqrt{n}$	V

Finally, as an illustration, some numerical results are shown to highlight the effect of correlation values and the number of elements in the parallel system - see Figure 9.15. This shows that the system failure probability rapidly decreases with the number of elements, but fades out for 3 to 4 elements. Also the value of the correlation coefficient has an important influence. The lowest system failure probability is found for cases with a low value of the correlation coefficient ρ.

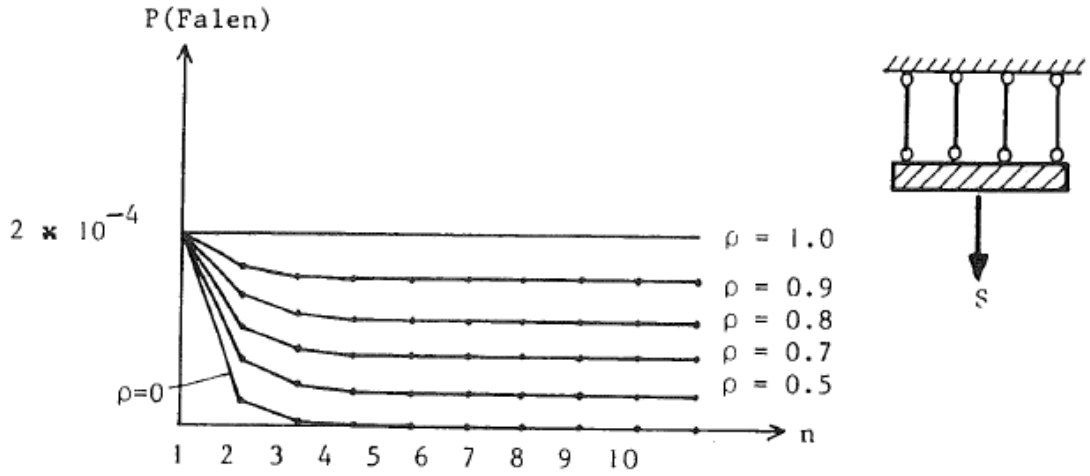


Figure 9.15: Failure probability for a parallel system with n elements with identical correlations with $\mu(R) = 2\mu(S)$ and $V(R_i)=0.1$, $V(S)=0.2$.

9.4 Approaches to analyse systems using logic trees

This section describes logic tree techniques to analyse the undesired events that can lead to system failure, and the corresponding failure probabilities. The following sections summarize a number of these techniques. After the system definition (9.4.1), undesired events are identified (section 9.4.2). Consequently, the corresponding failure probabilities can be determined using techniques such as event trees and failure trees (9.4.3 – 9.4.5) and appendix 9.1. Eventually, this information can be used to quantify and evaluate the risk – see section 3 for more information.

The presented techniques aim to provide insight in mechanisms of system failure and the associated failure probabilities. Techniques such as fault tree and event trees allow the combination of technical failures and human and organization failures in one approach. Insights of the risk analysis can be used to optimize system design and management, as insight is provided in the most important mechanisms and so-called "weak links". The outcomes of the process can thus serve as a tool for communication and management.

9.4.1 Systems analysis

The analysis of a system starts with a description of the functions of the system and the components of the system.

The process or system under consideration can usually be described as a so-called input-output element (see Figure 9.16). Here the system is assumed to be failing if no output takes place. Usually a system is divided into elements and subsystems, which can all be schematised as an input-output elements. By means of the internal relations the components and subsystems together form a configuration that is representative of the total system. The subdivision in components and subsystems goes up to a level where the probabilities of failure can be determined.

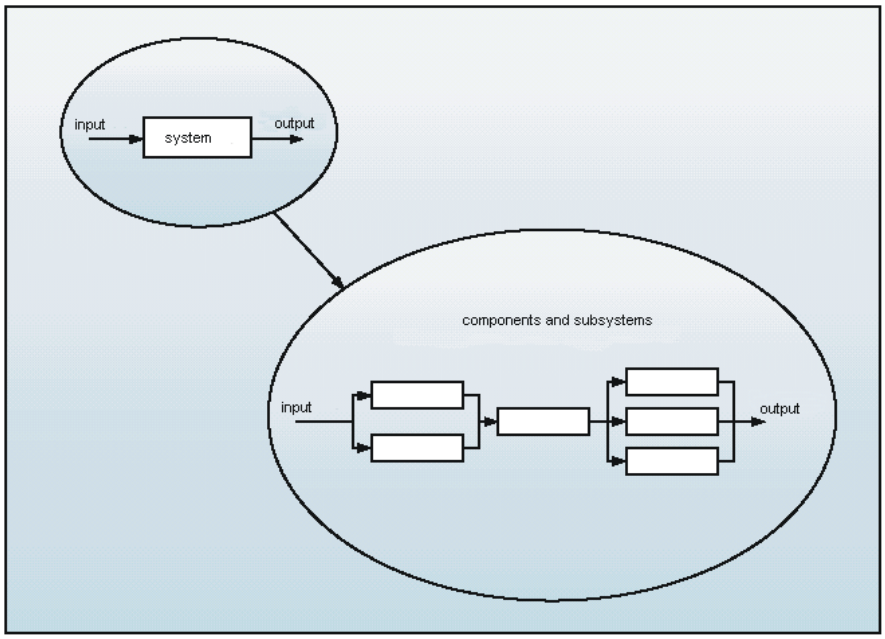
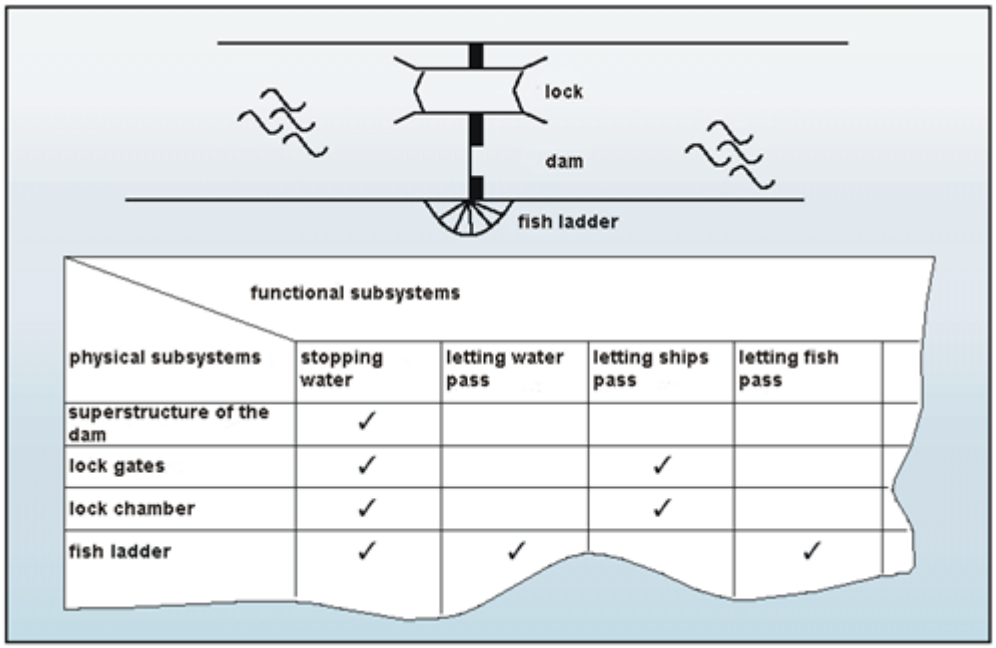


Figure 9.16: input-output system, components and subsystems.

In order to analyse systems failure, not only physical components, but also organizations and operators and users need to be considered (see e.g. Bea, 1998). This means that the reliability analysis can include failures of engineering elements (e.g. gates of a storm surge barrier that fail

due to overloading) and human and organizational elements (e.g. failure of closure of the barrier due to failure of the closure operation by the operators).
A system can have one or multiple functions, see a simplified example below for a lock complex (Figure 9.17). Different physical subsystems can be necessary for various functions. By drawing a matrix, in which these relationships are indicated, the functional requirements for subsystems can be determined.



physical subsystems	stopping water	letting water pass	letting ships pass	letting fish pass
superstructure of the dam	✓			
lock gates	✓		✓	
lock chamber	✓		✓	
fish ladder	✓	✓		✓

Figure 9.17: Relationships between functional and physical subsystems of a lock complex.

As part of the risk analysis, undesired events are identified that lead to failure of (one of the functions of) the system. Different subsystems can have a common basic event that leads to failure. In such a case common cause failure is involved. An example of this is an earthquake, which can lead to the collapse of various subsystems – see also section 9.5.

9.4.2 Failure modes and effects analysis (FMEA)

Several qualitative techniques are available to analyse which undesired events can lead to system failure. The Failure modes and effects analysis (FMEA) is a qualitative approach which aims to systematically identify the failure modes of the components and subsystems, and the associated consequences. The FMEA is usually carried out in the design phase of a system or process, with the objective of identifying the components and subsystems that require improvements to achieve a sufficiently reliable system. The general goal of the FMEA is to offer an overview that is as complete as possible of all the foreseeable unwanted events and consequences in a system or process, so that decisions concerning whether or not to undertake certain actions can be well-founded. Figure 9.18 gives a general flowchart for the FMEA.

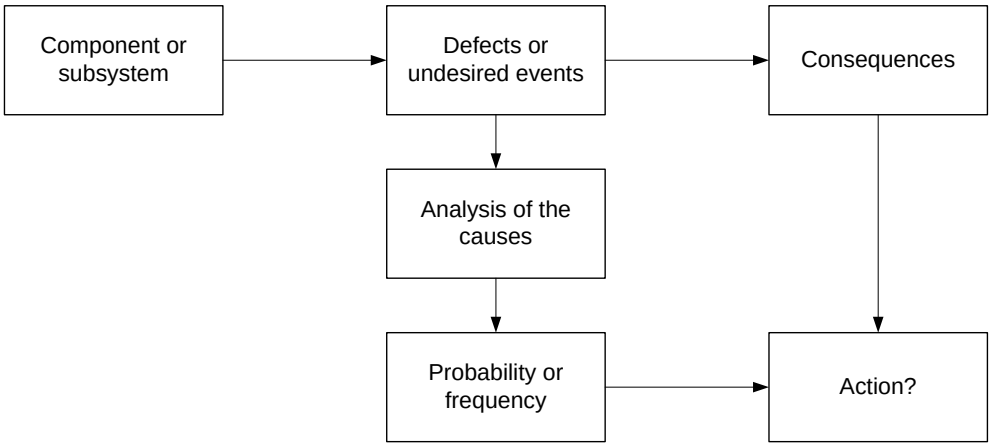


Figure 9.18: Flowchart for the FMEA

The FMEA can be executed bottom-up by starting at the level of the components and subsystems and by subsequently expanding the analysis to system level. The scale of the analysis greatly depends on the lowest level (component or subsystem). To reduce the often enormous scale, the availability of data concerning failure of components and subsystems can determine the lowest level, for which it is reasonable to start the analysis. The disadvantage of the bottom-up approach is that components and subsystems are analysed that may be of lesser importance for the functioning of the total system or process .

Another way to carry out the FMEA is the so-called top-down approach. This entails carrying out the analysis in two or more phases. The first phase is dividing the system into function blocks. A function block is a set of components or subsystems which together have one function. For every function block an inventory can be made of the possible undesired events and the resulting consequences, e.g. based on experiences with similar subsystems. Per function block the importance for the functioning of the total system can be identified. Depending on this, the decision is made whether or not to further analyse a function block in a following phase. With the top-down method a lot of time and effort can be saved, but failure modes can be overlooked.

Recording the results of an FMEA can be done in a structured way by means of a table. An example of an FMEA table for a system that supplies water to a production facility is shown in Table 9.4.

Table 9.4: Example of a simplified FMEA table for a water supply for a production facility.

Element: Water pipe Function: Supply of water			
Defect	Possible cause	Consequence	Action
No water	• Pump does not work • Pipe is broken • Valve does not work	• Stagnation of the production process	• Installing a second pump • Constructing a backup pipe
Too much water	• Pump doesn't turn off • Valve does not work	• Water problems • Waste of water	• Safety system for pump operation • Constructing a double valve

FMECA: Failure modes, effects and criticality analysis

An FMEA becomes an FMECA by adding a so-called criticality matrix. In this matrix the different failure modes and consequences are related to each other and the consequences are classified according to their severity. Furthermore, the frequencies of the different failure modes are estimated. This way a ranking is determined, which highlights the most frequent failure modes and the gravest consequences.

The data included in a criticality matrix differ per system. Possible factors are:

- λ_i is the frequency of failure mode I when in operation;
- t_i is the time span during which the system or subsystem is operational;
- P_{si} is the conditional likelihood with which the failure mode leads to the final effect;
- S_i is the damage factor, which expresses the gravity of the consequence.

With the help of these data the so-called criticality factor can be determined, (neglecting starting problems):

$$C_i = \lambda_i \, t_i \, P_{si} \, S_i \tag{9.50}$$

The magnitude of C_i determines the order of the failure modes. This order is subsequently decisive in the choice of the actions to be undertaken. The costs of an action to reduce the frequency of a failure mode or to limit the consequence can also constitute a part of the criticality matrix.

9.4.3 Event tree

The event tree is an aid in the analysis of the response of a system to one event. In a logical manner, the event tree relates this one "initial event" to all possible consequences, by making an inventory and an analysis of all the possible events that can follow the initial event. Figure 9.19 presents examples of the event trees for a series and a parallel system with three light bulbs.

For event trees it is common to indicate “failure” in the bottom branch and “functioning” in the top branch after a node. In the event trees below the pure series and parallel systems can be recognised instantly. The series system fails if one of the components fails (see section 9.2), the parallel system fails if all three elements fail (see section 9.3). Many systems are less simple, because they consist of combinations of series and parallel subsystems. An example of an event tree for a combination of such systems is given in Figure 9.30. With the help of the event tree a set of events can be defined that leads to failure or functioning. In Figure 9.20 the probabilities of "light" and "no light" are determined for the series system and the parallel system of Figure 9.19. The different combinations of events can also be recognized in the corresponding Venn diagram of Figure 9.8. The resulting probabilities of failure for the series and parallel system can be determined with the bounds derived.

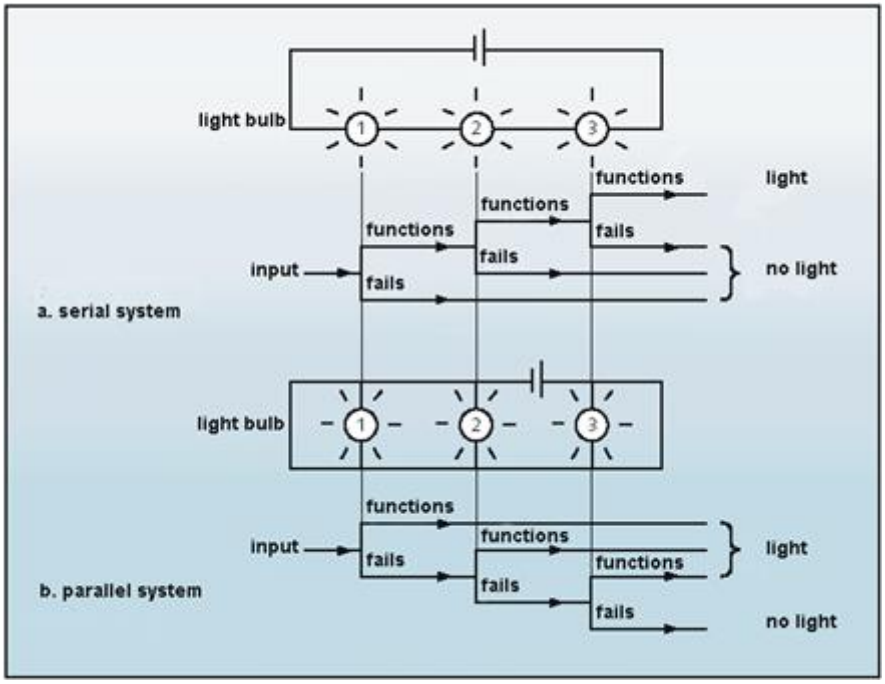


Figure 9.19: Event trees for a series system and a parallel system.

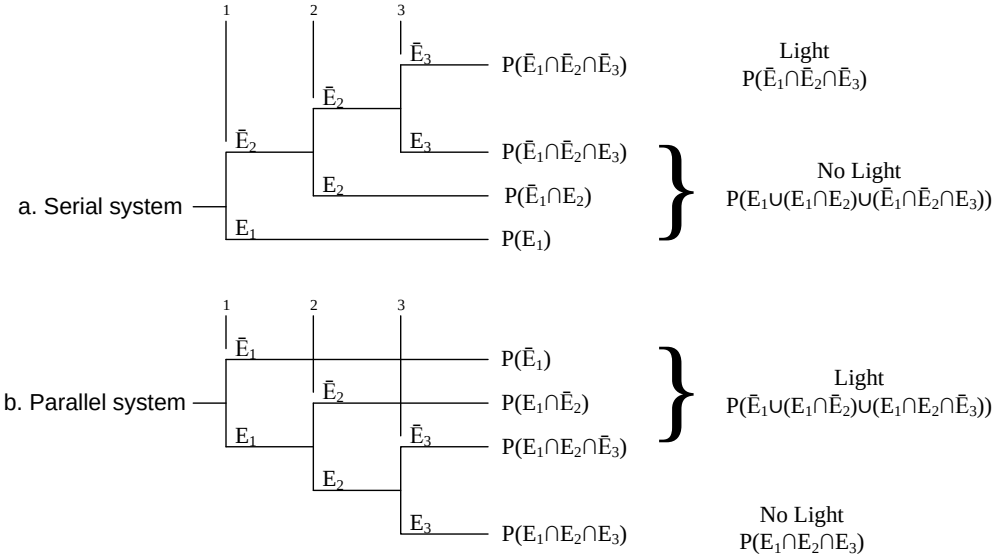


Figure 9.20: Probabilities of 'light' and 'no light' for a series system and a parallel system. Note that $P(E_i)$ denotes the probability of failure of element i, whereas $P(\bar{E}_i)$ denotes the probability of non-failure (i.e. survival) of element 1.

As a simple example from civil engineering, one can consider a dike with a revetment which is subject to wave attack during a storm. Failure, i.e. a complete breach in the dike, will occur if the revetment fails and if the wave attack lasts sufficiently long to lead to erosion of the earthen dike structure. Figure 9.21 below shows the simple fault tree only taking the failure mechanism of revetment failure into account. Note that there will be some correlation between both events as both are affected by the same phenomenon (i.e. wave attack).

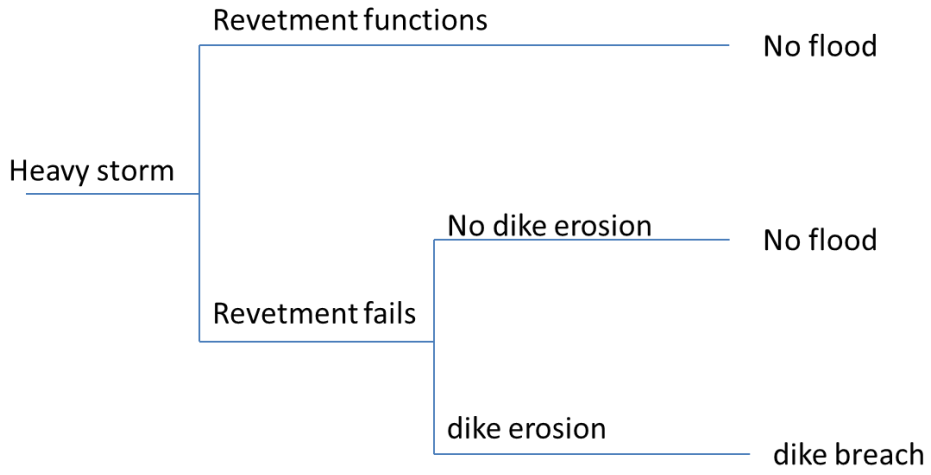


Figure 9.21: Simple event tree for a revetment

In making an event tree it is important not to include too many details to avoid a very large tree. The subsystems can best be analysed separately to prevent cluttering the event tree.
Since the branches of an event tree shows different types of combinations of failure events (and thus different scenarios) for which the consequences can be different, it can also be used to show the outcomes on the right side. Figure 9.22 shows a very simplified example for a flood prone area with a dike system for flood protection and a rescue operations system to prevent fatalities in case of flooding. The dike system protects the economic assets, whereas fatalities can be averted due to rescue operations. For more complex systems, different levels of economic damages and other outcomes can occur for different combinations. In that case the event tree can be used to create the probability mass (or density) function of consequences, as it shows the different scenarios with their probabilities and consequences.

