Quality Management in the Bosch Group | Technical Statistics

13. Methods of design for reliability, verification and validation

BOSCH
Invented for life

Contents

1. Introduction

Reliability, among many other product qualities, is often attributed particular value in customer surveys [1] p. 2. The guarantee of adequate product reliability is therefore a key factor for business success. To achieve this in these times of reduced development cycles yet increasingly exacting product requirements, new approaches are necessary. In addition to traditionally important experimental tests to demonstrate reliability, targeted design for reliability, combined with an arithmetical prediction of reliability and optimization, are increasingly at the center of development activities. This is reflected in the frontloading strategy, which requires work on reliability to be shifted to an earlier phase of product development, in order to reduce overall expenditure.

Furthermore, high customer expectations and global trends such as miniaturization, growing complexity and reduced weight, for example, are leading to greater performance densities and so to declining reliability reserves. This situation can often only be offset by employing more appropriate and accurate methods.

This volume is addressed to associates whose work involves design for reliability, verification and validation. It presents important methods that satisfy the requirements described above. The content should be read in conjunction with the BES Practice Description "Design for Reliability" [2], in which the basic principles of design for reliability are explained in more detail.

The layout of this volume is as follows:

- Chapter 2 introduces the basic terms and relationships, which are vital for an understanding of the explanations that follow.
- Chapter 3 is devoted to design for reliability. First, it explains the general concept and the reliability requirements. Next, it discusses quantitative and qualitative methods of design for reliability, prediction and optimization.
- Chapter 4 presents basic principles and methods of reliability verification and validation.
- Chapter 5 is concerned with questions about reliability in the field.
- The Annex contains some fundamental statistical concepts that are necessary in order to understand the relationships explained in this volume.

2. Basic principles

2.1. Definition of terms

Reliability is one of the most important product characteristics and is therefore an aspect and integral component of quality. For its part, reliability exerts an influence on further quality features, such as safety. The visual demarcations of the terms quality, reliability and safety are illustrated in Fig. 1.

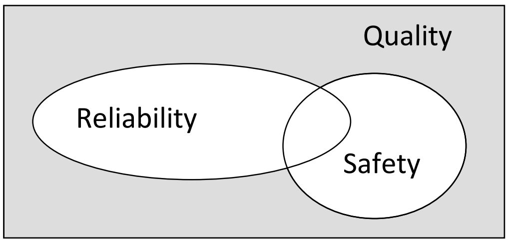

Fig. 1: Quality, reliability and safety

According to [3], reliability is an umbrella term that describes availability and its influencing factors performance reliability, maintainability and maintenance support.

In order to take account of the broad use of these terms, the hierarchical demarcations illustrated in Fig. 2 and based on the system shown in [4] are employed in [2] and in the present volume. Dependability can therefore be regarded as an umbrella term for reliability, availability, safety and security, all of which make up the attributes of dependability. This volume focuses on design for reliability for design elements and systems. Here, a "design element" denotes the smallest self-contained component that can fulfill one or more functions (Glossary BES-PE 3-2009/11). It may consist of one or several components.

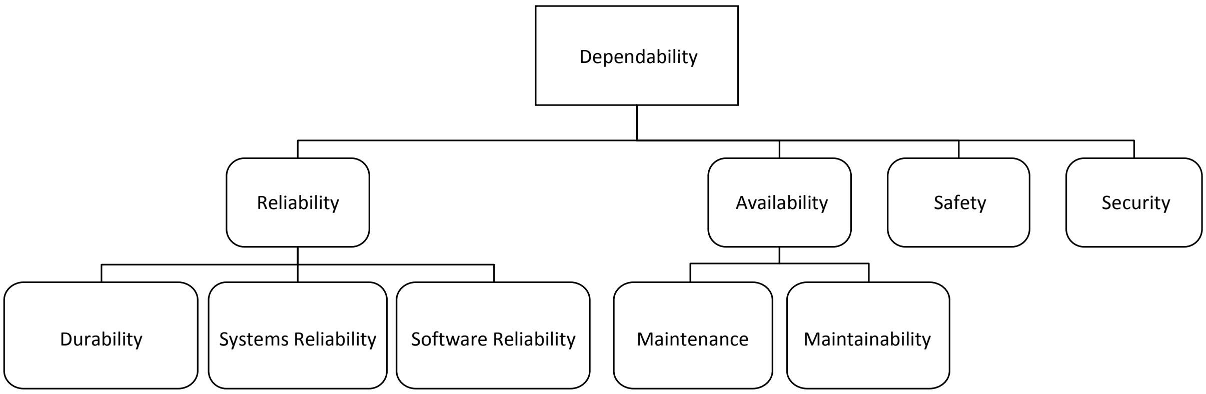

Fig. 2: Dependability and reliability

According to [2], reliability is the ability of the product to perform a required function in the predefined operating range over a defined useful time.

This useful time is divided into operating time and idle time. The operating time is the time during which the unit is functioning in accordance with its intended purpose. During idle time, on the other hand, the unit is not in operation.

The stress time is the time during which a damage mechanism (processes that lead to a gradual change in a unit's properties due to load) is acting on the unit. It generally consists of parts of the operating time and idle time, operating time does not necessarily concur with the stress time: for example, certain electronic components in a parked vehicle are still under stress or exposed to corrosion.

Lifetime is the time during which a unit can be exposed without interruption to a damage mechanism until failure. In general, developers endeavor to achieve a sufficiently long lifetime in terms of the stress time; the correlation to useful time or operating time, on the other hand, is not crucial. The ratio between lifetime and stress time can be expressed as safety versus failure. The above terms are graphically illustrated in Fig. 3.

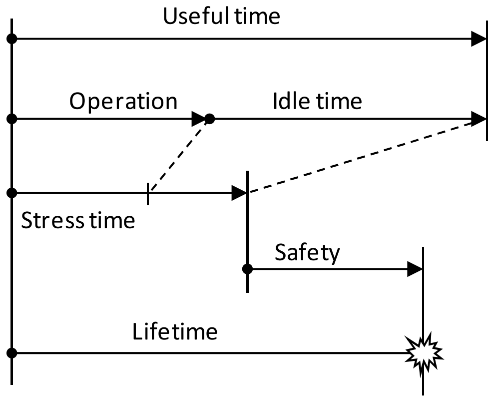

Fig. 3: Relationship between useful time, stress time and lifetime

For the sake of simplicity, hereinafter the words "time" or "duration" are used to represent all possible, damage mechanism-specific lifetime characteristics such as time (hours, months, years), travel, distance covered, mileage, cutting length of a tool, number of load cycles, actuations, switching operations, work cycles, revolutions, etc.

Design for reliability is a discipline in engineering science, which applies scientific findings to ensure that the design of the product possesses the product characteristic "reliability". This also includes incorporating the ability to maintain, test and support the product over the entire life cycle.

Where reliability is concerned, different objects can be examined:

- mechanical and electrical/electronic hardware,
- software,
- people,
- systems composed of the above units, and
- services.

This volume deals primarily with the reliability of design elements consisting of mechanical and electrical/electronic hardware in terms of durability, and the system reliability derived from these design elements, see Fig. 2.

2.2. Causes of component failure

2.2.1. Local and global approaches

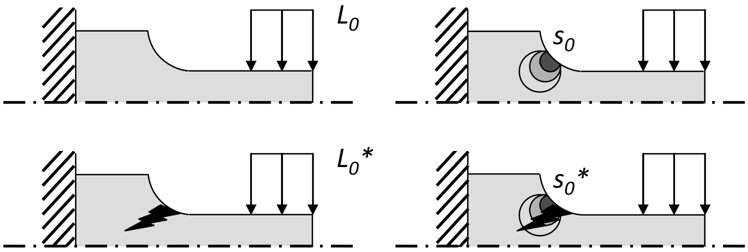

Fig. 4: Global (left) and local (right) reliability assessment concepts

Ostensibly, component failure occurs when a load exceeds the amount of load that can be withstood (load capacity) by the component. However, this statement is not really helpful, because it expresses the reasons for failure in a very superficial way and does not contribute to a deeper understanding of the causes of failure. Furthermore, we can soon observe that certain components may fail while others do not, under the same load. The cause of failure cannot therefore be found in load alone. Concepts that explain the failure of a component by comparing load and load capacity are referred to as global or load-based concepts.

A more in-depth method of observation starts with the premise that the external load at each location on the component generates a (locally variable) stress. Failure then occurs when the local stress exceeds the amount of stress that can be withstood at the location of failure (strength). Both stress and strength are dependent upon the nature of the load, as different types of load can give rise to different damage mechanisms, which are characterized by different parameters.

Concepts that are based on the local comparison of stress and strength are referred to as local concepts.

Example:

Let's take a look at a stepped bar with a fillet radius of R_0*, which is subjected to quasi-static load from an external force* L_0*, Fig. 4.*

In order to ascertain the safety of a component with respect to brittle facture by the help of a global concept, component experiments must be performed under different forces L*, and the force* L_0* *determined at which failure first occurs. This force represents the global load capacity. The quotient* L_0*/L_0 *constitutes a measure of the safety of the component in terms of failure.*

The results cannot simply be transferred to other components, however: the complete component experiments must be repeated for a different bar with a fillet radius of R_1*, in order to ascertain its load capacity* L_1**. Moreover, statements about safety cannot be made before a component is available.*

For reliability assessments using a local concept, first of all we have to find a local variable that characterizes stress as a result of the external force. In this example, this would be mechanical stress. Assuming a linear elastic characteristic, this stress can simply be calculated at the notch root (the point of the highest stress) as $s_0 = c(R_0)L_0$ *from the force* L_0*, whereby the transfer factor* $c(R_0)$ *depends on the fillet radius but not on the load. The variable* s_0 *represents the stress.*

Likewise, a local quantity can be derived that characterizes load capacity: $s_0* = c(R_0)L_0*$*, this being strength. The quotient* s_0*/s_0 *constitutes a local measure of component safety versus failure. This procedure has clear advantages: for assessing a design alternative, e.g. a bar with a fillet radius* R_1*, it would not be necessary to conduct further experiments, but simply to recalculate* $c(R_1)$*, as* $s_1*/s_1 = s_0*/(c(R_1)L_0)$.

At this point we have simplified the process by assuming that $s_1^* = s_0^*$*, i.e. all bars have the same local strength, which therefore represents a real material parameter. This does not necessarily have to apply to all damage mechanisms, however; the fundamental laws defining the dependence of strength on various influences have to be known. This is a key element of a local assessment concept.*

The above example clearly illustrates the advantages and disadvantages of the two assessment concepts. So, for load-based (global) concepts:

- The advantage is high accuracy, as all experimentally determined variables were ascertained on the component itself.
- On the other hand, transferability to other components cannot simply be assumed, and relatively little can be learned about the causes of failure. What is more, conducting the component experiment with complex load is often difficult or impossible, or too expensive.

As for local concepts, on the other hand:

- Their advantage is their suitability for use in earlier stages of the product creation process, or where component experiments are difficult, too expensive or unfeasible. In addition, they can be used to build up a deeper understanding of the causes of failure.
- As a disadvantage of local concepts, insufficient accuracy can sometimes be mentioned. Furthermore, a lot of expertise has to go into the assessment concept.

As is often the case, the solution here is to use a combination of both approaches: the strength that is to be employed with a local concept must be determined in an experiment, if possible on the component. This approach covers all influences that may not explicitly be included in an assessment concept due to insufficient knowledge.

2.2.2. Damage characteristic, S/N diagram

The deliberations above always focused on the level of load/load capacity or the local variables stress/strength. This only result in a meaningful procedure if a load is applied that either immediately leads to failure or can be withstood for an infinitely long time. Such cases are seldom of any practical meaning, particularly in the context of examining reliability.

Failure often does not occur immediately, even under constant or periodic load, but only after a certain duration, Fig. 5. Examples are metal creep at temperatures above 40 % of the melting temperature, or crack propagation in brittle materials. This model assumes that the load provokes a certain damage that increases over time (and possibly depending on location), which may lead to failure. This must be characterized by a damage parameter that conforms to the physical laws of the local processes taking place.

Thus, the load is a two-dimensional quantity; it features both a (possibly variable) level, and a duration. Consequently, when stating the level of load the duration of load must also be mentioned, and if the level is variable the complete load-time curve (often referred to as a load-time series) must be indicated.

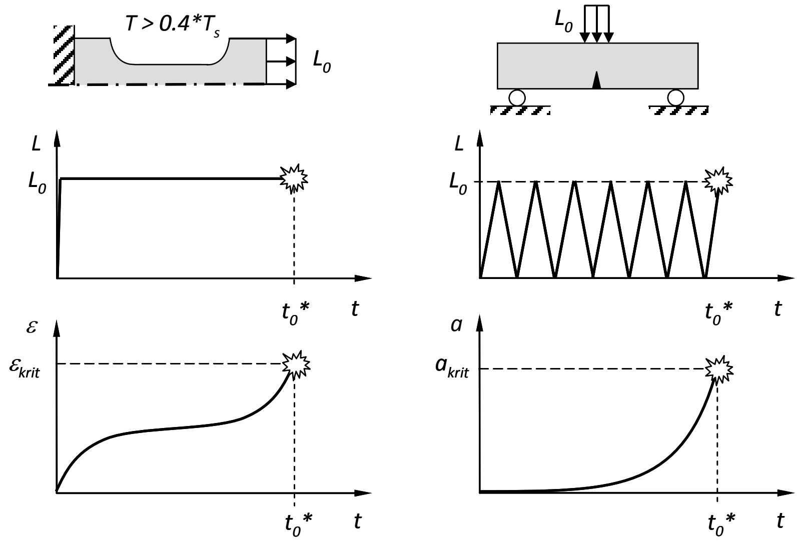

Fig. 5: Delayed fracture as a result of progressive damage. Left: Metal creep under constant load (damage parameter elongation ε). Right: Crack propagation (damage parameter crack length *a*). *t** indicates the lifetime.

Different lifetimes can be achieved with different load levels. To put it simply, the information about the dependence of lifetime on the load level, particularly where load is uniform, can only be illustrated in a *t-L* diagram by means of corresponding lifetime points (t^*,L_0), whereby the nature of the load must be described by other means (e.g. constant, uniform, cyclic with zero underload, etc.). The lifetime points can also be connected via a curve (assuming continuous behavior), to illustrate the dependence of lifetime on the load level, Fig. 6. This dependence is often described by means of a power function, which appears in logarithmic form as a straight line. Under uniform cyclic loading, this type of representation is traditionally referred to as a Wöhler curve (or S/N diagram) .

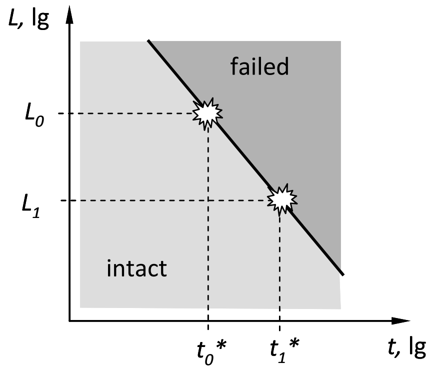

Fig. 6: Lifetime points in an S/N diagram.

It is extremely important to clearly differentiate the S/N diagram from an illustration of damage over time by means of a damage parameter. The latter is crucial for visualizing and understanding the physical processes at work. In practical reliability assessments, however, it is the Wöhler representation of lifetime points that plays a key role. This is also because knowledge of the nature of a damage parameter and the concrete course of damage over time is not necessarily available for every damage mechanism.

The S/N diagram does *not* illustrate a time series, but rather a boundary between the zones "unit under observation intact" (bottom left) and "unit under observation failed" (top right). Both the "in load direction", for a constant lifetime feature, and the "in lifetime direction", at a constant load or stress, can be examined.

2.2.3. Distributed characteristics, failure probability

As already mentioned above, a concrete component fails if its load exceeds its load capacity, or its local stress at the location of failure exceeds its strength.

If we now conduct a series of experiments with several components under the same load, different lifetimes will result, despite the same test conditions and (nominally) identical components. Lifetime is therefore a statistically distributed, not a deterministic, quantity. What this means is that due to randomly fluctuating material and manufacturing conditions, the lifetime of a component cannot be exactly predicted. The same frequently applies to its load, as a result of fluctuating conditions of use. As the local variables stress and strength are derived from global load and load capacity, these are also distributed variables. They are characterized by their statistical distribution. Details on distributions and their characteristic values can be found in section 6.1.3 of the Annex. The fact that lifetime is a distributed variable is sometimes implied symbolically by entering a distribution density function above the lifetime points, Fig. 7.

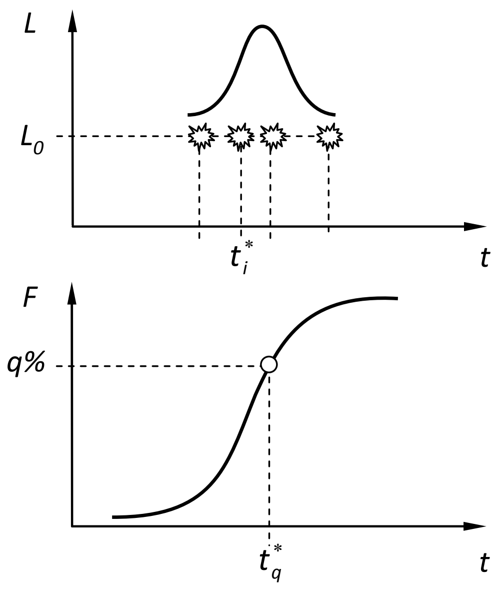

Fig. 7: Distributed lifetime. The points indicate the variable lifetime t_i* of the *i*-th, nominally identical part at a constant load quantity L_0 (top). Failure probability *F* dependent on the lifetime characteristic *t* (bottom).

The distributed characteristics approach opens up an entirely new perspective regarding the question as to why components fail: as neither the exact stress nor the exact strength of a concrete component is known in advance (but can only possibly be recorded or measured afterwards), the question as to whether a concrete component will fail will not be answered deterministically. In response to this question, we can only state a failure probability. In other words, we can only determine what proportion, from the population of all components, can fail, but not whether a particular component will fail.

As a rule, t_q* (also occasionally known in the literature as B_q) signifies the lifetime until which the proportion q of a population of products has failed. Common values for q are e.g. 10 %, 2 % or 1 %. Expressed mathematically, t_q* is the q-quantile of distribution. t_{10}*, for example, denotes the lifetime up to which 10 % of products of a population have failed. The median of the distribution is the 50% quantile t_{50}*. Where generally skewed distributions are concerned, the median differs from the mean, also known as MTTF (mean time to failure); in most cases, however, a relationship can be illustrated between two variables. t_q* can be determined easily with the aid of the failure probability curve $F(t)$ (possibly as a straight line in a suitable probability paper). To this aim, the point where the *q%*-horizontal intersects with the failure probability curve is determined and the associated lifetime characteristic read on the t-axis, Fig. 7.

Reliability is described as the survival probability $R(t)$, which can be determined from the failure probability:

$$R(t) = 1 - F(t). \qquad (2.1)$$

2.3. Failure characteristic over time, failure rate and "bathtub curve"

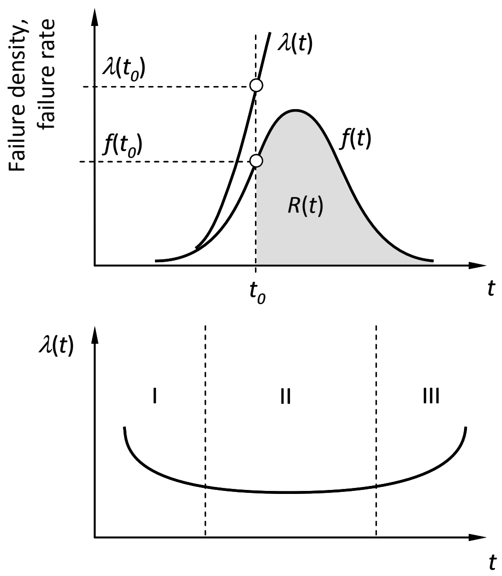

Fig. 8: Failure rate and "bathtub curve" according to [1]

In the field of quality assurance, in particular, it is common to express the failure characteristic by means of the failure rate λ:

$$\lambda(t) = \frac{f(t)}{R(t)}, \qquad (2.2)$$

whereby $f(t)$ signifies the failure probability density function and $R(t)$ denotes the survival probability at the time t, Fig. 8. The failure rate at a particular point in time can be estimated empirically, by dividing the number of failures per time unit by the sum of not-yet-failed components. This estimation is known as the failure quota.

The failure rate can be interpreted as a measure of the probability that a component will fail, if it has not failed up to this point in time, [1] p. 23. The failure rate is a *relative* quantity: the failures are divided by the number of still intact units. Furthermore, the rate is a *time-referenced* quantity, as the name suggests: the failures are stated per time unit (period, year, hour, etc.). Consequently, the statement "X-ppm" is *not* a failure rate.

The failure rate curve over the duration of service of components has a characteristic shape. Due to its similarity to the image of a longitudinal section through a bathtub, it has the name "bathtub curve". This curve generally has three segments, although not every component portrays this typical behavior:

- Segment I shows a falling failure rate and is typical of early failures due to manufacturing and assembly errors. An early failure characteristic means that the probability of failure in the next time segment is high at the start of stress, and declines over time. These failures can typically be averted by quality controls or experiments on a pilot series.
- Segment II shows a constant failure rate, which is typical for random failures due to operator error or contamination, for example. Here, the probability of failure in the following time segment does not depend on the product's previous history. The same proportion of the products still intact at the beginning of each time segment always fails within time intervals of the same length. Electronic components may demonstrate this characteristic, due to cosmic radiation, for example. In this case, failures can be avoided by correct usage.
- Segment III features a rising failure rate and is typical of degradation, wear and fatigue. In this area, it becomes increasingly probable that an as-yet intact product will fail within the following time segment. (Premature) failures in this area can only be avoided by the correct design.

The failure rate must not be equated with the absolute number of failures, as shown by the following two examples.

EXAMPLE:

A manufacturer has produced 100 units and records the following failure numbers in the following 4 time periods:

Period	1	2	3	4
Failures in period	10	9	8	7

Do you think the failure rate here is falling, constant or rising?

The failure quota can be used as the basis for estimating the failure rate. The table below shows that despite falling failure numbers, the failure quota remains roughly constant. This effect is due to the fact that the total number of still intact units falls over time.

Period	Intact units at the start of period	Failures in period	Failure quota, [1/period]
1	100	10	0.10
2	90	9	0.10
3	81	8	0.10
4	73	7	0.10

EXAMPLE:

A manufacturer produces 100 units per period over 4 time periods and then records the following failure numbers:

Period	1	2	3	4
Failures in period	10	19	27	34

Do you think the failure rate here is falling, constant or rising?

Again, the failure quota is used as the basis for estimating the failure rate. The following table shows that despite rising failure numbers, the failure quota remains roughly constant. This effect is due to the fact that the total number of intact units rises over time, because newly produced units are constantly being added.

Period	Intact units at the start of the period from production period				Failures in this period from production period				Total intact units at the start of period	Total failures in period	Failure quota, [1/per.]
	0	1	2	3	0	1	2	3			
1	100				10				100	10	0.10
2	90	100			9	10			190	19	0.10
3	81	90	100		8	9	10		271	27	0.10
4	73	81	90	100	7	8	9	10	344	34	0.10

As part of product development, endeavors are made to increase the reliability of a product systematically, i.e. to avoid early failures as far as possible (quality assurance), to widen the range of "random failures" (usage period; useful time) and to delay the beginning of wear and aging failures as far as necessary, Fig. 9. The failure characteristic of products is investigated in lifetime tests accompanying the development process. These are evaluated by applying Weibull’s theory.

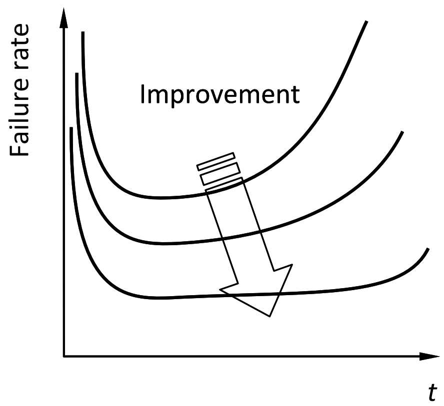

Fig. 9: Improving reliability during the course of product development (schematic)

The failure rates of numerous electronic standard components are listed in Failure Rate Catalogs, e.g. Military handbook 217 on predicting the reliability of electronic components [10]. Starting with the basic, temperature-dependent failure rates of components stated therein, failure rates for concrete usage conditions can be calculated, if simple models are assumed and load factors are taken into consideration. The load factors cover mechanical, climatic and electrical stress, for example.

However, measured and predicted values of reliability of electronic components (calculated on the basis of various manuals) may differ by a factor of up to 100. The calculation methods mentioned must therefore be critically examined:

- One of the principal problems is that the data employed for the assessment is old. The manuals are not always updated regularly (the most recent issue of [10] is from 1995, for example), so that the information they contain do not conform to today's state of the art.
- Secondly, vital predictors such as temperature change, temperature gradient, shock, vibration, switch-on/off processes, manufacturing quality and aging cannot be predicted, or not realistically.
- One key objection arises from the fact that these calculation models always operate with constant failure rates and can therefore only apply to random failures (zone II of the bathtub curve). This is because in this case, calculating the system failure rate from the failure rates of the individual elements is especially simple, see section 3.4.5. In the electronic components of today, however, degradation plays an important role. Degradation leads to rising failure rates and cannot be covered using models such as exponential distribution, which assume constant failure rates.

For the reasons described above, quantitative analysis of the reliability of electronic components based on manuals should be employed with caution and understanding.

FIT (Failures In Time) is a unit of measurement showing the failure rates of electronic components in the area of random failures, λ = constant:

$$1FIT = \frac{1}{10^9 \mathrm{h}} . \qquad (2.3)$$

Example:

If 7 out of 500 components fail within an operating time of one year, the failure rate is

$$\lambda = \frac{7}{500 \cdot 8760\mathrm{h}} = 1{,}598 \cdot 10^{-6} \mathrm{h}^{-1} ,$$

which is synonymous with λ = 1598 FIT.

Conversely, the expected number of failures can be calculated from a known failure rate (e.g. from a tabular value). If 2000 components complete a 1200-hours test with the failure rate stated above, approximately $2000 \cdot 1200\mathrm{h} \cdot 1598 \cdot 10^{-9} \mathrm{h}^{-1} \approx 4$ *failures can be expected.*

3. Design for reliability

3.1. General concepts

A range of approaches exists for achieving reliability in a product. A common feature they all share is that the achieved reliability must then be verified and validated.

1. Test-based design. In this approach, the primary focus during the development process is on designing the functionality. Weak points affecting reliability are exposed in tests and then corrected by a design measure. Further tests follow, to verify the effectiveness of the measure. The "Design⇨Test⇨Fix⇨Test" cycle is discontinued when the required reliability has been demonstrated by a test.
2. Experience-based design. The designer can draw on empirical and qualitative experience to produce a design that will be reliable. This approach requires knowledge of the reliability characteristic of earlier products and of whether this can be transferred to the current product. Such knowledge typically stems from earlier test, standard and/or overdesign-based approaches and from the analysis of field data.
3. Overdesign-based approach. The design is conceived such as to clearly surpass reliability requirements. Here, the aim is to achieve a particularly large safety reserve between stress and strength. The selection of this safety reserve is not made on the basis of knowledge of the distributions of stress and strength, but is derived from experience, or from standards and sets of rules.
4. Standard and rule-based design. Design takes place in accordance with the specifications or design guidelines in standards and sets of rules.
5. Design based on Design-for-Reliability principles. In this approach, reliability is knowingly and actively incorporated in the design element during the design phase. In this case, the aim is to achieve a high degree of accuracy, thereby avoiding both "underdesign" and "overdesign".

On the basis of the damage mechanisms to be expected as a result of load, the design element is conceived such that the stresses occurring during the service time stipulated by the reliability requirements remain smaller than its strength. The design process entails the selection of the optimum system structure and partitioning, solution principle, design element geometry (dimensioning), positioning, material and production process. Reliability can subsequently be predicted with the aid of design concepts based on the arithmetical superposition of stress and strength.

All these approaches have their advantages and disadvantages. The most important of these are listed in Table 1. Which approach to employ for conceiving a design element must be examined and established for each individual design element. The principal criteria are:

- level of innovation,
- cost effectiveness, i.e. development costs (test outlay, series launch costs, applicability of existing load, stress and strength data as well as design concepts to successive generations), production costs, potential quality costs,
- speed (time to market),
- future strategic importance of the design element, and
- legal and customer requirements.

Approach	Principal advantages	Principal disadvantages
Test-based	No previous knowledge needed for design, verification on the design element	Effort (tests, repeat tests), time required, limited transferability, of limited suitability when level of innovation is high
Experience-based	Effort (fast)	Experienced staff required, often not transferable, only feasible for evolutionary development
Overdesign-based	Effort (fast), (legal) certainty	Higher planned manufacturing costs due to overdesign, only feasible for evolutionary development
Standard-based	Effort (fast), (legal) certainty, acceptance	Higher planned manufacturing costs due to overdesign, only possible with standardized components and applications
Design for reliability	Transferability, accuracy as based on an understanding of the laws of physics, also suitable for new developments	Preliminary development work needed for determining load, stress, strength and design concept

Table 1: Principal advantages and disadvantages of the 5 approaches to achieving reliability.

Table 2 contains a relative evaluation of the 5 approaches and their main criteria. Different approaches may be employed depending on the weighting of the criteria for the specific design element concerned.