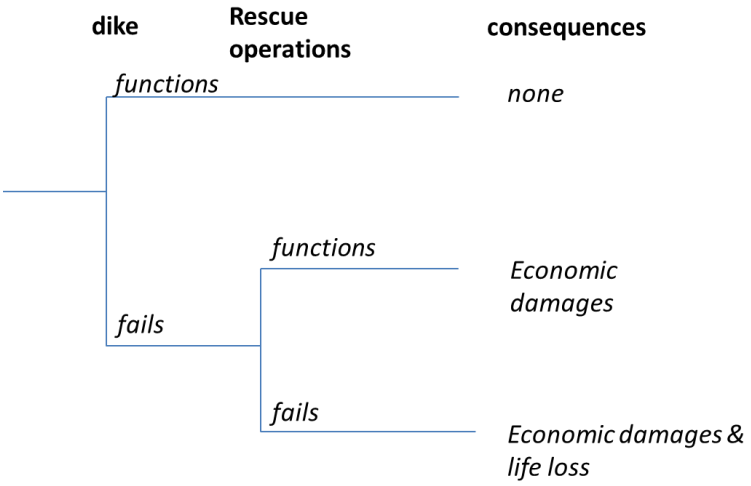


Figure 9.22: Event tree with consequences of failures, for a combined dike – rescue operations system.

9.4.4 Fault tree

The fault tree gives a logical succession of all events that lead to one undesired "top event" at the top of the tree. Fault trees were developed in the 1960's for applications to defence and aviation.

In the 1970's fault trees were implemented and further developed in the nuclear industry. After that, other fields, such as the chemical industry and civil engineering (e.g. for the design of the Eastern Scheldt barrier) started to use fault trees. Nowadays, techniques such as fault trees and event trees are used in all engineering fields.
Figure 9.23 gives the fault trees for the series and parallel system introduced in Figure 9.19. The events E_1, E_2 and E_3 are base events. E_1 is the event: "light bulb 1 fails". The node above the base events shows the condition that has to be met for the overall top-event located to occur. This condition is called a gate. Figure 9.23 gives two of these conditions: the "and-gate" – used for parallel systems - and the "or-gate" – for series systems. For the and-gate all underlying events have to take place for the top event to occur. For the or-gate at least one of the underlying events has to take place to pass the gate in the fault tree.

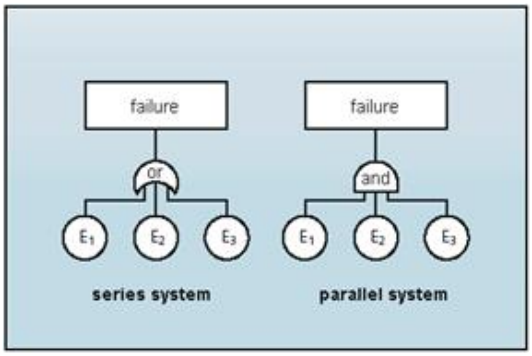


Figure 9.23: Fault trees for a series system and a parallel system both consisting of three elements.

The failure probabilities of both systems can be determined with the bounds determined for the series and parallel system in earlier sections of this chapter. As a recap, the main bounds are summarized in Figure 9.24.

For the series system of n elements the total failure probability is found as follows

$$P(F) = P(E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n) \tag{9.51}$$

For the parallel system the following is found:

$$P(F) = P(E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n) \tag{9.52}$$

$P_{f,system}$ (with n components):

system	gate	operator	components		
			mutually exclusive	independent	fully dependent
series	OR	$\cup$	$\sum_{i=1}^{n} P_i$ (upper bound)	$1-\prod_{i=1}^{n}(1-P_i)$	$\max\{P_i\}$ (lower bound)
parallel	AND	$\cap$	0 (lower bound)	$\prod_{i=1}^{n} P_i$	$\min\{P_i\}$ (upper bound)

Figure 9.24: Summary of values for the system failure probability for various cases (lecture by T. Schweckendiek).

Other types of gates

Beside the "or"-gate and the "and"-gate there are other variants of these gates:

- the "voting gate";
- the "inhibit gate";
- the "priority and gate";
- the "exclusive or gate".

The "**voting gate**" requires a minimum number of underlying events to occur to pass the gate in the fault tree. An example of this is a power plant with a system of three generators, of which at least two must function to be able to supply enough power. When two generators fail, the entire system fails. The fault tree for this example is shown in Figure 9.25.

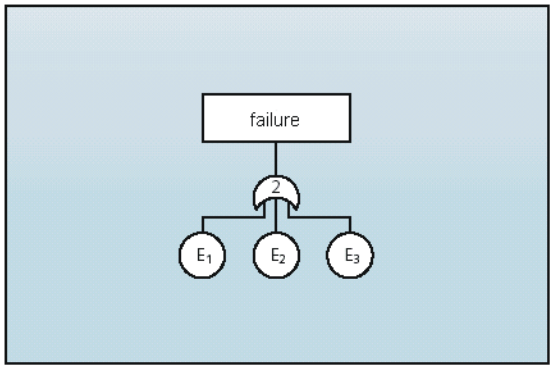


Figure 9.25: Fault tree with voting gate. In this case failure occurs if at least two elements fail.

If the failure of the elements is statistically independent and the probability of failure of an element is equal for all elements, the Binomial distribution can be used. The following is valid:

$$P_f = \sum_{k=m}^{n} \frac{n!}{k!(n-k)!} (P_{f_e})^k (1-P_{f_e})^{n-k} \tag{9.53}$$

in which:

- P_{fe} is the probability of failure of an element;
- n is the total number of elements;
- m is the number of elements that have to fail to let the system fail.

If the failure probabilities of the elements are not equal, the failure probability has to be calculated for every combination that leads to failure. The probability of failure is then the sum of the probabilities of the combinations. In the case of the voting gate in Figure 9.25 at least two elements need to fail. The combinations $(E_1 \cap E_2 \cap \bar{E}_3)$, $(E_1 \cap \bar{E}_2 \cap E_3)$, $(\bar{E}_1 \cap E_2 \cap E_3)$ and $(E_1 \cap E_2 \cap E_3)$ lead to system failure. If the events are independent, the probability of failure is:

$$\begin{aligned} P_f = {} & P(E_1)\,P(E_2)\,(1-P(E_3)) + P(E_1)\,P(E_3)\,(1-P(E_2)) \\ & + P(E_2)\,P(E_3)\,(1-P(E_1)) + P(E_1)\,P(E_2)\,P(E_3) \end{aligned} \tag{9.54}$$

The "**inhibit gate**" is a replacement of the and-gate for cases where one of the underlying events is a conditional event. Figure 9.26 shows an example of this. The event next to the "inhibit gate" is a conditional event, which can only occur if the initial event has taken place.

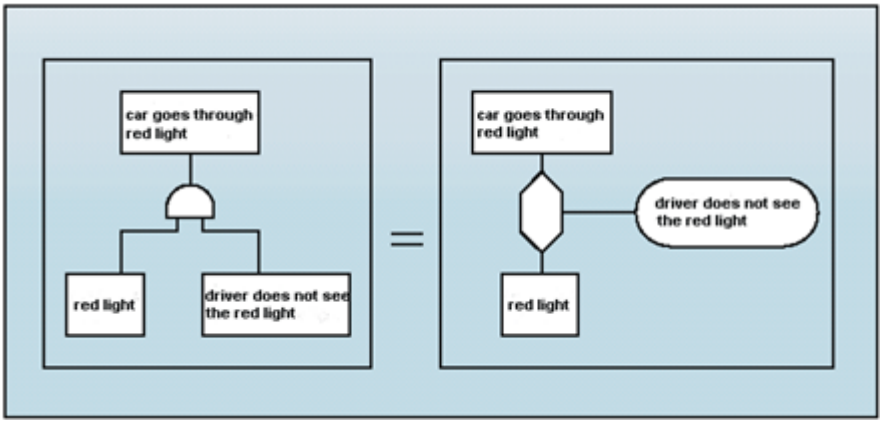


Figure 9.26: Fault tree with an "inhibit gate".

The "**priority and gate**" is an enhancement of the normal "and"-gate. For this gate all underlying events must take place in a given order, from left to right, to lead to the top event. An example is given in Figure 9.27. In this case consequent failure of the guardian, sleeper and dreamer will lead to flooding of the village.

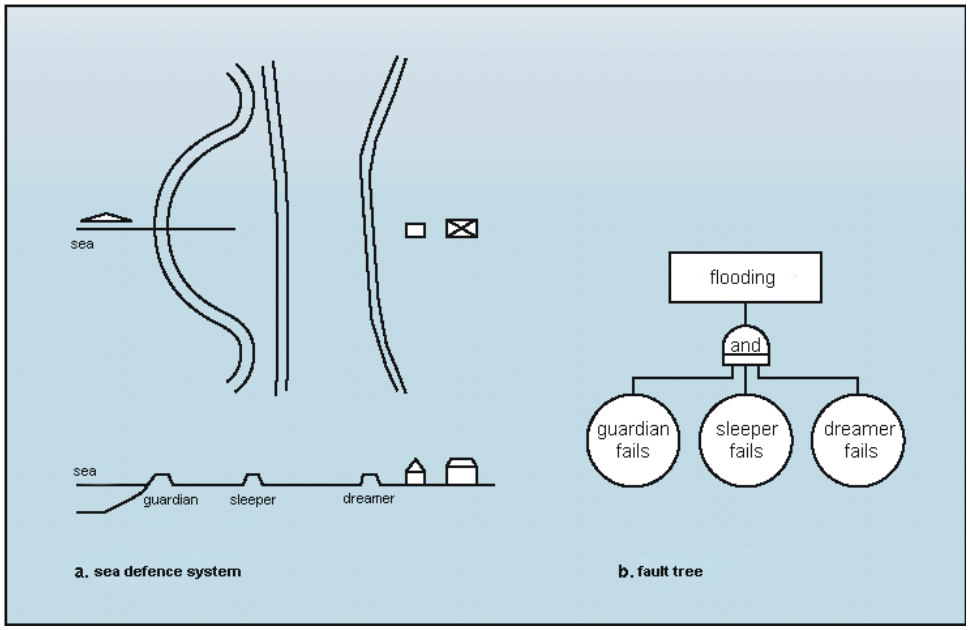


Figure 9.27: Fault tree of a sea defence system.

The probability of failure of the system is:

$$P_f = P(E_1)\,P(E_2|E_1)\,P(E_3|E_1 \cap E_2) \tag{9.55}$$

in which:

- E_1 is the guardian fails;
- E_2 is the sleeper fails;
- E_3 is the dreamer fails.

Finally, there is the "**exclusive or gate**" as an enhancement of the "or"- gate. In the case of an "exclusive or gate" the top event (only) takes place if exactly one of the underlying events occurs.

Consider a packing system as an example of such a system. A packing system can be split into a system that supplies the product to be packed, a system that supplies the packing material.
Suppose that failure of the system is defined as: "The system delivers unpacked goods or empty packing material". In that case not delivering a product is not considered the top event, because no direct consequential damage occurs. The failure space of the system is determined by:

$$(E_1 \cup E_2) - (E_1 \cap E_2) \tag{9.56}$$

in which:

- E_1 is no supply of the product;
- E_2 is no supply of the packing material.

The probability at failure can be determined simply with level III methods. These methods can simply include the statistical dependence of the events E_1 and E_2 in the simulation or integration procedures.
Table 9.5 gives a short overview of the different gates and the way they are drawn in a fault tree.

symbol	meaning
	"and"- gate
	"or"- gate
	voting gate
	inhibit gate
	exclusive or gate
	priority and gate

Table 9.5: Symbols for the various gates in a fault tree.

Beside the different gates, a lot of different events also appear in the fault tree. Table 9.6 gives seven symbols that indicate the different events.

symbol	meaning
	base event
	event that is not developed further
	compound event I (consequence)
	compound event II (consequence)
	conditional event (used for inhibit gate)
	house event
	reference symbol

Table 9.6: Events used in a fault tree.

The base event is an initial event, which is situated at the base of the fault tree. Such an event usually concerns the failure of a system component.
For the event that is not developed further, no thorough analysis of the underlying base events has been carried out. When quantifying the probability of failure of the total system, the probability of the event that is not developed further is sometimes neglected.
The compound event I is a consequential event that occurs if the condition defined in the underlying gate is met. The compound event II is a variant of the base event, for which the event occurs if the numerical condition shown in the box is met. This gate is useful for defining the failure space. Figure 9.28 gives an example. Theoretically, the compound event is a combination of a gate and an event.

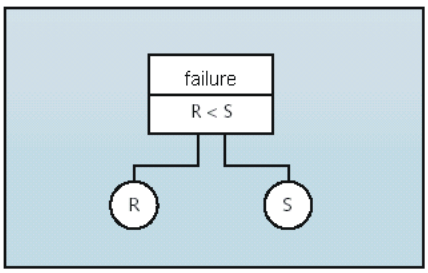


Figure 9.28: Compound event II.

The conditional event only occurs in combination with the "inhibit gate" and can only occur if the initial event under the gate has taken place. The house event is an event, which always takes place. Even if a fault tree is drawn up under the supposition that a certain, non-house, event has taken place, this event is often indicated with the symbol for a house event. The reference symbol is used in large fault trees to split the tree into sections and to refer to sections presented elsewhere. With the given formulae for the various gates the probabilities of the compound events can be calculated and combined to find the probability of the final top event.

Closing remark

Fault trees are very useful for an efficient analysis of the reliability of a system. This is also demonstrated by their wide application in many fields. The fault tree provides insight in the contribution of various failure modes and mechanisms to system failure probability. For complex systems, fault trees can grow to very large sizes. A disadvantage of a fault tree is that the analysis focuses on a single "top event" and only on the failure probability. When multiple failure scenarios with different consequences can occur, the analyst has to make multiple fault trees, or define the top event in terms of range of consequences. The event tree provide some more opportunities to quantify different scenarios with varying consequences in one analysis. Also, a cause consequence chart can be used to incorporate different undesired events with varying consequences. Since this technique is less widely used in civil engineering, it is treated in appendix 9.1.

Example for a combined system

The previous sections have shown how event and fault trees can be composed for pure series and parallel systems. However, in practice, a system will consist of different components and subsystems with both parallel and series configurations. As an example Figure 9.29 shows a system that consists of five elements. A number of combined events (I…IV) has been identified for the analysis. For this system both an event tree (Figure 9.30) and a fault tree (Figure 9.31) have been made. This shows that both the event and fault tree can be used to represent system failure.

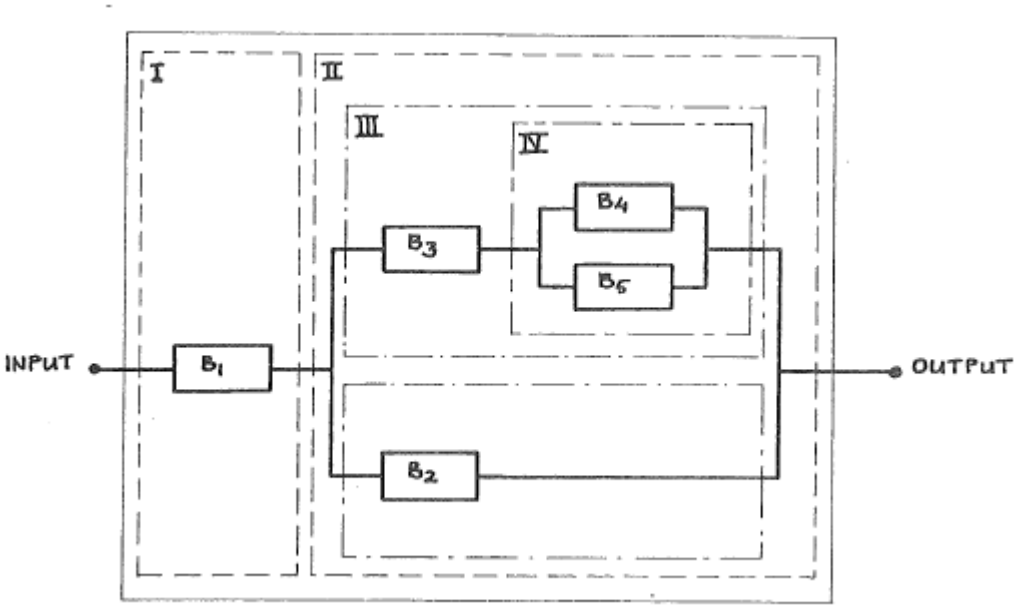


Figure 9.29: Example of a combined system consisting of series and parallel subsystems

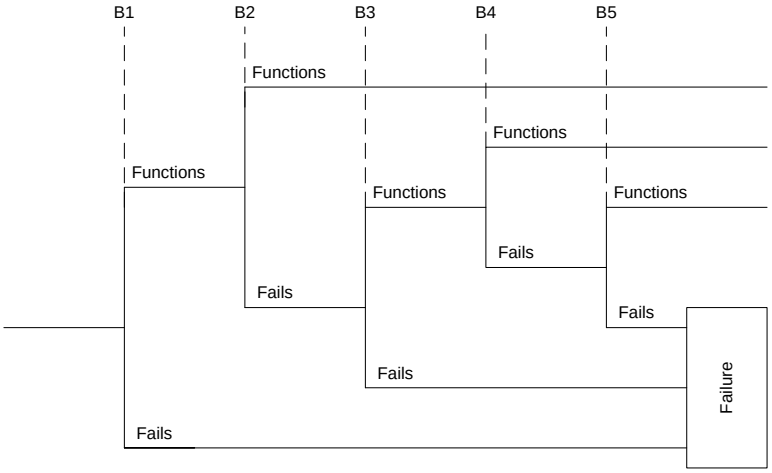


Figure 9.30: Event tree for the example of the combined system.

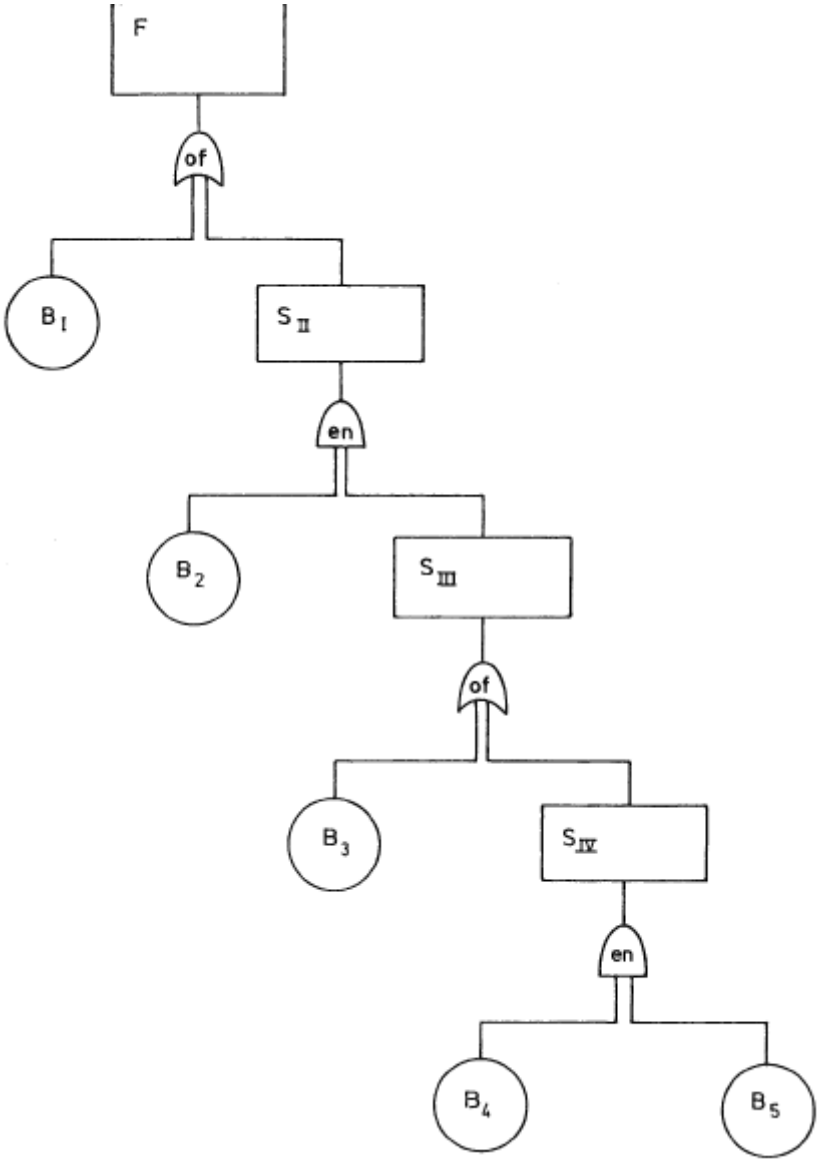


Figure 9.31: Fault tree for the example of the combined system.

The probability of system failure can be formulated. In this case, it is most convenient to use the fault tree and to start at the top. The combined events (I…IV) are used to assess the probability of system failure $P(F)$.

$$\begin{aligned}
&P(F) = P(B_1 \text{ or } S_{II}) \\
&P(F) = P(B_1 \text{ or } (B_2 \text{ and } S_{III})) \\
&P(F) = P(B_1 \text{ or } (B_2 \text{ and } S_{III})) \\
&P(F) = P(B_1 \text{ or } (B_2 \text{ and } (B_3 \text{ or } S_{IV}))) \\
&P(F) = P(B_1 \text{ or } (B_2 \text{ and } (B_3 \text{ or } (B_4 \text{ and } B_5))))
\end{aligned} \tag{9.57}$$

Consequently, the system failure probability can be determined – assuming independence of the events. This is done in a bottom up approach.

$$\begin{aligned}
&P(S_{IV}) = P(B_4) \cdot P(B_5) \\
&P(S_{III}) = 1 - (1 - P(B_3)) \cdot (1 - P(S_{IV})) \\
&P(S_{II}) = P(B_2) \cdot P(S_{III}) \\
&P(F) = 1 - (1 - P(B_1)) \cdot (1 - P(S_{II}))
\end{aligned} \tag{9.58}$$

The example is also used to introduce the concept of **minimal cut sets**. These are the unique combinations of component failures that can cause system failure[1]. Within the example, the following cut sets can be observed (see also Figure 9.32): $(B_1); (B_2, B_3); (B_2, B_4, B_5)$. Any combination of these failures will result in system failure.

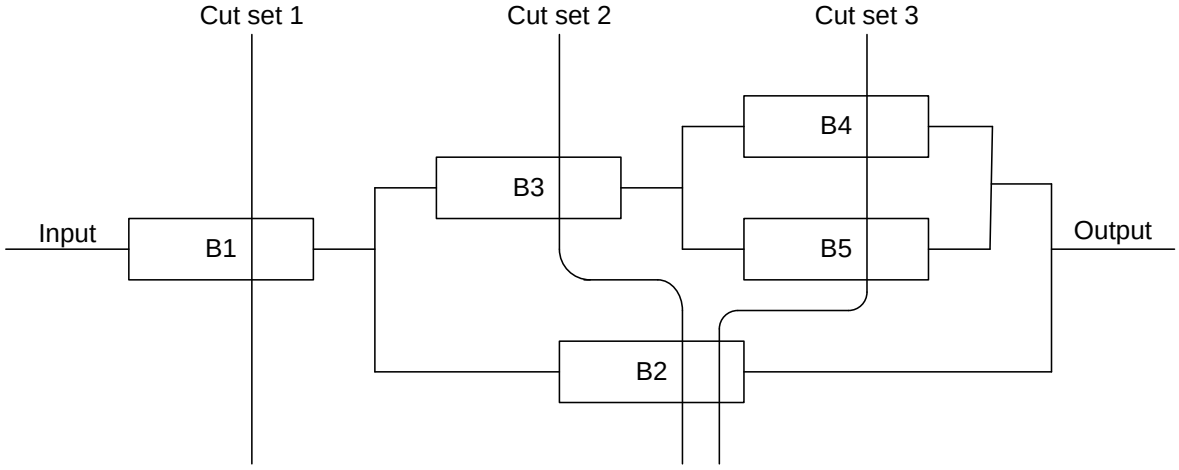


Figure 9.32: Example with various cut sets

9.4.5 Bayesian networks

Will be added in a next version of the lecture notes

9.5 System design considerations

The optimal system design and configuration will be case specific. Based on the previous sections a number of general considerations for system design can be given.

[1] http://reliawiki.org/index.php/Fault_Tree_Diagrams_and_System_Analysis#Minimal_Cut_Sets, accessed August 6, 2015

Firstly, it is important that a high-reliability system has sufficient **redundancy**. A system is redundant, if critical elements are duplicated with the intention of increasing reliability of the system (source: Wikipedia). Redundancy can be achieved by implementing elements and subsystems in a parallel configuration. It is then very important that these failures of these parallel elements are independent or mutually exclusive. In the case of full dependence the system failure probability reduces to that of the strongest component, i.e. $P(F) = \min(P_i)$.

In designing a parallel system, it is important that it shows ductile (and not brittle) failure behaviour. If multiple elements (e.g. columns in a building) show brittle behaviour, the total strength will be smaller than the sum of the resistances of all the elements.

It is also important to take into account the occurrence of common cause failures. These occur when multiple elements fail due to the same (shared) cause. One example, could be the inclusion of an unsound part in two subsystems. Another way a common cause failure could occur, is the occurrence of joint events, such as an earthquake or the loss of power, which will lead to failures of multiple elements. Such common cause failure could greatly increase the failure probability of the parallel system. For example, consider a system consisting of two elements with a failure probability of $P_1 = P_2 = 0.1$. In the case of independence the system failure probability equals $P(F) = 0.01$; When there would be common cause failure and both failures are dependent, the system failure probability would become $P(F) = \min(P_1, P_2) = 0.1$.

Adding elements and failure mechanisms in a series configuration will lead to an increase of the failure probability, unless the elements are fully dependent. It is therefore better to prevent a system design with multiple mechanisms or elements with a similar failure probability, as these probabilities will start to add up. In several cases, a series configuration can hardly be avoided at reasonable cost. Consider for example a dike ring system with multiple elements and failure mechanisms. Then, the question becomes how the acceptable failure probability can be distributed over the different failure mechanisms. Once the failure probability that is assigned to an element and mechanism is known, the required "strength" of the elements can be determined. Figure 9.33 gives an example of a possible assignment of a probability of failure for a series system with three independent elements.

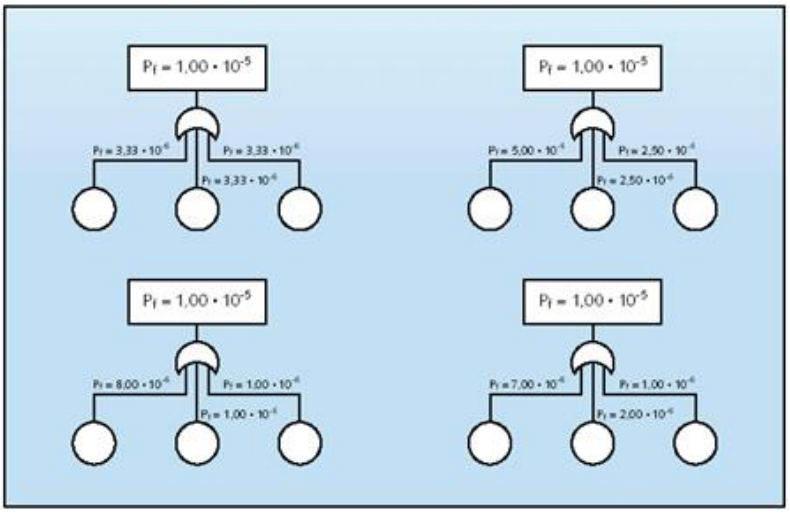


Figure 9.33: Possible distributions of the probability of failure for a series system with three independent elements.

This example shows that in theory a large number of distributions of the probability of failure over the various elements is possible. In practice, the distribution over failure mechanisms and elements will be determined by various aspects such as:

- Costs of reducing the failure probability for a given failure mechanism
- Feasibility of the interventions
- Societal preferences, in some cases it is desired to have low failure probabilities for specific mechanisms, because the consequences are more severe than for other mechanisms. In such cases, it could be argued whether all the mechanisms should be associated with the same top event in the fault tree.

These type of problems can also be subject to optimization. In such an optimization, it can be investigated which distribution over mechanisms leads to the cheapest structure for the given failure probability level. For example, a dike designer can allow more failure probability for overflow (leading to a higher dike) or allow a larger failure probability for geotechnical mechanisms (leading to a wider, but slightly lower dike).
In addition to the distribution over known mechanisms, it is good practice to assign a reserve of 20% for yet unidentified failure modes and other setbacks. As the design and insight progresses, the failure probability analyses become more detailed. In such phases of a project, a redistribution of the assigned probabilities of failure is possible and often appropriate.

References

- Bea, R., 2010. Human & Organizational Factors : Risk Assessment & Management of Engineered Systems Technical report, California: University of Berkeley.
- Bea, R., 1998. Human and organization factors: engineering operating safety into offshore structures. Reliability Engineering & System Safety, 61(1-2), pp.109–126. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0951832097000586 [Accessed November 20, 2013].
- Ditlevsen, O., Fundamentals of second moment structural reliability theory. International Research Seminar on Safety of structures, Trondheim, 1977.
- Hohenbichler, M. and R. Rackwitz, First-order concepts in system reliability. Structural Safety 1, Elsevier Science Publishers B.V., Amsterdam, 1983.
- Vrouwenvelder, A.C.W.M. and J.K. Vrijling, Probabilistic design (in Dutch: "Probabilistisch ontwerpen"). Delft University of Technology, Faculty of Civil Engineering, Delft, September 1987.

Appendix 9.1 Cause consequence chart

One characteristic of the fault tree is that only the probability of failure (or non-failure) of the entire system is addressed. The consequences of different combinations or scenarios of events cannot be represented in the fault tree. Therefore the cause consequence chart has been developed, which combines some characteristics of the fault tree and the event tree. It enables the analyst to show multiple sets of consequences (similar to the event tree). Figure 9.34 gives an example of a cause consequence chart for a sea barrier system, consisting of the guardian, the sleeper and the dreamer. The consequences of failure of different lines of defence will differ. For example only agricultural losses in case of failure of the first two defences, but larger economic losses of all three defences fail and the village gets flooded. By combining the probabilities and consequences of the events, the probability density function (or distribution) of consequences can be compiled and the risk can be determined (see also chapter 3 for further background). Although the cause consequence chart is already an improvement relative to the mentioned trees, it still doesn't offer the possibility of considering the consequence of a random variable.

Within the cause consequence chart, failure can also be assessed as a function of a number of continuous random variables representing the load and strength (see Figure 9.35). For example, the hydraulic loads S due to surge on the dike can be described by means of statistical distribution. Also, (model) uncertainty in the overtopping resistance R can be taken into account. Level II and III calculation methods can be used to determine the probability of failure.

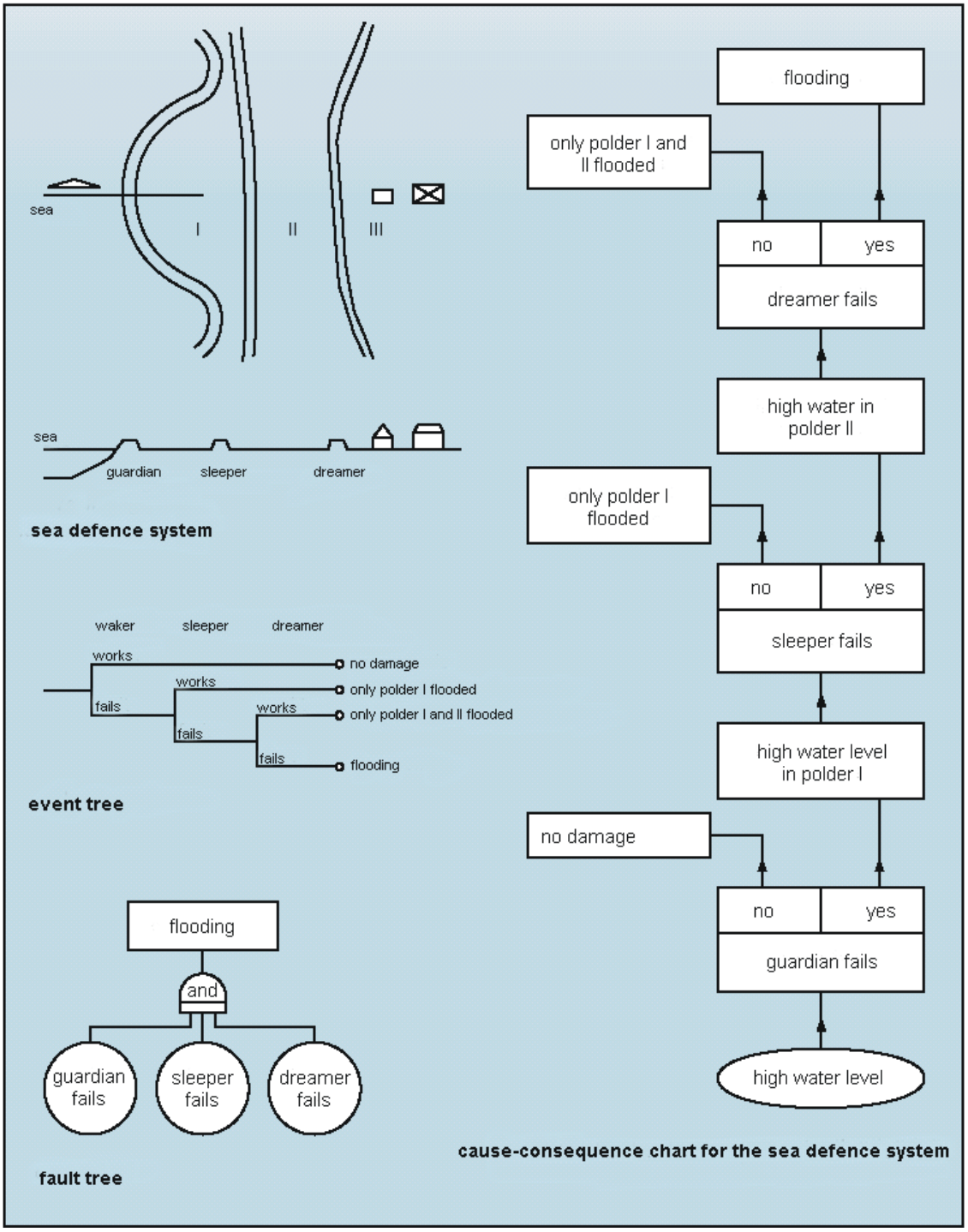


Figure 9.34: Event tree, fault tree and a cause consequence chart for a sea defence system consisting of multiple lines of defence.

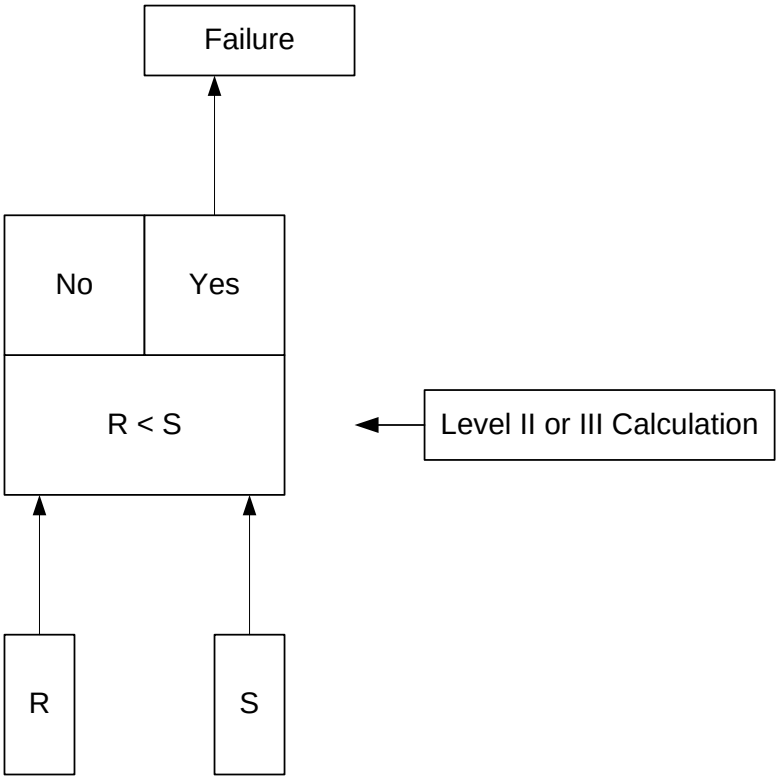


Figure 9.35: A simple cause consequence chart for a limit state.

Appendix 9.2 Systems reliability for variables described by various copulae

In section 9.2.2, the failure probability of a series system of two identical elements as a function of the value of the correlation coefficient was discussed, under the assumption that the joint distribution function is described by the Gaussian copula. This does not always have to be the case. In fact, when the joint distribution is described by another copula, for example the t-copula or the Clayton copula, the failure probability as a function of the correlation coefficient can show very different behaviour, see Figure 9.36.

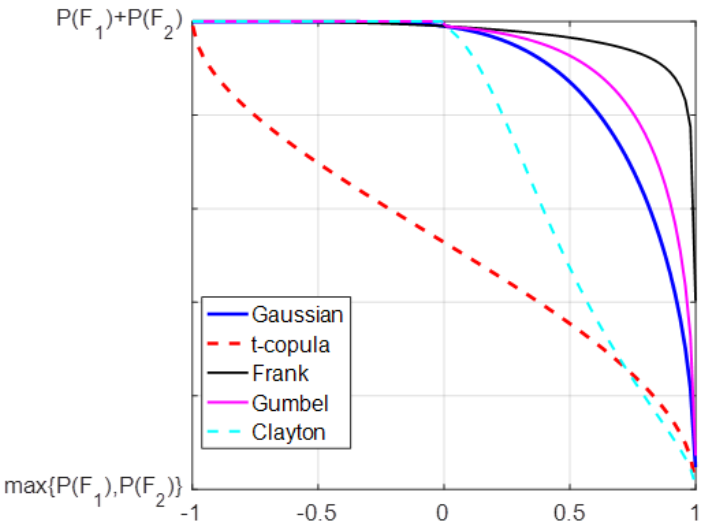


Figure 9.36 Failure probability of a series system of two identical elements as a function of the value of the correlation coefficient according to different copulas (schematic representation for small failure probabilities).

Part III – Applications

Chapter "Application and design codes in civil / structural engineering"
Author: R.D.J.M. Steenbergen

Large parts of this chapter are based on 'Risk Analysis of Construction Processes' by L. Taerwe and R. Caspeele (UGent lecture notes), the examples of the structural failures

10 APPLICATIONS AND DESIGN CODES IN CIVIL / STRUCTURAL ENGINEERING 234

10 Applications and design codes in civil / structural engineering

This section focuses on reliability concepts and applications concepts used in civil and especially structural engineering. First, some general background in the structural safety is given (section 10.1). Consequently, design concepts and codes are introduced (10.2) and target values for reliability in civil engineering (10.3). Subsequently, some general approaches for modelling the load and strength are presented (10.4). The final section deals with variable loads and actions on structures (10.5).

10.1 Background

A difficult question that a designer sometimes has to answer can be "Is this construction now really safe?". This question cannot simply be answered by "yes" or "no". First of all, safety is a relative concept that always has to be evaluated in the context of a certain reference framework for risk evaluation (see chapter 3). Secondly, such a question relates in general to the global structure or system, but problems that are considered often only concern a limited number of elements or stages of the construction process.. Human and organizational errors are crucial in (preventing) failures and in risk management. Therefore, it should be emphasized that integral quality management and the continuous education (at all levels) is an absolute necessity. Control procedures will inevitably stay an important aspect in the construction process, but also self-discipline (for example by a good education) can be very effective in this regard. Especially the lack of reliable communication is one of the major causes of structural accidents in the construction industry. It appears that the failure and collapse of structures is strongly influenced by human mistakes, errors and shortcomings, in the design phase, the construction phase and during the service period. This problem is illustrated in the following by discussing a number of important and remarkable structural accidents and collapses that are summarized in Appendix 10.1.

The failure and collapse of structures can be summarized with the term '**structural accidents**'. The probability of dying due to the consequences of structural accidents is in the order of magnitude of 10^{-7} per year (CIRIA, 1977) and thereby rather small compared to other risks (e.g. 10^{-4} per year for traffic accidents and 10^{-2} per year for mountaineering – see chapter 3). Considering that the time which is spent inside or on structures or infrastructures is quite large, the corresponding risk is rather small, which proves that structures are in general designed in a safe way.

Table 10.1 gives an overview of structural accidents (not necessarily lethal) according to the type of structure. This overview is obtained from a study performed by Matousek which is summarized in (Hauser, 1979) and relates to more than 800 structural accidents in Europe. The establishment of such statistical information is rather difficult because not all cases are reported and information about it is in general not clearly and uniformly reported.