The design for reliability approach represents the best compromise when design elements have to be developed whilst simultaneously overcoming increasing stress (increasing loads and/or performance densities), ever shorter development times, increasing pressure on product costs and more exacting requirements. These days, this combination is a frequent occurrence. What is more, only the design for reliability is suitable when there is a high level of innovation. Further, only this approach is sustainable over the long term, as it enables an understanding of the failure behavior of the design element based on the laws of physics, allowing a knowledge base applicable to future generations to be built up. This knowledge base is then available for future developments as empirical knowledge that is reliable and can be extrapolated. Firstly, this enables advance development investment to be offset over time; secondly, it allows the speed advantage of an experience-based approach to be exploited.

Thus, use of design for reliability is recommended for the reasons stated above.

Approach	Suitability for new developments	Development costs	Manufacturing costs	Risk of quality costs	Speed (time to market)
Test-based	o	-	o	o	-
Experience-based	-	+	o	-	+
Overdesign-based	-	+	-	+	+
Standard-based	-	+	-	+	+
Design f. reliability	+	o	+	o	o

Table 2: Relative evaluation of the 5 approaches for ensuring the reliability of design elements ("+" good, "-" poor, "o" average suitability).

3.2. Basic task and principles

The task of design for reliability is to conceive the design element such that its stress during its service time as specified by the requirements remain smaller than its strength.

This is achieved, firstly, by designing the stress by means of

- system structuring and partitioning,
- solution principle,
- design element geometry, and
- design element position

and secondly by designing the strength accordingly by means of

- solution principle,
- design element geometry,
- material selection and manufacturing process.

A prerequisite for the above is a complete, quantitative description of the load acting on the higher-level system, and the internal loads originating from this load, which act on the free section of the design element.

The principles provide an answer to the following fundamental questions:

- What is the objective of design for reliability?
- What work methods are used to design reliability so that this objective is achieved?
- On what object is reliability being designed?
- When in the development process must reliability be designed?

For the developer, the principles act as a "corridor" within which he can achieve reliability in the design. For management, they provide the basis for controlling the aspects involved in design for reliability.

The principles are:

- Failure avoidance. The objective of design for reliability is to prevent the occurrence of failures during the service time, by not triggering failure mechanisms. If the triggering of failure mechanisms over the course of the service time is unavoidable, reliability must be designed such that the failure probability is kept sufficiently low with regard to the situation concerned. How small this failure probability must be is established individually in each case.
- Design with models and cause-effect relationships. Design for reliability is based on reliability models and cause-effect relationships. To be precise, knowledge is required of the cause-effect relationships between load and design, design and stress, design and strength, stress and damage mechanisms, reliability and stress/strength.
- Focus on design elements. The focus is on design elements because all products consist of them, and the same design elements are repeatedly employed in a variety of products. Design for reliability begins with system structuring and partitioning along the V-model. This creates ideal conditions for achieving reliability at design element level. The requirements from the higher level are applied to the level below. Verification starts from the design element and continues to the system, with particular consideration being given to interactions.

- Correctly timed design. Since design for reliability is achieved through the design concept, it must take place in those phases of the product creation process in which the design is specified. These are product innovation, project preparation or preliminary development and product/process conception. Design for reliability forms part of the product design process.

Reliability is verified and validated from the product/process development stage through to the start of serial production. During serial production, parameters relevant to reliability are monitored by means of statistical process control (SPC), and their effect on the product tested in experiments (reliability inspection). During operation, reliability is measured. Firstly, this step validates the accuracy of the reliability predictions and consequently the design concepts used; secondly, it helps to create a database (strength figures) for future developments.

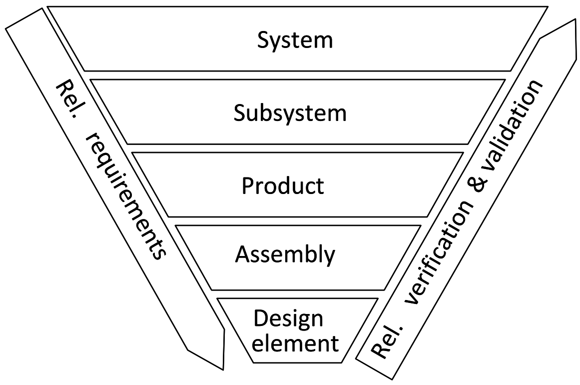

Fig. 10: V-model of product development. Requirements are broken down from the higher-level system to the design elements. Verification and validation take place in reverse order, from the design element to the system.

3.3. Reliability requirements

For a guarantee of sufficient reliability, first of all the reliability requirements have to be determined. These are characterized by the following aspects:

- For which functions must reliability be guaranteed, and what are the failure modes at which the function can no longer be accomplished? This statement allows a clear decision to be made later on as to whether failure has occurred or not. More details on functions can be found in [11], Annex 3.

- At what load must reliability be guaranteed? The load covers all externally acting mechanical, thermal, electromagnetic and chemical loads, and therefore also the ambient conditions. The amount of load is often variable. Its definition must therefore include part loads and the frequency with which they occur. Loads outside the intended use (misuse), which are not among the requirements, must be kept strictly separate from the above, however.

- For what useful time and operating time must reliability be ensured? This information is subsequently used to define the relevant stress time, which forms the basis for assessing the reliability of the component. Different characteristics may denote the "time/duration", depending on the damage mechanism in question: typically, these are the number of load changes in the case of fatigue, and the time in the case of corrosion. It must also be borne in mind that where complex load with several components is involved as well as the duration, the synchronization of the components over time is frequently of particular importance.

- What reliability measure must be guaranteed? This quantifies the required reliability in the form of a survival probability. This measure must be selected so as to ensure that verification through trials is also possible.

Lifetime variables such as MTTF and B_{10} are sometimes stipulated as a reliability measure, see section 2.2.3. This is generally not recommended, as these are not reliability requirements but rather lifetime stipulations, which considerably restrict design freedom during development on the one hand, and do not provide any information about the survival probability of the design element on the other hand, because the useful time and operating time have not been specified. A procedure such as this only makes sense if the useful time and operating time are unknown, as is typically the case with catalog goods, also see section 3.4.2. Furthermore, the stipulation of a certain lifetime quantile as a measure of reliability also neglects the variance in lifetime, which can have far-reaching consequences for reliability, as the following example illustrates.

EXAMPLE:

A manufacturer must choose between two suppliers of a part, who have provided the following information about the lifetime of the part: Supplier 1: B_{10} *= 5000 h, supplier 2:* B_{10} *= 6000 h. The manufacturer decides in favor of supplier 2 because of the better* B_{10} *values.*

During the warranty period of 1000 h, however, several complaints are received due to failure of the purchased part. Nevertheless, closer analysis of the failure numbers shows that the supplier had complied with his pledges, as shown in the following table:

Supplier	Weibull parameter		MTTF, [h]	B_{10}, [h]	Failures after		
	T, [h]	*b*, [-]			1000h	5000h	6000h
1	7275.4	6	6749.5	5000	7 ppm	10%	
2	26896.7	1.5	24280.9	6000	7143 ppm		10 %

What has happened? The manufacturer did indeed choose a supplier who pledged a maximum volume of failures of 10 % after 6000 h for his parts, but he forgot to take into account how quickly these failures are reached, i.e. whether 9 % of failures are achieved after as few as 1000 h or only after 5000 h, for example. In other words, the manufacturer neglected to take into proper consideration the variance in lifetime (described by the Weibull parameter b*).*

What's more: if his decision had been made on the basis of the supplier's statements regarding MTTF*, the result would have been the same. In this case, the difference in favor of supplier 2 is even clearer.*

The general requirements given may be on very different levels from one customer to another, which is why it can be particularly difficult to base reliability requirements on them:

- If the requirements originate directly from the end customer, they are very general and expressed "in the voice of customer". Functions are only specified for the uppermost system level (product level). Loads, such as useful time and operating times, must often be assumed on the basis of the intended use. The same applies to the reliability measure. Here, the road to be traveled until all the relevant reliability requirements have been determined is mostly especially long. A systematic use groups and use case analysis and the Quality Function Deployment (QFD) method can help to quantify the requirements step by step to the extent that loads, useful time/operating times and a suitable reliability measure can be specified firstly for the product and subsequently for the unit in question.
- For a system supplier, quantitative requirements may have been laid down by an OEM customer for the system to be delivered. These must then be examined to ensure that they are complete and reasonable. For example, the requirements for a fuel injection system will result from the conditions in the vehicle. Reliability requirements for the individual components are unknown in this case, but can be derived from the system requirements. If no requirements have been given by an OEM customer, these must first be formulated with him, as described in the case "requirements from end customers".

- In the case of tier 2 suppliers, functions, loads, useful time and operating time and the reliability measure may have been stipulated as requirements for the component, frequently based on the requirements for the higher-level system. Requirements for a gear will result from conditions in the transmission, for example.

The following sources may be employed for determining or adding to the requirements:

- Customer; the satisfaction of customer wishes is at the heart of every quality-oriented company policy. Here, in particular, variance due to different users (e.g. beginner, experienced, semi-professional), higher-level systems (e.g. applications in various vehicle types), use of parts including assembly, transport and service (e.g. drilling and turning) and operating conditions (e.g. use in different geographical locations) must be taken into consideration.
- Market; comparison with competitors is a decisive factor for successful positioning in the market.
- Legal regulations and the state of the art; these must generally be observed when deducing and establishing reliability requirements.

The customer satisfaction that one wishes to achieve is set against development and production costs. In addition, there is a direct relationship between the desired reliability and the quality costs, as well as the overall costs caused by the system (total life cycle cost). Therefore, the determination of target parameters for the reliability requirements is a business decision that brings about positioning in the market. The QFD method can provide considerable assistance at this point both in the definition of target values and in aligning product development with the customer's wishes.

3.4. Quantitative reliability assessment

3.4.1. Overview

The relationship between stress/strength and reliability can be summarized as follows, Fig. 11:

- The starting point for assessment are the loads that may lead to damage to the component.
- The stress is derived from these loads within the framework of a local concept. For this purpose, suitable variables must be found that characterize the damage mechanism under observation.
- Parallel to this, the strength is determined by means of an experiment. In the context of the local concept, this takes the form of a trial on specimen. The strength of the component is calculated from the load capacity of the sample using this concept.
- The failure probability F is calculated from the distributed stress and strength. The survival probability $R = 1 - F$ is a measure of reliability. It can be assessed by way of a comparison with the required reliability.

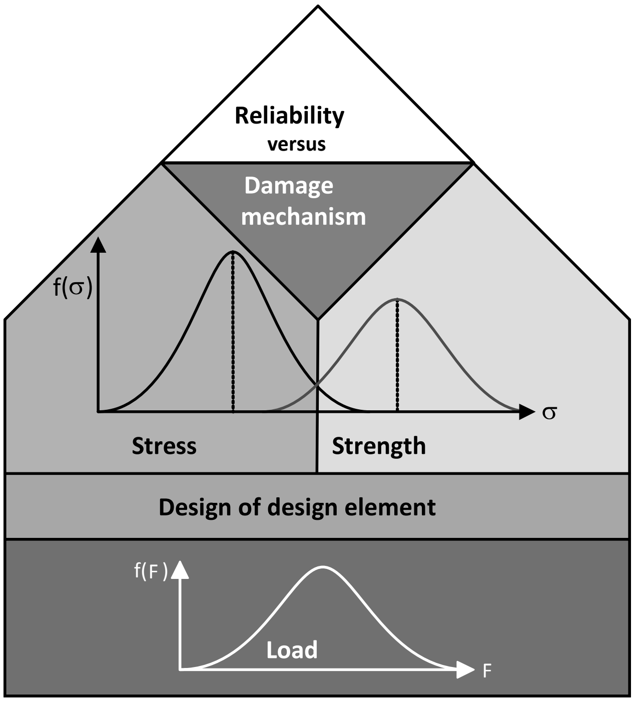

Fig. 11: House of Reliability: From load to failure probability

3.4.2. Determining stress

This step involves deducing the stress and the stress time from the requirements. Stress is a local parameter that characterizes the effects of the load on the design element in terms of the damage mechanism under observation.

Damage mechanism	Loads and boundary conditions	Stress parameter
Plastic deformation	Static mechanical load	Mechanical stress or strain
Creep	Static mechanical load and temperature	Mechanical strain and temperature
Fatigue right up to technical cracking	Variable mechanical load or variable temperature	Amplitude of mechanical stress or strain
Crack propagation	Static or variable mechanical load or constant or variable temperature	Stress intensity factor (SIF) or range of SIF
Wear (adhesion, abrasion, surface cracking, tribo-oxidation)	Mechanical load and temperature, composition of tribological system and relative movement of partners	Friction power density*
Electrochemical corrosion	Metal, conductive medium and temperature	Electrical potential*
Diffusion	Difference in concentration	Concentration gradient
Dielectric damage	Electric field, temperature	Electrical voltage, temperature
Electromigration (including thermal and voltage migration)	Electrical current, and possibly differences in temperature and a static mechanical load	Electrical current density, temperature, temperature gradient, gradient of hydrostatic mechanical stress

* Not yet established as an element of a local design concept

Table 3: Loads and associated damage mechanisms

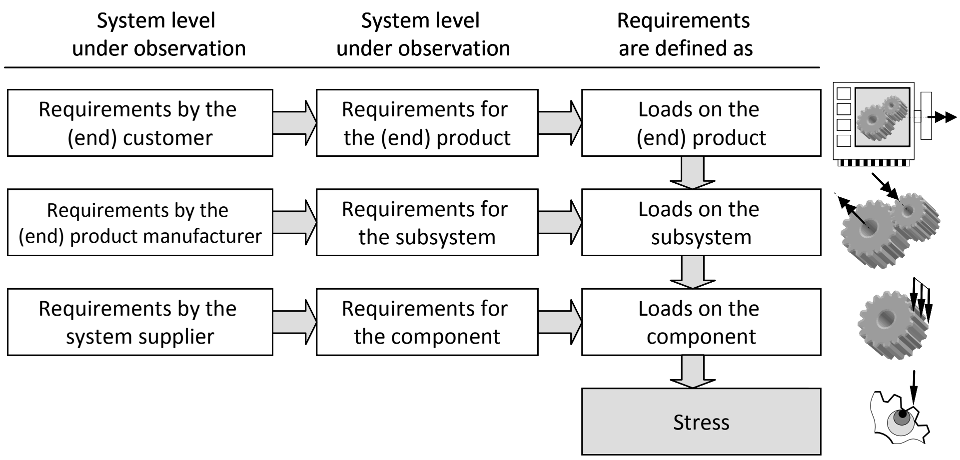

Fig. 12: From requirements through loads to stress

Table 3 contains some damage mechanisms with their associated stress parameters. Further details can be found in [12,13]. Combinations of damage mechanisms are also possible; a typical example is thermomechanical fatigue (fatigue, creep and oxidation).

The starting point for determining stress is always

- a concrete design,
- a load $L(t)$, which may consist of several components, and
- a stress parameter P as a local variable, which characterizes the damage mechanism under observation.

Determining the stress means establishing the local characteristic over time of this local variable $P(\mathbf{x},t)$ on the component as a consequence of load ($\mathbf{x}$ – position vector, t - time). If several damage mechanisms are stimulated for the load in question, different local variables and different stresses may also need to be taken into consideration. In this case, the complexity of the assessment grows enormously.

To determine the stress, first of all load-time characteristics are deduced from the requirements, and stress-time characteristics derived from these, Fig. 12. These may often need to be divided into

- part application,
- damage mechanism, and
- type of stress,

because the subsequent weighting, superposition and extrapolation of the individual stress components cannot simply be combined as desired. The direct addition of tensional and torsional load is not possible, for example, neither is the direct addition of stress from part application with a different duration. Likewise, different stress components, such as shear stress and normal stress, cannot be directly added together.

The lovel and duration of load can:

- be determined from the loads of the higher-level system, depending on the intended use. They will then typically take the form of load-time series or (classified) load collectives. Collectives are compact in form, but also have disadvantages. For example, the chronological sequence of the load is not defined in a collective. Furthermore, different components of a complex load cannot be synchronized over time using collectives. If these cases are relevant to the part application in question, load-time series should therefore be utilized.

- be derived from measurements on similar predecessor products. In this case, transferability from the old to the new product must be assured, or the loads must be re-evaluated.

The stress is typically calculated from the loads using simulation, but may also be deduced on the basis of analytical relationships. The technique used is often structural-mechanical calculations with the aid of the Finite Element Method (FEM), which is used to calculate local stresses from the external forces and moments - typically for assessing the damage mechanisms fatigue or crack propagation. However, local temperature characteristics can also be calculated from a given, possibly variable ambient temperature or heat source, for example, e.g. for assessing the damage mechanism creep. If a linear-elastic material characteristic is presupposed, quasi-static calculations can be carried out to determine the stress conditions for a unit load. Here, the chronological sequence of the stress parameter is achieved by scaling the load-time series $L(t)$ by the (location-dependent) transfer factor $c_{L=1}(\mathbf{x})$:

$$P(\mathbf{x},t) = c_{L=1}(\mathbf{x}) \cdot L(t) \ . \qquad (3.1)$$

Where the characteristic is not linear, P must be derived directly from the relationship $P = P(\mathbf{x},L(t))$, which is a great deal more complicated. The variable does *not* have to be a scalar for either the stress parameter or the transfer factor. In the case of fatigue, for example, mechanical stress is employed for both variables, which constitutes a 2nd-order tensor with 6 independent components.

Alternatively, in exceptional cases stress can also be determined in experiments, with the aid of strain gages, for example, often if there is insufficient knowledge of the cause-effect relationships between load and stress. This turns out to be difficult, however, as the accessibility of the measuring locations for determining local variables cannot be taken for granted.

As stress is a locally distributed variable and the location of failure is generally unknown in advance, a reliability assessment must be conducted for the complete (discretized) unit, and then the location with the smallest reliability named as the most probable failure location on the basis of the worst-case principle. This step can only be dispensed with if the failure location is known in advance from experience.

The determination of the stress time is often more difficult. This is the time during which the damage mechanism under observation is acting, measured in a damage mechanism-specific unit. It must be derived from the operating time and idle time, which are typically real time units.

The following procedure is recommended:

1. Firstly, determine the unit in which the stress time is to be measured. The two most common cases are
    a. real time units (hours, days, years, etc.) for time-dependent damage mechanisms such as corrosion and creep, and
    b. number of load changes for the damage mechanism fatigue.
2. Next, the useful time and pattern of use must be determined by means of use case and use groups analysis. The useful time is thus divided into operating time and idle time.
3. After this, the proportion of operating time and idle time during which the damage mechanism is acting (stress time) must be determined. Load change-dependent damage mechanisms such as fatigue will probably act primarily during the operating time, whereas time-dependent factors such as corrosion can act throughout the useful time (operating time and idle time). Unexpected effects may occur, however. For example, an electronic component in a vehicle may also be under voltage when the vehicle is not in operation, which may result in alternating thermal stress

and so in fatigue. Alternatively, the corrosion suffered by a component during operation in the vehicle can only be suppressed by an increased temperature, meaning that it is only damaged during idle times.

4. Finally, the total stress time is divided into the times under stress of the individual part application. For this purpose, higher-level assumptions about the relative frequency of the individual part applications are often required, which must be made as objective as possible. For example, the relative proportion of urban, interurban and freeway driving, or certain driving maneuvers, can be determined on the basis of statistical analyses of driving behavior. The same applies to the relative frequency of use at low or high outside temperature, humidity, etc., and weather statistics from the region in question may provide assistance in this connection.

The deliberations above demonstrate that it is not possible to formulate generally applicable rules for deducing the stress time, and that each individual case must be evaluated separately.

EXAMPLE:

A piston pump is operated via a triple cam on a drive shaft. The system to which the pump belongs is intended for 20 years of use and is in operation 8 h a day in a 1-shift cycle. The shaft rotates at a constant speed of 3000 rpm while the system is in operation. To determine the stress level in terms of fatigue suffered by a component in the pump, which is subjected to load due to pressure pulsations in the pump chamber, the following procedure can be applied:

First of all, we need to determine the unit in which the amount of stress is to be measured; in the case of fatigue this is the number of load changes. Next, we must ascertain the proportion of operating time and idle time at which the damage mechanism causes fatigue - in this case during the operating time only. This is not immediately obvious, however, but must be calculated from the useful time and pattern of use:

T_B = 20 years · 230 days/year · 8 h/day · 60 min /h = 2208000 min

Here, it is assumed that the system is in operation 230 days of the year. From the operating time, we can now calculate the number of piston strokes:

$N_K = T_B$ min · 3000 rpm · 3 = $19.872 \cdot 10^9$.

Since we already know from earlier pressure measurements that there are approx. 10 pressure oscillations (reverberations) per piston stroke, for the number of load changes:

$N = N_K \cdot 10 = 198.72 \cdot 10^9$.

A special procedure must be employed when determining stress in so-called catalog products. These are products with a wide range of application and for which the future use and, in particular, the useful time, are uncertain.

In such cases, the manufacturer must group potential uses into typical classes, and deduce loads and stresses within each class as described above. Here, however, the reliability assessment by the manufacturer is not for a stress time defined in advance, but rather is obtained indirectly by stating the lifetime of the product for the use class in question. In this way, after he has selected the use class that seems most apt to his particular use, the user can determine the stress time during which a certain reliability is guaranteed. This procedure is illustrated in Fig. 13, further details on presenting lifetime data in the probability plot can be found in section 4.3.2.

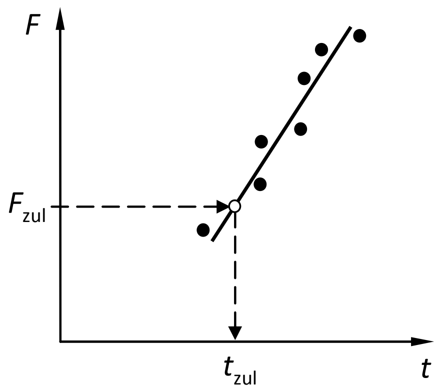

Fig. 13: Calculating the stress time from the permitted failure probability on the basis of lifetime for catalog products (representation on the probability paper)

Once all stress components have been ascertained, they are superposed on a scalar equivalent stress time series, then classified to form a stress collective. Classification is understood to be a process in which the characteristic over time of the stress-time characteristics is broken down into individual, damage-relevant parts, which are then grouped in classes. For the "fatigue" damage mechanism, the amplitudes of the changes in stress play a key role, but so does the mean stress at which the changes in stress take place. With a collective, which forms the basis for a damage assessment, the information about the mean stresses is lost. For this reason, the amplitudes must be transformed as "damage-equivalent" to a certain mean stress. In addition, it is necessary to ensure that the constant amplitude strength curve (Wöhler curve) is compatible with this, i.e. shows the same mean stress.

Finally, a damage assessment is performed, assuming that the strength in question is present. This procedure is set out in section 3.4.6.

3.4.3. Determining strength

Strength denotes the maximum stress that can be withstood by a design element for a defined stress time and for a given damage mechanism. It is calculated from the load capacity, which is deduced directly from an experiment. This experiment is typically conducted in such a way that the amount of load L_i is kept constant, and several specimens are subjected to load until they fail. The experiment in which the load is applied uniformly at different levels until failure is known as a Wöhler experiment. The time to failure represents the (variable) lifetime of the specimens t_i^*. Through statistical evaluation, parameters of the distribution, such as the median $t_{50}^*(L_i)$ are determined at a given load L_i. This procedure is then repeated for another load.

If the aim is to apply the data in a local concept, the load L_i can then be converted into a local variable at the failure location (of stress) s_i. Trials on specimens are usually conducted in this case. However, the Wöhler experiment can also be performed on components, in order to determine load capacities for load-based assessments. To ensure uniform treatment in both cases, a parameter P_i is next introduced, which can stand both for a load L_i and a stress s_i, Fig. 14.

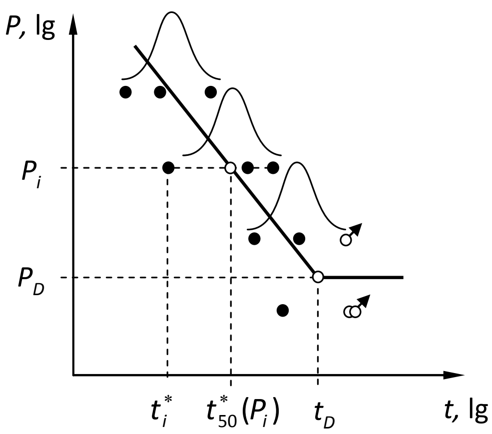

Fig. 14: Wöhler experiment and Wöhler curve. The shaded circles mark the time of failure with the parameter *P* at a certain level; the blank circles are experiments that ended without failure.

The mean lifetime increases as the load falls. When entered as a double logarithm, this relationship can typically be represented as a straight line. Below a certain threshold, the lifetime values rise so dramatically that the time that is realistically available for the experiment is exceeded and no failure can be identified. This threshold is the so-called endurance limit. The time span until the endurance limit is reached is referred to as the finite life. To simplify matters, the Wöhler curve is often assumed to run horizontally below the endurance limit, although for some material classes a very flat, but not horizontal, form can be demonstrated.

The Wöhler curve can be described mathematically by means of the following equation:

$$\frac{t}{t_D} = \left(\frac{P}{P_D}\right)^{-k} \text{ for } P \geq P_D \text{ and } t = \infty \text{ for } P < P_D, \tag{3.2}$$

whereby P_D designates the so-called endurance limit and t_D denotes the transition (turnaround) point from finite to infinite life. The Wöhler exponent *k* determines the slope of the Wöhler curve: the higher *k* is, the flatter the curve in the finite life range. *t* generally signifies the lifetime characteristic, which may look different for certain damage mechanisms. Where fatigue is concerned, for example, the variable *N* is traditionally used to denote the number of load changes. The variables are median values of load capacity or strength. The index "50" and the superscript "*" have been omitted for the sake of clarity.

Equation (3.2) is also referred to as the Coffin-Manson law, and is often used to described lifetime in the event of fatigue as a result of mechanical or thermal load. This relationship is also employed in numerous other cases, such as for characterizing the lifetime of insulators, for example, under exposure to an electrical voltage [12], p. 259.

In some cases, another relationship between *P* and *t* can also be deduced on the basis of physical and theoretical correlations. The following relationship is a typical example:

$$\frac{t}{t_0} = \exp\left(\frac{c}{P} - \frac{c}{P_0}\right), \tag{3.3}$$

which was calculated directly from the known Arrhenius equation to describe the speed of chemical reactions. Here, *c* signifies the ratio E_a/K of activation energy E_a and Boltzmann constant K = 8.6E-5 eV/K. However, *c* can also generally be expressed as a free constant, which has to be adapted in line with the experimental results. This equation is often employed to describe lifetime in the case of damage due to the effect of temperature, so that *P* = *T* (*T* absolute temperature in Kelvin).

The motivation for this is a model in which damage is a temperature-dependent, physicochemical process that only leads to failure once a critical material conversion is reached, with the result that the lifetime t is inversely proportionate to the reaction speed v, as described by the Arrhenius equation

$$v(T) = v_0 \exp\left(-\frac{c}{T}\right) \tag{3.4}$$

(v_0 signifies a material constant). Thus

$$t \sim \frac{1}{v(T)} \text{ becomes } \frac{t}{t_0} = \frac{v(T_0)}{v(T)} \tag{3.5}$$

and the generalization $T = P$ leads directly to equation (3.3). Further theoretical relationships between stress and lifetime are discussed in [12], section 7.4.

3.4.4. Component reliability

3.4.4.1. True failure probability

The failure probability is deduced from the distributions of strength and stress, Fig. 15.

It is assumed that the variable in question has log-normal distribution, i.e. the logarithm of the variable is normally distributed. $P = \lg(X)$ therefore designates the logarithm to base 10 of either the load, the stress or the stress time, and $P^* = \lg(X^*)$ denotes the logarithm of load capacity, strength or lifetime (the procedure is identical in all three cases). The distribution density functions of P and P^* are expressed by

$$f_P(P, \bar{P}, s_P^2), \quad f_{P^*}(P^*, \bar{P}^*, s_{P^*}^2). \tag{3.6}$$

$\bar{P}, \bar{P}^*$ signifies the mean values of P and P^*, and $s_P^2, s_{P^*}^2$ their variances. For the purpose of assessment, an auxiliary variable

$$z = P^* - P \tag{3.7}$$

is initially introduced. The normal distribution of P and P^* directly results in the normal distribution of z, in the form of

$$f_z(z, \bar{z}, s_z^2), \quad \bar{z} = \bar{P}^* - \bar{P}, \quad s_z^2 = s_P^2 + s_{P^*}^2. \tag{3.8}$$

Failures occur if $P > P^*$, i.e. if the stress is higher than the strength, which is identical with $z < 0$. The failure probability is the relative proportion of these cases among all possible cases, and can be expressed by

$$F = \int_{-\infty}^{0} f_z(z)dz. \tag{3.9}$$

The explicit evaluation of the above integral can be avoided by means of a transformation. We introduce the variable u,

$$u = \frac{z - \bar{z}}{s_z}, \tag{3.10}$$

so that it has standard-normal distribution (mean value 0 and standard deviation 1), and express the following with u_0:

$$u_0 = u(z = 0) = -\frac{\bar{z}}{s_z}. \tag{3.11}$$

For the failure probability, this results in

$$F(u_0) = \int_{-\infty}^{u_0} f_u(u) du \ . \qquad (3.12)$$

This integral can be calculated using spreadsheet programs such as Microsoft Excel, and is set out in tabular form in section 6.3.1 of the Annex. A graphical representation can be found in Fig. 15 below.