Table 10.1 Relative occurrence of structural accidents with respect to the type of structure (Hauser, 1979)

Residential and office buildings	52 %
Industrial buildings	22 %
Road constructions (bridges, tunnels, etc.)	11 %
Coastal engineering and hydraulics	7 %
Air-raid shelters	2 %
Other types	6 %

The construction phase during which an error occurred or during which a certain aspect was neglected, is indicated in Table 10.2. Design errors can be mainly allocated to a wrong notion about the behaviour of certain components, wrong dimensions and/or not recognizing certain risk causes.

Table 10.2 Construction phase during which an error occurred (Hauser, 1979)

Design	37 %
Constructon	35 %
Design and construction	18 %
Use	5 %
Other	5 %

From this research, it was observed that the most efficient way to increasing structural safety or attaining a predefined safety level is preventing large errors. Apparently, the current structural prescriptions and models are in general sufficiently safe and accurate. Furthermore, the quality of the applied materials is most often sufficiently high. However, a crucial point is to execute an efficient quality control on the different activities that take place in the construction process.

Possible negative consequences of structural accidents are most often allocated to both the designer and the constructor. Very often a joint liability is pronounced, because the contribution of both parties to the construction process are closely correlated and sometimes difficult to distinguish. Luckily, the consequences of this liability are less draconic nowadays compared to the time of Hammurabi, king of Babylon from 1728 to 1686 BC. The so-called Code of Hammurabi, engraved in a diorite stele which is currently on display in The Louvre, is considered to be the oldest known building law (Meyer, 1986). The Code contains among others the following extract Figure 10.1:

"If a builder builds a house for someone, and does not construct it properly, and the house which he built collapses and kills its owner, then that builder shall be put to death."

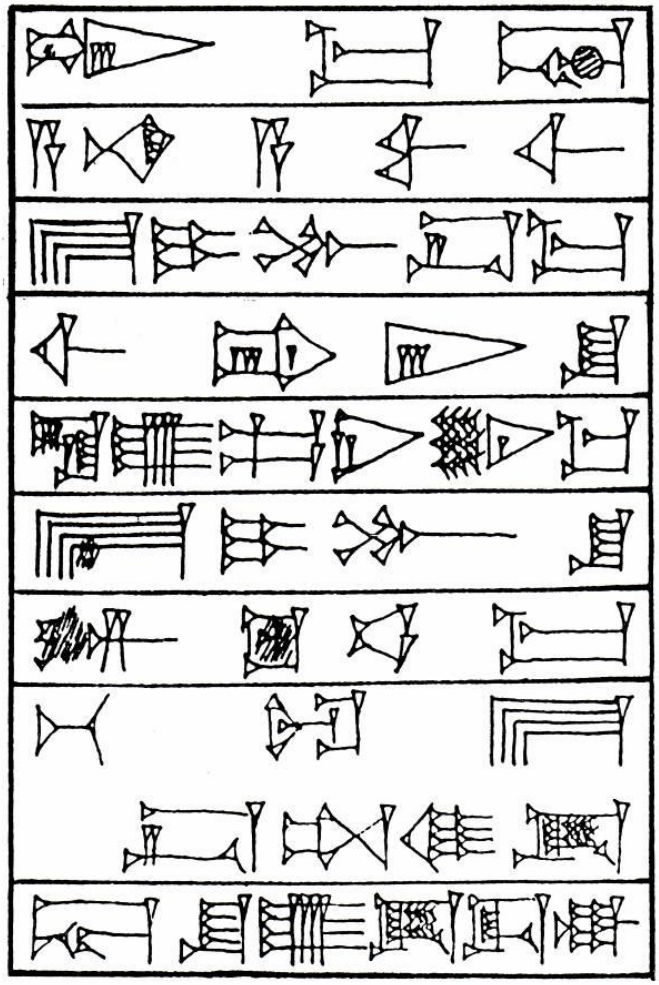


Figure 10.1 Extract of the Code of Hammurabi with respect to the liability of the master builder

Not only the dramatic aspect of this law is striking, but also the fact that the legislators of that time attributed the failure of structures completely to the shortcomings of the master builder. Hence, they completely neglected the existence of natural causes.

Because only a very limited number of structures fail, one can have the impression that most structures are absolutely safe. However, absolute safety is not realizable and it is moreover not desirable to aim for such absolute safety as this would lead to a probably unaffordable and large investment. After all, increasing safety has always some economic consequences. However, this does not mean that there is always exists a conflict between both aspects. One has to aim at 'fitness for purpose' and accept some calculated risks. Hence, the main objective of the design is to make sure that there is a low probability that the structure would be unfit for its intended purpose during its lifetime.

Most structures have to fulfil several performance criteria, most often expressed as limit states. Most often these limit states interact and the problem is more complex than specifying a single probability. As such, the concept of safety is currently considered in a broader context than was previously the case. A risk analysis not only consists of the calculation of a failure probability or

the application of appropriate partial factors, but one also has to consider optimization of the design with respect to the specified performance criteria, application of decision theories, modelling of accidental actions, quality control, human errors and the associated problems. This broader context requires the incorporation of a whole new set of concepts. However, often it is not easy to adjust the existing calculation models – of which the deterministic formulation is already rather sophisticated – towards an equivalent formulation in a probabilistic framework.

The design of more complex structures requires the verification of several characteristic design situations. The number of such design situations that have to be verified can increase quickly and make it very difficult to have a clear overview of the general behaviour, which can sometimes not be grasped by one person. As an example, the cryogen storage tanks of the Fife Ethylene Project in the neighbourhood of Edinburgh are mentioned, where in total twelve different load combinations were considered accounting for the different situations that can occur during the lifetime of the structure (construction phase, water test, cool down period, use, maintenance, accidental actions) (Wendrich, 1983). The structural aspects of the storage tanks is of course only one part of the project, that included also supply, gasification and distribution installations. Hence, with respect to safety considerations of the entire project, it was deemed necessary to incorporate a number of justified simplifications and to consider that some events take place independently.

The same complexity also occurs in nuclear power plants, offshore structures, high-rise buildings, flood defence systems etc. For all these cases, the information with respect to the different basic variables is sometimes rather limited, especially in case new materials are applied. Moreover, the knowledge with respect to the structural behaviour under complex load combinations is sometimes limited, i.e. the degree of uncertainty with respect to fundamental aspects is higher than usual and only a probabilistic approach can lead to a sound result.

10.2 Design concepts and codes

In this section some general concepts are introduced that are used in civil and structural engineering. Most concepts which are mentioned in this chapter are based on the European Standard EN 1990 entitled “Basis of structural design”.

10.2.1 N**ormative Reference Framework**

The European Standard EN 1990 (Eurocode 0) "Basis of structural design” contains

- the principles and requirements with respect to safety, use and durability of civil structures
- the basis for the design and control of buildings and civil engineering structures
- guidelines for structural safety aspects.

The idea is that this EN 1990 is used together with EN 1991 to 1999 in order to design buildings and civil engineering structures, including geotechnical aspects, fire design, seismic design and design of temporary structures. The entire set of Eurocodes is conceived as follows:

- EN 1990: Basis of structural design;
- EN 1991: Actions on structures;
- EN 1992: Design of concrete structures;

- EN 1993: Design of steel structures;
- EN 1994: Design of composite steel and concrete structures;
- EN 1995: Design of timber structures;
- EN 1996: Design of masonry structures;
- EN 1997: Geotechnical design;
- EN 1998: Design of structures for earthquake resistance;
- EN 1999: Design of aluminium structures.

EN 1991 is subdivided in :

- part 1-1: Densities, self-weight and imposed loads for buildings;
- part 1-2: Actions on structures exposed to fire;
- part 1-3: Snow loads;
- part 1-4: Wind actions;
- part 1-5: Thermal actions;
- part 1-6: Actions during execution;
- part 1-7: Accidental actions due to impact and explosions;
- part 2: Traffic loads on bridges;
- part 3: Actions induced by cranes and machinery;
- part 4: Actions on silos and tanks.

10.2.2 Basic Assumptions

The following basic assumptions are applicable:

- supporting systems are designed by people with proper education and training;
- the execution is realized by people with proper skills and experience;
- in factories and at the construction site adequate quality control is applied;
- the construction materials and products are used as prescribed in EN 1990 till 1999 or in the concerning material or product specifications;
- the structure will be suitably maintained;
- the structure will be used for the purpose it was intended for.

10.2.3 Fundamental Requirements

A structure must be designed and executed so that, during the anticipated lifetime:

- it withstands all actions and influences that can occur during construction and use;
- it remains suitable for the use it was intended for and this in a reliable and economical way.

The previous fundamental requirement basically means that the structure should be designed so that it has **sufficient strength, serviceability and durability**.

Furthermore, a structure also has to be designed and executed in such a way that it does not suffer disproportionate damage due to fire, explosion, impact or due to the consequences of human errors.

One can limit or avoid possible damage by considering the following measures in favour of robustness:

- prevent or limit potential hazards to which the structure is exposed;
- choose a construction type with low vulnerability with respect to the potential hazards under consideration;
- avoid as much as possible supporting systems that can fail without notification provide connectivity between different elements;
- choose a construction type that can withstand accidental failure of an element or a limited part of the construction, or that can withstand local damage of acceptable size.

The fundamental requirements are satisfied by:

- appropriate choice of materials;
- professional design and detailing prescription of control procedures for the design, production, execution and use of the structure.

10.2.4 Design Lifetime of the Structure

The design lifetime is the anticipated life expectancy of a structure. During the design lifetime one assumes that the structure can function for its intended use, considering reasonable maintenance and without large repair. Required values for the design working life are summarized in Table 10.3.

Table 10.3 Required values for the design working life (according to Dutch National Annex to EN 1990)

Class	Required lifetime (years)	Examples
1	2 to 10	Temporary structures
2	15	Agricultural and equivalent structures
3	50	Buildings: houses, community buildings, offices
4	100	Monumental buildings, bridges and civil engineering construction works

10.2.5 Design Situations

A design situation considers the variation in actions, environmental influences and structural properties which will occur during the design working life of a structure. Different design situations are distinguished, each representing a certain time interval with associated hazards or conditions:

- **persistent design situations**: these refer to the conditions of normal use and are generally related to the design working life of the structure;
- **transient design situations**: these refer to temporary conditions of the structure and refer to a time period much shorter than the design working life, e.g. during construction or repair;
- **accidental design situations**: these refer to exceptional conditions of the structure or of its exposure, e.g. due to fire, explosion, impact or local failure;

- **seismic design situations**: these refer to exceptional conditions applicable to the structure when subjected to seismic events.

10.2.6 Limit States

Definition

In order to verify the fundamental requirements for the design situation under consideration, one defines so-called limit states (already briefly introduced in chapter 3.2.3). A limit state is a condition beyond which the structure or part of the structure does no longer fulfil one of its performance requirements. Thus, for each of these performance requirements one or more limit states can be formulated.

The **Ultimate limit state**s (ULS) refer to the safety of persons and/or the structural safety or the protection of the content of a structure.

In the ultimate limit state the ultimate bearing capacity of the structure is defined. Beyond this limit state the entire structure or part of it fails. The following ultimate limit states can be considered:

- loss of static equilibrium of the structural system or one of its components, considered as rigid body (turn over, sliding, push up, etc.);
- fracture or excessive deformation in critical sections of the structural system or its connections;
- fatigue or other time-dependent phenomena;
- formation of a mechanism of a structural system or a part of it (collapse) ;
- instability, divergence of equilibrium of the system or its components (buckling, lateral buckling, aero-elastic instability, …).

Serviceability limit states (SLS) refer to:

- the performance of the structure and its components during normal use;
- the comfort of the users;
- visual aspects.

When crossing a serviceability limit state, one or more of the requirements with respect to the functionality is not anymore fulfilled. The verification of serviceability limit states is related to:

- **deformations** that can have an influence on visual aspects, comfort of users and the functionality of the structure, including installations (e.g. equipment) or damage caused to finishing or non-bearing elements;
- **vibrations** that can result in discomfort of users or influence the functionality of the structure;
- **damages** that can have a negative effect on visual aspects, the durability and the use.

One can distinguish between two types of SLS (Figure 10.2) :

- **irreversible SLS** where the critical value stays permanently crossed after removal of the load that caused the first passage (e.g. permanent local damage, permanent deflection);
- **reversible SLS** where the critical value is no longer crossed after removal of the load that caused the first passage (e.g. cracks in prestressed concrete elements, temporary deflections, excessive vibrations).

In case of irreversible SLS, the design criteria are similar to those for ULS. The time until the first passage is determinant. In case of reversible SLS, the first passage not necessarily results in the failure or unfitness for use. Alternative serviceability requirements can be formulated depending on the acceptability of crossing a limiting value, their frequency and duration. These aspects can be associated to different load combinations (see further).

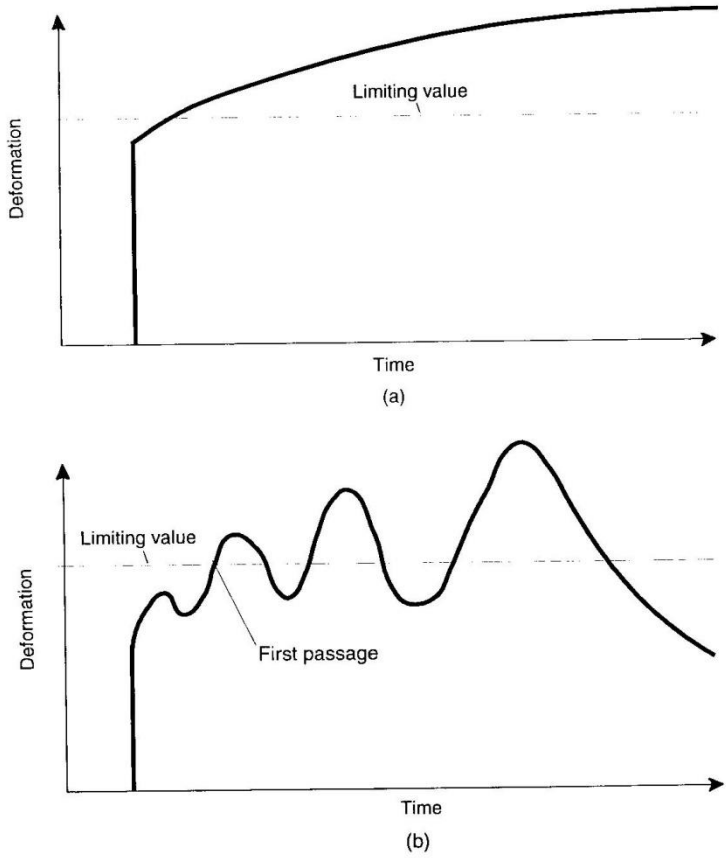


Figure 10.2 Different types of SLS: irreversible (top figure) and reversible (lower figure).

10.3 Target Values for the Reliability Index in Civil Engineering

10.3.1 General

In order to determine how safe a structure or system should be, an acceptable level of risk needs to be defined. A previous chapter (3) of these lecture notes has described how risk assessments can be used to derive an acceptable failure probability for a system. For standard applications and systems that are frequently constructed, codes are available that define acceptable safety levels.

The Eurocode EN 1990 (Annex C) defines target values for the reliability index. In Table 10.4 target values for β are mentioned for reference periods of 1 and 50 years, designated β_1 and β_{50} respectively. These values correspond to a reliability class RC2 (see below).

Table 10.4 Target values for β (from Annex C of EN 1990).

Limit state	t_{ref} = 1 year	t_{ref} = 50 years
Ultimate limit state	4.7	3.8
Serviceability limit state	2.9	1.5

For most verifications that have been proposed in the framework of EN 1990 to EN 1999 the following assumption for the distribution type have been considered:

- lognormal or Weibull distributions for material properties, strength and model uncertainties
- normal distributions for self-weight
- extreme value distributions for variable actions (sometimes normal distributions as a simplification).

When the main source of uncertainty can be related to actions of which the yearly maxima are independent, the following relationship can be used to convert β values in relation to different reference periods:

$$\Phi(\beta_n) = [\Phi(\beta_1)]^n \tag{10.1}$$

with β_n the reliability index for t_{ref} = n years and β_1 the reliability index for t_{ref} = 1 year. Considering equation (10.1) can be rewritten as:

$$P_{s,n} = P_{s,1}^n \tag{10.2}$$

based on:

$$\begin{aligned} &P[\text{survival in n years}] \\ &= P[\text{survival in year 1}].P[\text{survival in year 2}]...P[\text{survival in year } n] \\ &= P[\text{survival in a period of 1 year}]^n \end{aligned} \tag{10.3}$$

Alternatively, the probability of failure over the lifetime ($P_{f,n}$) can be derived:

$$P_{f,n} = 1 - P_{s,n} = 1 - P_{s,1}^n = 1 - \left(1 - P_{f,1}\right)^n \tag{10.4}$$

For example, for a system with a yearly failure probability of $P_{f,1}$=0.01 and a lifetime of 10 years, the failure probability of the lifetime becomes $P_{f,10} = 1 - (1\text{-}0.01)^{10} = 0.095$

10.3.2 Reliability differentiation

Considering the importance of the failure consequences, Annex B of EN 1990 provides the Consequence Classes as mentioned in Table 10.5.

Table 10.5 Definition of Consequence Classes

CC	Description		Examples
	Consequences with respect to loss of human lives	Economic, social and environmental consequences	
CC3	High	Very large	Tribunes, public buildings with high consequences of failure (concert hall, …)
CC2	Moderate	Considerable	Home and office buildings, public buildings with moderate consequences of failure (offices, …)
CC1	Low	Small or negligible	Agricultural building where people do not normally enter (depositories, greenhouses, …)

The three consequence classes CC1, CC2 and CC3 correspond to three reliability classes RC1, RC2 and RC3, respectively. In Table 10.6 the recommended minimum values for β are given with respect to ultimate limit states.

Table 10.6 Recommended minimum values for β (ULS) (Eurocode EN 1990)

Reliability class	t_{ref} = 1 year	t_{ref} = 50 years
RC3	5.2	4.3
RC2	4.7	3.8
RC1	4.2	3.3

The values mentioned in Table 10.4 correspond to RC2. Also, the Eurocode distinguishes different levels of design supervision and inspection level, see Appendix 10.2.

10.4 Characterising the strength, and loads and actions (-effects) on structures

10.4.1 Material and geometrical properties

Material and product properties are represented by characteristic values, denoted with X_k or R_k. These correspond to a prescribed non-exceedance probability in the theoretical distribution of an in principle infinite sample size. Unless specified otherwise, characteristic values are defined in EN 1992 to EN 1999 as the 5-quantile in case of strength parameters and mean values in case of stiffness characteristics. For certain situations, one can define low as well as high characteristic values which then correspond respectively to the 5- and 95-quantile of the distribution under consideration.

Material properties are determined by standard tests which are executed under specified circumstances. Conversion factors are used to transform the results into in-situ material strength of the structure or the soil.

Geometrical properties are represented by a characteristic value in case the distribution function is known, or by a design value otherwise.

10.4.2 Classification of loads, actions and environmental influences

A distinction is made between the external loads (e.g. wind and waves) and the actions that they cause within a structure. An action is:

- a group of forces acting on a structure (direct action)
- a group of imposed deformations or accelerations e.g. caused by variation in temperature or humidity, differential settlements or earthquakes (indirect action)

Based on their variation in time, actions can be classified as follows:

1) permanent actions *G*: actions that are present during the entire duration of the reference period. Their change in function of time is negligible with respect to their mean value or the change is constant until a limiting value is reached (e.g. the water level at a spillway). Examples are self-weight, permanent equipment, floor or road finishing, shrinkage, differential settlements, …

2) variable actions *Q*: actions that are not present during the entire duration of the reference period. Their change in function of time is not negligible with respect to their mean value. Examples are live loads, traffic loads, snow load, wind load, ….

3) accidental actions *A*: actions that have a low probability of occurrence during the reference period, but can have an important influence on structural calculations due to their magnitude (e.g. impact, explosion, fire, earthquake, …).

Actions that act together and reach their maximum at the same time are most often considered as one action in practical calculations. Actions that are only weakly correlated can be considered to act independently and occur separately in the calculations.

Actions can also be classified with respect to the structural response:

1) Static loads: these do not produce significant accelerations in the structural system or its components;
2) Dynamic loads: these can produce significant accelerations in the structural system or its components.

The effect of dynamic loads is often be calculated by multiplying the static response by a dynamic factor. In case this is not possible, a specialized calculation has to be performed in order to determine the dynamic response of the structural system. For example, in the field of earthquake engineering the most simple analysis concerns a pseudo-static method. Within this approach a static force equivalent to the seismic loading is inserted in the force balance calculation of the structure. A more detailed dynamic analysis of the structure can be made using finite element models that allow dynamic calculations.

Finally, one can also differentiate between fixed loads and free (or mobile) loads.

In this framework loads are considered as random variables with a certain distribution type. However, modelling of variable and accidental actions as a stochastic process is sometimes more appropriate.

10.4.3 Characteristic values of permanent actions

In case the variation of permanent action G during the design working life is negligible (coefficient of variation V_G = 5 % to 10 % depending on the type of structure) the characteristic value G_k corresponds to the mean value. The self-weight is calculated taking into account the nominal dimensions and the mean values of the densities (EN 1991-1-1).

In case the variation of G is not negligible, 2 characteristic values can be determined. This is also the case when the integrity of the structure is very sensitive to variations of G even when δ_G is small. Considering a normal distribution, the 5- and 95-quantiles can be obtained as follows:

$$\begin{aligned} G_{k,\text{inf}} &= \mu_G - 1.645\,\sigma_G \\ G_{k,\text{sup}} &= \mu_G + 1.645\,\sigma_G \end{aligned} \tag{10.5}$$

10.4.4 Statistical and model uncertainties

One should differentiate between 3 major types of uncertainty, namely physical, statistical and model uncertainties. In the sections above we discussed the **physical or intrinsic uncertainties** which are related to the random nature of the variable under consideration. They are accounted for through a density function or a stochastic process.

Statistical uncertainties are due to the fact that the parameters of distribution functions cannot be estimated exactly based on a sample of limited size. In case one has obtained a sample set of n observations and an additional sample set is taken, a new set of values will be obtained as well as new sample characteristics. Each time a different estimate of the distribution parameters is calculated. Hence, the estimations should also be considered as random variables of which the variation can be described by a distribution function. In this case the predictive distribution (i.e. global distribution, incorporating parameter uncertainties) of the basic variable X with parameter vector $\underline{\theta}$ is given by:

$$f_X(x) = \int f_X\left(x|\underline{\theta}\right) f_{\underline{\theta}}\left(\underline{\theta}\right)\, d\underline{\theta} \tag{10.6}$$

Model uncertainties are related to the choice of the applied models. This can be related to the probabilistic model (wrong distribution function, neglecting correlation,...) or the deterministic mechanical model that describes the limit state under consideration. Frequently simplifications are made which result in more or less large deviations between the calculated response and the real response of a structure (e.g. when applying a linear elastic analysis in case of an intrinsically non-linear behaviour). Limit states with important model uncertainties are for example related to shear, deformations, crack widths, etc. The model uncertainty of a model can be expressed by a distribution function of an additional random variable X_m defined as:

$$X_m = \frac{\text{real response}}{\text{predicted response based on a model}} \tag{10.7}$$

This variable can be incorporated in an additive or multiplicative way, considering most often a normal or lognormal distribution. In case the average response of the model is well-predicted, $\mu_{Xm} = 1$. In case of a conservative strength model $\mu_{Xm} > 1$ and in case of an "unsafe" strength model $\mu_{Xm} < 1$.

10.5 Variable loads

10.5.1 Return period

In a time-dependent problem one is most often interested in the number of time intervals until the first occurrence of a certain event. In case the event can only occur once in an elementary time interval ΔT and in case subsequent intervals are stochastically independent (cf. experiments), the time between two subsequent random occurrences of an event is equal to the time until the first occurrence. In case one assumes that the occurrence in an elementary time interval can be represented as a Bernoulli experiment (with probability of occurrence p), the time T between the subsequent occurrences of an event follows a geometric distribution. The mean value of this time is called the **return period** T_r. Considering the properties of a geometric distribution it holds that $T_r = 1/p$.

When designing structures for exposure to extreme events such as high waves or storm winds, one has to assess the probability that a critical value x_{crit} of the variable under consideration (wave height, wind speed) will be exceeded. Therefore, the distribution of the annual maxima is considered. Considering $p = P[X > x_{crit}]$, the average time between years in which x_{crit} is exceeded is equal to the return period $T_r = 1/p$. The value x_{crit} is called the T_r value of the variable, for example, the 100 year wave height equals H_s=4m. Hence, this value corresponds to the $(1\text{-}1/T_r)$ quantile or in other words the value of the annual maximum with an exceedance probability of $1/T_r$. As an example, coastal engineering and offshore structures are often designed based on the 100 years wave (T_r = 100 years) corresponding to p = 0.01 in 1 year.

The probability that an event with probability p occurs (at least once) in a timeframe T is equal to:

$$P\left[\text{occurrence in } T\right] = 1 - \left[1 - p\right]^T \tag{10.8}$$

A specific case is considered when one assesses the probability that the return period event (with $p=1/T_r$) occurs within that period $T=T_r$. For example, a client can ask a question what the probability is that the 10 year wave height occurs within a period of 10 years. In case T_r = 10 (p = 0.10) a probability of 0.651 is obtained, which is quite large. Using the binomial law, the complementary probability can be calculated as:

$$(1-p)^{T_r} = 1 - T_r\, p + \frac{T_r\,(T_r-1)}{2!}\, p^2 - \frac{T_r\,(T_r-1)\,(T_r-2)}{3!}\, p^3 + \ldots \tag{10.9}$$

In case of large values for T_r (and hence small values for p) the right side of the equation becomes:

$$e^{-T_r\, p} = e^{-1} = 0.368 \tag{10.10}$$

Hence:

$$P\left[\text{occurrence in } T_r\right] \approx 1 - 1/e = 0.632... \tag{10.11}$$

In case of rare events (T_r large) the probability of occurrence within the return period T_r can be approximated as 0.632. In case of T_r = 10 the approximation achieves a deviation of only 3 % compared to the exact result.

10.5.2 Characteristic value Q_k

In case of variable actions the characteristic value (Q_k) corresponds to:

- a high or low quantile value that corresponds to a specified low resp. high maximum probability of exceedance;
- a nominal value when insufficient statistical information is available.

In EN 1990 the **characteristic value Q_k** of a variable action is in general determined as the value that corresponds to an exceedance probability of 2% during a reference period of 1 year. Hence, the return period is 50 years.

In general the probability of not exceeding Q_k in a certain reference period t_{ref} can be calculated as follows in case of t_{ref} > 1 year (e.g. the design working live):

$$P\left[Q \le Q_k \,\middle|\, t = t_{ref} \right] = \left(1 - \frac{1}{T_r} \right)^{t_{ref}} = 1 - p \tag{10.12}$$

where p is the exceedance probability within the design working life. In Figure 10.3, equation (10.12) is illustrated as:

$$p = 1 - \left(1 - \frac{1}{T_r} \right)^{t_{ref}} \quad \text{or} \quad T_r = \frac{1}{1 - (1-p)^{1/t_{ref}}} \tag{10.13}$$

For example for an event with a return period of T_r = 100 years and a working life of t_{ref} = 5 years, it is found that p=0.049. So there is approximately a 5% probability of exceeding the 100 year event within the working life. The upper limit of 63.2 % was already calculated in section 10.5.1. In case of t_{ref} = 10 years the exceedance probability of Q_k is equal to 18.3 %.

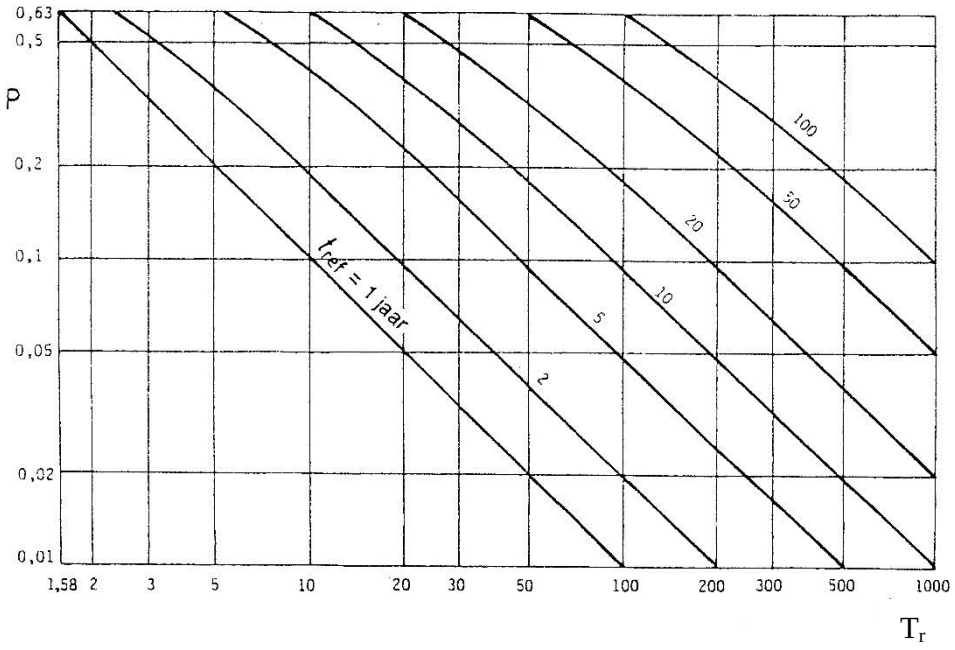


Figure 10.3 Exceedance probability p of a maximum value corresponding to a return period T_r during a reference period t_{ref}

10.5.3 Extreme distribution for the maxima (Gumbel distribution)

General

Extreme value distributions have been introduced in section 2.6.6.. These types of distributions are generally used to model the distributions of extreme loads, such as wind loads. This section focuses on the application of these type of distributions.

The random nature of the variable load X is modelled according to an extreme value distribution of type I for the maxima, i.e. the Gumbel distribution. The CDF and PDF are respectively:

$$F_Y(y) = \exp\left[-e^{-\alpha(x-u)}\right] \qquad \text{for} -\infty < x < +\infty \tag{10.14}$$

$$f_Y(y) = \alpha \exp\left[-\alpha(x-u) - e^{-\alpha(x-u)}\right] \tag{10.15}$$

The relationship between the parameters u (mode), α and (μ_Y, σ_Y) is as according to table (2.3):

$$\mu_x = u + \gamma / \alpha \cong u + 0.5772 / \alpha \tag{10.16}$$

$$\sigma_x = \frac{\pi}{\sqrt{6}\,\alpha} \cong \frac{1.282}{\alpha} \tag{10.17}$$

Considering the normalized variable:

$$W = (X - u)\,\alpha \tag{10.18}$$

the CDF becomes:

$$F_W(w) = \exp(-e^{-w}) \tag{10.19}$$

Application of extreme value distributions

Extreme value distributions are often used to determine the distributions for the yearly maxima and for the maxima during the reference period (e.g. 10, 50 or 100 years). This is elaborated below.

Firstly, according to the definition given in section 10.5.2 of this chapter, the characteristic value corresponds to an exceedance probability of 2% during a reference period of 1 year.

$$F_W(w) = \exp(-e^{-w}) = 0.98 \tag{10.20}$$

Q_k can be formulated as a function of μ_1 and σ_1 (the subscript designates the basic reference period t_1 which is often equal to 1 year):

$$w = -\ln(-\ln(0.98)) = 3.902 = (Q_k - u)\,\alpha \tag{10.21}$$

Or:

$$3.902 = \left(Q_k - \mu_1 + \frac{0.5772}{1.282}\,\sigma_1 \right) \frac{1.282}{\sigma_1} \tag{10.22}$$

from which:

$$Q_k = \mu_1 + 2.593\,\sigma_1 \tag{10.23}$$

Consequently, the distribution of the maxima for a return period $t_{ref} = n \cdot t_1$ can be determined. Assuming that the annual maxima occur independently, the following holds:

$$F_r(x) = \left[F_1(x) \right]^n \tag{10.24}$$

with $F_1(x)$ the CDF of the annual maxima. Elaboration yields:

$$F_r(x) = \exp\left[-n\, e^{-\alpha(x-u)} \right] \tag{10.25}$$

This distribution is again of the type I for the maxima (i.e. Gumbel). Equation (10.25) can be rewritten as:

$$F_r(x) = \exp\left[-e^{-\alpha\left(x-u-\frac{\ln n}{\alpha} \right)} \right] \tag{10.26}$$

which yields the u, μ, and σ values for a reference period of n:

$$\begin{aligned} u_n &= u_1 + \frac{\ln(n)}{\alpha} \\ \mu_n &= \mu_1 + \frac{\ln(n)}{\alpha} \\ \sigma_n &= \sigma_1 \cong \frac{1,282}{\alpha} \end{aligned} \tag{10.27}$$

Hence, the standard deviation remains unchanged but the mode and the mean value shift to higher values (Figure 10.4), this is called the so-called “Gumbel shift”. In case one draws $F_1(x)$ and $F_r(x)$

on appropriate probability paper, two parallel lines are obtained with (horizontal) distance ln*(n) /α* from each other. Based on the aforementioned calculations equation (10.23) can be rewritten as:

$$Q_k = \mu_n - \frac{\ln(n)}{1,282}\sigma_n + 2,593\,\sigma_n \tag{10.28}$$

and for $n = 50$ ($t_1 = 1$ year ; $t_{ref} = 50$ years) one obtains:

$$Q_k = \mu_n - 0,4585\,\sigma_n \tag{10.29}$$

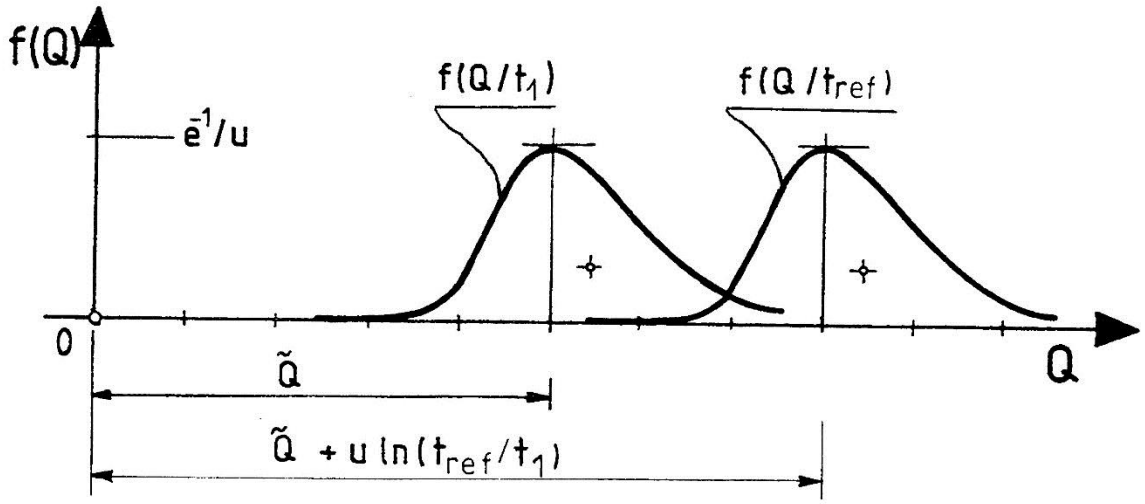


Figure 10.4 Distribution of the maximum over a time period t_1 and $t_{ref} = n \cdot t_1$

Example 10.1

The distribution of the yearly maxima of the wind pressure can be described with a Gumbel distribution:

$$p_{wind,1}(x) = \exp\left(-\exp\left(-\alpha_1\left(x - u_1\right)\right)\right) \tag{10.30}$$

For this example, we assume u_1 = 20 kN/m^2 and α=0.53m^2/kN

The distribution of the 50 year maxima of the wind pressure can now be determined as follows:

$$p_{wind,50}(x) = \exp\left(-\exp\left(-\alpha_{50}\left(x - u_{50}\right)\right)\right) \tag{10.31}$$

$$\begin{aligned} a_{50} &= a_1 = 0.53\ m^2\ /\ kN \\ u_{50} &= u_1 + \frac{\ln(50)}{a_1} = 27.38 kN\ /\ m^2 \end{aligned} \tag{10.32}$$

10.5.4 Other representative values (*not part of the examination materials)*

In Figure 10.5 a schematic representation of the time-dependent behaviour of a variable action *Q* is illustrated during the reference period. The distribution $F_Q(q)$ is related to the point-in-time value. Further, also the distribution of the annual maxima $F_{Qmax}(q)$ is shown.

a) The **combination value** $\psi_0 Q_k$ is used for load combinations in order to account for the simultaneous occurrence of extreme values of different independent variable actions. As a first approximation once can consider the following:

$$P[Q_1 + Q_2 > Q_{1k} + \Psi_0 Q_{2k}] \cong P[Q_1 > Q_{1k}] \cong P[Q_2 \geq Q_{2k}] \qquad (10.33)$$

The rare load combination (see further) is used in case of ULS or irreversible SLS when no exceedance of a limiting value is allowed.

b) The **frequent value** $\psi_1 Q_k$ has a limited exceedance probability during the reference period. In case of buildings this probability is commonly chosen 1 %. The frequent value is used for the verification of ULS in which accidental actions occur and in case of reversible SLS for which exceedance of a limiting value is acceptable during a certain time span or with a certain frequency.

c) The **quasi-permanent value** $\psi_2 Q_k$ is chosen so that it is exceeded during a significant time of the reference period. Commonly this fraction is taken as 50% of the reference period (cf. median). One can also consider the mean value as the quasi-permanent value. The quasi-permanent value is used for the verification of ULS where also accidental actions can occur and in case of reversible SLS when the exceedance of a limiting value during longer time periods is accepted.

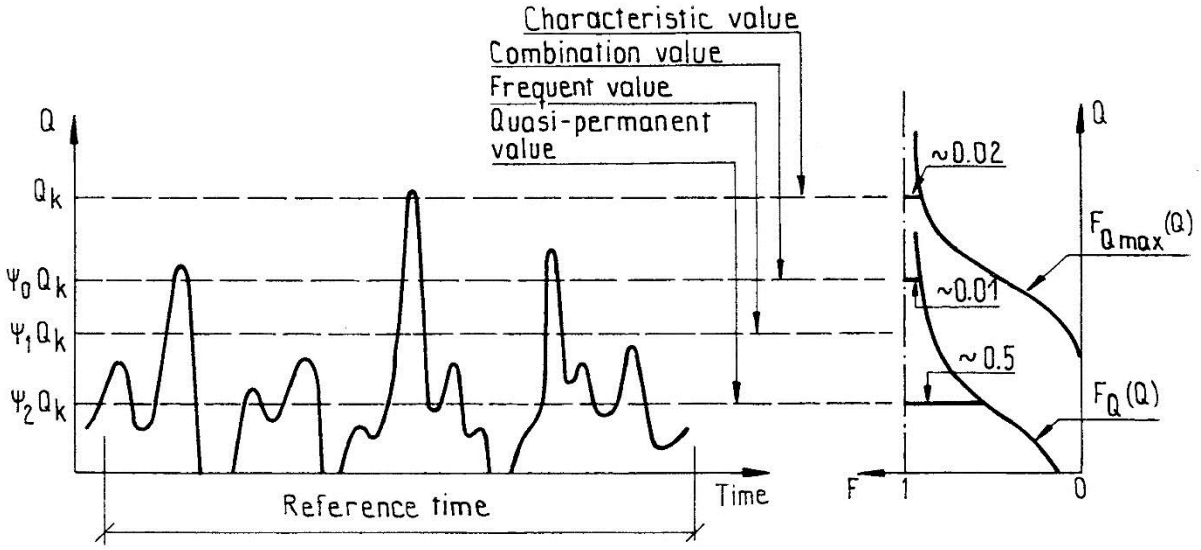


Figure 10.5 Representative values of variable actions

References

- CIRIA Report 63, "Rationalisation of safety and serviceability factors in structural codes", July 1977.
- Eurocode (2002) Eurocode Bases of structural design EN 1990: 2002.
- Hauser R., "Lessons from European failures", Concrete International, December 1979, pp. 21-25.
- Hundt J., Tzschaetzsch M., "Die Kongresshalle in Berlin-Tiergarten – Schlussfolgerungen aus einem Schadensfall", Die Bautechnik 6/1983, pp. 185-189.
- Maquoi R., Massonnet Ch., "Leçons à tirer des accidents survenus à quatre grands ponts métalliques en caisson", Tijdschrift der Openbare Werken van België, jaargang 1971-1972, nr. 2, pp. 69-84.
- Meyer H., "Qualitätssicherung – eine neue Aufgäbe im Bauwesen ?", Bautechnik, Juli 1986, pp. 217-223.
- Pidgeon N., Blockley D., Turner B., "Design practice and snow loading-lessons from a roof collapse", The Structural Engineer, Volume 64A, No. 3, March 1986, pp. 67-71.
- Proske D., "Katalog der Risiken", Dresden, 2004
- Ratay R., "Building around a building", Civil Engineering, April 1987, pp. 58-61.
- Taerwe L., "Mastering global quality by an interactive concept", Proc. IABSE Symposium Tokyo 1986 : Safety and Quality Assurance of Civil Engineering Structures, pp. 317-325.
- Vandepitte D., "Collapse of the Bridge at Pulle", Proc. IABSE Workshop, Rigi 1983, IABSE Reports, Volume 47, pp. 187-190.
- Wendrich J., "Cryogenic storage facilities with their structural analysis", Second Int. Conf. on Cryogenic Concrete, Amsterdam, 1983.
- Sharpe ME (1998) Yi Si, "The World's Most Catastrophic Dam Failures: The August 1975 Collapse of the Banqiao and Shimantan Dams," in Dai Qing, "The River Dragon Has Come!", M.E. Sharpe, New York, 1998

Appendix 10.1: Structural accidents and collapses

The bridge of Pulle

On November 12, 1966, a bridge of the highway 'Boudewijn' (E313) suddenly collapsed around midnight. The bridge was located in Pulle (Belgium) and crossed the canal Nete (Figure 10.6). The bridge was built in 1958 and was one of Belgians typical prestressed bridges with 3 spans and varying height. The abutments as well as both pillars were built on pile foundations. It is remarkable that the piles beneath the pillars were only 2,5 m long and hence were located 3 m above the lowest level on the canal cross-sectional profile (Figure 10.7). Further, apparently no bank revetment was placed.