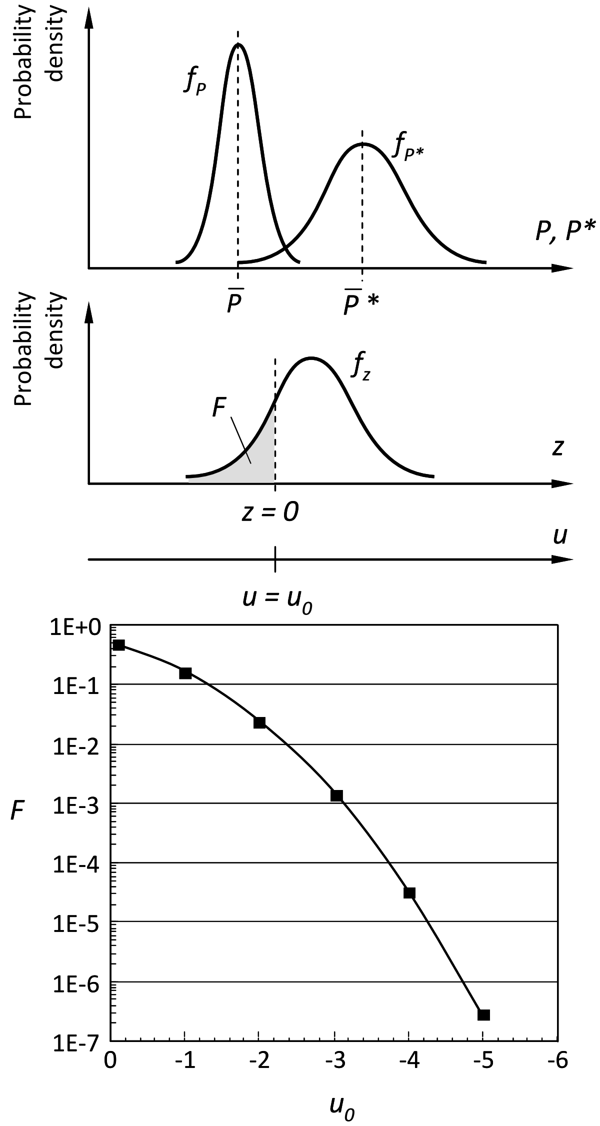

Fig. 15: Failure probability for normally or log-normally distributed variables

In summary, the failure probability calculation can be described as follows:

1. The mean values and variances of P and $P*$ are expressed by $\overline{P}, s_P^2, \overline{P}*, s_{P*}^2$. Calculate

$$u_0 = -\frac{\overline{P}* - \overline{P}}{\sqrt{s_{P*}^2 + s_P^2}} \ . \qquad (3.13)$$

2. Read the value $F(u_0)$ of the standard normal distribution from section 6.3.1, this is the failure probability. Reliability is $R = 1 - F$.
3. Assess whether the given reliability exceeds the required reliability ($R \geq R_{erf}$).

3.4.4.2. Simplified failure probability

The distribution of stress is often not known. Therefore, in simplified form a fixed value *P*, which always occurs, is assumed as the stress, Fig. 16.

In this case, the assessment is simplified to form

$$\bar{z} = \bar{P}^* - P\,, \tag{3.14}$$

$$s_z^2 = s_{P^*}^2\,, \tag{3.15}$$

$$u_0 = -\frac{\bar{P}^* - P}{s_{P^*}}\,. \tag{3.16}$$

However, we can demonstrate that this is only a conservative approach if *P* is a high quantile of the (unknown) distribution of the stress (e.g. 99%), otherwise, the calculated, simplified failure probability is smaller than the true probability. Details can be found in [5], p. 311f.

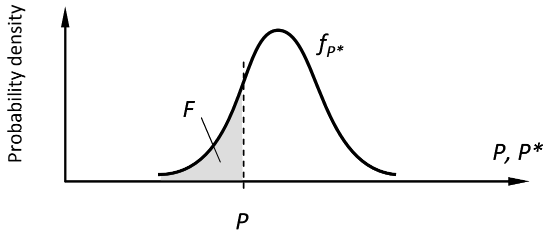

Fig. 16: Simplified failure probability

3.4.4.3. Safety margins

Alternatively, to assess reliability, a safety margin

$$j = \frac{X_{50}^*}{X_{50}}\,, \tag{3.17}$$

can be used in place of survival probability. Here, X_{50}^* denotes the median (value with a 50% failure probability) of the log-normally distributed variables load capacity, strength or lifetime, and X_{50} denotes the median of the log-normally distributed variables load, stress or stress time (*not* the mean values of these variables). Assessment is achieved by comparing *j* with a required safety j_{erf}, and is successful if $j \geq j_{erf}$. The safety margin is often presented graphically on a logarithmic scale as a distance between the logarithms of the median values, Fig. 17.

The expression of a safety margin constitutes an alternative to an assessment of reliability on the basis of survival probability. Indeed, this is often used, but has a major disadvantage in that a relatively arbitrary required safety has to be stipulated. Guidelines or empirical values can be used for determining this figure. However, this very rarely concurs with real failure numbers, so that the connection between field behavior and design is played down.

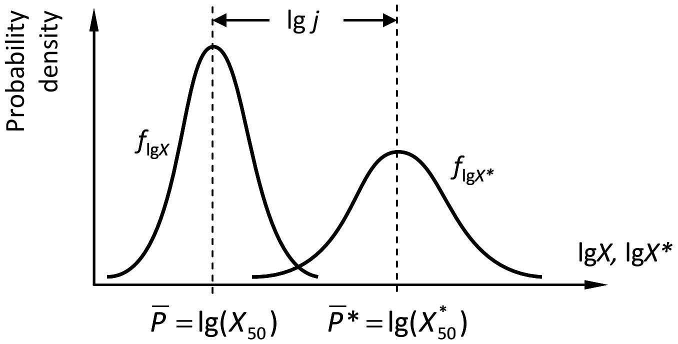

Fig. 17: Representation of the safety margin

When the required safety is based on statistics, however, it is possible to show that the two procedures are equivalent. The relationship between the safety margin and failure probability can be deduced as follows.

By expressing logarithms of equation (3.17) and taking account of the relationship (see [6], p. 71)

$$\bar{P} = E(\lg(X)) = \lg(X_{50}), \qquad (3.18)$$

$$\bar{P}^* = E(\lg(X^*)) = \lg(X^*_{50}), \qquad (3.19)$$

equations (3.8) and (3.11) show that

$$\lg j = \bar{P}^* - \bar{P} = \bar{z} = -u_0 s_z \ . \qquad (3.20)$$

Here, E(.) signifies the expected value of a random variable. Equation (3.20) produces

$$j(u_0) = 10^{-u_0 s_z} \ . \qquad (3.21)$$

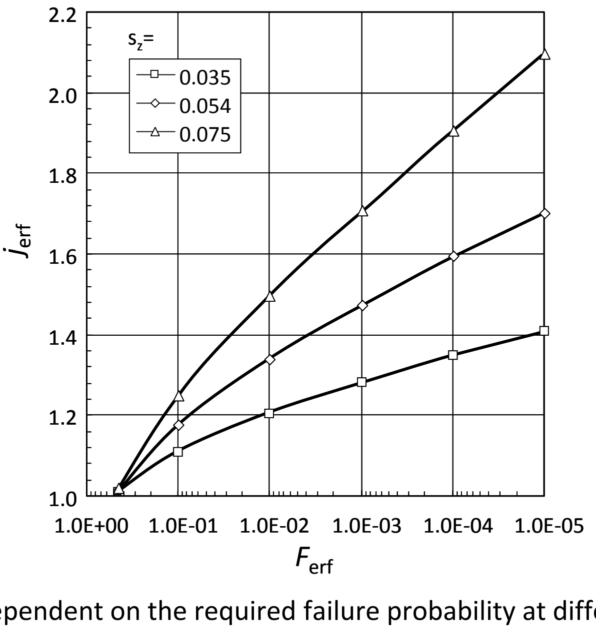

Fig. 18: Required safety dependent on the required failure probability at different values for s_z.

On the other hand, according to equation (3.12), the failure probability F also depends upon u_0 , i.e. for a given variance s_z a relationship exists between $j(u_0)$ and reliability $R(u_0) = 1 - F(u_0)$. If therefore, with a known and predefined variance s_z, the required safety is derived from $u_{0,erf}$ in accordance with equation (3.21):

$$j_{erf} = 10^{-u_{0,erf} s_z} \qquad (3.22)$$

and $u_{0,erf}$ is calculated from the required reliability R_{erf} in accordance with equation (3.12):

$$1 - R_{erf} = \int_{-\infty}^{u_{0,erf}} f_u(u)du \,, \qquad (3.23)$$

the two procedures are equivalent. The relationship is illustrated in Fig. 18.

REMARKS:

The measure of variation $1/T = X_{R=10\%} / X_{R=90\%} > 1$ *frequently used for fatigue can be converted to s by means of the relationship*

$$s = \frac{\lg(1/T)}{2.56} \qquad (3.24)$$

$X_{R=10\%}$ or $X_{R=90\%}$ refers to the quantile of any varying log-normally distributed variable, e.g. stress, strength or lifetime, with a survival probability of R=10% or R=90%.

3.4.4.4. General distribution functions

The sections above are based on the use of normally or log-normally distributed variables. But this assumption does not always make sense. Below, we go on to describe how a procedure may look when variables are generally distributed.

The distribution of load, stress or stress time is expressed by the distribution density functions f_P, the distribution of load capacity, strength and lifetime by f_{P*}. The failure probability is generally the proportion of expected failures in relation to all possible cases.

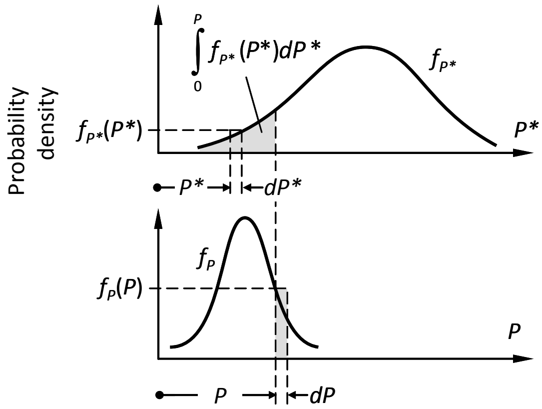

Fig. 19: Failure probability based on the distributed load and load capacity or stress and strength

First of all, let us take a look at a concrete value for P, Fig. 19. Here, failures can be expected for all components with $P^* < P$, i.e. load capacity is smaller than the load, or the strength is smaller than the stress, or the lifetime is smaller than the stress time. Components in the area $(P^*, P^*+dP^*]$ have the probability of occurrence $f_{P*}(P^*)dP^*$, see section 6.1.3.1. To determine the failure probability for all $0 < P^* < P$, these components must be added together. As this population is infinitely large (characteristics are described by distributions), this adding process is achieved by forming integrals:

$$\int_0^P f_{P*}(P^*)dP^* \,. \qquad (3.25)$$

The above integral represents the so-called conditional failure probability for a given *P* (load, stress or stress time).

In order to calculate the (total) failure probability for all possible *P*, in accordance with the theorem of total probability (see section 6.1.1 in the Annex), each conditional failure probability must be weighted with its probability of occurrence, and added together over all possible cases. The loads/stresses/times under stress in the area (*P,P+dP*] have a probability of occurrence of $f_P(P)dP$, adding is again achieved by the formation of integrals. Finally, the following results for the failure probability:

$$F = \int_0^{\infty} f_P(P) \left[\int_0^{P} f_{P^*}(P^*) dP^* \right] dP \; . \tag{3.26}$$

The above formula is a general means of expressing any distributions. In practice, the focus is mostly on special cases that permit simplified calculation, so that the calculation of the above integrals can be avoided.

3.4.4.5. Arithmetical lifetime at variable stress

In order to experimentally determine the lifetime of components at variable loads, experiments will repeatedly be necessary. The question arises as to whether this lifetime cannot be arithmetically deduced from the Wöhler curve (constant amplitude load capacity) and the load collective. Similar considerations apply to the local observation of a stress collective and a strength Wöhler curve. This is possible to a certain extent, if we apply the concept of damage accumulation, which is described below. To ensure the uniform treatment of local and global variants, the parameter *P* is employed, as usual, which can stand both for a load *L* and a stress *s*. *t* generally signifies the relevant lifetime characteristic. In the case of fatigue, as here, this is the number of load changes *N*.

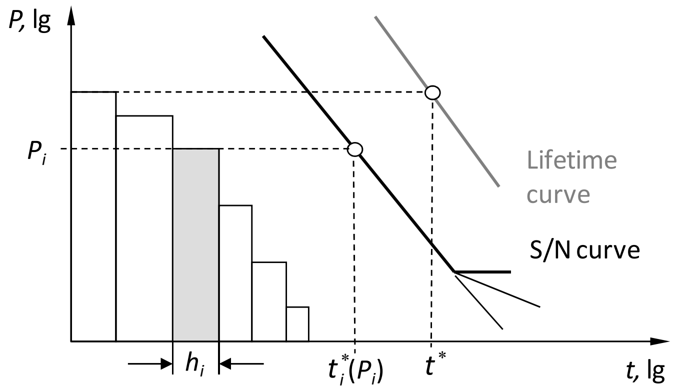

Fig. 20: From the Wöhler to the lifetime curve for variable loading through damage accumulation

With a constant amplitude load of a particular level P_i, the point where the Wöhler curve intersects with this load represents the lifetime t_i^*, Fig. 20. The fact that this is a statistically distributed variable is initially disregarded. If at this load level the amount is not t_i^* but $h_i < t_i^*$, failure will not occur. The model assumes that only a part of the lifetime is "used up", namely, the h_i/t_i^*-th part. This part may also be understood as partial damage D_i. If further part collectives with different load levels P_i occur in a load collective, to determine the total damage of the collective *D*, their partial damage D_i must be added together:

$$D = \sum_i D_i = \sum_i \frac{h_i(P_i)}{t_i^*(P_i)} \; . \tag{3.27}$$

For assessing the part collectives below the so-called endurance limit, at which $t_i^*(P_i)=\infty$, there are different approaches, also referred to as Miner's rule:

- "Elementary Miner's rule": These part collectives contribute exactly as much to the total damage as the others. The slope of the Wöhler curve continues below the endurance limit.
- "Original Miner's rule": These part collectives do not contribute to the total damage, and are disregarded. A horizontal curve is assumed below the endurance limit.
- "Modified Miner's rule": These part collectives contribute less to the total damage than the others. This is because a flatter gradient than that of the Wöhler curve is assumed below the endurance limit.

From a practical point of view, the third approach has proven the most useful and is often employed in the "Miner-Haibach" model, although this is not discussed any further here, see [5] p. 191ff.

In the case of a constant amplitude load, failure evidently occurs at $D = t_i^*/ t_i^* = 1$. At the same time, when the amount of load is variable it initially appears reasonable to assume that failure occurs at $D = 1$. If the collective damage equals $D < 1$, the collective can be repeated $1/D$ times until $D = 1$ and failure is therefore reached. Consequently, with a collective amount h_0, the achieved lifetime is

$$t^* = \frac{1}{D} h_0 = \frac{1}{D} \sum_i h_i \qquad (3.28)$$

whereby D can be determined from equation (3.27).

This strategy is known as linear damage accumulation, and constitutes a certain simplification, as in reality there is no physical reason for this linear approach. Numerous effects are known that result in a non-linear total damage characteristic. Thus, endurance tests have shown that actual failure often does not occur at $D = 1$, but rather at amounts of damage between 0.1 and 10, and sometimes beyond. Nonetheless, this approach is frequently the only practicable solution for obtaining an arithmetical estimate of lifetime under variable load based on constant amplitude strength. A certain pragmatic improvement in accuracy can be achieved by determining the actual total damage, known as the critical total damage $D_{50\%} \neq 1$ for certain collective forms, and using this as the basis for further calculation instead of $D_{50\%} = 1$. Consequently, for the lifetime assessment:

$$t^* = \frac{D_{50\%}}{D} \sum_i h_i \ . \qquad (3.29)$$

In order to be able to make use of the approaches described in the preceding sections to calculate reliability, we need to deduce the exact distribution of lifetime according to equation (3.29). This is no trivial matter, however. It is possible to show that with log-normally distributed stress h_i and strength t_i^*, the partial damages D_i are also log-normally distributed. However, this applies *no longer* to the total damage D (as the sum of log-normally distributed partial damages), and consequently also not to the lifetime t^*. For the purpose of assessment, we could conceive of an approach in which the equations (3.29) and (3.27) are regarded as a functional correlation between the lifetime and the random variables h_i and t_i^*:

$$t^* = f(h_i, t_i^*) \ . \qquad (3.30)$$

Thus, the lifetime may well be a random variable with an unknown distribution, but its characteristic values such as mean and variance can be derived from the characteristic values of the distributions of h_i and t_i^*, see section 6.1.3.3 in the Annex. From a pragmatic viewpoint, however, a log-normally distributed lifetime t^*, with the same spread as the Wöhler curve, is assumed.

For lifetime safety j_t, the following is finally produced with equation (3.29):

$$j_t = \frac{t^*}{\sum_i h_i} = \frac{D_{50\%}}{D}\,. \qquad (3.31)$$

3.4.5. System reliability

The reliability of a system can be evaluated on the basis of the failure characteristic of its components with the aid of the Boolean systems theory. To this aim, certain important preconditions must be met [1], p. 78:

- The system is either irreparable or the assessment is taking place upon the first system failure (for reparable systems).
- The system elements can only assume the states "functioning" or "failed".
- The system elements are mutually independent, i.e. the failure characteristic of an element is not influenced by the failure characteristic of other elements.

The first step in assessing system reliability consists in simulating the reliability structure of the system in a Reliability Block Diagram (RBD). In this diagram, the effect of failure of an element on the system is modeled, whereby the mechanical structure does not necessarily have to be modeled. Thus, in the Reliability Block Diagram, elements that are only physically present once in the structure can occur at several locations.

The diagram features an input node *I* and an output node *O*. The system is functioning if a path between the input and output nodes is formed by the functioning elements. If not, the system has failed.

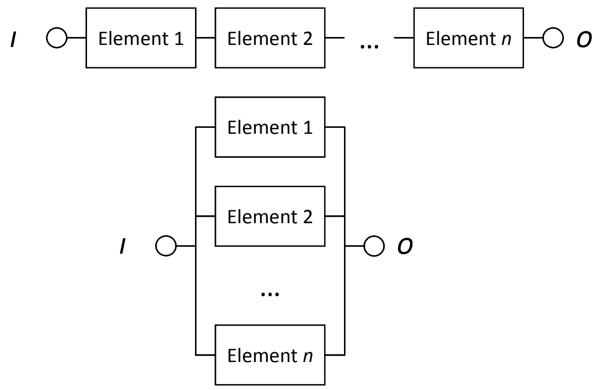

Fig. 21: Reliability Block Diagram of a serial structure (top) and a parallel structure (bottom)

Three elementary structures can be employed for constructing the block diagram, Fig. 21:

- The serial structure, Fig. 21 top. Here, the system can only function if all elements are functioning, i.e. the failure of *one* element immediately results in the failure of the structure.
- The parallel structure, Fig. 21 bottom. Here, the system functions if at least one of the elements is functioning, i.e. only the failure of *all* elements results in the failure of the structure.
- The *k*-out-of-*n* structure. Here, the system functions if at least *k* of a total of *n* elements are functioning.

REMARKS:

Thus, a serial system is an n*-out-of-*n *system, and a parallel system is a* 1*-out-of-*n *system. The block diagram of a* k*-out-of-*n *system can be represented as a combination of serial and parallel elements.*

The Reliability Block Diagram simulates the reliability structure and must not be confused with the physical structure of the system, as the following example demonstrates.

EXAMPLE:

An electric component consists of a parallel circuit of three identical electric resistors R_i=3Ω, Fig. 22 left. According to the specification, the resistance of the component must not exceed 1.2 Ω. What does the Reliability Block Diagram look like? According to the law for resistors connected in parallel, the total resistance of the component at n=3 is:

$$\frac{1}{R} = \sum_{i=1}^{n} \frac{1}{R_i} = 1\Omega \,. \tag{3.32}$$

If 1, 2 or 3 of the resistors fail, the total resistance figures of the component are as follows:

Failures	n	R, Ω
1	2	1.5
2	1	3
3	0	∞

The above demonstrates that for the component to function, all resistors must be functioning, which is described by a Reliability Block Diagram with a serial *structure, Fig. 22 right.*

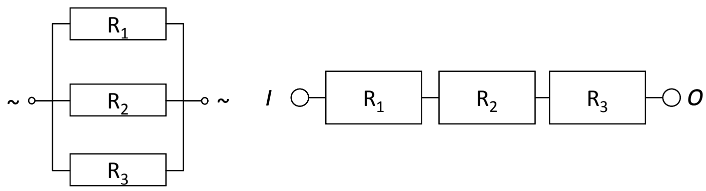

Fig. 22: Electric circuit diagram of a component with three resistors connected in parallel (left), Reliability Block Diagram of the component (right)

The failure or survival probability of these elementary structures is derived from the failure or survival probabilities of the elements as follows, [1], p. 81.

According to the product law, the survival probability of the **serial structure** based on *n* elements is

$$R_S = R_1 R_2 \ldots R_n = \prod_{i=1}^{n} R_i \,, \tag{3.33}$$

whereby $R_i < 1$ dictates that $R_S < \min(R_i)$, i.e. the survival probability of the structure and therefore also its reliability, is *smaller* than the reliability of each individual element. If, however, clear differences exist in the reliability of the individual elements, and to the extent that only one element has a reliability of $R_0 < 1$ and all others $R_i \approx 1$, the element with the smallest reliability will determine the reliability of the system, and $R_S \approx R_0$.

If the failure characteristic of the elements in a serial structure can be described by means of an exponential distribution, the calculation of system reliability is especially simple. In this case (also see section 6.2.2), the reliability of the elements is expressed by

$$R_i(t) = \exp(-\lambda_i t) \,. \tag{3.34}$$

According to equation (3.33) for system reliability

$$R = \prod_{i=1}^{n} \exp(-\lambda_i t) = \exp(-\lambda t)\ , \quad \lambda = \sum_{i=1}^{n} \lambda_i \tag{3.35}$$

whereby λ and λ_i denote the failure rates of the systems and the elements. Accordingly, with a serial structure, the failure rate of a system is simply the sum of failure rates of the elements (with an exponentially distributed failure characteristic). However, we should make clear once again at this point that exponential distribution can only describe random failures (constant failure rate), not early failures (falling failure rate) or degradation, fatigue or wear (rising failure rate).

On the other hand, the product law dictates that for the **parallel structure** based on *n* elements

$$F_P = F_1 F_2 \ldots F_n = \prod_{i=1}^{n} F_i = \prod_{i=1}^{n} (1 - R_i)\ , \tag{3.36}$$

$$R_P = 1 - F_P = 1 - \prod_{i=1}^{n} (1 - R_i)\ , \tag{3.37}$$

whereby $F_i < 1$ dictates that $F_P < \min(F_i)$ and therefore $R_P > max(R_i)$, i.e. the survival probability of the structure and thus also its reliability is *greater* than the reliability of each individual element.

If *n* independent elements of the system all have the survival probability R_0, the survival probability of a ***k*-out-of-*n* system** is expressed by (see [13], p. 20):

$$R = \sum_{i=k}^{n} \binom{n}{i} R_0^{\,i} \left(1 - R_0\right)^{n-i} . \tag{3.38}$$

Complex systems can be formed from a combination of the above-mentioned elementary structures. The quantitative assessment of a system’s survival probability can be deduced from the survival probabilities of the elements by searching for the said elementary structures in the hierarchical system structure from the inside out, and evaluating them according to the formulae described above until the entire system has been processed.

The following example illustrates the considerable influence the system structure has on reliability.

EXAMPLE:

A complex product (a car, for example) consists of 500 parts, in which one failure leads directly to the failure of the product as a whole. We wish to estimate the system reliability of the product if the individual parts have a reliability of 99% or 99.99%.

As the failure of any individual part results directly in the failure of the system, a serial structure is appropriate for modeling the reliability of the product. Thus, the survival probability of the system is $R_{S;99\%} = 0.99^{500} = 0.66\%$*, or* $R_{S;99.99\%} = 0.9999^{500} = 95.12\%$.

Although one could assume that at 99%, the individual parts have a high survival probability, this is obviously not sufficient to guarantee sufficient system reliability in complex systems. This also explains the extremely exacting requirements for the reliability of elements in complex systems.

3.4.6. Summary and example

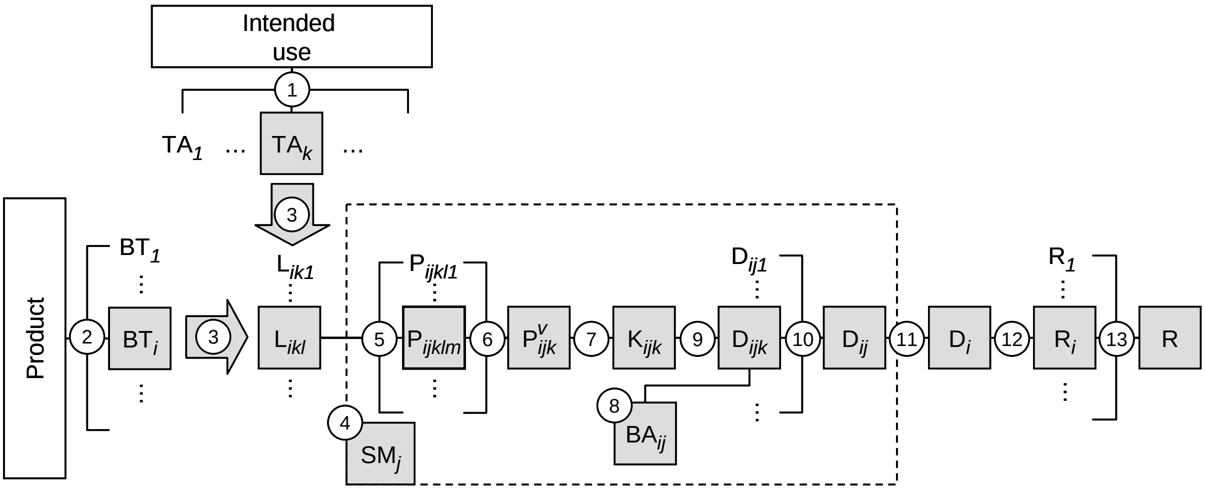

Fig. 23: From loads to system reliability. Abbreviations: TA - part application, BT - component, SM - damage mechanism, L - load, P - stress parameter, P^v - equivalent stress parameter , K - stress collective, BA - strength, D - damage, R - reliability.

The path from the loads to proof of reliability can be summarized as follows, Fig. 23:

1. Divide the application into part applications. In the automotive field, this could be driving routes e.g. urban/interurban/expressway, driving maneuvers e.g. cornering/driving straight ahead, or geographical and climatic conditions e.g. Northern Europe/Africa or winter/summer. Different applications are treated as different applications; in a worst-case approach, the application relevant to design (the one with the smallest reliability) is singled out.
2. System partitioning of the product into components.
3. Calculate the load-time characteristics $L_{ikl}(t)$ of the *i*-th component for the *k*-th part application for a representative period from the system loads, typically by means of multibody simulation (MBS) or measurement, possibly on a predecessor product.
4. Determine the relevant damage mechanisms.
5. Determine the stress parameter for the *j*-th damage mechanism. Plot the stress-time characteristics $P_{ijklm}(t)$ for the *ikl*-th load in respect of the *j*-th damage mechanism, typically through calculation using the Finite Element Method (FEM) or measurement using strain gages.
6. Calculate the equivalent stress-time characteristic $P^v_{ijk}(t)$ by weighting and superposing all stresses from the loads of the *i*-th component, *j*-th damage mechanism and *k*-th part application by means of the design concept. (In the case of fatigue: employ the relevant multiaxial approach.)
7. Classify $P^v_{ijk}(t)$ to form a collective K_{ijk}. If necessary, transform the collective with the aid of the design concept. (In the case of fatigue: transform to the stress ratio of the Wöhler curve.)
8. Determine the strength curve (Wöhler curve) BA_{ij} for the *i*-th component and the *j*-th damage mechanism, typically by means of a trial with subsequent transfer to the component using the design concept to take account of the differences between the sample and the component. (In the case of fatigue: e.g. surface quality and stress gradient.)
9. Calculate the total damage D_{ijk} of the *i*-th component for the *j*-th damage mechanism and the *k*-th part application from the respective collectives and Wöhler curves.

10. Calculate the total damage D_{ij} of the i-th component for the j-th damage mechanism by scaling (extrapolation) and adding the damage totals D_{ijk} over all part applications. This step may be skipped if necessary, as then rare events that are treated as a part application (e.g. restricted driving to the next service station), can be taken into consideration during Step 13 by applying the theorem of total probability and superposing the reliabilities of all part applications.

11. Calculate the total damage D_i of the i-th component by weighting and adding the damage totals D_{ij} over all damage mechanisms with the aid of the design concept. This typically presents a major difficulty, because in this case a linear damage accumulation is often not appropriate. The superposition of fatigue and creep in the case of high-temperature fatigue can be mentioned as an example.

12. Calculate the reliability R_i of the i-th component from its damage total D_i.

13. Calculate the system reliability R from the reliabilities of its components R_i. Verify reliability by comparing R with the required system reliability R_{erf}. Alternatively, target values can be defined after system partitioning for the required reliability of the components $R_{i,erf}$, at which the verification of system reliability is assured. In this case, reliability can be verified earlier on at component level, by comparing R_i with $R_{i,erf}$ for the i-th component. If reliabilities of individual part uses with known probabilities of occurrence are present, these may be superposed to form a total reliability by applying the theorem of total probability (see section 6.1.1), taking their probabilities of occurrence into consideration.

The path from the requirements to the reliability assessment is represented schematically by means of a (fictitious) drill chuck of a hammer drill for home use. It must be borne in mind that this is a schematic presentation that cannot incorporate all aspects of a reliability assessment in the necessary depth, and that illustrates some relations in simplified and abridged form.