The cause of the collapse was erosion of the east bank in the vicinity of the pile foundation. This caused a lateral movement of the feet of the east pillar which ultimately was swept away, resulting in the collapse of the bridge. The following lessons were drawn from this accident (Vandepitte, 1983):

a) When the responsibility with respect to the construction work lies in the hands of several parties, the potential risk exists that not all available information is exchanged.
b) The necessity to dispose of a clear plan that indicates the situation as-built. In case of the bridge of Pulle only original design plans were available, where the pile foundations were indicated to start much deeper.
c) The necessity of regular inspection and maintenance.

Figure 10.6 This bridge over the highway 'Boudewijn' in Pulle collapsed (1966) due to erosion of the east bank in the vicinity of the pile foundation at the side of the canal Nete

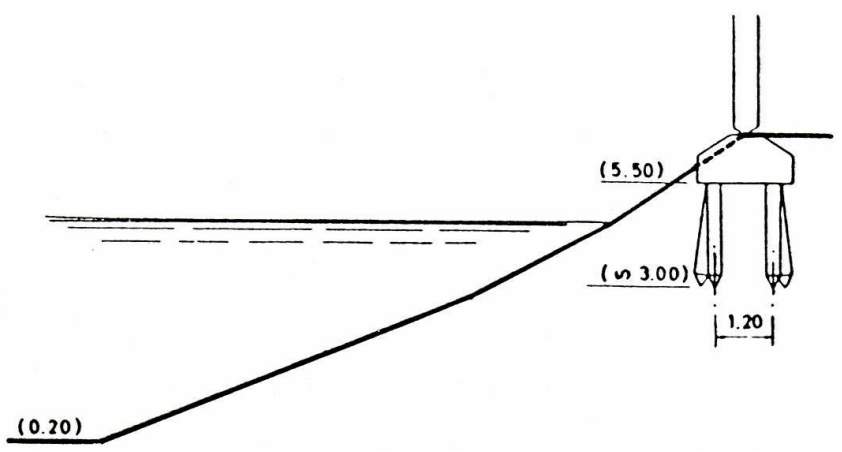


Figure 10.7 The east pillar of the bridge in Pulle that was swept away due to the erosion of the east bank in the vicinity of the pile foundation, resulting in the collapse of the bridge (Vandepitte, 1983).

The Congress Hall of Berlin

The combination of negligence during the design and errors during the construction of specific structural details can have disastrous consequences, as can for example be seen from the spectacular collapse of the Congress Hall of Berlin in 1980, 23 years after being constructed (Hundt, 1983).

The architectural shape of the building was quite original for that time, applying a hypar shell which connected 2 cantilever edge beams in the shape of arches which were clamped in two joint abutments (Figure 10.8). This elegant, but structural daring shape was only capable of withstanding symmetrical loading conditions and is not appropriate to withstand asymmetrical wind or snow loads. Additional supports or heavier abutments were not desired from an architectural point of view and as such a ring beam was constructed in the plane of the roof which was supported by concrete walls and columns that were integrated in the walls of the auditorium (Figure 10.9). From the outside, these very important structural elements were almost not visible. Hence, the roof had the shape of a shell, but did not at all work like one.

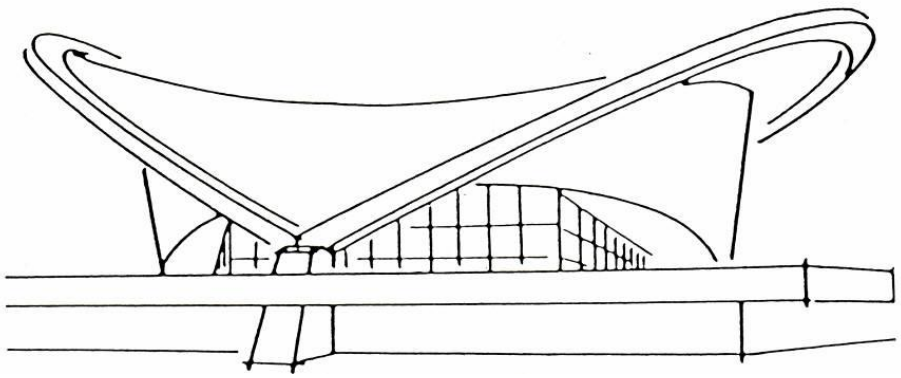


Figure 10.8 Side view of the Congress Hall in Berlin that collapsed in 1980,

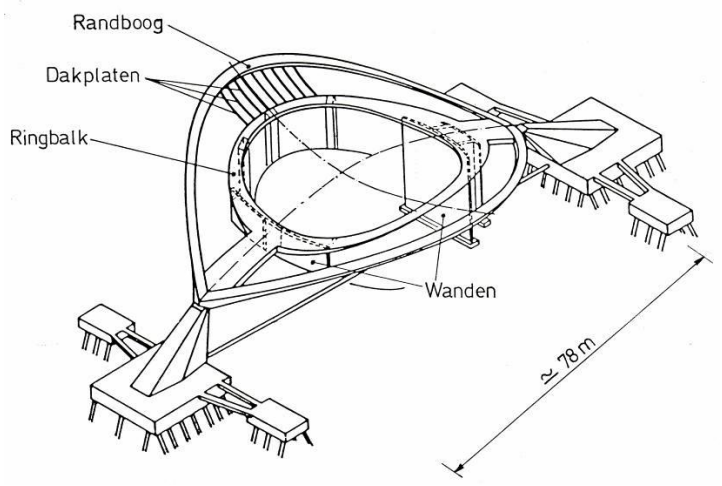


Figure 10.9 Constructive elements of the Congress Hall in Berlin (Hundt, 1983)

The part of the roof between the edge arches and the ring beam consisted of pre-stressed plates with a thickness of only 7 cm. These pre-stressed plates were anchored at one side in the arches and at the other side in the ring beam.

Due to corrosion, the tendons failed in the neighborhood of the joints at the edge arches and the ring beam and resulted in the collapse of the southern edge arch and the adjacent part of the roof. After investigation, it was found that the corrosion was due to an insufficient protection of the tendons (with concrete cover and injection mortar), cracks in the roofing and a excessive bending of the tendons due to all kinds of side effects which were not accounted for in the design.

After investigation with respect to the cause of the accident, the following problems were indicated:

- Insufficient collaboration between architects and engineers, because several structural tricks had to be performed in order to avoid deviation from the original architectural design;
- Time pressure during the design as well as during the construction phases, which interfered with a thorough study of the details and prevented thorough quality inspection at the construction site;
- The lack of regular inspection during the service period (visual inspection, non-destructive testing…) in order to detect certain deficiencies in time and to prevent failure of components.

With respect to the latter item, it should be emphasized that in Belgium the Ministry of Public Transport performs regular inspections of all infrastructures.

Tubular bridges

Between November 1969 and November 1971 a series of structural accidents have occurred during the construction of four large tubular bridges with orthotropic bridge decks. In total about 50 people lost their lives (Maquoi, 1971). It concerns the bridge over the Donau in Vienna, the bridge in Milford Haven (Great Britain), the West Gate Bridge in Melbourne and the bridge over

the Rhine in Koblenz. These four bridges all had multiple spans of which the largest span with a length that varied from 210 m (Vienna) to 336 m (cable-stayed bridge in Melbourne). Most often these bridges were erected using the cantilever construction method (sometimes with additional temporary pillars). The accidents occurred when the central span (or half of it) was almost finished. Local instability occurred, more specifically buckling of the steel tube at the location of the connection of adjacent parts (Figure 10.10). As a result, a large part of the bridge collapsed or was subjected to large deformations, except for the first case were two hinges were established.

The following possible causes were indicated (Maquoi, 1971):

- errors or oblivion in the study of the load cases (different construction phases, temperature effects,…) ;
- errors with respect to the structural concept (type and connection of the stiffeners, transversal joints in the steel plates,…) ;
- lacunae with respect to the calculations (initial deformations, random eccentricities, residual stresses, shear lag,…).

This series of accidents illustrates that when constructing more innovative structures, sometimes the boundaries of the current design principles are reached. The simple application of existing design guidelines without considering the inherent constraints can lead to structural accidents, although the designer originally felt comfortable (Pidgeon, 1986).

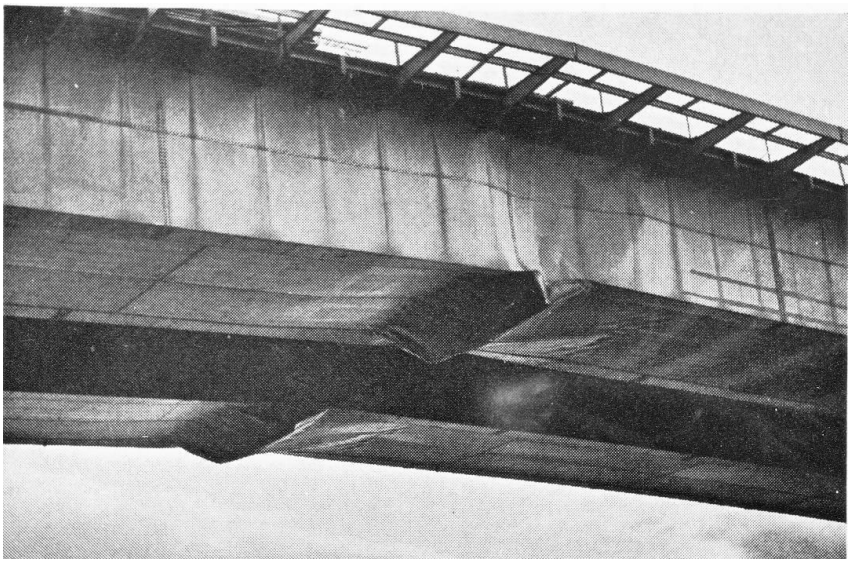


Figure 10.10 The steel bridge over the Donau in Vienna with a length of 210 m, experiencing buckling problems during the coupling of two adjacent parts (Maquoi, 1971)

Temporary structures

An important number of accidents is related to the failure of temporary structures such as supports of formworks etc. Most often, larger risks are accepted when designing such structures and consequently reduced partial factors are applied. Temporary structures are frequently composed of reusable components that can particularly be subject to wear due to transportation, assembly, loading, disassembling, storage, etc. On smaller construction sites one moreover frequently

improvises with all useful material that can be found on site in order to build temporary supports and struts.

The collapse of a temporary structure can cause the death on numerous people. In 1982, for example, 12 workers lost their lives and 15 got injured during the collapse of the support of a viaduct in East Chicago (prestressed concrete) (Ratay, 1987). The cause was the failure of a concrete support, although days before the collapse cracks appeared and were reported. This warning was however not taken into account by the authorised inspectors. Especially during economic difficult times some people prefer not to create unnecessary problems. Most of the workers that lost their lives had moreover no experience on construction sites of this nature.

Failure of dams (Proske, 2004)

On August 7, 1975, the Banqiao and Shimantan dams collapsed in the province Henan in Central China after more than 26 hours of heavy rainfall due to an unusual strong typhoon. A tsunami of 6 m height subsequently travelled with a speed of around 50 km/h towards the lower valleys and destroyed once more 61 dams and numerous dikes. In total 600 million m³ of water was released. Within 24h after the dam failure 85 000 people lost their lives. Due to starvation and diseases, the following period 145 000 additional people lost their lives. This dam failure is known as the largest technical catastrophe of all times.

Figure 10.11 The Banqiao dam after failure due to heavy rain (Sharpe, 1998)

A more recent example of a dam failure is the collapse of the Zeyzoun dam in North-Syria on June 4, 2002, during which 71 million m³ of water was released. This disaster caused the death of 22 people and made more than 3800 people homeless.

Katowice International Fair

On January 28, 2006, the roof of one of the exhibition halls of the 'Katowice International Fair' in Katowice (Poland) collapsed due to the snow load on the roof. 65 people lost their lives, more than 170 people were injured (Wikipedia). The investigations with respect to the cause of the collapse pointed out that both the managers and the architects were responsible for the accident. On the one side the managers did not make sure that snow and ice was regularly removed from the roof (although there was a budget reserved for this), which resulted in an overload of more than 100% on the roof. On the other side, also design and construction errors which can be

attributed to the architects contributed also to the cause of the collapse. Moreover, in 2000 already part of the roof collapsed due to snow load, after which the roof was repaired without any quality inspection and without any additional test or calculation with respect to the structural bearing capacity (Wikipedia).

Figure 10.12 Collapse of the roof of the Katowice International Fair due to snow load (Wikipedia)

On January 2 of the same year also the roof of the indoor ice skating rink in Bad Reichenhall (Germany) collapsed during a heavy snow shower. The accident caused the death of 15 people, 32 people were injured.

The I-35W bridge over the Mississippi (Wikipedia)

On August 1, 2007, the I-35W Mississippi River Bridge in Minneapolis (USA) collapsed during the evening rush hour. Due to the collapse 13 people died and more than 100 people were injured. The 8 lanes wide bridge carried daily 140 000 vehicles of the highway I-35W across the Mississippi river. In total the bridge was 580 m long, consisting of 14 spans of which 5 for the southern approach, 6 for the northern approach and 3 for the central span. The central span of the bridge consisted of a 140 m long steel truss arch. The other 2 adjacent spans consisted of the same type of truss with a span of 81 m each.

On January 15, 2008, the National Transportation Safety Board explained that the cause of the collapse was due to a design error. The gusset plates that realized the connection in the nodes of the truss were under-sized with respect to the load conditions of the bridge.

Figure 10.13 Collapse of the steel truss arch bridge I-35W over the Mississippi (Wikipedia)

Figure 10.14 Under-sized gusset plate of I-35W over the Mississippi (Wikipedia)

Leaning Tower of Pisa

The failure of a structure does not necessarily mean that the entire structure or part of it will collapse. In general it is related to the observation that the structure cannot fulfil one or more of its intended functions. Hence, failure can also be related to phenomena such as excessive deformations, local damages or uncomfortable vibrations.

A remarkable example is the leaning Tower of Lisa of which construction started in 1173 (Figure 10.15). Apparently, the failure probability is currently much larger than what is found acceptable in any standard. However, the tower can still fulfil its intended function as ‘campanile’. Because it has become a major tourist attraction due to the local weakness in the soil conditions, the use of the tower is many times larger than what the medieval builders had ever dreamt of. This exceptional case illustrates the relativity of safety and serviceability and shows how our value judgment can sometimes change in an unexpected way.

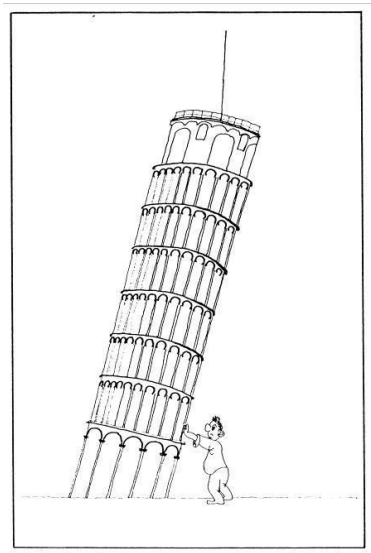


Figure 10.15 The relativity of safety ! (Taerwe, 1987)

Appendix 10.2: Classes for design supervision and inspection

In addition to a differentiation based on consequence class, one can also introduce a differentiation with respect to the design supervision (quality control). In EN 1990 (Annex B), one considers three design supervision levels (DSL) as indicated in Table 10.7.

Table 10.7 Design supervision levels (DSL)

DSL	**Description**	**Minimum recommended requirements for checking of calculations, drawings and specifications**
DSL3	Extended Supervision	Third-party checking: checking performed by an organisation different from that which has prepared the design
DSL2	Normal supervision	Checking performed by different persons than those originally responsible and in accordance with the procedure of the organisation
DSL1	Normal supervision	Self-checking: checking performed by the person who has prepared the design

In principle, DSL1, DSL2 and DSL3 correspond to RC1, RC2 and RC3 respectively. One can also obtain a classification of designers and supervisors based on their competence and expertise.

With respect to execution also different inspection levels (IL) can be introduced as indicated in Table 10.8.

Table 10.8 Inspection levels

IL	Description	Requirements
IL3	Extended inspection	Third-party inspection
IL2	Normal inspection	Inspection according to the procedures of the own organization
IL1	Normal inspection	Self-inspection

In principle IL1, IL2 and IL3 correspond to RC1, RC2 and RC3 respectively.

Chapter "Planning and Maintenance"
Authors: J.K. Vrijling, S.N. Jonkman

This chapter has been compiled using materials from the CUR book (1997; also version 2015).

Contents

11 Planning and Maintenance

11.1 Introduction

This chapter concerns the application of some of the theory introduced in previous chapters. First it will be demonstrated how estimates of uncertainties in project costs and durations can be made (section 11.2). Consequently, the effects of maintenance and inspection on system reliability will be treated in section 11.3. In these cases it is possible to change or correct the loads or strength of a system during the lifetime and such systems are referred to as "correctable". Examples of correctable systems include the management and realisation of projects and the execution of maintenance in an arbitrary system to limit the risks.

11.2 Probabilistic budgeting and Time planning

For further background information on this topic, see the paper by Vrijling and van Gelder on this topic that will be included as additional reading materials on blackboard.

11.2.1 Introduction

For virtually all realisation processes the required time span cannot be determined exactly beforehand. The amount of time required is the result of a statistical process that is determined by different uncertainties. In a statistical schedule the available or desired duration of the project is considered as the strength and the duration of the activities is considered the load.

The probability distribution and the accompanying parameters of the duration of an activity can, for example, be determined from historical data or by making estimates based on experience. The probability distribution that describes the time span is bounded on the left side, because there is always a minimum realisation time. In practice, the probability that an activity takes longer than planned seems to be greater than the probability that the activity takes less time, as many infrastructure projects are characterized by cost and budget exceedances (Flyvbjerg, 2003). The probability density function of the required time span is therefore often skewed to the right.

If the probability distribution of the duration of an activity cannot be determined from statistical material, the triangular distribution is often used to describe the duration of an activity (see section 2). The three parameters of the distribution can be determined by an optimistic, a pessimistic and a most realistic estimate of the duration of an activity (see Figure 11.1).

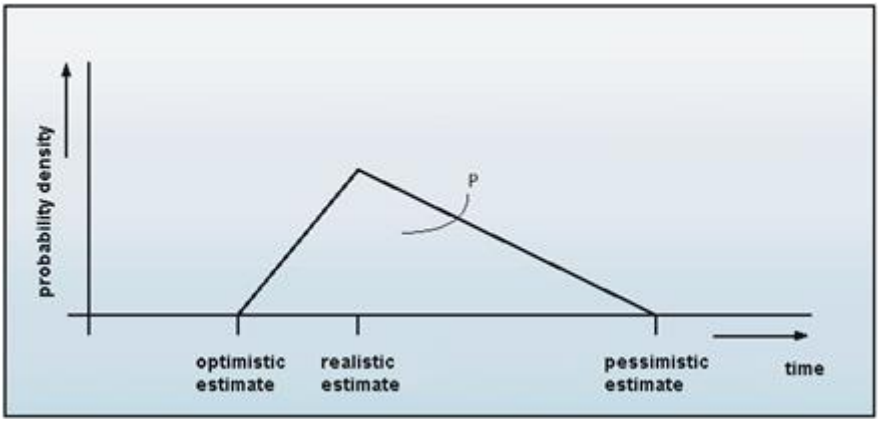


Figure 11.1 Triangular distribution for the description of the activity duration.