Example:

According to definitions by the "home use" market segment, first of all customer requirements for the hammer drill product are ascertained through surveys, and technical product characteristics developed on this basis using the QFD method. A comparison of the competition reveals that the product must cover the two following part applications with a survival probability of $R_{erf} = 99\%$ *:*

Part application 1: 20 min hammer drilling in defined concrete with pressing force of $F_1 = 100$ N.

Part application 2: 40 min drilling in defined masonry with pressing force of $F_2 = 100$ N.

These requirements are aimed at the uppermost system level (product use). The pressing forces on the handle of the hammer drill may have been determined by trials on similar predecessor products, for example. Following further detailing of the design, loads for all subsystems or components, such as transmission, impact mechanism, etc. result from this system load at the uppermost system level. For an SDS drill (component 1), the following load-time characteristics can be deduced from the loads F_1 *and* F_2 *by means of a 1-minute multibody simulation of the (hammer) drill process:*

Load-time characteristic 111: Load from impact force $F_{111}(t)$, t = 0,...,1min,

Load-time characteristic 112: Load from torsional moment $M_{112}(t)$, t = 0,...,1min,

Load-time characteristic 121: Load from torsional moment $M_{121}(t)$, t = 0,...,1min.

The supplier of the tool needs this data in order to perform successful design reliability. As knowledge of the complete system is required to derive this data, these requirements can only be provided by the product manufacturer (OEM).

From the load-time characteristics 111, 112 and 121, once a tool design is available the tool supplier can now calculate stress by means of a structural-mechanical FEM simulation. From experience, he knows that fatigue can be considered as a damage mechanism (SM 1) for the tool, and therefore defines mechanical stresses in the tool as a local stress variable. Assuming a linear material characteristic, a linear-elastic FEM simulation delivers the (location-dependent) transfer factors $c_{F=1}(\mathbf{x})$ *and* $c_{M=1}(\mathbf{x})$ *for a unit load or a unit moment. These can then be used to calculated the (location-dependent) stress-time characteristics. If* $c_{F=1}$ *and* $c_{M=1}$ *stand for the coefficients at the* failure location, *the following stress-time characteristics result:*

Stress-time characteristic 1111: Normal stress $\sigma_{1111}(t) = c_{F=1} \cdot F_{111}(t)$, t = 0,...,1min,

Stress-time characteristic 1112: Shear stress $\tau_{1112}(t) = c_{M=1} \cdot M_{112}(t)$, t = 0,...,1min,

Stress-time characteristic 1121: Shear stress $\tau_{1121}(t) = c_{M=1} \cdot M_{121}(t)$, t = 0,...,1min.

The individual stress-time characteristics can now be superposed. Stress-time characteristics 1111 and 1112 concern the components of a multiaxial stress which, depending on the multiaxiality approach used, can be weighted and superposed to form a equivalent stress distribution (more details in this connection can be found e.g. in [5], p. 162ff):

$$\sigma_{111}^{V} = \sigma_{111}^{V}(\sigma_{1111}(t), \tau_{1112}(t)), t = 0..1\text{min},$$

$$\sigma_{112}^{V} = \sigma_{112}^{V}(\tau_{1121}(t)), t = 0..1\text{min}.$$

The scalar equivalent stress characteristics can now be classified into collectives:

$\sigma_{111}^{V} \rightarrow K_{111,1min}$,

$\sigma_{112}^{V} \rightarrow K_{112,1min}$,

thus producing mean-stress transformed collectives at

$K_{111,1min} \rightarrow K_{111,1min,tr}$,

$K_{112,1min} \rightarrow K_{112,1min,tr}$.

These transformed collectives represent the actual stress we are looking for. For a damage assessment of the "fatigue" mechanism, the strength of the drill needs to be determined in the form of a Wöhler curve for a cyclic pressure load:

BA_{11}.

Thus, for damages arising from 1 minute of use:

$D_{111,1min} = D_{111,1min}(K_{111,1min,tr}, BA_{11})$,

$D_{112,1min} = D_{112,1min}(K_{112,1min,tr}, BA_{11})$.

These partial damages can now be projected and added for the required time of part use:

$D_{11} = 20 \cdot D_{111,1min} + 40 \cdot D_{112,1min}$.

Alternatively, the collectives can be upscaled with subsequent damage calculation. However, this method of damage calculation is considerably longer, whereas upscaling the damage total requires only multiplication by one factor, making this procedure much more efficient.

D_{11} *is the damage total of the only damage mechanism, "fatigue", under observation, thus* $D_1 = D_{11}$. *The reliability of the tool* R_1 *can now be expressed from the damage total* D_1 *as*

$$R_1 = R_1(D_1, D_{50\%}),$$

whereby $D_{50\%}$ *signifies the critical damage total, which is known from previous endurance tests.*

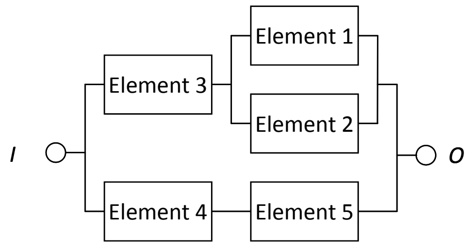

***Fig. 24:** Reliability diagram of the drill chuck/drill system*

In order to verify sufficient reliability of the overall "drill chuck/drill" system, the reliability of the remaining system components must also be assessed in the same way. Let us assume that the system consists of 5 elements: 1 – SDS drill, 2 – SDS drill bit, 3 – SDS drill chuck, 4 – additional hexagonal drill chuck, 5 – hexagonal drill. The survival probability of the individual elements is R_1 *to* R_5. *The system reliability is determined using the Boolean system theory on the basis of the reliability of the elements, using the reliability diagram of the structure, Fig. 24. Elements 1 and 2 form a parallel structure, as both SDS drills must fail in order for the "SDS drill" subsystem to not function. The subsystems 12 and 3 form a serial structure, as both the "SDS drill" subsystem and the "SDS drill chuck" subsystem are required for a functioning SDS drill chuck/drill subsystem. The same applies to the hexagonal drill chuck/drill subsystem comprising elements 4 and 5. In order for the overall drill chuck/drill system to function, either the 123 or 45 subsystems must be functioning, thereby forming a parallel structure.*

To find the solution, first of all we look for the elementary structures in the block diagram. Elements 1 and 2 form a parallel structure, therefore

$$R_{12} = 1 - (1 - R_1)(1 - R_2)\,.$$

In the following illustrations, elements 1 and 2 are eliminated and replaced by a fictitious element 12 with the above survival probability. Elements 3 and 12 form a serial structure, as do elements 4 and 5, therefore:

$$R_{123} = R_3 R_{12} = R_3\big(1 - (1 - R_1)(1 - R_2)\big),$$

$$R_{45} = R_4 R_5\,.$$

Elements 3 and 12 are eliminated and replaced with the fictitious element 123, the same applies to elements 4 and 5, which are replaced with element 45. Finally, elements 123 and 45 have a parallel structure, therefore the survival probability of the system is

$$R = R_{12345} = 1 - (1 - R_{123})(1 - R_{45}) = 1 - \big(1 - R_3\big(1 - (1 - R_1)(1 - R_2)\big)\big)\big(1 - R_4 R_5\big)\,.$$

The reliability of the drill chuck/drill system is regarded as verified if

$$R \geq R_{erf} = 99\%\,.$$

3.5. Qualitative methods

This section discusses several methods that provide a qualitative means of assessing or increasing reliability.

3.5.1. Failure Mode and Effects Analysis (FMEA)

The FMEA is an analytical method of preventive quality assurance. It is used to identify potential weak points, to recognize and evaluate their significance and to introduce measures in good time to avoid or disclose them. The systematic analysis and elimination of weak points minimizes risks, lowers failure-related costs and improves reliability.

The FMEA is a method for analyzing risks posed by single failures. In this process, individual risks are weighted against one another to reveal the main focal points. The FMEA does not provide information about the absolute level of the risk of a failure. Fault tree analysis is more suited to examining combinations of failures.

Targeted application of the FMEA brings the following advantages:

- Avoidance of failures in design and development.
- Fewer subsequent product modifications and so reduced costs.
- Avoidance of repeat failures through the systematic acquisition of specialist/failure knowledge about the product or process.

The time spent on avoiding failures at the start of the product creation process is more than justified, as it eradicates the considerably more substantial follow-up costs that would be incurred at a later point.

The FMEA is a systematic method. Its basic principle is to ascertain all conceivable types of failure for any systems, subsystems and components. At the same time, it uncovers possible failure consequences and causes. Next, the risk is evaluated and optimization measures determined. The FMEA can thus be performed in 5 steps:

1. Structuring
2. Functional analysis
3. Failure analysis
4. Analyses of actions
5. Optimization

Details on performing an FMEA can be found in [9].

3.5.2. Design Review Based on Failure Mode (DRBFM)

DRBFM is a systematic, preventive approach in the product creation process, which has the aim of avoiding potential problems before they occur and completely satisfying technical requirements. While the FMEA reveals all conceivable weak points through systematic analysis, and is therefore a broad assessment, DRBFM focuses on particularly relevant areas that are then dealt with "in depth".

The most important benefits of DRBFM consist in:

- tracking down potential problems even in the draft phase, taking fundamental issues into consideration in suitable fashion during the design process and so avoiding problems.

- pushing ahead with detailed work, gaining an understanding of cause-effect relationships and influencing parameters.
- systematically revealing gaps in knowledge.
- enabling associates to pursue objective debate.
- including supervisors in design reviews and encouraging them to provide leadership in terms of content.
- reducing failure rates and failure-related costs.

DRBFM is a method that focuses on concerns raised by modifications to a reference design. This reference is regarded as understood, robust and proven in the field. It may originate from predecessor products or from a different area of application.

The DRBFM method basically consists of two elements:

- analysis and improvement of the design parallel to the development process, and
- the design review.

To put this method into practice, the development process must be "front-loaded", i.e. the main focus must be on the early stages, and technical experts must be involved. The method is incorporated into everyday development work.

DRBFM follows the GD^3 philosophy:

- **Good Design:** A good design is the basis for a product that is outstanding, competitive and simple to produce. A good design is always a robust design. A good design is not modified without a compelling reason or without in-depth analysis.
- **Good Discussion:** Good discussions are already held in the early stages of development. Different points of view are applied involving the SE team, technical experts, Production, suppliers, etc. These discussions are open, constructive, focus on causes and are conducted with respect for all parties.
- **Good Design Review:** Thorough and detailed examination of the draft by an interdisciplinary team of experts and executives. The developer(s) explain(s) his/their draft and invite(s) discussion. Drawings, actual parts, experimental results and field data are employed for this purpose.

Further details on this subject are contained in the Committed BES Practice, "Design Review Based on Failure Mode (DRBFM)".

3.5.3. Fault Tree Analysis (FTA)

Fault tree analysis (FTA) is a deductive, analytical method that is used to find combinations of failures (primary events) of individual functional elements of an object under observation, which lead to an undesired event, or "top event". Based on Boolean algebra with the help of corresponding symbols, a graphical representation, the fault tree, is created. This tree expresses the logical interaction of primary events. When probabilities of single events are also expressed, it can be used to calculate the probability of the top event. As such, this method enables not just qualitative but also quantitative assessments to be made.

In a second step, the fault tree is transformed into logical expressions, with the help of which all repetitions of events are eliminated. Then the "shortest" event path, which leads to the "top event", can be determined (minimal cut set; minimal set of subsystems whose failure causes the system to fail).

An FTA consists of the following steps:

- Prepare an inclusive system analysis
- Define the undesired event (top event)
- Establish the response variables of analysis
- Produce the fault tree (qualitative description)
- Qualitative evaluation
- Determine the reliability characteristics of the basic events (quantitative description)
- Quantitative evaluation
- Determine the need for action, measures, monitor success
- Documentation.

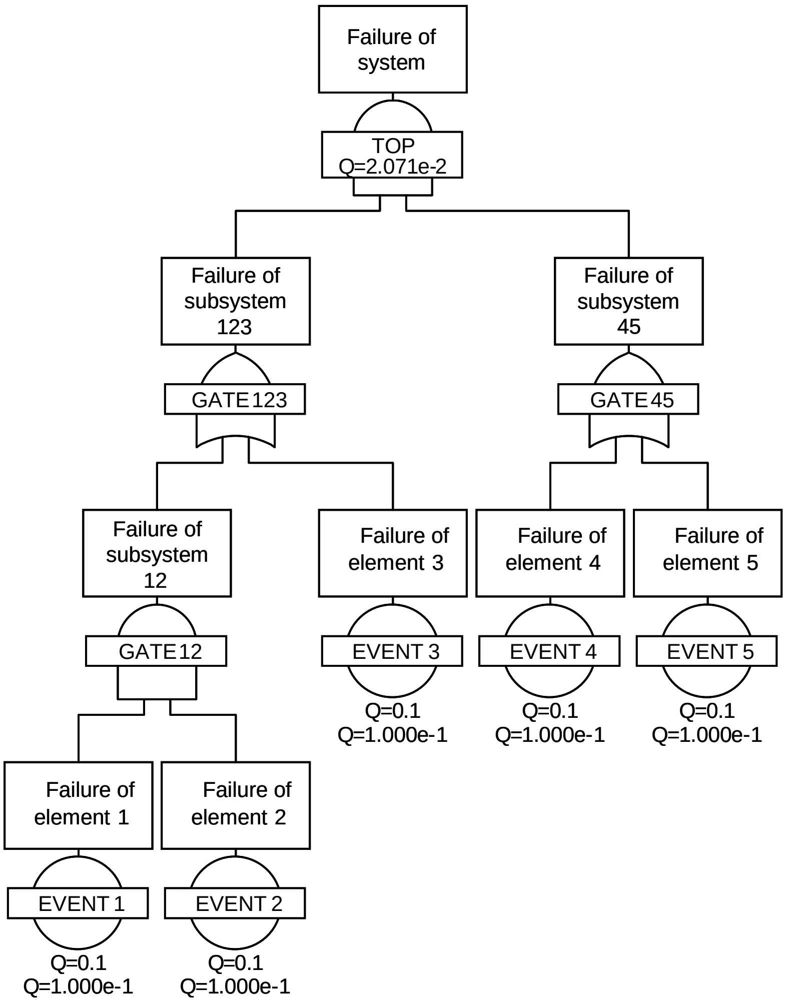

Fig. 25: Fault tree for the example in section 3.4.6, generated using the FaultTree+ tool. Failure probability Q of the top event "system failure", as a function of the failure probabilities of the individual elements $Q_i = 0.1$.

4. Verification & validation

4.1. Trials and quality

The quality and reliability of products must be ensured on the launch of series production, especially in the automotive industry, which is no mean feat given the current reduction of development times to 3 years and the corresponding pressure on costs. Virtually all relevant standards, rules, directives, etc. e.g. [7,8] require products to be subject to trials during their development.

Here, a trial is understood to be a test conducted on a technical product to verify and validate quality features, including reliability.

According to DIN EN ISO 9000, verification is confirmation, through the furnishing of objective proof, that set requirements have been met. Validation, on the other hand, is said by the above standard to be confirmation, through the furnishing of objective proof, that the requirements for a specific use or application have been met.

Thus, trials are intended to confirm that the set requirement has been satisfied and the reliability of the product has been proven. The latter also explicitly takes account of the aspect of product liability. Furthermore, trials are intended to provide knowledge about various influences (e.g. environmental) on products and processes and, in particular, to increase data certainty.

To quote [7]: "Confirming or improving concepts after design through experiments lays the foundations for a quality product." The trial is thus understood to be a core component of quality assurance. On the other hand, the shortened development times and increased cost pressure mentioned above have led to a certain paradigm shift: trials are no longer regarded as the only means of proof. [7] therefore recommends that "… trials focus on new components or technologies. The scope of trials conducted until now must be critically analyzed and reduced if necessary … Then, through constructive optimization, trials can be avoided in many cases." However, now as ever it remains clear that "trials are necessary wherever required technical features cannot be determined satisfactorily or at all during development using theoretical methods" and that "components that remain high risk require the added certainty of a trial" [7].

Trials are based on the following principles:

- Trials must be started as early as possible. It is clear, however, that trials cannot be conducted at very early stages of product development when no prototypes are as yet available. The result is an iterative procedure, whereby the latest prototype version at any time undergoes trials.
- The specimens themselves, and the conditions of the trial, must correspond as far as possible to the established requirements that need to be fulfilled (verification). The result is a procedure whereby the latest prototype version at any time undergoes trials under the relevant trial conditions, such as loads or media. Finally, the evaluation of field data of products already in series production enables conclusions to be drawn about the relevance of the established requirements (validation).
- An important basic principle is working with standard test procedures, which are governed by specific sets of rules, e.g. DIN IEC 68 "Basic environmental testing procedures", or VDI 4005 "Influences of environmental conditions on reliability of technical products". However, the individual, product-specific test severity is agreed between the principal and the customer. Test procedures related to release are set out in a testing sheet.

Two trial philosophies with their respective advantages and disadvantages apply to the above principles.

- "Test to pass" refers to a procedure whereby a particular trial must be passed without adverse effects on the product in order for the test to be rated as positive. Only a lower threshold for the reliability of the product can be deduced from the results of this test, however; the actual distance between the load and the load that can be withstood remains unclear, which harbors considerable risks for subsequent use. This approach is closely related to a development philosophy whereby trials are seen as the satisfaction of customers' wishes, for trials of this kind are always set by the customer. Such requirements are often based on higher-level sets of rules, so that the product can be regarded as "fit for standard" after it has passed the test. The advantages of this procedure are wide acceptance of the results and a degree of certainty, as one can always invoke the existence of a prescribed trial. The disadvantages are obvious, however: the field relevance of such tests is frequently unclear, and problems can occur in the field despite such a trial being passed.
- "Test to failure" refers to a procedure whereby the limits of the product are determined in a more stringent trial, the end-of-life test, which will generally result in clear adverse effects or even failure. This type of trial is understood here as an important development activity, which is intended to flag up risks and open the door to improvements. "Assessment" of the product is not achieved directly as a result of the trial, but in a separate investigation, which incorporates the results of the trial and the field-relevant conditions of use. In this approach, the existence of prescribed tests is of lesser importance: these may be important in order to satisfy the customer, but are of limited conclusiveness. As field relevance has great importance in this approach, following a positive rating the product can be classed as "fit for use". A clear advantage here is that when correctly applied, field problems cannot occur because the limits of the product and the field-relevant conditions of use are known, and the distance between the two has been classed as adequate. The greatest difficulty lies in determining the field-relevant conditions of use, which must generally be agreed with the customer.

Generally speaking, we can conclude that the advantages described above point to a clear shift from the first to the second of these two trial philosophies. Consequently, "fitness for intended use" is the quality objective to be aspired to for all products, beyond compliance with agreed specifications, and this includes the development of test methods for relevant loads in the field.

Detailed instructions on the trial process are contained in process procedures. A test plan is used to systematically establish the necessary tests, with information on

- the type of test (vehicle testing, trial on customer's premises, static/climatic/dynamic/other tests, endurance tests on test benches),
- the test duration (e.g. in the form of a testing sheet), and
- the intended number of samples.

As a rule, a selection sheet listing the relevant types of test for the product area in question is used for drawing up the test plan.

4.2. Test planning

4.2.1. Number of specimens and test duration for a success run test

To determine the number of required specimens, the following procedure described in [1], p. 267f, based on binomial distribution, can be used.

n identical specimens that have the survival probability $R(t)$ at test time t are examined. The probability that *all* specimens survive this time is $R(t)^n$ according to the product law. Thus, the probability of at least one random failure by the time t is $P_A = 1 - R(t)^n$.

By reversing this calculation, we can establish that if no failure occurs during the test of a sample of size n by the time t, the minimum survival probability (and thus reliability) of a specimen equals $R(t)$, with a confidence level of P_A. Thus

$$R(t) = (1 - P_A)^{1/n} \tag{4.1}$$

and accordingly

$$n = \frac{\ln(1-P_A)}{\ln(R(t))}. \tag{4.2}$$

This type of test planning is referred to as "success run", because the test involved is without failures. The dependence of the number of specimens on the required reliability with a given confidence level is illustrated in Fig. 26. It is clear that experimental proof becomes increasingly complex for higher levels of reliability.

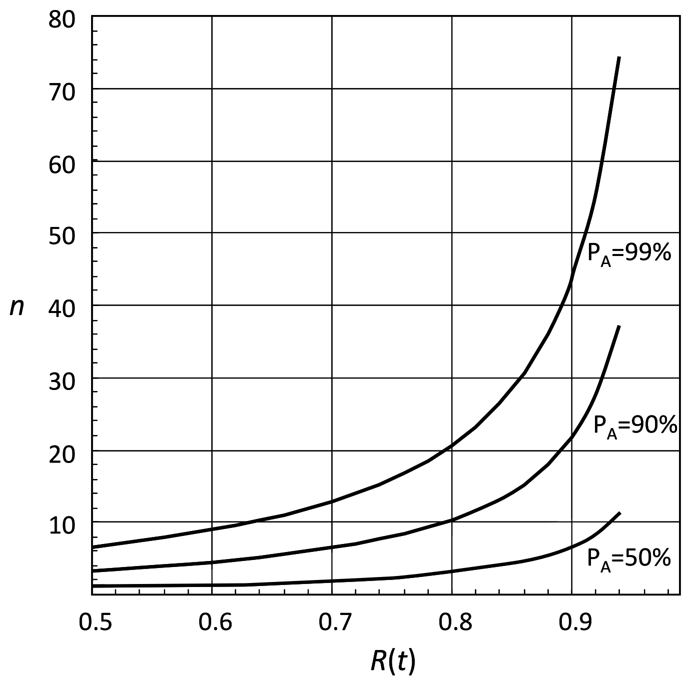

Fig. 26: Number of specimens n for a success run test, as a function of the reliability of a specimen to be demonstrated $R(t)$ and the required confidence level P_A.

If a statement now needs to be made about reliability over a time $t_0 \neq t$, and a Weibull distribution with parameters b, T is assumed, for the two times t and t_0

$$R(t) = \exp\left(-\left(\frac{t}{T}\right)^b\right), \quad R(t_0) = \exp\left(-\left(\frac{t_0}{T}\right)^b\right). \tag{4.3}$$

Following the logarithmizing and division of the two equations, the result is (also see [1], p. 268):

$$\frac{\ln(R(t))}{\ln(R(t_0))} = \left(\frac{t}{t_0}\right)^b = L_V^b \,, \tag{4.4}$$

with the lifetime ratio $L_V = t/t_0$. Consequently $R(t) = R(t_0)^{L_V^b}$. Applying this in equation (4.2) results in:

$$n = \frac{\ln(1-P_A)}{L_V^b \ln(R(t_0))} \,. \tag{4.5}$$

Therefore, at a constant reliability to be demonstrated of $R(t_0)$ and a confidence level of P_A, increasing the test duration *t* leads to a reduction in the number of specimens *n* and vice versa.

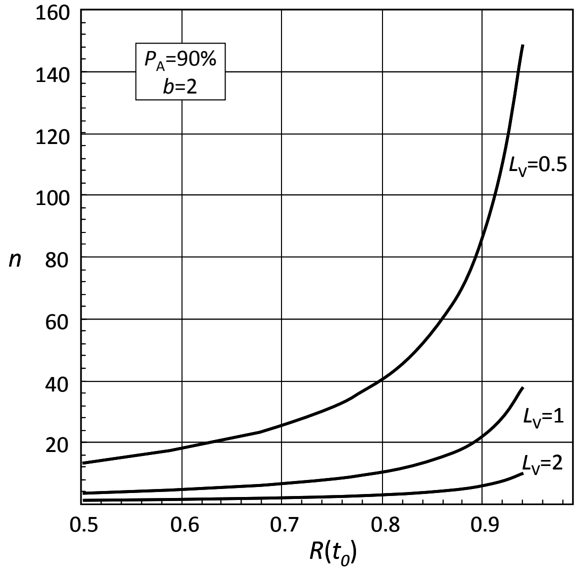

Fig. 27: Number of specimens *n* for a success run test, as a function of the reliability of a specimen to be demonstrated $R(t_0)$ at differing lifetime ratios L_V (confidence level P_A = 90%, *b* = 2).

4.2.2. The consideration of failures

In the equations above, failures during the test time can be taken into consideration as described below. The confidence level is generally subject to binomial distribution as per [1], p. 272:

$$P_A = 1 - \sum_{i=0}^{x} \binom{n}{i} (1-R(t))^i R(t)^{n-i} \,, \; R(t) = R(t_0)^{L_V^b} \,, \tag{4.6}$$

with *x* - number of failures during the test period *t* and *n* - number of specimens. Generally speaking, this equation cannot be explicitly resolved by *n*, but must be determined iteratively. This can be avoided by means of a graphical evaluation using the Larson nomogram. Details on this are not contained here, but can be found in [1].

4.2.3. The consideration of prior knowledge

Bayesian statistics (Thomas Bayes, 1702-1761) deal with estimation and test methods under the consideration of prior knowledge. Normally, one estimates the unknown parameters μ (mean) and σ (standard deviation) of a normal distribution through the variables $\bar{x}$ and *s*, which are determined from the results of a sample. Confidence intervals (by $\bar{x}$ or *s*) can be stated for both parameters, which cover the unknown values μ and σ, see section 6.1.3.4. The widths of these intervals depend upon the sample size and the chosen confidence level. Inasmuch as "suitable prior knowledge" of μ

and σ is available, estimate values can be calculated on the basis of the Bayesian theorem, the confidence levels of which are smaller than those attained without consideration of prior knowledge.

Generally speaking, combining prior knowledge (in the form of an "a priori distribution") about the characteristic of a random quantity with additional knowledge derived from experimental results delivers improved information ("a posteriori distribution") about the random variable.

The above-mentioned example is only one of the many application examples of the Bayesian theorem, but it already shows its problematic aspects: what does "suitable prior knowledge" mean?

If applied to the problem of estimating the characteristic lifetime T and the shape parameter b of the Weibull distribution, the prior knowledge about the failure characteristic of a product could be used to improve the estimate of b and T using the results of a current lifetime test. Provided that the prior knowledge ensued from experiments on a sample of specimens that are comparable with the specimens currently under test, the Bayesian theory would, as expected, deliver the same result as if both samples were grouped in one unit.

If, on the contrary, the prior knowledge was based on a different sample version or on a similar product, its applicability would depend upon the degree of similarity of the objects under test. The subjectivity of this kind of decision is one of the reasons why Bayesian theory is the subject of debate in the professional world and why its practical applicability is controversial. A detailed discussion of this topic can be found in [17].

4.2.4. Accelerated lifetime test

In many cases, the test times required to demonstrate compliance with reliability requirements under "normal" stress are unacceptably long in relation to the overall time available for the development of a product. Endeavors are therefore made to employ methods that curtail or limit the test times. Here, a distinction is drawn between the following subgroups:

- Parallel testing: If the equipment and technology permit, the simplest procedure is the simultaneous testing of specimens. Here, it must be borne in mind that the actual failure times of parts tested in parallel must be recorded, and the failure of an individual part must not exert any influence on the failure characteristic of the remaining parts. Since these requirements are sometimes difficult to meet, it may be appropriate to combine parallel testing with evaluation strategies for incomplete tests.
- Incomplete tests: Methods whereby the test time or the number of failures is established prior to the test (e.g. sudden death test, degradation test), and whereby the "incompleteness" of the recorded lifetimes is accounted for by a special evaluation.
- Physical acceleration: Method whereby conclusions on lifetime under "normal" load are drawn based the results of tests under an increased load on the basis of a physical model. Such a conclusion is obviously only justifiable on the assumption that the increase in load does not provoke any change in the failure mechanism.

No further explanation is required for the performance of parallel tests. The evaluation of incomplete tests is described in more detail in section 4.3. Physical acceleration, however, requires more in-depth discussion.

The simplest method of physical acceleration consists in increasing the load intensity without changing the load level, e.g. omitting idle times, raising load frequencies, etc. However, caution is recommended, because unexpected phenomena may occur, such as marked heating when the load frequency is raised, for example, which will contribute to a shortened lifetime.

Physical acceleration can also be achieved by increasing not the load but the stress, by varying a control factor (e.g. wall thickness). Here, too, undesirable side effects (e.g. changed surface properties after modification of the wall thickness) must be taken into consideration.

If a stress P (e.g. temperature, acceleration, voltage) is applied, after a certain time t (general value of the lifetime characteristic, e.g. operating time, number of load changes, of pulses, etc.) the object to which this is applied will demonstrate a certain degree of damage. The relationship between P and t can frequently be expressed by

$$t \cdot f(P) = const. \qquad (4.7)$$

If equation (4.7) is applied once for t,P and once for t_0,P_0, and both equations are divided, the relationship can be expressed as follows:

$$\frac{t}{t_0} = \frac{f(P_0)}{f(P)}. \qquad (4.8)$$

If such a relationship is valid, we would expect that an increased stress P within the shortened time t will provoke the same damage (same failure probability) as the original stress P_0 within the original time t_0.

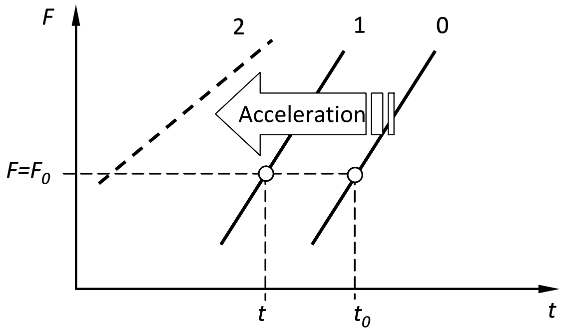

Fig. 28: Physical acceleration on the probability paper (schematic). Line 0 – initial failure probability distribution, Line 1 – distribution after a permitted physical acceleration, Line 2 – distribution after an impermissibly high physical acceleration, which causes the failure mechanism to change.

Examples of the relationships expressed in equation (4.7) are the Arrhenius equation (temperature T as stress parameter P)

$$t \cdot \exp\left(-\frac{c}{T}\right) = const. \qquad (4.9)$$

and the Coffin-Manson formula (number of load changes N as lifetime characteristic t)

$$N \cdot P^k = const., \qquad (4.10)$$

whereby c and k denote constants.

Based on the Coffin-Manson relationship, equation (4.8) means that

$$N = N_0 \cdot \left(\frac{P_0}{P}\right)^k, \qquad (4.11)$$

which is also the typical representation of a Wöhler curve as is used to describe fatigue failures, for example, see section 3.4.3. Thus, the increased stress $P > P_0$ gives rise to a lifetime $N < N_0$ reduced by the factor

$$\left(\frac{P_0}{P}\right)^k =: \frac{1}{\kappa_{CM}}. \qquad (4.12)$$

The factor κ_{CM} is known as the Coffin-Manson acceleration factor, and is frequently used for the physical acceleration of lifetime tests under mechanical load.

In physical chemistry, the temperature dependency of a reaction rate is expressed by the Arrhenius equation:

$$v = v_0 \exp\left(-\frac{c}{T}\right). \tag{4.13}$$

Here, c denotes the ratio of activation energy and Boltzmann constant E_a/K, although c can also be expressed as a free constant that has to be adapted in line with test results. v_0 signifies a reaction constant.

If a model is developed whereby a physicochemical process takes place during damage through exposure to temperature, which only leads to failure once a critical "material conversion" is reached, the time until this material conversion is attained can be interpreted as lifetime t. Thus, the lifetime is inversely proportionate to the reaction rate $t \sim 1/v$, and under consideration of equation (4.13), the direct result is equation (4.9).

If equation (4.9) is applied once for t,T and once for t_0,T_0, and these equations are divided, the relationship is as follows:

$$t = t_0 \exp\left(\frac{c}{T} - \frac{c}{T_0}\right). \tag{4.14}$$

Accordingly, at a higher temperature $T > T_0$, the lifetime $t < t_0$ is reduced by the factor

$$\exp\left(\frac{c}{T} - \frac{c}{T_0}\right) =: \frac{1}{\kappa_A}. \tag{4.15}$$

κ_A refers to the Arrhenius acceleration factor, and can be used very effectively for the purpose of time-acceleration.

It must be noted that an acceleration factor is based on a concrete failure mechanism, i.e. different failure mechanisms can display different constants c and k, so that their acceleration factors will also differ. A physical acceleration can be taken into consideration during test planning by replacing the quantity L_V with $\kappa \cdot L_V$ in equation (4.5) or (4.6), with $L_V = t/t_0$, t – test duration (after physical acceleration), t_0 – period of time in the field.

4.2.5. Reliability Growth Management (RGM)

RGM is a concept for planning, tracking and observing the reliability of products and serves as a visual aid during the development phase. This strategy, which was developed by Duane in the 1960s, is based on the idea that the reliability of a product increases as a result of efforts undertaken during development. Here, Duane established that such an increase cannot grow in an unlimited fashion, as experience shows, but moves within certain limits (possibly subject to organizational units).

Tracking whether the current level of reliability lies within these aforesaid limits can be employed as an early indicator of possible problems. Reliability is measured by the statistical characteristic MTTF (Mean Time To Failure). Duane's approach assumes that double logarithm plotting of MTTF versus the cumulative testing time will produce a straight line with a positive slope when reliability is increasing.

The procedure can be described as follows:

- In the planning phase of a reliability growth program, a target lifetime is established (in the diagram: $MTTF_F$=3000h), and a planned cumulative total testing time T (in the diagram: T = 10000 h), after which this target value should be reached. These two values define an end point in the RG diagram (the indices I und F stand for Initial and Final).
- The initial reliability , $MTTF_I$, is the mean lifetime reached at the end of an initial testing phase or an estimate value based on earlier developments (starting point).
- The straight line between the starting point and end point in the RG diagram establishes the desired growth rate for product reliability. During the course of development, this rate must be achieved, at least, when the set reliability target has to be attained within the planned time and with the available resources. The two straight lines above and below the target line correspond to a growth rate of 0.3 (lower line) and 0.6 (upper line) respectively, and indicate a typical growth rate.
- The current reliability determined through tests is plotted in the RG diagram. A comparison with the straight lines established during the planning phase renders any deviation from the planned growth obvious.

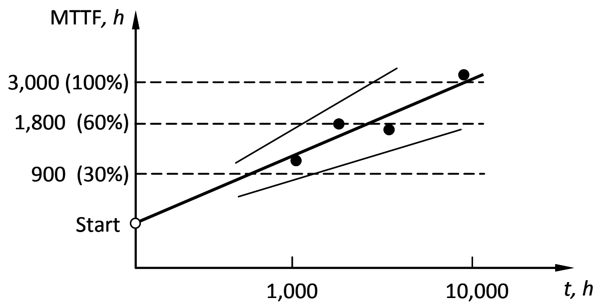

Fig. 29: RGM diagram

The RGM approach adds the elements planning, tracking and visualization of development progress to the performance and analysis of lifetime tests. It suffers major disadvantages, however:

- It conveys the impression that the reliability of a product ("reliability growth") rises automatically and purely due to the fact that products are being tested. In reality, only current reliability can be measured or proven in tests. The improvement of reliability, on the other hand, is based on an improvement to the design during the development process.
- The RGM approach does not take into consideration the different efforts of a successful and a less successful development process. Furthermore, the empirical values for typical growth are based on experience from the 1960s. However, development times and methods have changed dramatically since then.
- This method suggests that long development cycles with sophisticated tests using the "trial and error" approach are inescapable. However, in reality an optimum is achieved when a largely mature design is developed based on knowledge of cause-effect relationships within the framework of a preventive approach, so that this design must then only be verified in experiments.

For this reason, the RGM method should be treated with some skepticism, and employed solely as a planning and tracking instrument.

4.3. Evaluation of lifetime data

The evaluation of lifetime data entails adapting the data to a particular distribution and determining its distribution parameters. The question generally arises as to how the parameters of the population, from which this data originates, can be determined from the data of a sample. The most important graphical and analytical procedures are presented below.

4.3.1. Complete and und incomplete (censored) data

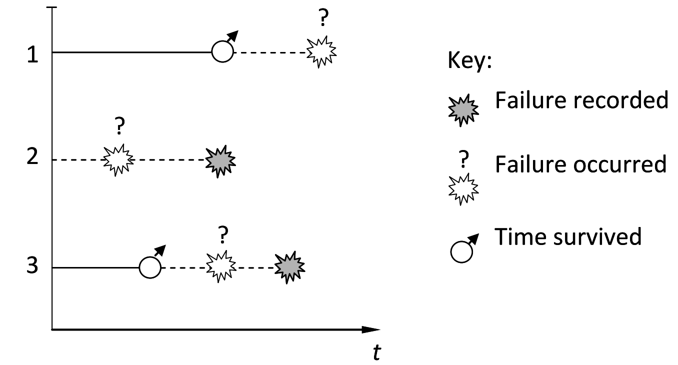

Fig. 30: Types of censoring: Right (1), left (2) and interval censoring (3).

If a sample of *n* parts is tested in a lifetime test until the failure of all specimens, this is referred to as a complete sample. If, on the other hand, some specimens are removed from the test before failure, the test is discontinued before the failure of all objects or the precise time of failure is unknown, this is known as an incomplete sample. In this connection, we refer to censored samples.

We generally distinguish between three types of censoring, Fig. 30:

- Right censoring: Here, we know that a specimen has survived up to certain point in time. However, the (later) time at which this specimen failed (or will fail) is unknown.
- Left censoring: Here, we know that a specimen has already failed at a certain point in time. But we do not know exactly at what (past) time this failure occurred. This case arises, for example, when no automatic failure monitoring has taken place, and failures that may have occurred at any time after the start of the test are revealed upon the first manual inspection.
- Interval censoring: This type of censoring is a combination of the first two types. Here, the failure of a specimen may have taken place in a certain interval between two moments in time (e.g. between two inspections), but the exact time is unknown.

The right censoring method has the greatest practical relevance. Here, in turn, we distinguish between three types, Fig. 31:

- Type I censoring: The test is discontinued at a fixed point in time *t*. The number of failed specimens r, $0 \leq r \leq n$ up to this time is a random variable.
- Type II censoring: The test is discontinued after the failure of a predefined number *r* of specimens. *r* = 1 is the so-called sudden death test. In this case, the number *r* is known before the start of the test, but not the time *t* at which failure occurs. Here, the total test time is a random variable.
- Mixed censoring: Objects under test frequently have to be withdrawn from an experiment before failure. Unlike with type I or II censoring, in which the surviving runouts are all removed from the experiment at the same time, here there may be different times of "removal". This is the case when products are being tested for a particular type of failure (e.g. in the electronics), and different types of failure occur (e.g. mechanical).

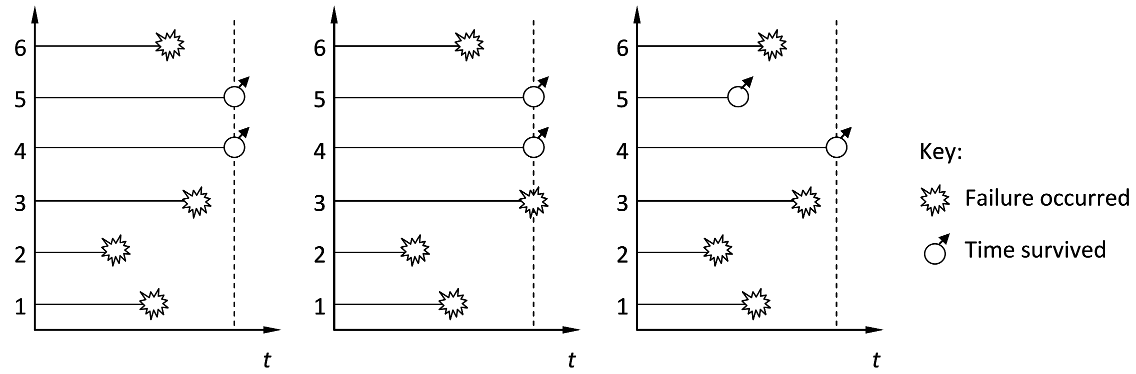

Fig. 31: Types of right-censored data. Type I censoring (left): Experiment stopped after a prescribed time. Specimens 4 and 5 reach the end of the experiment without failure. Type II censoring (center): Experiment stopped after a prescribed number of failures, in this case 4. Specimens 4 and 5 are still intact at the time of the most recent failure. Multiple censoring (right). Specimen 4 is removed from the experiment early, specimen 5 reaches the end of the experiment without failure.

Special mathematical techniques are employed to evaluate incomplete samples. Although no failure time exists for the specimens that have been removed from the test, the test time completed without failure constitutes vital information that must be considered during the evaluation.

4.3.2. Graphical evaluation in the Weibull plot

A double logarithmic expression of the distribution function of the Weibull distribution

$$F(t) = 1 - \exp\left(-\left(\frac{t}{T}\right)^b\right), \quad (4.16)$$

produces the following

$$\ln\left(\ln\left(\frac{1}{1-F(t)}\right)\right) = b \cdot \ln(t) - b \cdot \ln(T) \ . \quad (4.17)$$

This is a linear equation of the form $y = a_0 + a_1 \cdot x$, with $x = \ln(t)$. In a graph that has been transformed accordingly ("Weibull plot"), the distribution function of the Weibull distribution would appear as a straight line. By entering observed failures in the Weibull plot and through subsequent approximation with a straight line, estimate values of the two Weibull parameters b and T can be defined.

Although these days graphical evaluation no longer seems in keeping with the times and can be accomplished by standard software such as Weibull++ from the Reliasoft company, this representation is important for an understanding of the relationships. Moreover, most of the techniques presented here are also implemented in commercial software packages (possibly in modified form).

4.3.2.1. Complete data

This section describes the procedure for graphically determining estimate values of the parameters characteristic lifetime *T* and shape parameter *b* for complete data , see [1], p. 203f.

Procedure:

1. Order the *n* failure times t_i by size, the smallest value with rank 1, the largest with rank *n*.
2. Calculate the failure probability F_i for each rank number *i* with

$$F_i = \frac{i-0.3}{n+0.4}. \quad (4.18)$$

3. Enter the pairs of variates (t_i,F_i) in the Weibull probability paper.
4. Determine a best-fit line through the points (graphically or analytically). If the plotted points lie on the best-fit line to a sufficient extent, the failure characteristic can be adequately described by means of a Weibull distribution.
5. The point at which the best-fit line intersects with the horizontal 63.2 % line has the characteristic lifetime *T* on the time axis, which can be read there.
6. Enter a parallel line to the best-fit line through the pole point. The shape parameter *b* can only be read on the right-hand shape parameter scale (graphical evaluation). In the case of an analytical evaluation, the shape parameter *b* reflects the slope of the best-fit line.

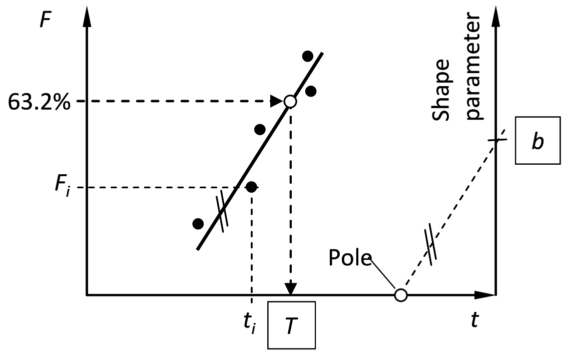

Fig. 32: Graphical definition of Weibull parameters in the probability plot (schematic)

REMARKS:

Equation (4.18) is an approximation of the exact relationship, which can be deduced on the basis of the binomial distribution, see section 6.1.3.5 in the Annex.

If we drop a perpendicular line from the point where the best-fit line intersects with the horizontal 10% line on the time axis, the lifetime t_{10} *can be read. The same applies to the 50% line and all further quantiles.*

The scaling of the logarithmic time axis (t-axis) can be adapted to the occurring failure times by multiplying it by a suitable scale factor. If we select the factor 10h (10 hours), for example, then

number	*1*	*on the t-axis corresponds to the value*	*10*	*hours,*
number	*10*	*on the t-axis corresponds to the value*	*100*	*hours,*
number	*100*	*on the t-axis corresponds to the value*	*1000*	*hours.*

Instead of for the time, scaling can also be carried out for the mileage, the number of load changes, switching operations or work cycles, for example.

Example:

A sample of 10 electric motors was examined on a test bench after a predefined test cycle. Until their failure, the specimens achieved the following numbers of cycles (in ascending order and in units of 10^5 cycles):

Motor no.	1	2	3	4	5	6	7	8	9	10
Cycles until failure, $/10^5$	0.41	0.55	0.79	0.92	1.1	1.1	1.4	1.5	1.8	1.9
Failure probability, %	6.7	16.3	26.0	35.6	45.2	54.8	64.4	74.0	83.7	93.3

Read: Characteristic lifetime $T = 1.3 \cdot 10^5$ *test cycles, shape parameter* b=2.1.

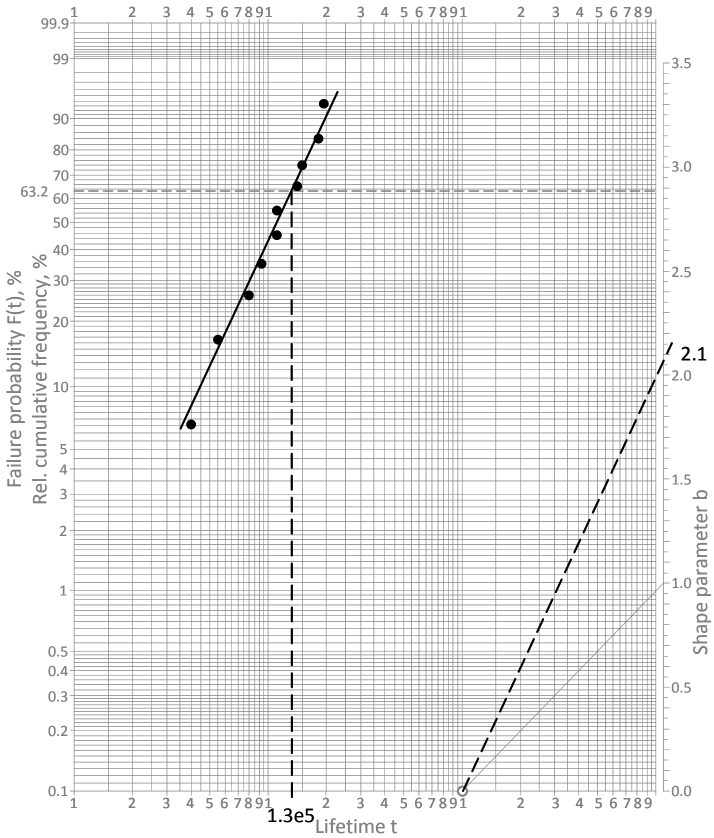

Fig. 33: Example in the probability paper

4.3.2.2. Incomplete data

Incomplete data is what is obtained when the failure times of some specimens are unknown. The test time that has elapsed for specimens that have not failed but were removed from the test, or for specimens that have not yet failed constitutes a vital piece of information that must be taken into consideration during evaluation, as shown by the following example.

Example:

A manufacturer has produced and sold 100 units. The following failures are recorded in the first three months after their sale:

Month	No. of failures
1	1
2	1
3	2

What is the mean lifetime? A naive (and incorrect!) evaluation would base the mean on the evident lifetimes, without taking into consideration the fact that the overwhelming majority of units has survived for at least 3 months:

$$\bar{t} = \frac{1 \cdot 1 + 1 \cdot 2 + 2 \cdot 3}{4} = 2.25\,. \qquad (4.19)$$

The mean lifetime obviously cannot be 2.25 months, for at this time it is not true that nearly half of the units has failed, as one would expect from the mean lifetime.

The sections that follow go on to describe suitable methods for the consideration of incomplete data.

Censoring types I and II

In both cases (types I and II censoring), the evaluation process is similar to that of a complete sample, see section 4.3.2.1. The failure probabilities can be calculated for entry in the Weibull plot using the approximation formula

$$F_i = \frac{i - 0.3}{n + 0.4} \text{ for } i = 1,2,\ldots,r \qquad (4.20)$$

(an exact relationship is shown in section 6.1.3.5 of the Annex).

The fact that n - r specimens have not failed is expressed by n, not r, occurring in the denominator in the approximation formula.

EXAMPLE (CONTINUED):

So, correct evaluation of the example from the previous section must also consider the fact that 96 parts have survived for 3 months (type I censoring). Thus, equation (4.20) delivers the following failure probabilities:

Failure no.	Time	F_i,[%]
1	1	0.7
2	2	1.7
3	3	2.7
4	3	3.7

A graphical evaluation is contained in Fig. 34. The following characteristic values were read: Characteristic lifetime T *= 35 months, shape parameter* b *= 1.4, time at which half of parts will have failed (median):* t_{50} *= 27 months.*

NOTE:

During the evaluation of experiments with type I or II censoring, it is mostly necessary to extrapolate the best-fit line in the Weibull plot beyond the time of most recent failure in order to estimate the characteristic lifetime T*. This is basically problematic, inasmuch as further failure mechanisms cannot be ruled out. In case of doubt, therefore, a statistical conclusion on the failure characteristic can only be based on the time between the smallest and greatest observed lifetimes.*

The note above can be generalized thus: With both complete and incomplete samples, evaluation in the Weibull plot provides information about the failure characteristic that is limited to the interval between the smallest and largest value of the lifetime characteristic.

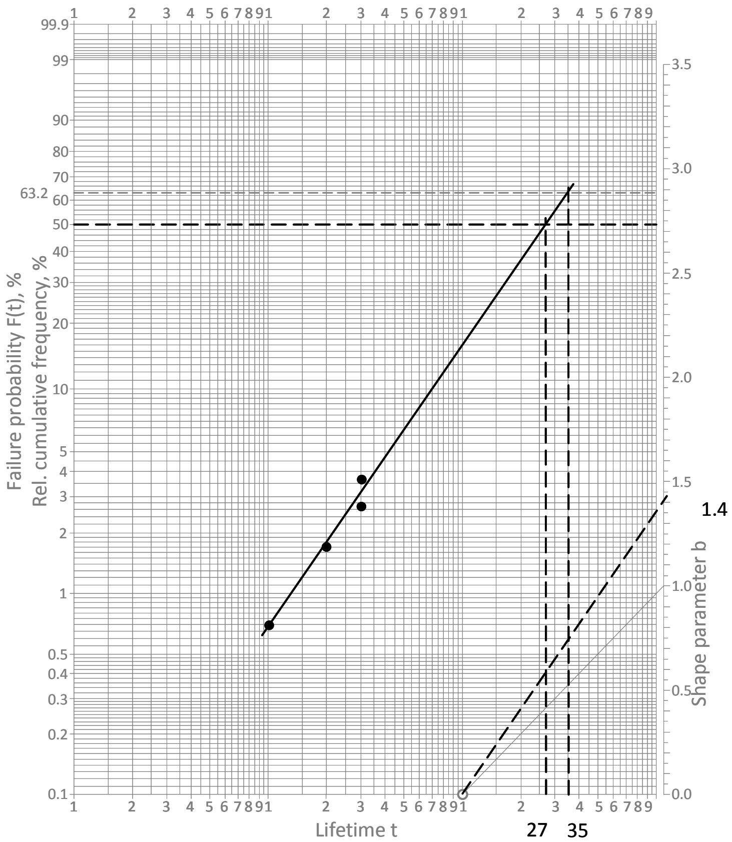

Fig. 34: Example of type I censored data in the probability paper

Multiple censoring

Multiple censoring describes the case where specimens are removed from the test before it has come to an end, in a destructive test, for example, or if field parts were intact at the time of inspection and after that point no more information about them is available. In another typical case, mixed failures from two different failure mechanisms have occurred. During evaluation, the failures of the other mechanism in each case are not neglected, but must be rated as runouts.

Johnson evaluation

[8] describes a procedure based on Johnson that is used for evaluating experimental results with multiple censoring. The background to this procedure can be explained as follows.

Let us assume the following experimental data:

Time	Event
t_1	Failure 1
t_2	Runout 1
t_3	Failure 2
t_4	Runout 2

The question now arises as to which ranks the two failures can be assigned to. The first failure obviously has rank 1, as no event happened previously. With the second failure, the situation is somewhat more difficult. Failure 2 may have either rank 2 or 3 depending on whether runout 1 would have failed before or after failure 2. The idea is to allocate a (non-integer) rank number between 2 and 3 to the second failure. This rank number may be based on the consideration as to which cases are possible with the available data:

Rank	Possibility 1	Possibility 2	Possibility 3
1	Failure 1	Failure 1	Failure 1
2	Failure of runout 1	Failure 2	Failure 2
3	Failure 2	Failure of runout 1	Failure of runout 2
4	Failure of runout 2	Failure of runout 2	Failure of runout 1

According to the above, failure 2 would have the rank number 3 in one case, and rank number 2 in two cases. Thus, its mean rank number would be:

$$i = \frac{1 \cdot 3 + 2 \cdot 2}{3} = 2.33\,. \tag{4.21}$$

With this mean rank number, the procedure described in section 4.3.2.1 from step 2 onwards can continue to be applied. If the quantity of data becomes larger, the case differentiation may become complex but can be expressed adequately using a combination of relationships.

NOTE:

With this method, only the number of runouts between two failures is taken into consideration, not the actual time experienced by a runout. Thus, situations in this example in which t_2 is slightly greater than t_1 or somewhat smaller than t_3 would be treated absolutely identically. Intuitively, however, we know that the lifetime in the first case must tend to be smaller than in the second case. These circumstances reflect a disadvantage in this method, to the extent that where data is greatly censored (fewer failures, many irregularly distributed runouts) alternative approaches not based on rank numbers, such as maximum likelihood, may well be preferable.

Procedure:

1. Enter n events (failures and runouts) in column a, the associated event times t_j in column b ($j = 1,\dots,n$).
2. Sort the events and event times by t_j.
3. Enter consecutive numbers j in column c, starting with $j = 1$.

For each line j:

4. Calculate the growth in rank number N_j and enter in column d. If the current event is a failure, the following applies:

$$N_j = \frac{n+1-i_{j-1}}{n+2-j}\,,\ i_0 = 0, \tag{4.22}$$

   with the previous rank number i_{j-1}. If the current event is a runout, then $N_j = 0$.

5. Calculate the mean rank number i_j from the previous rank number i_{j-1} and the current growth N_j and enter in column e:

$$i_j = i_{j-1} + N_j\,,\ i_0 = 0. \tag{4.23}$$

6. Continue with step 2 from section 4.3.2.1.

Nelson evaluation

A method suggested by Nelson [8] is described here as an alternative to the Johnson procedure. This method is explained below.

j is the number of an event (failure or runout) in the set order of the relevant times t_j. At the moment at which a failure occurs, $r_j = n - j + 1$ units are undoubtedly still intact, including the "just failed" unit. Earlier runouts, for which uncertainty exists as to whether they are still intact at this time, are omitted from the calculation. In other words, it is assumed that runouts between the last and the current failure time have failed now, at the latest. The number of still intact units r_i is sometimes confusingly referred to as "inverse rank". The ratio

$$\hat{\lambda}_j = \frac{1}{r_j} \quad (4.24)$$

is by definition an estimate value for the momentary failure rate, which is defined by

$$\lambda(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{R(t)} \quad (4.25)$$

The definition of the failure probability density

$$f(t) = \frac{dF}{dt} = \frac{d(1-R)}{dt} = -\frac{dR}{dt} \quad (4.26)$$

gives rise to

$$\lambda(t) = -\frac{1}{R}\frac{dR}{dt} \text{ and } \lambda(t)dt = -\frac{1}{R}dR . \quad (4.27)$$

Further calculations are based on the following relationships for the cumulative failure rate $H(t)$, which is produced by integrating the preceding equation:

$$H(t) = \int_{-\infty}^{t} \lambda(\tau)d\tau = -\ln(R(t)) = -\ln(1-F(t)) . \quad (4.28)$$

For the failure probability, this results in

$$F(t) = 1 - \exp(-H(t)) . \quad (4.29)$$

For the purpose of practical calculation, we recommend systematically entering the intermediate steps in a table with columns a to h.

Procedure:

1. Number lines $j = 1,2,\ldots,n$ and enter in column a.
2. Enter event times in order of magnitude in column b.
3. Enter the data belonging to the failure times, "A" for failure or "H" for removal, in column c.
4. Enter the number of still intact units $r_j = n - j + 1$ in column d.

Further entries are made only for the failures ("A"):

5. Calculate the reciprocals of the numbers from column d and enter them in column e.
6. Calculate the total of numbers from column e from the first to the *j*-th line and enter in the *j*-th line in column f. The values $\hat{H}_j$ may be greater than 1.
7. With the help of the values $\hat{H}_j$ from column f, calculate the exponential powers $\exp(-\hat{H}_j)$ and enter in column g.
8. Subtract values from column g from 1 and enter the results in column h.
9. Continue with step 3 from section 4.3.2.1.

The example below should further clarify this procedure:

EXAMPLE:

The result of a field test on 40 products delivers a result of t_j *complete operating hours (lifetimes). The failure of the products occurred due to both mechanical and electronic defects. In this example, we assume that an evaluation should be undertaken to examine electronic failures ("A" = failure in the table). Thus, the products that failed due to mechanical defects can be treated as though they had been removed from the experiment prior to the occurrence of the electronic defect in question ("H" = removal in the table). We know that they were electronically intact until the mechanical failure at the time* t_j. *Of course, the mechanical failures can also be evaluated in addition to the procedure described here.*

The calculations are summarized in Table 4.

The observed lifetimes (failures "A") are entered in the Weibull plot with the failure probability from column h, Fig. 35 shows the result of evaluation for this example. The results read were b≈1.4 and T≈8000h. The unshaded circles represent the points that would arise if we ignored the lifetimes of the removed products and evaluated the 8 observed failures just like a complete sample. In this case one would wrongly ascribe a substantially smaller characteristic lifetime as the result. The fact that the "H" parts were intact up until their removal is an essential piece of information and must be considered during evaluation.