The probability density function of an activity is not always continuous. For example, if a special undesired event occurs, an extra effort and thus extra time may be necessary. The occurrence of the special event and the resulting extra effort are usually random variables. An example of an activity with a discontinuous probability density function is a process that is inspected after completion. If the activity is approved, the process is completed. If the activity is disapproved (special event), extra time must be spent. An example of such a probability density function is sketched in Figure 11.2, in which P is the probability that the product is rejected.

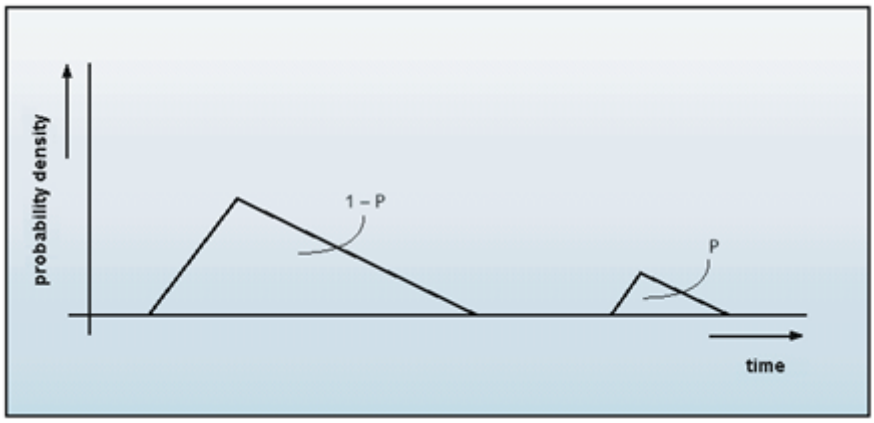


Figure 11.2 Discontinuous probability density function for the time span.

Because uncertainties in the execution period of realisation processes almost always have an effect on the costs, the approaches to assess uncertainties in project schedules and costs are similar.

The uncertainty in time can be assessed using a probabilistic schedule. Before the start of the project, the scheduling offers an insight into the probability distribution of the total duration of the project. During the project the scheduling can be used to test the progress and to undertake timely corrective measures.

To draw up a schedule, a project is divided into partial activities. The level on which the project is elaborated into partial activities, depends on the project phase. As the project progresses, a revision of the schedule can be made. If the schedule is developed in greater detail, the uncertainties concerning the total project duration decrease. However, a spread (or variation) remains around the estimated total duration of the project.

In a deterministic schedule the total duration of a project is determined by the sum of the execution periods of all activities on the critical path. An activity is on the critical path if a change of the execution period of this activity has a direct effect on the total duration of the project.

In its most simple form, a schedule consists of a succession of a number of realisation processes. Figure 11.3 shows four activities that have to be executed consecutively to complete the total project.

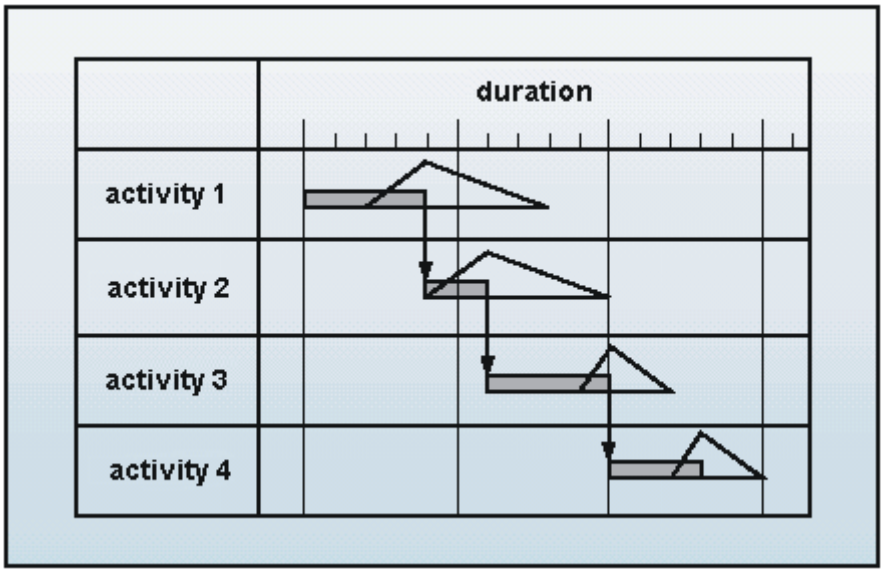


Figure 11.3 Schedule with four series activities.

The calculation rules for a serial system lead to the total duration of the project, which equals:

$$T_{project} = T_{act_1} + T_{act_2} + T_{act_3} + T_{act_4} \tag{11.1}$$

Calculating the total duration of the project becomes less simple if certain activities are undertaken simultaneously. According to the schedule sketched in Figure 11.4, activities 2 and 3 start simultaneously, after activity 1 has been completed.

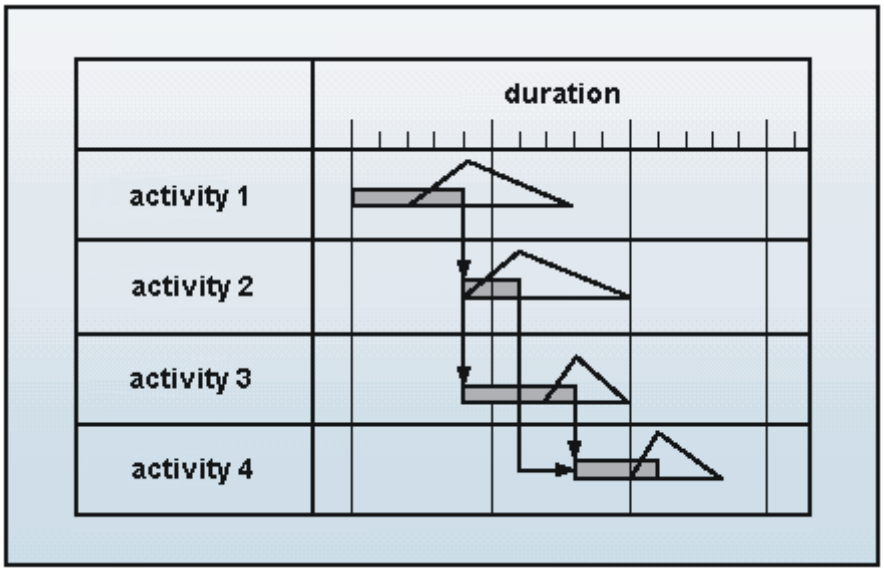


Figure 11.4 Schedule with serial and parallel activities.

Unlike in a deterministic schedule, the critical path in a probabilistic schedule cannot be designated beforehand, due to the spread of the duration of the activities. Each path can be the critical path with a certain probability.

The total duration of the project follows from a combination of the calculation rules for serial systems and for parallel systems, in this case:

$$T_{project} = \max\begin{pmatrix} T_{act_1} + T_{act_2} + T_{act_4} \\ T_{act_1} + T_{act_3} + T_{act_4} \end{pmatrix} \quad (11.2)$$

The total project is completed before a certain time ($T_{desired}$) if all possible paths are completed before this point in time. The probability of exceedance of a certain duration can be calculated for every path. The probability that a project is completed within a certain time span, determined beforehand, is the probability that all paths are completed prior to that particular final date:

$$P(T_{project} \le T_{desired}) = P(path_1 \le T_{desired} \cap path_2 \le T_{desired} \cap ... \cap path_n \le T_{desired}) \quad (11.3)$$

When there is dependence of the various paths, the Monte Carlo simulation is most suitable for the calculation of this probability (see chapter 5 on level III methods such as Monte Carlo simulation).

Figure 11.5 below gives an example of the outcomes of a probabilistic cost estimate. It shows the cumulative distribution function of the estimated costs for a given project. In general, the $p70$ value is used for budgeting purposes. It is estimated that in 70% of the cases, the project costs will be below the $p70$ value. The probability of exceedance of the associated project costs is 0.3.

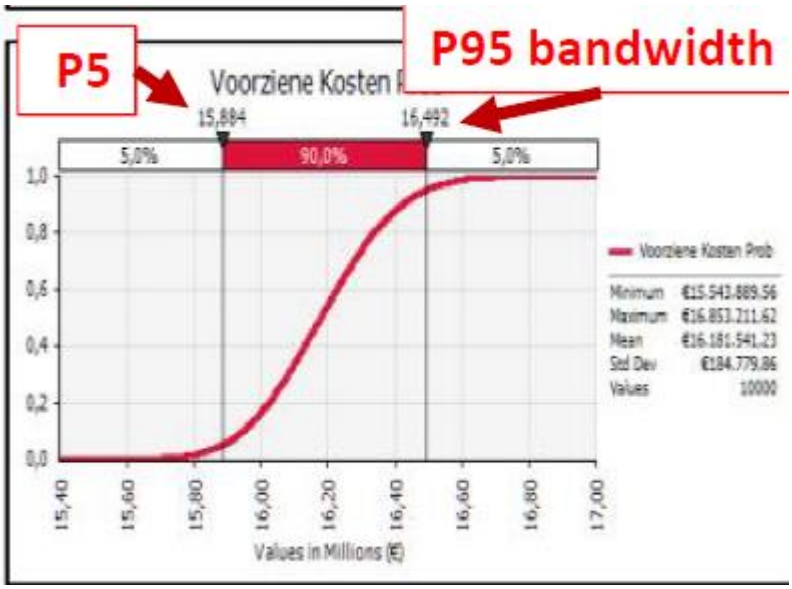


Figure 11.5 Example of a probabilistic cost estimate of a project.

11.2.2 Influence of corrective measures on duration and costs

If the available time for a certain activity is greatly exceeded, corrective measures will have to be taken to prevent exceeding of the total available time. The faster realisation of the subsequent activities, for instance by assigning a double shift or extra equipment can offer a solution. To include corrective measures in the schedule the "corrective" measures have to be included beside the "normal" activities. For the calculation of the probability distribution of the total duration of a project, that includes corrective measures, one can also use the Monte Carlo method.

Example 11.1

One wants to determine the probability distributions of the duration and of the costs of a project. The project consists of three partial activities. Depending on the execution period of activity 1, two scenarios are possible for activity 2. If activity 1 does not exceed a boundary value established beforehand, a normal execution time is kept for activity 2. If the boundary value is exceeded, activity 2 is undertaken more quickly.

The required period of time per activity is modelled using three estimates, an optimistic, a pessimistic and a most realistic estimate. It is assumed that the probability density for the activities has a triangular distribution.

If activity 1 is completed in less than 6 days, the normal execution period is applied for activity 2. If activity 1 takes longer, activity 2 is executed faster. The project schedule is shown in Figure 11.6.

Furthermore, the costs per activity are determined according to:

$$K_{act_i} = K_{fixed_i} + K_{\mathrm{var}_i} T_i \qquad (11.4)$$

in which:

- $K_{act\,i}$ are the costs of activity i ;
- $K_{fixed\,i}$ are the fixed costs of activity i ;
- $K_{var\,i}$ are the variable costs of activity i ;
- T_i is the duration of activity i .

Table 11.1 gives an overview of the fixed costs and the variable costs. The project is to be realised within a maximum duration of 14 days and a maximum cost price of EUR 300. The effect of a faster execution of activity 2 is to be investigated. The probability distributions of both the time span and the costs are estimated by simulation of random variables (see Figure 11.7) for the cases with and without faster realisation of activity 2.

This figure shows that the probability of time exceedance decreases, but that the probability of exceeding the costs increases if activity 2 is executed more quickly and when activity 1 takes longer than six days.

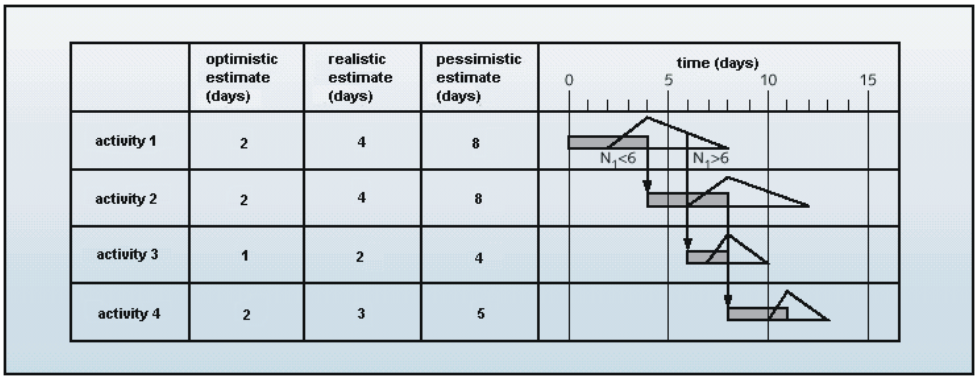


Figure 11.6 Project schedule with corrective measures.

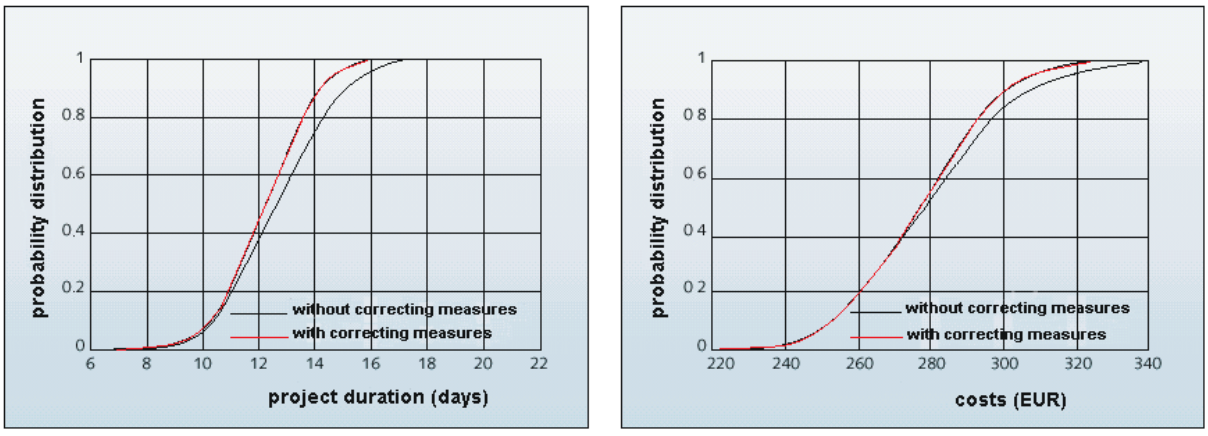


Figure 11.7 Probability distribution function of time span and costs.

Table 11.1 Fixed and variable costs per activity.

Activity	K_{fixed} (EUR)	K_{var} (EUR)
activity 1	40.-	10.-
activity 2 normal execution period	30.-	10.-
activity 2 accelerated execution period	30.-	25.-
activity 3	80.-	10.-

The occurrence of calamities during the realisation of a project usually has a great effect on the project costs and the duration of the project. Because calamities have a small probability of occurrence, they are absent in the "normal" schedule. The probability distribution calculated on grounds of this schedule is therefore a conditional probability distribution; it gives the probabilities of non-exceedance of the duration of the project if no calamities occur. The occurrence of calamities and the possible consequences for time and costs have to be considered separately, for example in an event tree, in which separate schedules are associated with the different events.

11.3 Maintenance

11.3.1 Introduction to maintenance strategies

Maintenance includes all activities aimed at maintaining of or returning to the technical state which is considered necessary for the system to fulfil its function. To minimise the maintenance costs, the optimal maintenance strategy is sought. The following classification of strategies is taken from mechanical engineering maintenance theory:

1. Curative maintenance: - failure-dependent maintenance.
2. Preventive maintenance: - work-dependent maintenance;
   - state-dependent maintenance.

According to **failure-dependent maintenance** an object is not replaced or repaired until it can no longer fulfil its function (see Figure 11.8). Thus, repair takes place after failure. The life span of the object is completely utilised. It is accepted that the object fails and may not be available during repair. Often this form of maintenance is not acceptable, for example because the consequences of failure are great. However, this type of maintenance can be applied for non-integrating parts (parts that do not contribute to the probability of failure of the system as a whole). The consequences of failure are then minor (provided repairs or replacement are not delayed too long).

According to **work-dependent maintenance**, maintenance is carried out after a period with a certain number of work units, established beforehand. The costs of maintenance and the risk generally determine the length of this period. The life span of the object is not fully utilised.

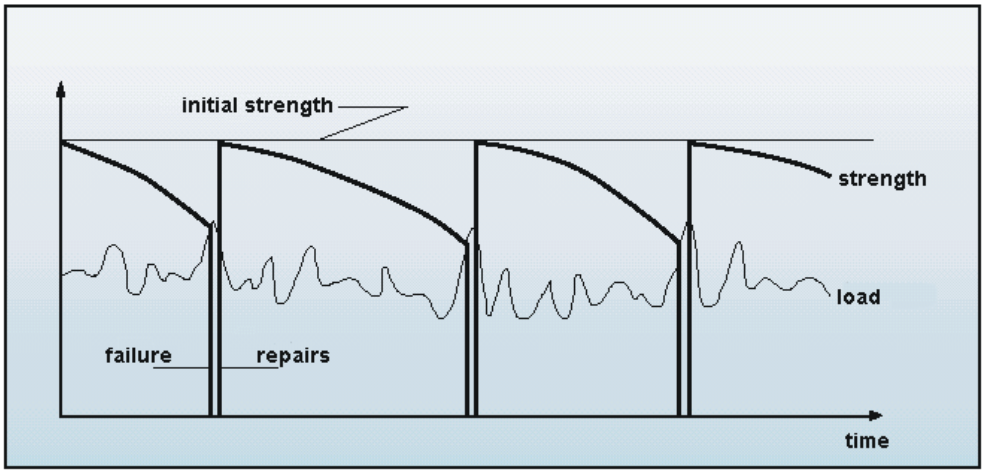


Figure 11.8 Possible course of the strength for failure dependent maintenance.

In mechanical engineering this type of maintenance is applied if the work units can be registered easily, for example with a mileage meter, product counter et cetera.

In general it is not possible to register all loads for a system. In this case the loads over a period are considered random variables. Subsequently, an estimate is made of the life span and a time for repairs is determined, which corresponds with a sufficiently low probability of failure and with minimal costs. It is therefore better to speak of time dependent maintenance (see Figure 11.9). This then involves a given time interval for maintenance activities, e.g. every week, month or year.

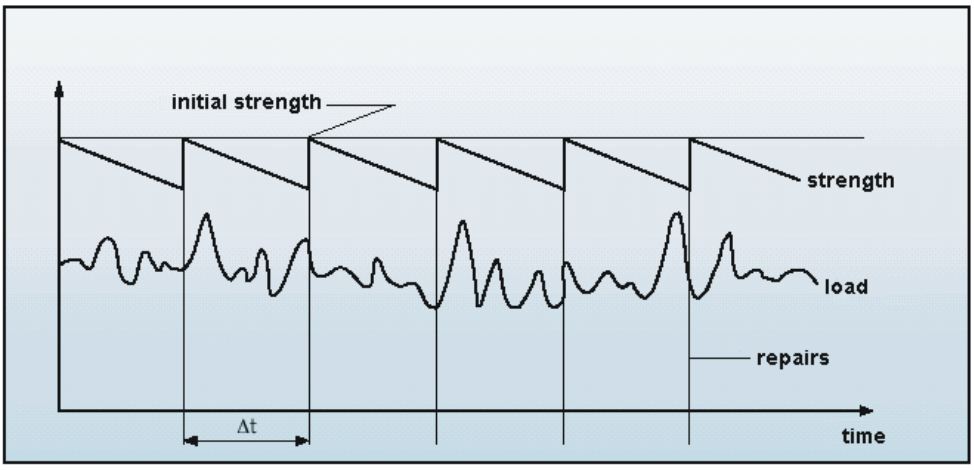


Figure 11.9 Possible course of the strength for time dependent maintenance.

If the loads which cause deterioration are registered, one can decide to carry out maintenance after an extreme load or after a certain amount of load (cumulative load). This involves load dependent maintenance and a threshold or standard for critical loads.
For a case where cumulative loads play a part (fatigue) a possible course of the strength and load is shown in Figure 11.10.