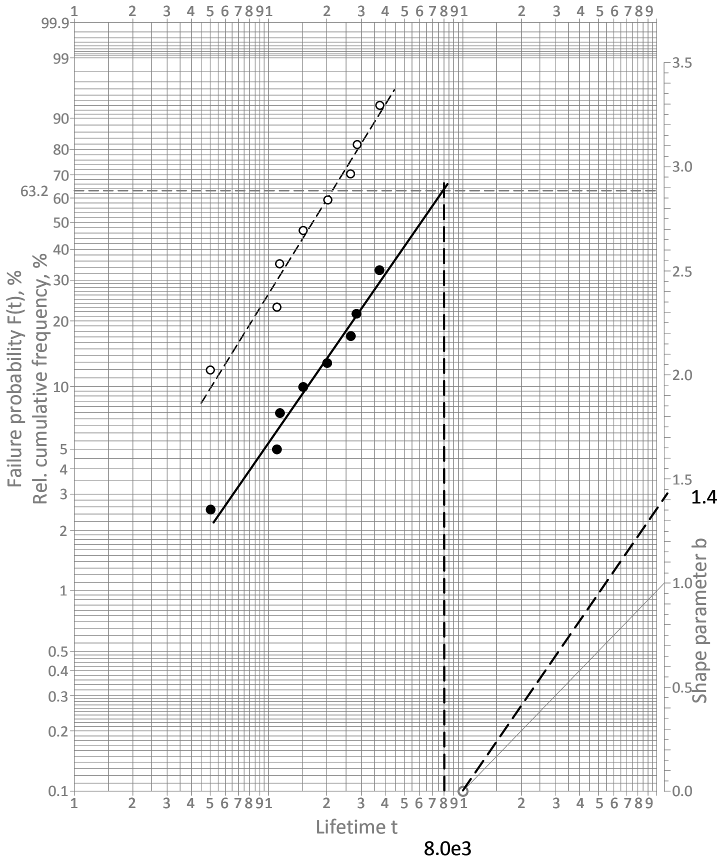

Fig. 35: Illustration of the Nelson evaluation on the probability paper (shaded circles). The unshaded circles represent an (incorrect) evaluation in which the lifetimes of the removed products were ignored.

a	b	c	d	e	f	g	h
j	t_j, [h]	A/H	r_i	$\hat{\lambda}_j = \frac{1}{r_j}$	$\hat{H}_j = \sum_j \frac{1}{r_j}$	$\hat{R}_j = \exp(-\hat{H}_j)$	$\hat{F}_j = 1 - \hat{R}_j$
1	513	A	40	0.0250	0.0250	0.9753	2.5%
2	1118	A	39	0.0256	0.0506	0.9506	4.9%
3	1186	A	38	0.0263	0.0770	0.9259	7.4%
4	1468	H	37				
5	1520	H	36				
6	1529	A	35	0.0286	0.1055	0.8998	10.0%
7	1688	H	34				
8	1907	H	33				
9	1919	H	32				
10	2094	A	31	0.0323	0.1378	0.8713	12.9%
11	2170	H	30				
12	2201	H	29				
13	2239	H	28				
14	2280	H	27				
15	2345	H	26				
16	2356	H	25				
17	2386	H	24				
18	2483	H	23				
19	2641	A	22	0.0455	0.1832	0.8326	16.7%
20	2724	H	21				
21	2736	H	20				
22	2739	H	19				
23	2796	H	18				
24	2865	A	17	0.0588	0.2421	0.7850	21.5%
25	2912	H	16				
26	2990	H	15				
27	3196	H	14				
28	3353	H	13				
29	3355	H	12				
30	3425	H	11				
31	3472	H	10				
32	3522	H	9				
33	3575	H	8				
34	3696	H	7				
35	3723	A	6	0.1667	0.4087	0.6645	33.6%
36	4319	H	5				
37	4343	H	4				
38	4563	H	3				
39	5360	H	2				
40	7497	H	1				

Table 4: Evaluation pattern for "multiple-censored data" according to Nelson

4.3.3. Analytical evaluation

Three analytical procedures can be used for determining the parameters of a distribution derived from experimental data:

- Method of moments: Here, characteristics of the population (so-called moments) are estimated on the basis of sample characteristics, and include elements such as expected value, variance, skewness, etc. Next, the parameters of the distribution are calculated from these estimated moments. This dependence between parameters and moments may be known theoretically, but is not that simple in some cases, such as in the 3-parameter Weibull distribution, for instance.
- With the regression method, the lifetime data is entered in the probability paper, then approximated to a straight line through a process of regression. The parameters of the distribution can now be estimated on the basis of this straight line.
- In the maximum likelihood approach, the parameters of a distribution are determined such that the probability of observing the lifetime values currently under observation is maximized. This probability is expressed by the so-called likelihood function.

These methods are discussed in further detail below.

4.3.3.1. Method of moments

In the method of moments, the sought-after parameters of a distribution are calculated from their moments. The following key figures are designated as a moment of the *k*-th order m_k:

$$m_k = \int_{-\infty}^{+\infty} t^k \cdot f(t)dt \text{ , } k = 1, 2, \ldots \qquad (4.30)$$

The following central moments also exist:

$$m_{kz} = \int_{-\infty}^{+\infty} (t - m_1)^k \cdot f(t)dt \text{ , } k = 1, 2, \ldots \qquad (4.31)$$

The best known moments are:

- expected value $\mathrm{E}(t) = m_1 = \int_{-\infty}^{+\infty} t \cdot f(t)dt$, (4.32)
- variance $\mathrm{Var}(t) = m_{2z} = m_2 - m_1^2$, (4.33)
- skewness $S_3(t) = \frac{m_{3z}}{\sqrt{m_{2z}^2}}$. (4.34)

The moments now have to be estimated from the empirical moments of a sample:

$$\mathrm{E}(t) \approx \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i \text{ ,} \qquad (4.35)$$

$$\mathrm{Var}(t) \approx s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2 \text{ ,} \qquad (4.36)$$

$$S_3(t) \approx \frac{n}{(n-1)(n-2)}\frac{1}{s^3}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^3 \text{ ,} \qquad (4.37)$$

and associated with the characteristic values of a distribution. For the Weibull distribution, for example, it can be shown (see [1] p. 240) that

$$\mathrm{E}(t) = (T - t_0)\cdot\Gamma\left(1+\frac{1}{b}\right) + t_0\,, \tag{4.38}$$

$$\mathrm{Var}(t) = (T - t_0)^2\cdot\left[\Gamma\left(1+\frac{2}{b}\right) - \Gamma^2\left(1+\frac{1}{b}\right)\right], \tag{4.39}$$

$$S_3(t) = \frac{\Gamma\left(1+\frac{2}{b}\right) - \Gamma^2\left(1+\frac{1}{b}\right)}{\Gamma\left(1+\frac{3}{b}\right) - 3\Gamma\left(1+\frac{2}{b}\right)\cdot\Gamma\left(1+\frac{1}{b}\right) + 2\Gamma^3\left(1+\frac{1}{b}\right)}\,, \tag{4.40}$$

whereby Γ(.) denotes the gamma function. From this relationship, the parameters T, b and t_0 can be deduced: first b from the third equation, then T and t_0 from the first and second equations.

NOTE:

Only complete samples can be assessed using the method of moments.

4.3.3.2. Regression method

Linear regression analysis is a method that enables best-fit lines to be calculated for a presumed linear relationship between two variables x and y, if the experimental data is available in the form of pairs of variates (x_i, y_i). Let us assume, for example, that n of these pairs are available:

$$(x_i, y_i)\,,\ i = 1,\ldots,n\,. \tag{4.41}$$

For $y = f(x)$ we suspect a linear relationship, from which the varying response y deviates only minimally and randomly by the amount ε:

$$y_i = a_0 + a_1\cdot x_i + \varepsilon_i\,. \tag{4.42}$$

During regression, it is now a matter of calculating the coefficients a_0 and a_1 in such a way as to enable the best possible prediction of y. Here,

$$\hat{y}_i = a_0 + a_1\cdot x_i \tag{4.43}$$

designates the value predicted from the best-fit line y at the point x_i. The "quality" of the prediction is assured using the familiar method of least squares, whereby the sum of squared deviations between the test value and the prediction at all n points x_i is minimized:

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - (a_0 + a_1\cdot x_i))^2 \to \min\,. \tag{4.44}$$

The necessary condition for minimizing the above function is that the first partial derivations in terms of a_0 and a_1 must disappear. The terms

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\,,\quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\,, \tag{4.45}$$

can be used to calculate the coefficients as follows (see [1], p. 243):

$$a_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\,,\quad a_0 = \bar{y} - a_1\cdot\bar{x}\,. \tag{4.46}$$

Under the precondition that

- the variable y is normally distributed for each setting of x,
- a linear relationship exists between x and the expected value of y, $E(y) = a_0 + a_1 \cdot x$,
- the deviations between this linear relationship are only random in nature, and
- the standard deviation characterizing this random spread is constant throughout the experimental x space,

it is possible to demonstrate that the estimate values calculated by means of linear regression for the straight line coefficients conform to expectations, i.e. their mean values over numerous tests concur with the actual values.

This strategy can be used for deducing the parameters of a distribution from the calculated coefficients of the best-fit line. For this purpose, a suitable transformation must be carried out, which can transform the distribution function into a straight line. For the two-parameter Weibull distribution, for example, the following relationships arise: Through a double logarithmic expression of the distribution function

$$F(t) = 1 - \exp\left(-\left(\frac{t}{T}\right)^b\right), \tag{4.47}$$

this expression is produced

$$\ln\left(\ln\left(\frac{1}{1-F(t)}\right)\right) = b \cdot \ln(t) - b \cdot \ln(T) \ . \tag{4.48}$$

This is a linear equation in the form $y = a_1 \cdot x + a_0$. The parameters of the best-fit line are produced from the above equation as

$$y_i = \ln(-\ln(1-F_i)), \ x_i = \ln(t_i) \ . \tag{4.49}$$

The parameters of the distribution can be deduced directly from the parameters of the best-fit line:

$$b = a_1 \ , \ T = \exp\left(-\frac{a_0}{a_1}\right) . \tag{4.50}$$

NOTE:

Regression analysis also enables the evaluation of incomplete samples. The failure probability F_i *required for the evaluation is estimated using the same approaches as described for the graphical evaluation, see section 4.3.2.*

In some software tools, x *is regarded as a dependent variable on which regression is based.*

4.3.3.3. Maximum likelihood method

The Maximum Likelihood Estimation (MLE) can be used very effectively to estimate parameters of a distribution from sample data, without the prior failure probabilities having to be estimated beforehand, as in the regression method. The technique is explained briefly below, further details can be found in [1], p. 246ff.

Let us assume there are n failure times t_i, $i = 1,...,n$. The unknown failure density function of the population, from which the observations originate, contains k unknown parameters ψ_j, $j = 1,...,k$, e.g. parameters b and T of the Weibull distribution, and is denoted by $f(t,\psi_1,...,\psi_k)$. The probability that the i-th failure takes place in the interval $(t_i, t_i+dt]$, i.e. precisely where it was observed, is by definition $f(t_i,\psi_1,...,\psi_k)dt$. Consequently, the probability that all failures simultaneously take place where they were observed (logical "and" operator) is a product of the probabilities of the individual failures:

$$P(t_1,...,t_n,\psi_1,...,\psi_k) = \prod_{i=1}^{n} f(t_i,\psi_1,...,\psi_k)dt \ . \qquad (4.51)$$

With the given failure times t_i, this probability obviously depends upon the choice of distribution parameters ψ_j. The principle behind Maximum Likelihood Estimation is to select the distribution parameters in such a way as to maximize this probability. To do so, first of all the likelihood function

$$L(t_1,...,t_n,\psi_1,...,\psi_k) = \prod_{i=1}^{n} f(t_i,\psi_1,...,\psi_k) \qquad (4.52)$$

is introduced. As dt is a constant that does not depend upon the variable parameters, the probability will be at its maximum at the same point as the maximum of the likelihood function. As the next step, the above function is expressed as a logarithm, which considerably simplifies further procedure. Here, too, as the logarithm is a monotonous function, the maximum of $\ln(L)$ is at the same point as the maximum of the function L. The result is:

$$\Lambda(t_1,...,t_n,\psi_1,...,\psi_k) = \ln[L(t_1,...,t_n,\psi_1,...,\psi_k)] = \sum_{i=1}^{n} \ln[f(t_i,\psi_1,...,\psi_k)] \ . \qquad (4.53)$$

$\Lambda(t_1,...,t_n,\psi_1,...,\psi_k)$ represents the logarithmic likelihood function. Its maximum (in terms of the parameters ψ_j) is at the point where all partial derivations disappear:

$$\frac{\partial \Lambda(t_1,...,t_n,\psi_1,...,\psi_k)}{\partial \psi_i} = \sum_{i=1}^{n} \frac{1}{f(t_i,\psi_1,...,\psi_k)} \frac{\partial f(t_i,\psi_1,...,\psi_k)}{\partial \psi_i} = 0 \ , \quad i = 1,...,k \ . \qquad (4.54)$$

The k equations contained in (4.54) can only be used to determine the k unknown parameters of the distribution. In general, however, these equations can be non-linear in the parameters, so that as a rule, software with a suitable numerical optimization process is required to solve them.

The process can also be applied without problem to censored data. To do so, the definition of the likelihood function in equation (4.52) is extended as follows, [14], p. 309:

$$L(t_1,...,t_n,\psi_1,...,\psi_k) = \prod_{i=1}^{n} L_i(t_i,\psi_1,...,\psi_k) \ , \qquad (4.55)$$

whereby $L_i(t_i,\psi_1,...,\psi_k)$ denotes the following terms:

$f(t_i,\psi_1,...,\psi_k)$, if t_i represents a failure, (4.56)

$F(t_i,\psi_1,...,\psi_k)$, if t_i represents a left-censored observation, (4.57)

$R(t_i,\psi_1,...,\psi_k)$, if t_i represents a right-censored observation, (4.58)

$F(t_i^R,\psi_1,...,\psi_k) - F(t_i^L,\psi_1,...,\psi_k)$, if t_i represents an interval-censored observation. (4.59)

These definitions directly reflect the properties of the likelihood function described above. The procedure for certain censoring types and concrete distribution functions such as the normal, exponential and Weibull distribution are not discussed any further at this point; further details can be found in [14], p. 309ff.

NOTES:

Unlike the regression method, MLE does not need a failure probability and is therefore independent from rank distribution. Censored data is taken into consideration with its exact time of observation, not only by the number of pieces of data, as with the rank regression method.

MLE can result in a systematic deviation of the estimation from the true value (bias), especially where smaller samples are concerned.

The above remarks lead us to recommend the use of this method for larger samples with more than 30 failures and strongly censored data (many runouts, possibly irregularly distributed). For small samples or incomplete data, the rank regression method is preferable.

4.3.4. The consideration of finite sample sizes, confidence intervals

If a parameter of the distribution of a population is estimated from a sample, some uncertainties always abound. If a mean of a sample is used as an estimate for the expected value, for example, in the event of multiple sampling the individual values and the mean calculated from them will necessarily fluctuate. This fluctuation will increase

- the greater the variance of the population, and
- the smaller the sample size.

The idea of the confidence interval is used to take into account this uncertainty when estimating the parameter on the basis of the (subjectively selected) sample size.

The term 'confidence interval' is understood as an interval calculated from sample values, which covers the true but unknown parameter of the distribution of a population with a predefined probability known as the confidence level P. For example, the selected confidence level may be $P = 90\%$. This probability states that when sampling is frequent, the confidence intervals calculated from it will cover the true parameter in 90% of cases, and will not cover it in only *10 %* of cases. The *10 %* is expressed as the probability for a type I error $\alpha = 10\%$, so that $P = 1 - \alpha$. Details on the various techniques for calculating confidence bounds are contained in the Annex.

REMARKS:

P *is frequently interpreted as a probability with which the true parameter of the population lies within the calculated confidence interval. This is incorrect, because the true value may be unknown but is not a random variable – it is a constant. The sample values and confidence intervals calculated from these are subject to fluctuations, however.*

A distinction is made between one and two-sided confidence intervals, depending on whether a lower/upper threshold or an interval estimation is required for the parameter of the distribution, Fig. 36. One-sided intervals are mostly employed for reliability assessments, as only a one-sided limit is needed for the parameters (e.g. *minimum* average lifetime).

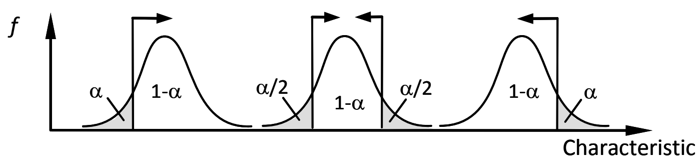

Fig. 36: One-sided (left and right) and two-sided (center) confidence intervals. The true parameter of the population is covered by the interval marked with an arrow with a probability of $P = 1 - \alpha$.

Confidence intervals can be stated for any estimated parameter. In a probability paper representation of failure data, the confidence interval of the failure probability function is a two-dimensional interval enclosed by curved lines (confidence bounds) above and below the calculated best-fit line, Fig. 37, right. All the straight lines in this interval correspond to distributions of the population that are possible but have varying probability, and which may have led to the failures observed in the sample. The confidence bounds lie closer together the greater the sample size (with the same confidence level) and the lower the confidence level (with the same sample size).

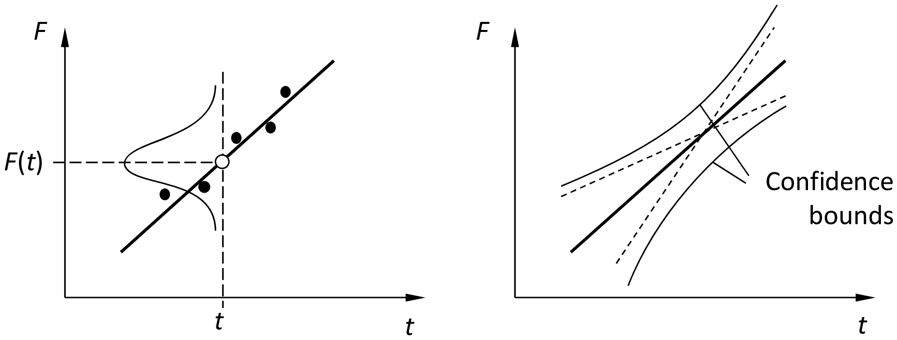

Fig. 37: Confidence interval (two-sided) in the probability paper.

A confidence level is typically taken into consideration in the evaluation of lifetime data by reading the quantile of the lifetime t_q not on the median (50 % value) but on the boundary of a given confidence interval, Fig. 38.

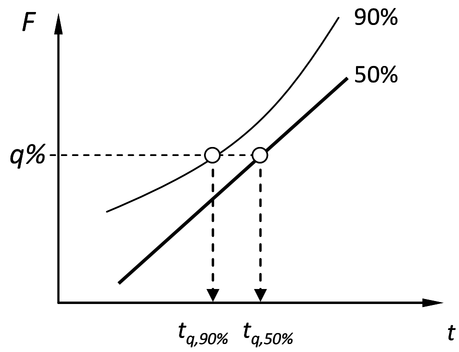

Fig. 38: Lifetime quantile t_q with confidence level 50 % or 90 % (one-sided)

If the uncertainty of a limited sample size is not taken into consideration, there is reason to fear far-reaching consequences, as the following example shows.

EXAMPLE:

A manufacturer must decide between two suppliers who have provided the following information on the reliability of the required part for a useful time of 3 years:

Supplier A	Supplier B
R = 99.4%	R = 99.9%

The manufacturer decides in favor of supplier B, as his product demonstrates a higher reliability. However, the supplied part suffers several instances of failure. What has happened?

An analysis has shown that the above data is based on the median (confidence level 50 %), as the manufacturer did not demand anything else. However, supplier B based his statements on a trial with only 5 parts, which by chance had a relatively high reliability. Supplier A, on the other hand, tested 25 parts. If the manufacturer had demanded reliability figures for a higher confidence level, e.g. P=*95 % (one-sided), his choice would have been different:*

Supplier A	Supplier B
$R_{95\%}$ = 97.1%	$R_{95\%}$ = 69.9%

4.4. Special experiments

This section discusses some special experimental techniques. Further details can be found e.g. in [15], section 11.

4.4.1. Lifetime test

Unlike the constant amplitude Wöhler experiment (S/N test), in which the load is kept uniform at one level, experiments must often be conducted with variable loads, as in reality uniform loads are rare. These experiments are known as lifetime tests.

Variable loads are typically represented as load collectives. A load collective is formed by first breaking down the load characteristic over time into a sequence of damage-relevant single events (e.g. load amplitudes with the damage mechanism fatigue), then classifying them. In each class, these single events are described by the load parameter P_i, which has a constant level within a class. The classes are referred to as part collectives. They are arranged in blocks, from the more to the less damaging parts, to represent the total collective. The size of the collective h_0 represents the sum of all part amounts h_i of the individual collective stages. Typically, the curve shown in Fig. 39 results. Load collectives are chiefly advantageous because they permit a much more compact representation than the original load-time series. Further details on different types of classification can be found in [5], section 2.

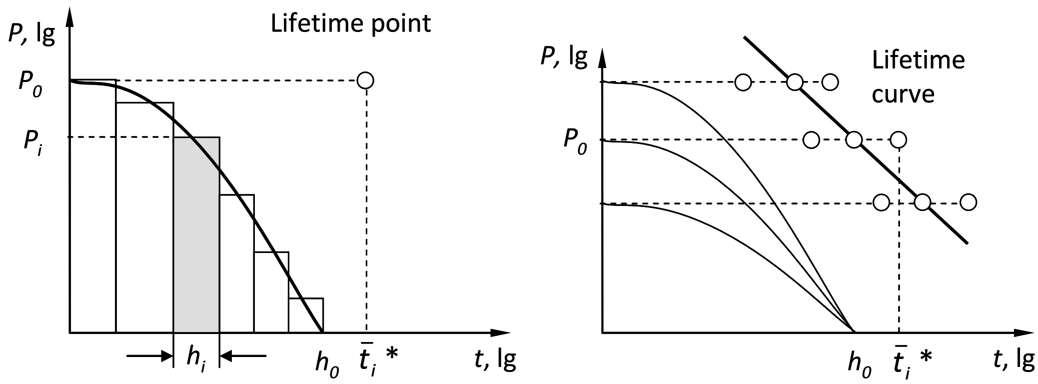

Fig. 39: Load collective and lifetime point (left). Lifetime curve (right).

The lifetime test is performed by repeatedly applying the variable load until failure. However, the application of the load over time does not have to correspond to the order of the block sequence mentioned above. Rather, the load is applied through a random sequence of loads of differing levels, the frequency of occurrence of which corresponds to the part frequency in the load collective. This is because applying the loads in blocks produces effects that do not occur in reality.

A lifetime point ($\bar{t}_i^*, P_0$) represents the lifetime $\bar{t}_i^*$ for a certain load (that varies over time), whereby the representation takes place at the level of the highest collective stage P_0, Fig. 39 left. The associated load collective is typically expressed in addition to enable the load to be visualized.

A lifetime test is repeated several times under the same load and with (nominally) the same specimens in order to record lifetime variance. The lifetime data is evaluated using the procedure described in section 4.3. To ascertain the lifetime at a different load level, all part stages of the load collective are changed proportionately such that the highest stage reaches the desired level. As with the Wöhler curve, the lifetime curve for variable loading, Fig. 39 right, is produced. In double logarithmic form, this often appears as a straight line.

4.4.2. Sudden death test

The sudden death test (test to the first fracture) is a time-reducing method for testing the failure characteristic of parts in test bench experiments.

The available total quantity of objects under test is divided into *m* groups of the same size (test lots), each with *k* specimens. *k* is the number of available test benches, or the number of specimens that can be simultaneously tested on one test bench. The first group (test lot) is subjected to load in an endurance test until the first specimen in this group fails. The time until the first failure is recorded.

Next, the second group is subjected to load in the same way until the first failure. This process is repeated for all *m* groups. In this way, *m* values of first failures are obtained.

Since the experiment is stopped after the first failure (first fracture) within a respective test lot, a special case of type II censoring with $r = 1$ is employed for this method.

4.4.2.1. Simplified graphical evaluation

[1], p. 221 suggests a simplified graphical evaluation, which we now go on to describe.

The lifetime values are arranged in ascending order and entered in the Weibull probability paper, as with the classic graphical Weibull evaluation. Each of these first failures is assigned the median of rank sizes based on the approximate relationship

$$F(t_i) = \frac{i - 0.3}{m + 0.4} \tag{4.60}$$

Alternatively, the exact formula from section 6.1.3.5 in the Annex can be employed. The straight line of the first failures is produced in the probability paper.

We can now show that the shape parameter of a Weibull distribution is the same both for a subset of a sample and for the entire sample. In other words, the shape parameter of the first failures corresponds to the shape parameter of the overall distribution; thus, the sought-after shape parameter *b* can be determined directly from the straight line of the first failures.

The straight line must be corrected, however (moved to the right) in order to precisely represent the failure characteristic and to determine the scale parameter. This correction is necessitated by the fact that a failure probability

$$F_1^* = \frac{0.7}{k + 0.4} \tag{4.61}$$

can be assigned to the first failure of a test lot, and the median (50 % value) of the calculated first failures is taken as a representative value for this failure. A vertical line is therefore drawn through the point of intersection of the 50% probability curve and the straight line of the first failures. The point where this intersects with the F_1^* line produces the straight line of overall distribution. The straight line of the first failures must be shifted parallel to this point, Fig. 40. The characteristic values of the distribution can now be read on the shifted straight line in the usual manner.

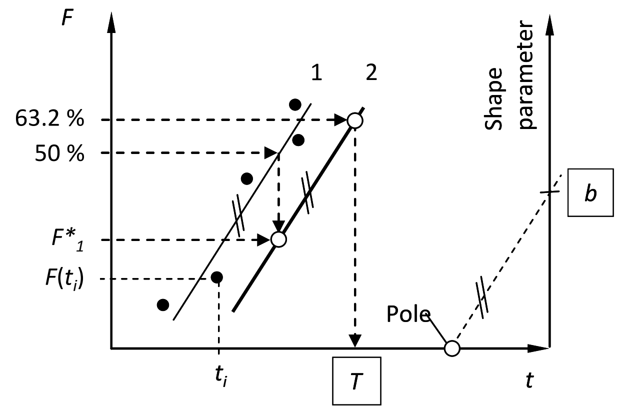

Fig. 40: Simplified graphical evaluation of a sudden death test (schematic). Line 1 – straight line of first failures, line 2 – straight line of total sample

4.4.2.2. Nelson evaluation

Below we discuss the Nelson method of evaluating sudden death data (example from [8]).

EXAMPLE:

In a laboratory experiment, n = 54 *parts were tested using the sudden death method. As 6 test bench spaces were available,* m = 9 *test lots of 6 parts each* (k = 6) *were tested in parallel. After each first failure, the entire test lot was replaced by the next 6 parts. In the following table, the lifetimes of the first nine failed parts are listed in ascending order (column b). In column c, A stands for failure and H for removal.*

The evaluation is carried out using the steps described in section 4.3.2.2.

a	b	c	d	e	f	g	h
j	t_j, [h]	A/H	r_i	$\hat{\lambda}_j = \frac{1}{r_j}$	$\hat{H}_j = \sum_j \frac{1}{r_j}$	$\hat{R}_j = \exp(-\hat{H}_j)$	$\hat{F}_j = 1 - \hat{R}_j$
1	10	1 A 5 H	54	0.0185	0.0185	0.9816	1.8%
7	14	1 A 5 H	48	0.0208	0.0393	0.9615	3.9%
13	16	1 A 5 H	42	0.0238	0.0631	0.9388	6.1%
19	18	1 A 5 H	36	0.0278	0.0909	0.9131	8.7%
25	20	1 A 5 H	30	0.0333	0.1242	0.8832	11.7%
31	23	1 A 5 H	24	0.0417	0.1659	0.8471	15.3%
37	25	1 A 5 H	18	0.0556	0.2215	0.8013	19.9%
43	28	1 A 5 H	12	0.0833	0.3048	0.7373	26.3%
49	32	1 A 5 H	6	0.1667	0.4715	0.6241	37.6%

Table 5: *Evaluation pattern for a sudden death test (example from [8])*

4.4.3. Highly accelerated test methods

The methods described below are elements that add to the trial or reliability test portfolio. They are not suitable for verifying the reliability of systems or components for which real operating load must be taken into consideration, and thus are no substitute for the quantitative reliability assessment.

4.4.3.1. Step-stress method (load increase)

In the step-stress method, the load applied to the specimen is increased in predetermined time intervals. This approach has the objective of shortening the failure times (accelerated life testing), under the assumption that the increase in load does not cause any change in the failure mechanism.

Where there is a fixed shape parameter *b*, each load level corresponds to a straight line in the Weibull plot. Therefore, calculation must be based on the failure straight lines belonging to the "basic load" on the basis of the failures observed at the higher load levels. The result is a recursive method that can, however, only function if a relationship is known or is assumed to be known between the load level and the characteristic lifetime *T*.

Due to these required assumptions, the method described here appears problematic, to say the least. The principle is presented in Fig. 41.

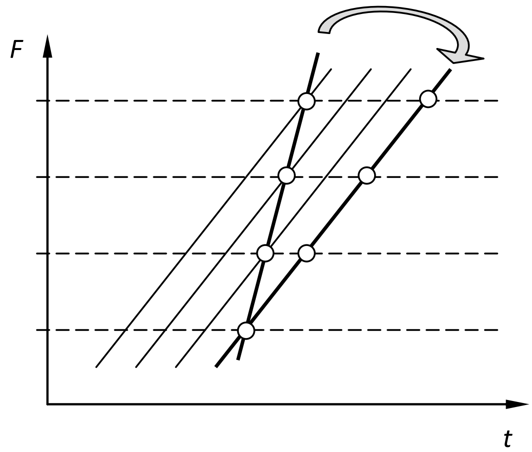

Fig. 41: Step-stress method in the Weibull probability paper (schematic)

4.4.3.2. Highly Accelerated Life Testing (HALT)

HALT stands for Highly Accelerated Life Testing, a method of accelerated testing discussed in the literature since 1993. Use of this method is preferable with electrical and electronic components, less so with mechanical components [1], p. 284f.

The basic principle behind this test method is, similarly to the step-stress method, to determine the relevant failure mechanisms in the shortest possible time by increasing a single load quantity step by step until failure, starting with the nominal load of the product. The step duration is constant.

However, as HALT works with loads that are considerably higher than the relevant load level in operation, this is a qualitative method for determining operating limits; it cannot be used to ascertain statistically relevant reliability values. This is also this method's principal difference from the step-stress approach. If failures occur, one must first examine very critically whether they can also occur under the intended "normal" operating conditions.

The HALT process commences with the analysis of possible loads, e.g. electrical, mechanical, thermal, etc. These must be established individually for each product. The necessary steps are run through iteratively below:

- Test with step-by-step load increase
- Analysis of test results, determination of root causes
- Optimization (of the design, the material, the supplier, etc.).

The functional and destruction limits are determined in these experiments. For this reason, it is important that the specimens are in operation and monitored during the tests. If the functional limits are exceeded, the object under test irreversibly becomes faulty; beyond the destruction limits, it is irreversibly (permanently) destroyed.