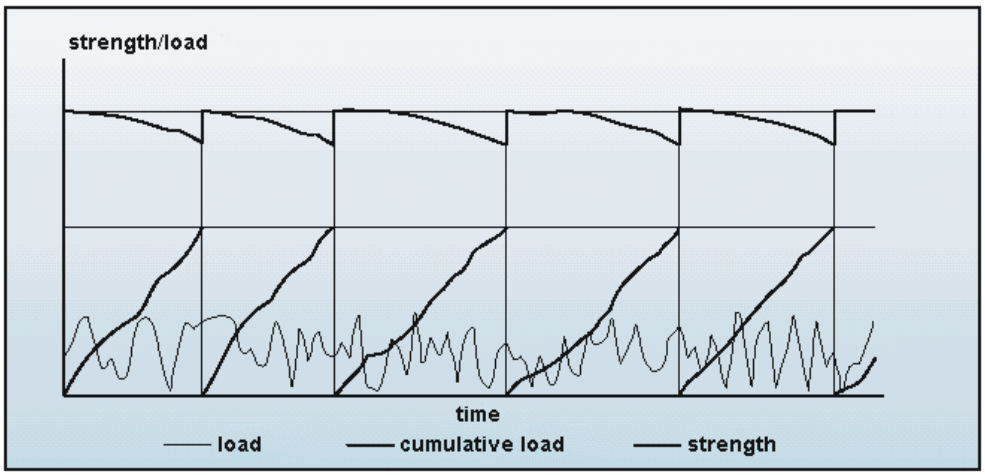


Figure 11.10 Possible course of the strength for load dependent maintenance.

For **state-dependent maintenance** the condition of the object is determined by regular inspections. Based on the observations repairs are decided necessary or not. The inspection intervals can be constant or dependent on the condition of the object. There must be condition parameters that can be observed and reflect the condition of the system.

The probability of failure during the period between inspections must be sufficiently small. Generally, the life span can be better utilised than in a work-dependent maintenance strategy, but the costs of the inspections have to be taken into account.
This form of maintenance yet again requires setting thresholds. These thresholds relate to (see Figure 11.11):

- a limit state that leads to the decision to increase the inspection frequency (warning limit);
- a limit state that leads to the decision to carry out repairs (action limit).

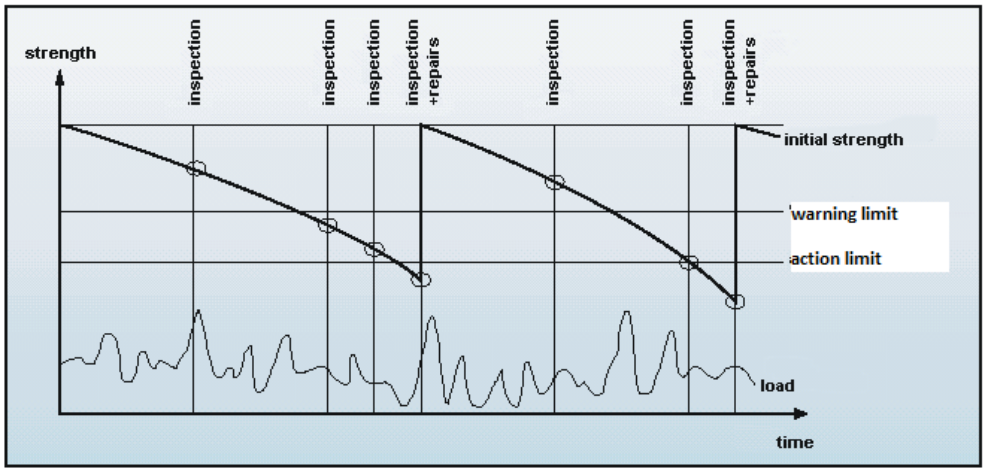


Figure 11.11 Possible course of the strength for state dependent maintenance.

This limit in fact concerns standards for the strength. These are the result of an optimisation of the maintenance or they are associated with a socially accepted probability of failure in a year.

The choice of which maintenance strategy to use depends on factors such as:

- predictability of the life span of the object;

- consequence of the failure of the object;
- costs of replacement or repairs;
- costs of inspection;
- Whether condition of the structure can be observed (damage or deterioration).

A first comparison of the different strategies concerning the applicability can be made with the help of Figure 11.12.

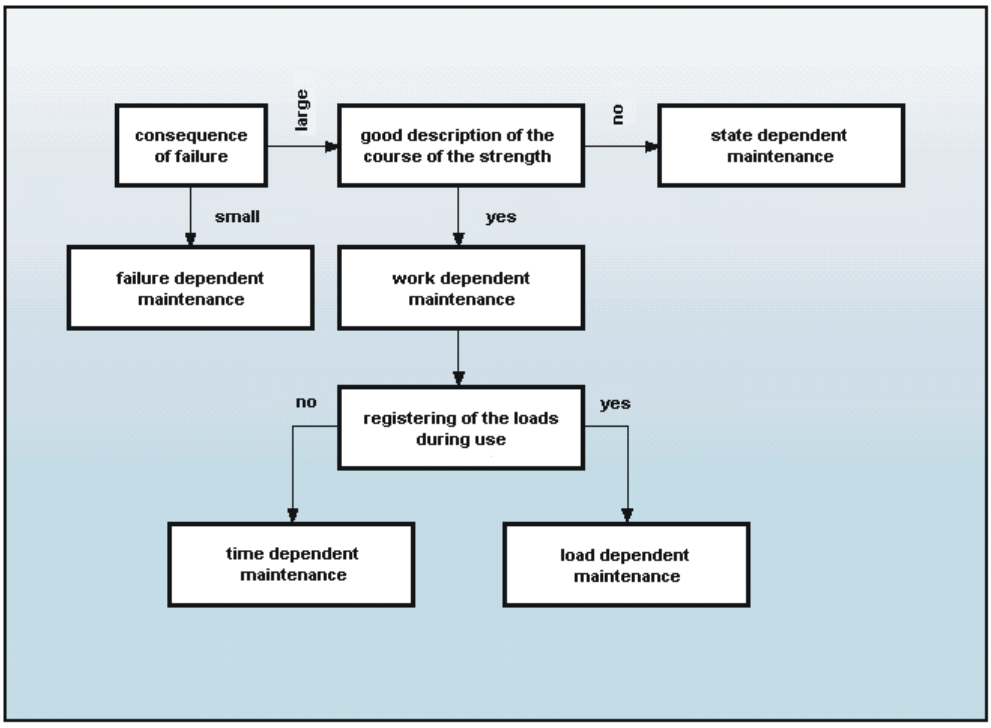


Figure 11.12 Rough choice of the maintenance strategy.

Often a combination of two or more strategies offers a better result than simply applying a strategy selected according to Figure 11.12. For instance, a schedule can be made according to time dependent maintenance strategy. It can be adjusted on the basis of the observed loads, while the decision to carry out maintenance depends on the inspected strength. Based on this it can be said that the boundaries of the application areas of the different strategies are not strict.
Choosing between time dependent and state dependent maintenance is less simple. Completely time dependent maintenance will be applied if inspection is not possible or if inspection is expensive relative to repairs. Completely state dependent maintenance will be used if absolutely no prognosis can be made for the course of the strength in time or if inspection is very simple and thus cheap.

An important aspect of state dependent maintenance is the collection of data concerning the course of the strength in time. This allows increasingly better planning of maintenance or inspections.

11.3.2 Effects of maintenance on the risk

The influence of the maintenance on the course of the strength in time is simple if maintenance restores the strength to its initial level. Figure 11.8 to Figure 11.11 give the effect of a number of

maintenance strategies maintenance on the strength. Figure 11.9 shows a constant maintenance interval, which means that the time interval between the maintenance periods is constant. The deterioration process is probably the same in every interval and thus there is no reason to vary the duration of the maintenance interval.

It can be the case that the deterioration process is not equal in the different intervals, for example because the degree of deterioration depends on the maintenance that is carried out. A good example of this is the maintenance of a dike.

Raising a dike body results in settlements. The settlements depend on the extent to which the dike has been raised and on the stresses in the subsoil. The stresses depend on previous maintenance. Figure 11.13 schematically presents the course of the settlements after carrying out maintenance.

The line drawn in Figure 11.13 shows the course of the expected value of the crest height of the dike. The dotted lines show the spread around the average crest position. This spread is a result of the uncertainties in the soil parameters and uncertainties in the method used to calculate the settlements. The figure shows that the spread also depends on the maintenance. This originates in the fact that calculations can be calibrated to the observations by adjusting the deterioration model. The way in which the deterioration model can be adjusted is described in the following section.

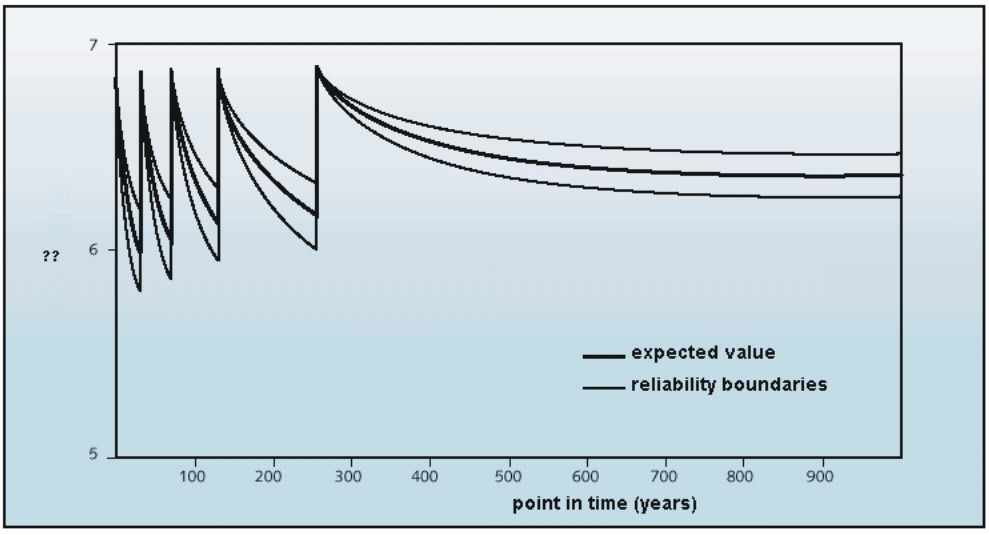


Figure 11.13 Strength reduction as a result of settlements.

Clearly, decreasing the maintenance interval reduces the risk but increases the maintenance costs. Usually an optimum is sought, for which the sum of the discounted value of the maintenance costs and the risk is lowest.

11.3.3 Effect of inspections

The quantitative part of this section will not be part of the examination.

Inspections can help determine the strength of a system at a given point in time. For this reason condition parameters are defined. Essentially, the condition parameters are no different from resistance parameters, as defined for the deterioration model. However, the condition parameters must be measurable with a reasonable accuracy. The strength, which is derived from the

condition parameters, may not contain an inaccuracy that is too great either. If the required accuracy cannot be achieved, inspection does not provide a better insight into the condition of the system. In that case better measurement methods or different condition parameters must be sought. Figure 11.14 illustrates this by means of two inspection methods (A and B) which can be used to update the original (or a priori) estimate of the strength according to an ageing model. The probability density of the strength reveals that method B hardly yields improvement of the knowledge concerning the strength, as the standard deviation is still large. Method A does offer a better insight into the strength of the system.

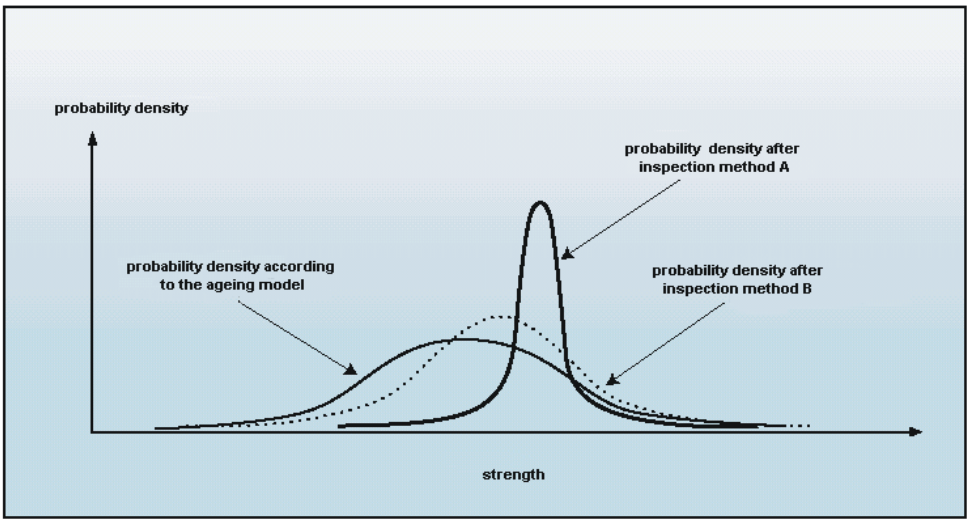


Figure 11.14 Probability density of the strength according to the deterioration model and two different inspection methods.

Clearly, inspection has no influence on the deterioration process, but is does influence the available knowledge regarding the modelling of the process. On these grounds one can decide to adjust the estimate of the condition and the deterioration model. If adjusting the deterioration model should is really the right decision, strongly depends on the type of deterioration process.

In the previously mentioned settlement process of a dike body, the placement and reading of a settlement measuring device can provide the necessary information to calibrate the parameters in the settlement formulae. The observation does not only provide information on the the strength of the dike (at the time of reading the device), but it can improve the deterioration model as well. This leads to smaller uncertainties in the model, which have to manifest themselves in a smaller variance of the strength of the model.

If the deterioration model is based on statistical data, a special updating technique can be used to adjust the model. Suppose that at a certain time it is decided to carry out an inspection. The probability density of the strength at that moment is still merely determined by the deterioration model. This probability density is called the "a priori" probability density of R(t). After inspection it is possible to create a better description of the strength by using the inspection results. The method according to which this is done is known as the "posteriori analysis". The probability density function of the strength that is adjusted to the inspection results is known as the a posteriori probability density.

The a posteriori probability density is determined with:

$$f_R^{II}(R) = N \cdot L(X \mid R) f_R^I(R) \tag{11.5}$$

in which:

- $f^{II}{}_R(R)$ is the a posteriori probability density;
- $f(R)$ is the a priori probability density;
- $L(X \mid R)$ is the likelihood function;
- X is the vector with the inspection values;
- N is the standardisation constant $= \dfrac{1}{\int_{-\infty}^{\infty} L(X \mid R) f_R^I(R) dR}$

The likelihood function gives the probability density function of the vector X for a given value of R. In formula this is:

$$L(X \mid R) = f_I(X_1 \mid R) f_I(X_2 \mid R) \ldots f_I(X_n \mid R) \tag{11.6}$$

in which:

- $f_I(X_i \mid R)$ is the probability density function of observation X_i for a given strength R;
- X_i is the observation of the strength.

After determining the a posteriori probability density, the strength at the moment of inspection can be estimated better. Sometimes it is then possible to adjust the deterioration model, as is demonstrated in Figure 11.15. This can best be clarified with an example.

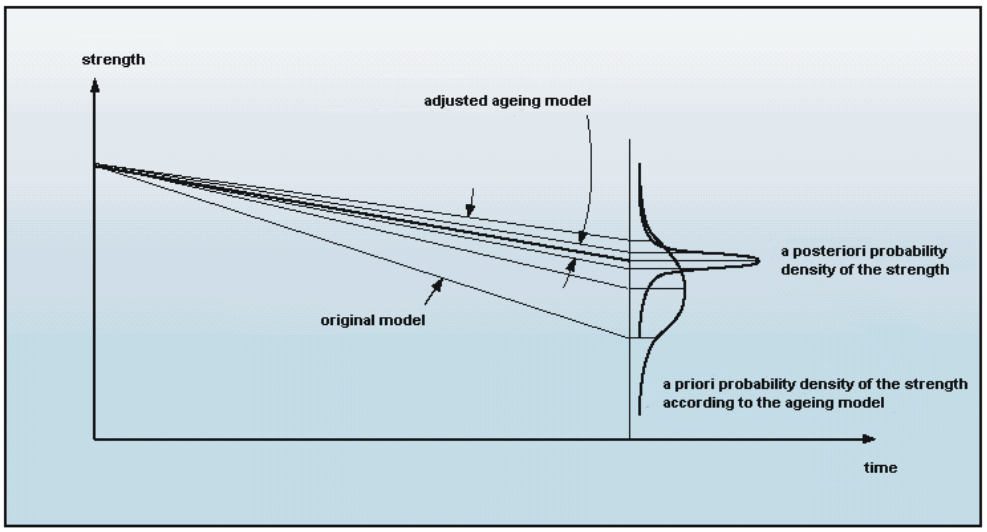


Figure 11.15 Adjustment of a deterioration model.

Example 11.2

Suppose that the strength of a steel element in a structure is more or less known at time $t = 0$. On average this is 130kN with a standard deviation of 1 kN.

As a result of corrosion the strength decreases with an average rate of 10 kN/year. The standard deviation of the reduction of the strength is $\sqrt{5/3}$ kN/year. It is assumed that the corrosion is independent of the time and the earlier corrosion. The parameters which are of importance for the description of the corrosion are assumed to be constant, which means that the corrosion rate is also constant.

After three years the expected value of the strength is: $\mu_R(3)=130-3\times10=100\text{kN}$. The standard deviation of the strength is:

$$\sigma_R(3) = \sqrt{1+\left(3\sqrt{5/3}\right)^2} = 4\,\text{kN} \tag{11.7}$$

It is assumed that both the initial strength and the rate of reduction of the strength are normally distributed. This implies that the strength at time t is also normally distributed. The a priori probability density of the strength after 3 years is:

$$f_R^I(R) = \frac{1}{4\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{R-100}{4}\right)^2} \tag{11.8}$$

After three years the strength of the system is inspected using an unspecified method. The method does not entirely exclude measurement errors. Thus, the strength cannot be determined exactly. For a given strength, the inspection method can be used to find values that are normally distributed around the actual strength. For this it is assumed that the values are obtained from independent measurements. The standard deviation of the distribution is 2kN. The strength is measured five times. The measurements are respectively 103, 102, 105, 104 and 103 kN. The a posteriori probability density is:

$$\begin{aligned} f_R^{II}(R) &= \frac{1}{N}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\left(\frac{1}{2}\frac{(103-R)^2+(102-R)^2+(105-R)^2+(103-R)^2}{4}+\frac{(R-100)^2}{16}\right)} \\ f_R^{II}(R) &= \frac{1}{0.87287\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{(103.24-R)}{0.87287}\right)^2} \end{aligned} \tag{11.9}$$

This is a normal probability density function with:

$$\mu_R(4) = 103.24\,\text{kN} \text{ and } \sigma_R(4) = 0.87287\,\text{kN} \tag{11.10}$$

The variance of the a posteriori probability distribution is therefore significantly lower than the variance of the a priori probability distribution.

This example assumed that the inspection observations were independent. This is not always the case. In such cases more inspections do not yield a proportional improvement of the knowledge of strength.

Recommended Literature and further reading

- Benjamin, J.R. and A. Cornell, Probability, statistics and decision for civil engineers. McGraw-Hill Book Company, New York, 1970.
- BUNT, B.P. VAN DEN, Relation between Sand Content of Coastal Profiles and Postions of High water and Low Water Lines (in Dutch: "Verband tussen zandinhoud van kustprofielen en posities van hoog- en laagwaterlijnen"). Delft University of Technology, Delft, June 1990.
- Flyvbjerg B, Skamris Holm M.K., Buhl S.L. (2003) How common and how large are cost overruns in transport infrastructure projects? Transport Reviews Vol. 23 Issue 1, pp.71-88
- GYSBERS, F.B.J., Optimisation of Maintenance (in Dutch: "Optimalisatie van onderhoud"). TNO-IBBC report B-86-549, Delft, April 1987.
- VROUWENVELDER, A.C.W.M. en S. LENOS, Optimisation of Maintenance – part II (in Dutch: "Optimalisatie van onderhoud - deel II"). TNO-IBBC report B-87-717, Delft, December 1987.
- KOK, M., Oosterschelde Storm Surge Barrier, Optimisation of Inspection and Maintenance of Wet Structures (in Dutch: 'Stormvloedkering Oosterschelde, optimaliseren van inspectie en onderhoud natte werken"). WL|Delft Hydraulics, report Q 606, Delft, 1988.
- KOK, M., Maintenance, Methods for Rational Maintenance of Civil Structures (in Dutch:"Onderhoud, methoden voor rationeel onderhoud van civiele constructies"). WL|Delft Hydraulics, Delft, September 1990.
- KUIJPER, H.K.T., Mainenance in Hydraulic Engineering, Economically Sound Planning of Maintenance (in Dutch:"Onderhoud in de waterbouw, economisch verantwoord plannen van onderhoud"). Delft Universty of Technology, Delft, 1992.