HALT is used:

- to detect weak points in an early sample phase,
- to determine the design limits of the design,
- to determine the load limit at which the damage mechanism of individual machine elements changes (this limit must be taken into consideration in the definition of QALT and HASS - see below),
- to determine robustness compared with other products (this is an effective means of estimating the effects of modifications),
- to recognize subsequent misuse in the field on the basis of damage patterns within a findings endurance run.

It is not possible to quantitatively predict reliability on the basis of data from these experiments, however.

4.4.3.3. Quantitative Accelerated Life Testing (QALT)

A QALT is a test that demonstrates the durability of a design element in respect of a single failure mode. Tests are conducted with omission and/or increased load, to reach a known failure mode as quickly as possible. The total stress of the design element at this failure mode is determined from the load collective of the product. This total stress is applied uniformly to the machine element, although loads outside the specification may also be employed.

Various methods of time acceleration can be used, depending on our understanding and the nature of the failure mode:

- Omission of unused times
- Omission of non-damaging times
- Conversion of a low stress level to a higher level by means of a suitable acceleration factor, to achieve failures more quickly.

QALTs are used:

- for time acceleration when testing a specific design element,
- as a means of ascertaining acceleration factors as compared with long-term endurance runs, vehicle tests or the field,
- to verify/demonstrate the long-term stability of a design element if the acceleration factor is known.

4.4.3.4. Stress Screening and Highly Accelerated Stress Screening (HASS)

The typical stress screening definition for electronic components features 100% testing under stricter test conditions, with subsequent delivery of products that passed the test to the customer. The idea here is to remove faulty parts that suffer early failure in the first part of the bathtub curve. This procedure is only practicable if the delivered parts are not significantly damaged by the test.

The object under test is typically damaged in hydraulic/mechanical tests. Therefore, the typical stress screening definition for electronic components is interpreted for the testing of hydraulic and mechanical products as follows:

- HASS is a constituent of the reliability test.
- A HASS test is a QALT in which the known target value has been set such that reaching this target value enables a clear conclusion to be drawn about the quality of the design element under observation.
- If a target value is not achieved in the HASS, this generally suspends the parts.

HASS tests are used:

- for the rapid comparison of a specific machine element in products from ongoing production with the released sample phase,
- as a necessary prerequisite for the release of process modifications to a specific design element during serial production.

4.4.4. Degradation test

It is quite possible that no failures may occur during the available test time, so that no conclusions about reliability can be drawn. For some damage mechanisms, however, a quantity that increases over time, known as the damage parameter, can be ascertained and measured, and this characterizes the degree of damage. For example, the wear of a component can be determined on the basis of material abrasion, or the degree to which a solder connection is ruined on the basis of the electrical resistance.

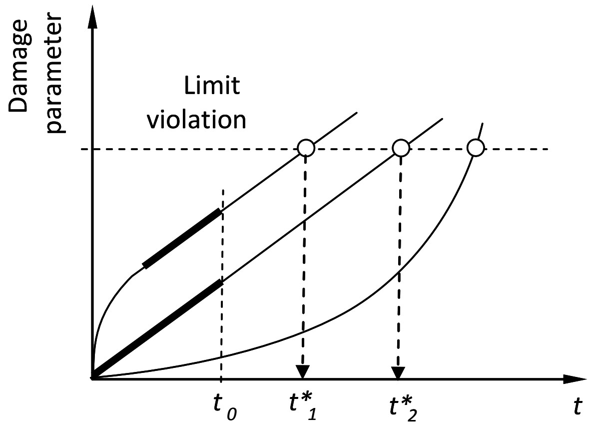

Fig. 42: Possible damage characteristics in a degradation test (schematic).

A helpful technique is now, on the basis of the damage characteristic determined in this experiment, to extrapolate this to the values at which it is assumed the object under test will fail. However, this is only possible if the type of characteristic (linear, exponential, etc.) is known either from physical relationships or from earlier tests. Otherwise, extrapolation cannot be permitted.

Different characteristics are possible, which must be taken into consideration, Fig. 42. t_0 denotes the aspired-to or maximum possible test duration, the line in bold signifies the sensible interval in which the damage parameter is recorded and can be extrapolated over the lifetime t^*. Extrapolation is often only feasible in the linear section of the characteristic.

Procedure:

1. Determine the functional relationship between the damage parameter Y and the lifetime characteristic t on the basis of physical relationships or several tests until failure. This relationship also depends upon k free parameters a_i , which must subsequently be experimentally adapted:

   $$Y = f(t, a_1, ..., a_k) . \quad (4.62)$$

2. Determine the limit violation Y^* from which failure can be expected. Establish the test duration t_0, measuring times $t_j < t_0$, $j = 1,...,n$ and number of specimens to be tested m.
3. For the i-th specimen, $i = 1,...,m$ conduct the trial and measure the damage parameter Y at the times t_j , $j = 1,...,n$.
4. From the experimental data of the i-th specimen (t_j, Y_j^i) , $j = 1,...,n$, determine the k free parameter in equation (4.62) by means of regression. Now, for the i-th specimen:

   $$Y^i = f(t, \hat{a}_1^i, ..., \hat{a}_k^i) . \quad (4.63)$$

5. Determine the lifetime t^i (time at which limit violation Y^* is reached) of the i-th specimen:

   $$t^i = f^{-1}(Y^*, \hat{a}_1^i, ..., \hat{a}_k^i) . \quad (4.64)$$

   To do so, the inverse function $t = f^{-1}(Y, a_1, ..., a_k)$ must be determined.

6. Repeat steps 3 to 5 for all specimens.
7. Evaluate the lifetimes t^i , $i = 1,...,m$ as described in section 4.3.

The example below further illustrates this procedure.

EXAMPLE:

A tire manufacturer knows from experience that the abrasion Y of his products has a linear dependence on the distance driven:

$$Y = a_1 \cdot t . \quad (4.65)$$

Therefore, for reasons of time, he does not test his tires until they are completely worn, but up to a distance of 10,000 km. The threshold wear for this manufacturer is 8mm. Higher values are prohibited due to legal requirements.

In a trial of a new product, 5 tires were tested. Their wear was measured every 1000 km.

The following data was recorded:

Distance driven	Abrasion in mm for tires				
km	1	2	3	4	5
1000	0.16	0.22	0.25	0.18	0.32
2000	0.36	0.49	0.61	0.35	0.57
3000	0.63	0.51	0.63	0.48	0.88
4000	0.83	0.99	1.12	0.86	1.10
5000	1.02	0.92	0.95	0.80	1.25
6000	1.24	1.43	1.62	1.19	1.55
7000	1.37	1.45	1.58	1.12	1.53
8000	1.71	1.69	2.02	1.48	2.34
9000	1.99	2.07	2.14	1.54	2.44
10000	2.27	2.07	2.09	2.00	2.39

Through linear regression, the following values of the parameter a_1 and lifetime $t = Y/a_1$ of each tire were determined, Fig. 43:

Parameter	Tire				
	1	2	3	4	5
a_1, mm/km	2.28E-04	2.15E-04	2.20E-04	1.87E-04	2.46E-04
t, km	3.50E+04	3.72E+04	3.64E+04	4.27E+04	3.25E+04

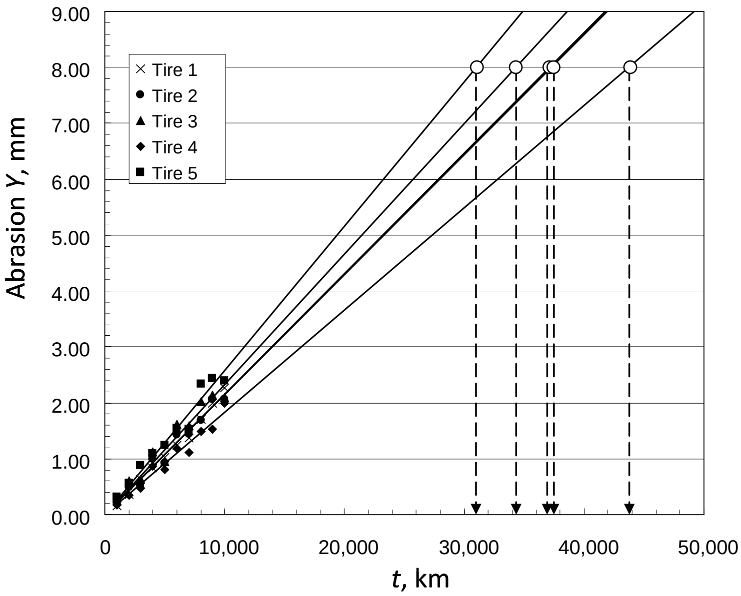

Fig. 43: Evaluation of the degradation test in the example.

4.5. Validation on the customer's premises

According to DIN EN ISO 9000, validation is confirmation, through the provision of objective proof, that the requirements *for a specific intended use or application* have been met. Consequently, a trial performed for the purpose of validation must take into account the specific use or intended application of the product in question.

The aim of validation on the customer's premises is to discover possible faults that may occur through interactions of a component with its environment (overall system, media, applications, etc.), and that were not foreseeable during the development phase. According to the state of the art, validating trials are the most effective method of revealing potential faults of this nature.

However, it is worth mentioning the limited statistical conclusiveness of these experiments, which can generally only be conducted with very few parts or for a limited duration. As a result, only obvious faults with a relatively high probability of occurrence can be disclosed, see Fig. 26.

The purpose of validation is to examine all uses of the part (e.g. different applications). To this aim, it may be necessary to divide the uses into groups, in which critical load cases are tested representatively for all uses in that group.

The following criteria may be used to select these uses:

- Level of innovation
- Sales volume
- Performance density.

The parts are subject to diagnosis regardless of whether or not they have passed the test. Here, particular attention must be paid to unexpected levels of wear, unusual fault patterns, etc.

The selection of tests to be performed must be based on the specific product concerned.

4.6. Tests during production

Reliability is an aspect of quality. Therefore, all activities performed within the framework of a QM system must also be applied to the fulfillment of established or set requirements in respect of reliability [8], p. 291. The regular testing of samples of products from serial production that are ready for delivery (reliability inspection) is intended to monitor whether these requirements are continuously being satisfied.

Specimens from production may be considerably less expensive and available in larger quantities than samples during the development phase; nevertheless, the scope of these tests must generally be limited due to restricted test bench capacity. The performance of normal (non-accelerated) endurance runs, in particular, would take far too long for the rapid detection of risks to product reliability.

A suitable approach here is to employ accelerated test methods such as HASS, to achieve rapid comparison with the verified, released sample phase, or at least to track modifications of features relevant to reliability.

Alternatively, so-called burn-in tests may be performed. The burn-in test is a form of 100% testing during the serial production of electronic components, in which potential early failures are triggered by a defined load (e.g. temperature, humidity, current), and the products in question removed. The test conditions must be chosen such as to ensure an adequate selection of "good" and "bad" products, without damaging the "good" ones. As a rule, however, these conditions are difficult to meet in the case of mechanical components.

5. Reliability in the field

5.1. Product liability

According to the German Product Liability Act, product liability means that a manufacturer is liable for damages suffered by third parties arising from non-conformities in his product. A product is faulty when it exhibits safety deficits, i.e. it does not provide the safety which the user of the product may rightfully expect. Safety expectations can be justified by all kinds of statements (advertisements, brochures, promotions). Therefore, these statements must be regarded with caution.

Product liability is primarily "strict liability". This means that the manufacturer is liable whether or not he is responsible for the non-conformity. The decisive factor is solely the occurrence of the non-conformity. Each unit (Development, Purchasing, Production and Sales) is responsible, within its scope of work, for avoiding safety deficits in products under all circumstances.

In order to satisfy product liability requirements, the manufacturer must examine his products on the market to ensure that:

- previously unknown non-conformities do not occur,
- the user is not exposed to danger through combining the product with popular accessories – including from other manufacturers – (e.g. add-on parts for the product), and
- intervention and modification are not frequently undertaken (e.g. changing the control limit in control units).

The customer must be instructed on how to use the product properly and without danger, and informed about residual risks. Installation instructions, instructions for use, warnings and safety information must be clear and comprehensible and aimed at the level of knowledge of the target group (professional/layman).

5.2. Field observation and Active Field Observation (AFO)

First and foremost, field or market observation means that the manufacturer of a product fulfils his obligations arising from the German Reform of the Law of Obligations in 2002 in respect of the German Equipment and Product Safety Act (GPSG) and the German Product Liability Act (ProdHaftG). In other words, in the event of a complaint the product in question must be examined in order to then permanently eradicate any defects.

Furthermore, the manufacturer is basically required to conduct market observations permitting the early detection of safety defects in connection with the product he has manufactured. This means that the manufacturer must follow up all notifications of complaints during the usual period of use if these give cause for concern that the product does not offer the safety that the general public may rightfully expect.

On the other hand, AFO refers to observations of product behavior in the field irrespective of complaints, i.e. tests are conducted on functioning products. The analysis of long-running functional products, i.e. throughout the product lifetime, provides valuable insight into the behavior of modules and machine elements of the product, which enables robustness to be assessed under real field load.

At the heart of the principle of AFO is the usability of the data acquired, i.e. the product-related gain in knowledge in terms of:

- the behavior of long-running products in the field in terms of function/safety/failure probability
- the comparison of product trials & tests with real field load

- a product design adapted to real field load
- finding out and verifying rationalization measures on the product

The procedure consists in virtually breaking the product down into modules or machine elements, in order to gain considerable knowledge for current and future products from the analysis of products in the field, including older generations. For this process to succeed, the reuse of modules over generations of product platforms is indispensable. Examples of this are the output stage designs of control units, material combinations for assembly and connection technology, the sealing of products and all materials and their loads, which are used in design elements.

According to this principle, the "AFO analysis" of the products in the laboratory of the relevant development department must not only be based on function. Rather, the dedicated examination of a module in combination with existing automotive data on experienced field loads should enable the robustness of the entire product to be determined.

5.3. Evaluation of field data

Methods for evaluating field data are described in detail in [16]. Here, just a few additional remarks are provided.

The principal difficulty inherent in the analysis of field data lies in their incompleteness. Data on failure behavior may be gathered from warranty cases, inspections, etc., but information about the behavior of products that have not failed is generally difficult to obtain. However, for the analysis of lifetime the incorporation of observations of intact units is vital, as was demonstrated in section 4.3. Only this permits evaluation using methods for censored data. For this reason, some special procedures have been elaborated, which are briefly presented here.

The simplest method consists in estimating the number of intact units from the number of units brought into circulation. If k failures were registered up to a certain time, for example, it can be assumed that if n parts have been brought into circulation, (n-k) parts have not failed up to this time. Evaluation is possible using the method for types I and II right-censored data . However, n is not necessarily equivalent to the number of sold units, as not all products immediately reach the end consumer. Here, product-specific relationships must be taken into consideration.

If the lifetime characteristic is not a time variable and the use of the product over time is clearly inhomogeneous, the above method cannot be directly applied. With a vehicle, for example, the lifetime characteristic is its mileage. This cannot be deduced simply from the fact that a certain number of vehicles were still intact after a year, even if the lifetime of the failed vehicles has been precisely determined. Two approaches may be employed to tackle this problem:

- The most accurate results can be expected when the performance distribution of the product (e.g. the mileage of a car) is known. In this case, the total number of units still intact at a certain time can be broken down, depending on the scope of the individual performance distribution classes. If we know, for example, that 20 %, 30 % and 50 % of vehicles have an annual mileage of 40,000, 25,000 and 15,000 km respectively, it follows that of 1000 vehicles that have not failed after one year, 200 will have driven 40,000 km, 300 25,000 km and 500 15,000 km without failure. Here, the analysis makes use of the process for multiple-censored data.
- A performance distribution is deduced from complicated and time-consuming use-case analyses. If these are not available, as a simplified process the number of intact units can be "distributed" over all failures. In other words, it is implicitly assumed that the performance distribution and the failure distribution are identical. Whether this assumption is justified must be examined from one case to another. Here, the evaluation can make use of the simplified method employed for a sudden death test.

A further difficulty lies in the complexity of the field data. The problem is not necessarily the quantity, but its nature: often, it is not immediately apparent exactly what has happened upon failure, and even whether something has failed. Thus, failure data can be distorted and may sometimes have unexpected effects. Examples are known where a manufacturer assumed a lack of reliability in his products due to larger numbers of returns during the warranty period. However, the problem lay in the fact that maintenance staff were paid by the number of replaced units, of which many were actually fully intact. Following a detailed analysis, the solution was to alter the performance incentive system for the maintenance staff, not the design of the product.

Another example is the case of an offroad car manufacturer in the 1960s who assumed an average speed of 50mph for his SUVs. After several failures had been recorded, an analysis revealed that 50mph was indeed reasonable for driving off road, but most SUVs were used as normal vehicles, which drive considerably faster. As a result, the requirements for future products were changed, instead of introducing quality assurance measures.

One situation that crops up with particular frequency is the occurrence of mixed failures, for example in a mechanical and an electronic component of a product, or two different failure mechanisms come into play in a single component. A "mixed" evaluation of both mechanisms in this case is prohibited, as is the "omission" of one of the mechanisms. Instead, correct evaluation necessitates treating the failures of the competing failure mode as a runout, and applying the method for multiple-censored data.

A similar situation occurs when field failures are recorded at a certain time, for which it is unclear whether these are early failures due to production faults, or wear and fatigue-induced failures. An inappropriate assumption during evaluation can lead to completely incorrect results. If, for example, 10% of produced units demonstrate a heat-treatment fault, field failures due to this fault must also be based on 10% of the production volume, which must be taken into consideration in the calculation of intact units.

The examples described above show that, in most cases, the detailed diagnosis and analysis of failures is indispensable. One should in no event rely on an estimate of the failure rate or the Weibull exponent to decide between competing failure modes. Further details on this subject can be found in [8.16].

6. Annex

6.1. Statistical basis

This section contains some fundamental statistical concepts that are necessary in order to understand reliability.

6.1.1. Definition and characteristics of probabilities

In everyday engineering practice, variations often occur that can basically be divided into two categories [15], p. 45ff:

- deterministic, their reason is known and their effect can be precisely predicted, and
- random, neither the cause nor the effect of this cause can be described without further work.

A typical example of the first case is the movement of a vehicle as the result of acting forces; an example of the second type is the strength of a part as a result of unknown effects during the manufacturing process. Since random variations cannot be accurately predicted, we turn to the concept of *random variables*, to which a certain probability can be ascribed. The strength of a part in production, for example, is described by a random variable that can assume certain values only with a certain probability.

Various terms can be used to define probability. In the "classic" definition by Laplace, the probability of an event is defined by the relationship

$$P := \frac{r}{n},\ 0 \le r \le n \tag{6.1}$$

whereby *r* signifies the number of most favorable cases and *n* the number of all cases. If, for example, we wish to calculate the probability of throwing a 4 with a dice, then this probability is 1/6, for the outcome would only be favorable in one out of six cases. This definition is often used to convert a relative proportion into a finite population and a probability: if a statement is made that 20 % of parts of a batch have failed, this also means that the probability of randomly "drawing" a failed part during a test is 20 %, and vice versa.

The above definition is only applicable when the number of events is finite, however. A remedy is provided here by the "statistical" definition of probability according to von Mises. This approach is based on an infinitely large population, from which a size *n* sample is taken. If the event takes place in *r* cases in this sample, the limit of relative frequency *r/n* for sample sizes tending to infinity is employed as the definition of probability:

$$P := \lim_{n\to\infty} \frac{r}{n},\ 0 \le r \le n\,. \tag{6.2}$$

Although equations (6.1) and (6.2) look similar, their interpretation is very different: In the first equation, *n* stands for the size of the finite population; in the second, *n* is the size of the sample of an (infinite) population. Thus, if 20 % of parts fail in a sample, this does *not* imply that the failure probability in the population is necessarily 20 %. Wholly different proportions may occur in the next sample taken. So, the variation in the samples also plays a decisive role.

Equation (6.2) is also not totally satisfactory, as it is strictly speaking an estimate, not a definition [1], p. 36. But it is fully adequate for considerations of reliability. We should also mention at this point, rather by and by, the "axiomatic" definition by Kolmogoroff, which does not define probability but considers it as a fundamental characteristic of events that satisfy certain axioms.

The probabilities of different events can be interrelated. Some of the most important relationships are explained here, further details can be found in [15], p. 25ff.

- The probability of an event A is expressed as $P(A)$. The probability that A does not take place is expressed by $1 - P(A)$.
- The conditional probability that event A takes place, assuming that event B has occurred, is expressed by $P(A|B)$. If the events are independent, $P(A|B)=A$ and $P(B|A) = P(B)$ applies.
- The probability that events A and B take place simultaneously is expressed by $P(AB)$ and is the result of $P(AB) = P(A)\cdot P(B|A)$. In the case of independent events, the "product rule" $P(AB) = P(A)\cdot P(B)$ applies.
- The probability that event A or B takes place is expressed by $P(A+B)$ and is the result of $P(A+B) = P(A) + P(B) – P(AB)$. In the case of independent events, $P(A+B) = P(A) + P(B) – P(A)\cdot P(B)$.
- Theorem of total probability: In the case of n mutually exclusive events B_i (i.e. $P(B_1 + ... + B_n) = P(B_1) + ... + P(B_n) = 1$), $P(A) = \sum_i P(A|B_i)\cdot P(B_i)$ applies.
- By transforming this last relationship, we obtain the Bayesian theorem: $P(A|B)= P(A)\cdot P(B|A)/P(B)$.

The theorem of total probability plays an extremely important role in considerations of reliability, as illustrated below.

EXAMPLE:

A manufacturer purchases parts from 3 different suppliers and wishes to calculate the failure probability of these parts. The quantities delivered by the suppliers and their respective failure probabilities are known:

	Supplier 1	Supplier 2	Supplier 3
Proportion supplied	20%	30%	50%
Failure probability	4%	2%	1%

The theorem of total probability can be used to calculate the failure probability. Here, A *denotes the event "failure" and* B_i *"delivery from suppliers 1 to 3":*

$P(A) = P(A|B_1)\cdot P(B_1) + P(A|B_2)\cdot P(B_2) + P(A|B_3)\cdot P(B_3) = 0.04\cdot 0.2 + 0.02\cdot 0.3 + 0.01\cdot 0.5 = 1.9\%$

The failure probability will therefore be considerably smaller than with supplier 1, because the latter delivers parts with the highest failure probabilities, but also delivers relatively small quantities, which is included as a "weight" in the above formula. It is important to bear in mind that the above scenario encompasses all suppliers, however, so that the sum of proportions delivered is 1.

6.1.2. Data series and their characteristic values

Even when an experiment is meticulously conducted, the results (e.g. the measurement of tensile strength) will differ – they will vary – if the test is performed more than once, despite conditions remaining constant. A data series

$$x_i, \; i = 1,\ldots,n, \tag{6.3}$$

with n values of the variable under observation can basically be expressed by the following characteristic values within the framework of descriptive statistics:

- Mean to describe the position of the values

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ , and} \tag{6.4}$$

- Variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 , \tag{6.5}$$

  or their square root, the standard deviation s, which characterizes the variance of the values around the mean.

Since the individual values are subject to variance, their characteristic values will also differ if the data series is measured several times. It is therefore important to recognize that the mean and the standard deviation are also subject to variance. However, the variance of the mean is considerably smaller than that of the individual values, and it becomes smaller the larger the sample size. We can therefore demonstrate (see [6], p.122) that

$$s_{\bar{x}}^2 = \frac{s^2}{n} . \tag{6.6}$$

Data series can be represented in histograms, which are produced when the values of the data series are classified. A histogram therefore presents frequencies of the occurring classes of attributes. Total frequencies can be deduced by adding the classes together.

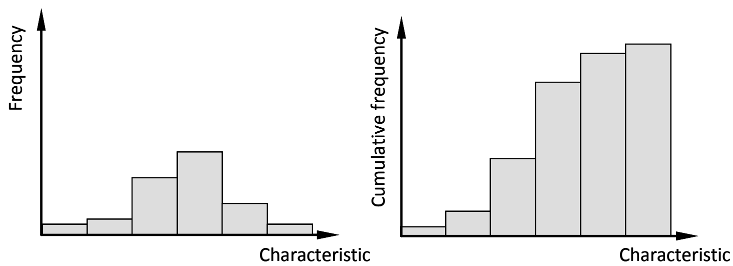

Fig. 44: Histogram and cumulative frequency curve (schematic diagram).

6.1.3. Distributions and their characteristic values

6.1.3.1. Distribution and distribution density, moments

Things are seen in a somewhat different light if the data is viewed as a sample from an infinitely large population. Thus

$$x_i, \; i = 1,\ldots,n, \tag{6.7}$$

is construed as manifestations of a random variable *X*. A theoretical model for describing the population of all possible values of the variable is provided by the distribution density function *f*(*x*). This assigns a number *f*(*x*) to the value *x* Fig. 45, left, similarly to the manner in which a relative frequency is assigned to the value *x* in a histogram. We can show that if the size of the sample under observation is very large and the class widths of the histogram are declining in size, the distribution density function can be deduced directly from the histogram, with the result that the latter can be regarded as an approximation of the density function. The probability of this random variable assuming *X* values in the interval (*x*;*x*+*dx*] is

$$P(x < X \leq x + dx) = f(x)dx \,. \tag{6.8}$$

From the cumulative frequency curve, a function can also be derived, the distribution function *F*(*x*), by crossing the limit to infinity. For each value *x*, this states the probability that a random variable *X* will assume a value lower than *x*, Fig. 45 right. The distribution function of a distribution is generally obtained by integrating its density function:

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(v)dv \,. \tag{6.9}$$

Here, the independent variable of the function *f*(*v*) is expressed as *v*, in order that it can be differentiated from the upper integration limit *x*. The function *F*(*x*) is therefore a function of the upper integration limit, while *v* is a variable by which integration takes place.

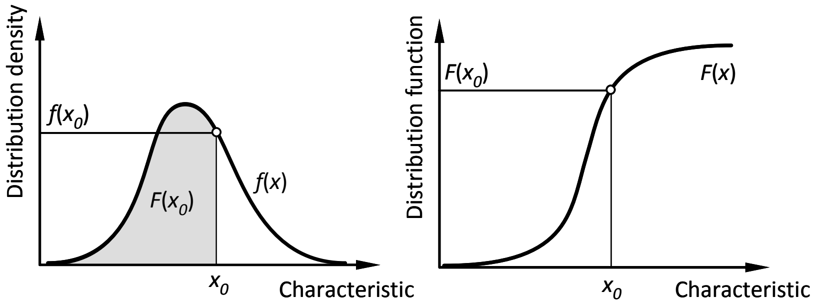

Fig. 45: Distribution density and distribution function (schematic diagram).

The distribution function assigns a probability to the value *x* of the variables *X*. Certain preconditions must therefore be satisfied for this to take place, so that a function can be employed as a distribution function:

- Within the limits, it must be the case that $F(-\infty) = 0$, $F(+\infty) = 1$, and consequently

$$\int_{-\infty}^{+\infty} f(v)dv = 1 \,. \tag{6.10}$$

- For all $-\infty < x < +\infty$ it must be the case that *F*(*x*) is a monotonously increasing function of *x*.

The density and distribution function of a distribution possess characteristic values (also known as moments), which determine the appearance of the distribution. The most important are:

- the mean (sometimes also referred to as the expected value)

$$\mathrm{E}(x) = \mu = \int_{-\infty}^{+\infty} x \cdot f(x)dx \tag{6.11}$$

- and the variance

$$\mathrm{Var}(x)=\sigma^2=\int_{-\infty}^{+\infty}(x-\mu)^2\cdot f(x)dx\,, \qquad (6.12)$$

which describe the position and width of the distribution.

6.1.3.2. Examples of distributions

The simplest means of illustrating the above issues is the example of normal distribution, which possesses the well-known "bell-shaped" distribution density function:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \qquad (6.13)$$

By integrating the density function, we obtain the s-shaped distribution function of normal distribution

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right)dv\,. \qquad (6.14)$$

As the integral cannot be further analyzed, the above equation cannot be simplified any further.

The equations show that with normal distribution, the mean and standard deviation characteristics occur as parameters in the distribution function; they are often also referred to as $N(\mu,\sigma)$. This makes working with them especially simple. The fact that this is not necessarily the case for every distribution, however, is demonstrated by the example of exponential distribution with the density function

$$f(x)=\lambda\cdot\exp(-\lambda x)\,,\ x\geq 0\,, \qquad (6.15)$$

in which a parameter λ occurs. The relationship between parameters and the above characteristic values is described as follows:

$$\mu=\frac{1}{\lambda}\,;\ \sigma^2=\left(\frac{1}{\lambda}\right)^2. \qquad (6.16)$$

Here, the relationship between distribution parameters and characteristic values is relatively simple, so that conversion can take place without problem. It becomes more complicated in the case of the Weibull distribution, which plays an important role in describing lifetime; there, conversion is no longer that straightforward. We do not go into further detail at this point, however.

6.1.3.3. Characteristic values of independent and dependent variables

If a variable Y indicates a linear dependence of n independent random variables X_i:

$$Y=a_0+\sum_{i=1}^{n}a_i\cdot X_i\,, \qquad (6.17)$$

then Y will also be a random variable. The distribution function of Y cannot generally be calculated without problem, particularly if the variables X_i have different types of distribution. Despite this, it is possible to calculate characteristic values of the Y distribution precisely as a function of characteristic values of the X_i distributions, see [6] p. 150:

$$\mu_Y=a_0+\sum_{i=1}^{n}a_i\cdot\mu_{Xi}\,, \qquad (6.18)$$

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \cdot \sigma_{Xi}^2 \ . \qquad (6.19)$$

In particular, for the mean of n normally distributed, independent variables

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ , \ i = 1,\ldots,n \ , \ X_i \in N(\mu,\sigma) \ , \qquad (6.20)$$

and this, likewise, is normally distributed with

$$\mu_{\bar{x}} = n \cdot \frac{\mu}{n} = \mu \ \text{ and } \ \sigma_{\bar{x}}^2 = n \cdot \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n} \ . \qquad (6.21)$$

If the relationship between Y and the independent random variables X_i, contrary to equation (6.17) is non-linear, then it is generally the case that

$$Y = g(X_1,\ldots,X_n) \ , \qquad (6.22)$$

a Taylor series development can be used to derive a first order estimate of mean and variance, [6] p. 150:

$$\mu_Y = g(\mu_{X1},\ldots,\mu_{Xn}) \ , \qquad (6.23)$$

$$\sigma_Y^2 = \left(\frac{\partial g(X_1,\ldots,X_n)}{\partial X_1}\right)^2 \sigma_{X1}^2 + \ldots + \left(\frac{\partial g(X_1,\ldots,X_n)}{\partial X_n}\right)^2 \sigma_{Xn}^2 \ . \qquad (6.24)$$

This linearized estimate is not a good approximation in the environs of extrema. In this case, a Monte Carlo simulation may deliver better results. To this aim, randomly larger samples are generated on the basis of the known distributions of the variables X_i. Next, a value for Y is calculated in accordance with equation (6.22) for each configuration of variables, so that in the end a sample of the same size is also available for Y. The characteristic values of the Y distribution can now be obtained from this sample using estimates based on equations (6.4) and (6.5). Further details are contained in [6], section 9.

6.1.3.4. Point and interval estimation of the mean, confidence interval

We are generally up against the question as to how to determine the unknown characteristic values of the population. It can be demonstrated that

- the mean of a sample $\bar{x}$ can be used as an estimate for the mean of the population μ, and
- the variance of a sample s^2 can be used as an estimate for the variance of the population σ^2.

In addition, both estimates conform to expectations, i.e. with many replications, they approximate the values of the population.

The quality of this point estimation is initially questionable, however, but can be evaluated using the following procedure. If a random variable X is normally distributed, the variable

$$u = \frac{X-\mu}{\sigma} \ , \ X \in N(\mu,\sigma) \qquad (6.25)$$

is also normally distributed with the mean 0 and the standard deviation 1. This distribution is known as "standard normal distribution". An estimate of the probability P with which X lies within a given interval $\mu - u \cdot \sigma \le X \le \mu + u \cdot \sigma$ can be made with the aid of the parameter u:

$$P(\mu - u \cdot \sigma \le X \le \mu + u \cdot \sigma) = 1 - \alpha \ . \qquad (6.26)$$

For example, the variable X will lie within the interval $\mu \pm 1 \cdot \sigma$ with a probability of 68.3 %. It is therefore clear that u equals the characteristic value of standard normal distribution with probabilities (1 - α /2) and α /2.

We will now examine the mean of n independent, normally distributed variables

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ , \ X_i \in N(\mu,\sigma) \ , \ i = 1,...,n \ , \ \overline{X} \in N(\mu, \frac{\sigma}{\sqrt{n}}) \ . \qquad (6.27)$$

The quantity

$$u = \sqrt{n}\,\frac{\overline{X} - \mu}{\sigma} \qquad (6.28)$$

is by definition standard, normally distributed. However, the standard deviation of the population is generally unknown, with the result that it must be estimated on the basis of a sample: $\sigma \approx s$. At the same time, we can shown that in this case, the quantity

$$t = \sqrt{n}\,\frac{\overline{X} - \mu}{s} \qquad (6.29)$$

is no longer normally distributed; its distribution is described by the so-called t-distribution (or student's t-distribution). For small values of the parameter n, the t-distribution manifests a greater width and edge definition than normal distribution; for the large values of n (n=50 and above), it approaches normal distribution.

An estimate of the probability P with which $\overline{X}$ lies within a given interval can be made with the aid of the parameter t. By the same token, an indirect conclusion can also be drawn regarding an interval

$$\overline{X} - t \cdot \frac{s}{\sqrt{n}} \le \mu \le \overline{X} + t \cdot \frac{s}{\sqrt{n}} \qquad (6.30)$$

that covers the unknown population mean. Here, the probability

$$P\left[\overline{X} - t \cdot \frac{s}{\sqrt{n}} \le \mu \le \overline{X} + t \cdot \frac{s}{\sqrt{n}}\right] = 1 - \alpha \qquad (6.31)$$

is usually referred to as the confidence level, and the interval as the confidence interval. The more certain one wishes to be that the interval covers the true mean, the higher the confidence level and the larger the parameter t that must be selected. Tabular values of t are contained in section 6.3.2.

In the same way, a confidence interval can also be stated for the variance, although this is not discussed any further here. Details can be found e.g. in [6].

6.1.3.5. Distribution-free estimation of failure probability, confidence intervals

During the evaluation of lifetime data, n observations (e.g. failures) $t_1 \le t_2 \le \ldots \le t_n$ taking place over time must often be described by an initially unknown, theoretical distribution. The question therefore arises as to which probability can be assigned to the various observations. Suitable methods of doing this are explained below. They are frequently also referred to as rank estimation methods, as they only take into consideration the order (rank) of the observations, not the actual times of observation.

Let us assume that 4 out of 10 specimens have failed at a certain time in an experiment. The question is how a certain failure probability F can be deduced from this observation. The simple approach F = 4/10 = 40 % would be naïve, as with a certain probability F, all failures between 0 and 10 would be possible, if also of varying likelihood, Fig. 46. Their probability can be calculated using the density function *p* of the discrete binomial distribution. This indicates the probability with which an event (e.g. a failure) will occur precisely *i* times in a sample with a size *n*, if the probability of this event is *F*:

$$p(i) = \binom{n}{i} \cdot F^i \cdot (1-F)^{n-i} . \tag{6.32}$$

Thus, from this observation of 4 failures, we do not immediately obtain a particular value for *F*, because 4 failures are possible with different values for *F*. However, this task can be formulated differently: A value for *F* must be selected that ensures that the probability of error is kept as small as possible, and the true but unknown failure probability is greater than the selected one. Thus, the probability for a non-conservative estimate must be kept small.

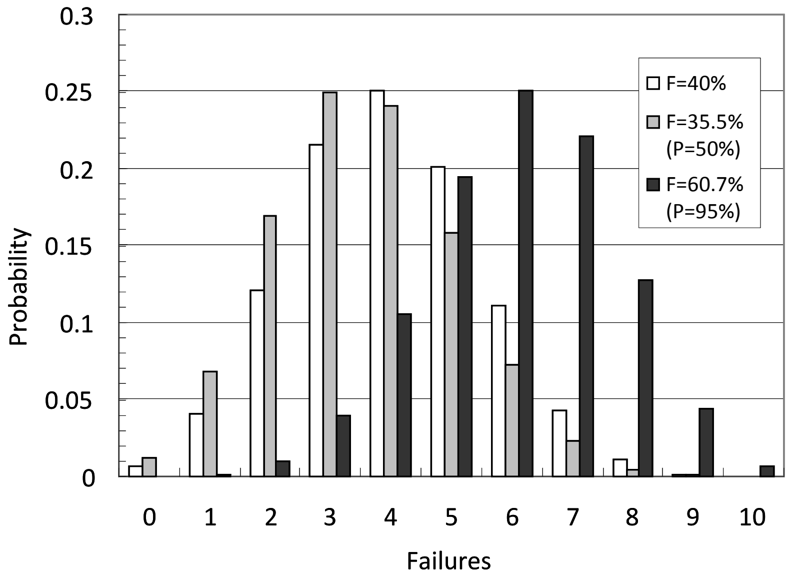

Fig. 46: Probability that 1 to 10 failures occur in a sample of *n*=10 with different failure probabilities *F*.

As Fig. 46 shows, fewer than 4 failures become ever more improbable as the failure probability *F* increases. In other words, if it is very improbable that fewer than 4 failures occur, but at the same time 4 failures are observed, the 4 failures cannot be random; they would indicate a failure probability lower than that which was estimated. The estimate would therefore be conservative, which is desirable. The probability that results as the sum of equation (6.32) for *i* = 0,...,3

$$\alpha = \sum_{i=0}^{x-1} \binom{n}{i} \cdot F^i \cdot (1-F)^{n-i} \tag{6.33}$$

can thus be interpreted as the residual risk of a non-conservative estimate (probability of a type I error α), which must be kept small. The probability

$$P = 1 - \alpha = \sum_{i=x}^{n} \binom{n}{i} \cdot F^i \cdot (1-F)^{n-i} \tag{6.34}$$

for its part, represents the confidence level.

The explanations above have assumed a given failure probability *F*. However, the task is to determine *F* for the *i*-th observation. The procedure for doing this is provided by equation (6.34): For a sufficiently small probability α, *F* can be deduced from (6.34). This equation can be solved either iteratively and numerically or, as proposed in [14], p. 79, converted to

$$F_{BB,1-\alpha} = \frac{i}{i + (n+1-i) \cdot F[2(n+1-i); 2i; 1-\alpha]} . \tag{6.35}$$

Here, $F[2(n+1-i);2i;1-\alpha]$ expresses the quantile of the Fischer distribution with degrees of freedom $2(n+1-i)$ and $2i$, which can be calculated using the "FINV" Microsoft Excel function, for example. The confidence bounds calculated using equation (6.35) $F_{BB.1-\alpha}$ are set out for α = 5% in tabular form in section 6.3.4. They define the one-sided confidence interval $F \le F_{BB,1-\alpha}$ with associated confidence level

$$P\left[F \le F_{BB,1-\alpha}\right] = 1-\alpha \ . \tag{6.36}$$

For α = 50%, equation (6.35) produces the exact median values of the failure probability, as explained in section 6.3.3. The much used equation

$$F = \frac{i-0.3}{n+0.4} \tag{6.37}$$

constitutes only a (relatively good) approximation.

Further estimators and associated confidence intervals exist, such as the Kaplan-Meier, the Fisher matrix and the likelihood ratio. These approaches are based only on asymptomatically correct approximations, and are not discussed any further here. Details can be found e.g. in [14].

EXAMPLE:

For the example from section 4.3.2.1, the median values and the one-sided confidence bounds (confidence level P = 90 %) of the failure probability are calculated using equation (6.35) as follows:

Motor	1	2	3	4	5	6	7	8	9	10
$F_{BB.50\%}$,%	6.7	16.2	25.9	35.5	45.2	54.8	64.5	74.1	83.8	93.3
$F_{BB.90\%}$, %	20.6	33.7	45.0	55.2	64.6	73.3	81.2	88.4	94.5	99.0

6.2. Distributions as a model of varying lifetime data

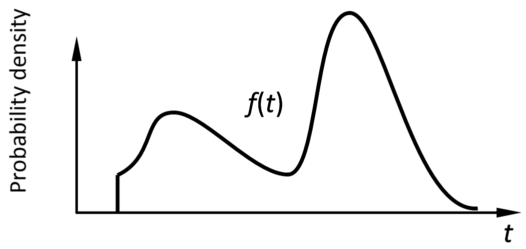

Fig. 47: Possible density function of the distribution of a real random variable (schematic).

Distributions should be understood as models of varying lifetimes. As with any model, here too, it makes no sense to ask whether the model is correct or incorrect, but rather whether or not it serves its intended purpose. Nevertheless, caution must be exercised when dealing with (simplified) models, as they do not reflect reality in all its complexity:

Random variables are not necessarily normally distributed. The Gaussian normal distribution assumes values between -∞ and +∞, but physical quantities are often only positive. Here, the use of the log-normal distribution may provide assistance. However, the normal distribution has a theoretical basis, for one can show that a random variable, which is subject to an infinite number of additive and independent influences, must be normally distributed (central limit theorem). Although this is a sensible model for the question at hand, a decision has to be made in each individual case. The same also applies to the log-normal and Weibull distribution.

Random variables do not even have to conform to a known analytical distribution, because they are subject to various complex influences and may therefore be skewed, multimodal or "cut" left and right by a selection, Fig. 47. Despite this, it can be helpful to use known distributions as a model. Although the random variable of pressure load obviously cannot assume values up to +∞, it makes sense to use a log-normal distribution to describe it. This is because deducing an upper limit for pressure is extremely difficult and would not really allow us to achieve our objective. The precise, "cut" distribution is very complex and scarcely manageable, a worst-case approach that takes account only of the maximum pressure but overlooks its very low probability of occurrence, resulting in overdesign. A "non-physical" distribution can provide assistance here as a simplified model.

Where such a model appears necessary, however, modeling must not be excessively simplified. If, for example, several mechanisms have obviously led to failure, these must also be taken into consideration with a reasonable outlay using the approaches for censored data, not approximated by means of a single distribution. It is necessary to examine on a case-by-case basis which distributions seem most fit for purpose for modeling. Further details can be found in [14], p. 49ff.

6.2.1. Weibull distribution

In the early 1950s, in very general conditions, Swedish engineer W. Weibull developed the universal cumulative distribution function

$$F(t) = 1 - \exp\left(-\left(\frac{t}{T}\right)^{b}\right), \qquad (6.38)$$

of which he was able to show, purely empirically, that in many cases it constitutes a good description of the characteristics of real objects, such as the tensile strength of steel, particle size of fly ash, fiber strength of cotton and fatigue life of steel.

Weibull himself was never of the opinion that this function must always apply. Nevertheless, Weibull’s distribution function has proven itself today and is inseparably connected with to the evaluation of lifetime tests. This is also due to subsequent work that was able to show the theoretical bases behind the distribution as a so-called extreme value distribution. These bases are outlined in brief below.

A body can theoretically be broken down into n parts, with corresponding lifetimes $t_1, \dots, t_n$. If the individual elements are assumed to possess a serial structure, the lifetime of the body is the same as the lifetime of its weakest member:

$$t = \min(t_1, \dots, t_n).$$

The lifetime of the body t therefore corresponds to the lowest failure time (first rank) of a size n sample. If this procedure is repeated for several bodies, the first rank will always be different – it will vary. The distribution that describes this variance is known as the extreme value distribution, since the first rank (like the n-th rank) constitutes an "extreme" rank. It is possible to demonstrate that the Weibull distribution always results as the distribution for $n \to \infty$, regardless of the type of distribution displayed by the lifetimes of the individual elements. Thus, the Weibull distribution describes in a theoretically sound way the failure characteristic of a body based on the principle of the weakest member.

With some failure causes, first of all a certain time t_0 must pass until a failure can occur as a result of this cause. Examples are wear (e.g. brush wear in electric motors), and corrosion. The three-parameter Weibull distribution that is obtained through a transformation $t \to t - t_0$ serves as a model for statistically describing this phenomenon.

The expected value (MTTF) of the Weibull distribution is defined as

$$\mathrm{E}(t) = \mathrm{MTTF} = \int_0^\infty t \cdot f(t)dt = \int_0^\infty T \cdot x^{\frac{1}{b}} \cdot e^{-x}dx = T \cdot \Gamma\left(1 + \frac{1}{b}\right), \qquad (6.39)$$

whereby $\Gamma(.)$ denotes the so-called gamma function. It is sensible to emphasize at this point that the MTTF value (or mean) of an asymmetrical distribution, such as is the case with Weibull, is *not* identical to the 50% quantile t_{50} (the median). The relationship between t_q and MTTF can be expressed as described below. Assuming that the lifetime is Weibull-distributed, the time t_q by which the proportion q of a population of equivalent products has failed can be expressed thus:

$$F(t_q) = 1 - \exp\left(-\left(\frac{t_q}{T}\right)^b\right) = \frac{q}{100}, \text{ with } 0 \le q \le 100. \qquad (6.40)$$

Consequently,

$$t_q = T\left[-\ln\left(1 - \frac{q}{100}\right)\right]^{\frac{1}{b}}. \text{ From } T = \frac{\mathrm{MTTF}}{\Gamma\left(1 + \frac{1}{b}\right)} \text{ it therefore follows that:} \qquad (6.41)$$

$$t_q = \mathrm{MTTF}\frac{\left[-\ln\left(1 - \frac{q}{100}\right)\right]^{\frac{1}{b}}}{\Gamma\left(1 + \frac{1}{b}\right)}. \qquad (6.42)$$

Table 6 illustrates some basic relationships underlying the Weibull theory. Fig. 48 shows the graphical representations of the density function, the distribution function and the failure rate. It is evident that the "appearance" of the Weibull distribution has strong variations depending upon the value of the shape parameter b.

Weibull probability paper

The Weibull probability paper is a diagram for plotting the cumulative failure frequencies versus the logarithms of failure time, Fig. 49. The y-axis in this diagram is selected to ensure that a straight line is formed with the representation of a Weibull distribution. A double logarithm of the distribution function produces:

$$\ln\left(\ln\left(\frac{1}{1 - F(t)}\right)\right) = b \cdot \ln(t) - b \cdot \ln(T), \qquad (6.43)$$

which corresponds to a linear equation in terms of the independent variable $\ln(t)$ with slope b and intercept $-b \cdot \ln(T)$. This enables a given data set with failure times of test objects to be examined graphically for Weibull distribution. If the recorded sequence of points lies approximately on a straight line, then we can deduce that the values are approximately Weibull-distributed. This representation can be used to determine estimate values for the parameters b and T, see section 4.3.2.

Two-parameter Weibull distribution	
$f(t)=\frac{b}{T}\left(\frac{t}{T}\right)^{b-1}\exp\left(-\left(\frac{t}{T}\right)^{b}\right)$	Failure density function. $f(t)dt$ is the probability that a unit fails within the time interval $(t,t+dt]$.
$F(t)=1-\exp\left(-\left(\frac{t}{T}\right)^{b}\right)$	Failure probability function. Probability that a unit fails by the time t.
$R(t)=1-F(t)$	Reliability function. Probability that a unit survives the time t (survival probability).
$\lambda(t)=\frac{f(t)}{R(t)}=\frac{b}{T}\left(\frac{t}{T}\right)^{b-1}$	Failure rate. $\lambda(t)dt$ is the probability that a unit from the remaining lot that has survived up to the time t fails within the time interval $(t,t+dt]$.
$\mathrm{E}(t)=\mathrm{MTTF}=T\cdot\Gamma\left(1+\frac{1}{b}\right)$	Expected value (mean).
$t_q=\mathrm{MTTF}\frac{\left[-\ln\left(1-\frac{q}{100}\right)\right]^{\frac{1}{b}}}{\Gamma\left(1+\frac{1}{b}\right)}$	q-quantile of the distribution. Median of the distribution for q = 50

Three-parameter Weibull distribution	
$F(t)=1-\exp\left(-\left(\frac{t-t_0}{T-t_0}\right)^{b}\right)$	Failure probability function
$f(t)=\frac{b}{T-t_0}\left(\frac{t-t_0}{T-t_0}\right)^{b-1}\exp\left(-\left(\frac{t-t_0}{T-t_0}\right)^{b}\right)$	Failure density function
$\lambda(t)=\frac{f(t)}{R(t)}=\frac{b}{T-t_0}\left(\frac{t-t_0}{T-t_0}\right)^{b-1}$	Failure rate

Parameters of the Weibull distribution	
t	Lifetime parameter
b	Shape parameter or Weibull exponent. Determines the slope of the straight line in the Weibull plot (this explains the German term *Ausfallsteilheit* (failure slope) and is characteristic of this type of failure.
T	Scale parameter or characteristic lifetime. Indicates the time up to which approx. 63.2 % of the products of a population have failed, as $F(T)$ = 0.632. Determines the relative position of the failure probability function in relation to the t-axis.
t_0	Time without failure

Table 6: Parameters and characteristic values of the Weibull distribution.

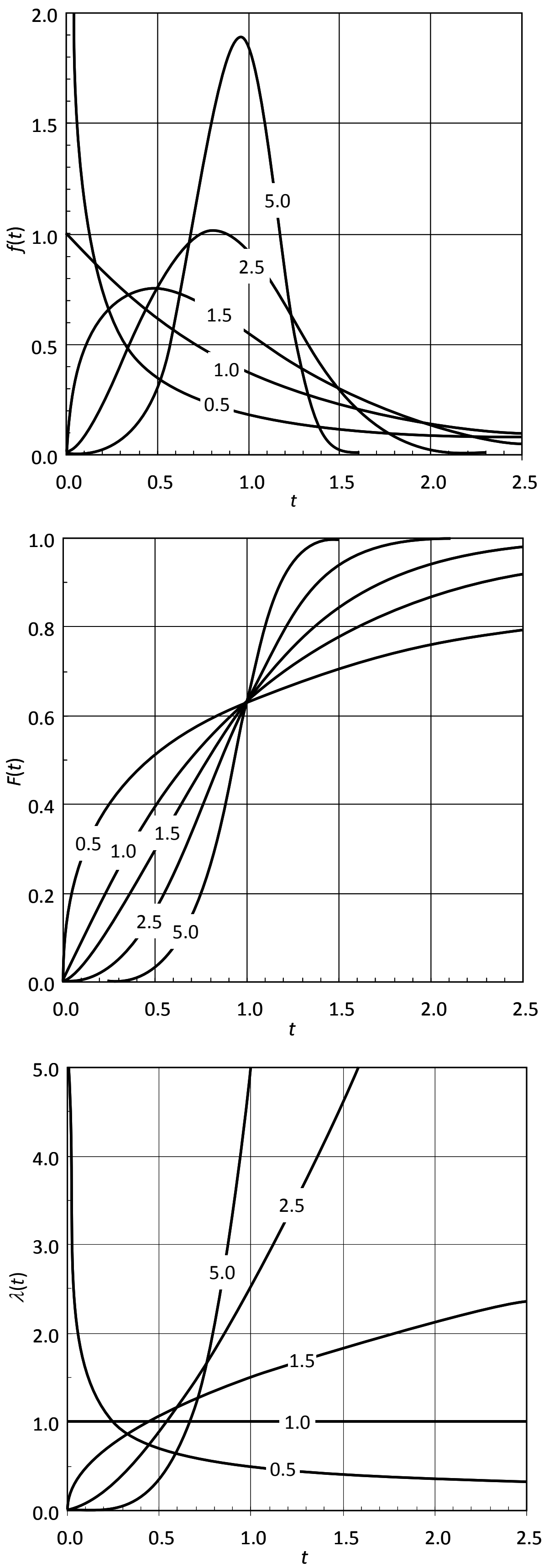

Fig. 48: Graphical representation of the density function (top), distribution function (center) and failure rate of the Weibull distribution (bottom) as a function of the shape parameter *b*.

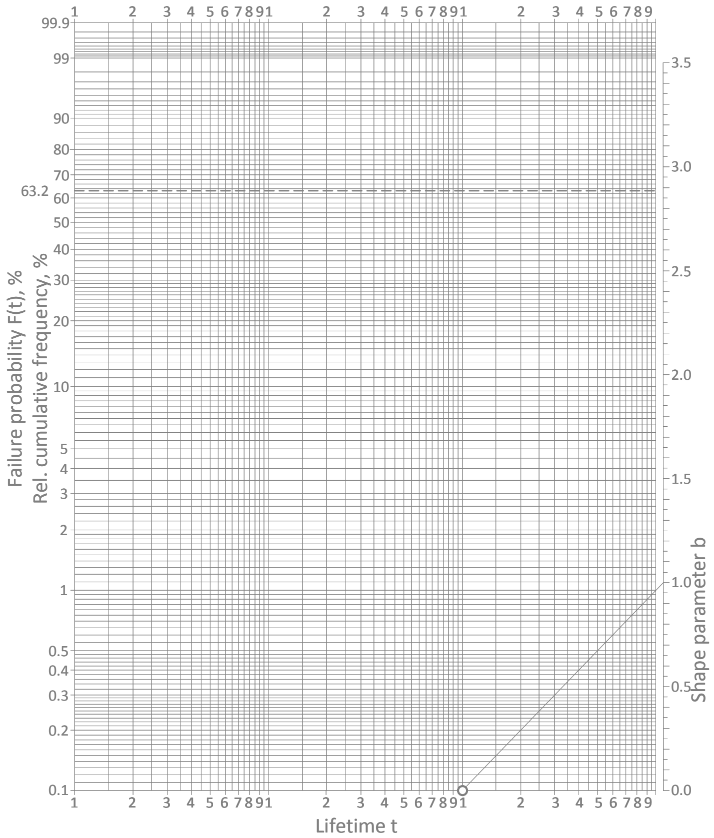

Fig. 49: Weibull probability paper

6.2.2. Exponential distribution

The exponential distribution is a model for describing random failures, which can often be applied when dealing with electronic components, for example. This distribution is a special case of the Weibull distribution, with shape parameters b = 1. Its distribution function is

$$F(t) = 1 - \exp(-\lambda t)\,, \qquad (6.44)$$

whereby λ stands for the parameter of the distribution. The distribution density function is obtained by deducing the distribution function according to t:

$$f(t) = \frac{dF(t)}{dt} = \lambda \cdot \exp(-\lambda t)\,. \qquad (6.45)$$

The failure rate of exponentially distributed failures is

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)} = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda\,. \qquad (6.46)$$

Thus, the parameter of the distribution is the same as the (constant) failure rate, so that the designation λ makes sense.

The exponential distribution has proven itself extremely effective for expressing radioactive decay (carbon dating using the ^{14}C method). The validity of the decay law has repeatedly been reconfirmed in numerous physical experiments. Where lifetime tests are concerned, however, the exponential distribution is only valid within a limited period of time (in section II of the "bathtub curve"), since after a sufficiently long time every technical product will eventually fail due to wear, aging and fatigue (the failure rate λ will then be time-dependent).

When evaluating a data set in the Weibull plot, the value b = 1 will practically never be exactly attained. For practical applications, we can also assume "random failures" in the case of $0.8 \le b \le 1.2$.

During the testing of electronic elements and components, in most cases a fixed number n of specimens and a fixed test time t are defined, and the number k of failed units up to this time determined (the individual failure times are unknown, and failed units are not replaced). An estimate value for λ is then:

$$\lambda \approx -\frac{1}{t} \cdot \ln\left(1 - \frac{k}{n}\right). \tag{6.47}$$

For an estimation of λ under other test and evaluation conditions (e.g. predetermined number of failures, type II censoring), see section 4.3.2.2.

The expected value of the exponential distribution (Weibull distribution with b = 1) is:

$$\mathrm{E}(t) = \mathrm{MTTF} = \int_0^\infty t \cdot f(t)dt = \lambda \int_0^\infty t \cdot \exp(-\lambda t)dt = -\frac{\exp(-\lambda t)}{\lambda}(\lambda t + 1)\Big]_0^\infty = \frac{1}{\lambda}. \tag{6.48}$$

As with the Weibull distribution, it also makes sense with the exponential distribution to determine the relationship between t_q and MTTF. t_q is the time up until which the proportion q of a population of equivalent products has failed:

$$F(t_q) = 1 - \exp(-\lambda t_q) = \frac{q}{100}, \text{ with } 0 \le q \le 100. \tag{6.49}$$

Thus

$$\ln\left(1 - \frac{q}{100}\right) = -\lambda t_q \text{ and } \mathrm{MTTF} = \frac{1}{\lambda} \text{ results in} \tag{6.50}$$

$$t_q = -\mathrm{MTTF} \cdot \ln\left(1 - \frac{q}{100}\right). \tag{6.51}$$

The table below shows some fundamental relationships.

Exponential distribution	
$f(t) = \lambda \cdot \exp(-\lambda t)$	Failure density function
$F(t) = 1 - \exp(-\lambda t)$	Failure probability function
$\lambda(t) = \lambda$	Failure rate
$\mathrm{E}(t)$ = MTTF = $1/\lambda$	Expected value (mean) of the distribution
$\mathrm{Var}(t) = 1/\lambda^2$	Variance of the distribution
$t_q = -\mathrm{MTTF} \cdot \ln\left(1 - \frac{q}{100}\right)$	q-quantile of the distribution. Median of the distribution for q = 50 %.

Parameters of the normal distribution	
t	Lifetime parameter
λ	Shape and scale parameter, $\lambda = 1/\mathrm{E}(t)$

Table 7: Parameters and characteristic values of the exponential distribution.

6.2.3. Normal distribution

When we refer to a normal distribution, we mostly associate this term with the Gaussian bell curve. The Gaussian curve expresses the probability density function of the normal distribution, Fig. 50:

$$f(t)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right). \tag{6.52}$$

The distribution function of normal distribution expresses the probability for each value *t* that the random variable *X* will assume a value between -∞ and *t*. It is obtained by integrating the above-mentioned density function, Fig. 50:

$$F(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left(-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right)dv\ . \tag{6.53}$$

F(*t*) denotes the area below the Gaussian bell curve up to the value *t*. The total area under the bell curve has the value one.

The normal distribution's special importance in statistics is explained by the central limit theorem. To explain this somewhat loosely, it states that the random interaction (addition) of numerous independent random variables produces a resulting random variable that is approximately normally distributed.

The normal distribution is clearly established by the two parameters μ and σ. The first of these determines the position of the distribution on the characteristics axis, the second its width, Fig. 50. One reason why working with the normal distribution is so simple is that these parameters are identical with the expected value (mean) or standard deviation of the distribution.

The transformation

$$\mathrm{U}=\frac{X-\mu}{\sigma} \tag{6.54}$$

converts a normally distributed random variable *X* with a mean μ and standard deviation σ into a random variable that is also normally distributed. The mean of *U* is zero, its standard deviation is one. The special normal distribution *N*(0,1) is known as standard normal distribution.

The values of the distribution function of the standard normal distribution $\Phi(u)$ are set out in tabular form in section 6.3.1. The tabular standard normal distribution can be used, for example, to calculate the proportions of any normal distribution that exceed predetermined limits.

Normal distribution plays only an indirect role in observations of reliability. This is due to the fact that here, a characteristic cannot fall below (e.g. surface roughness, eccentricity, roundness) or exceed (e.g. delivery rate, hardness) a certain limit. For example, a load or a load capacity can only assume values greater than zero (and the same applies to stress and strength).

Table 8 shows some fundamental relationships.

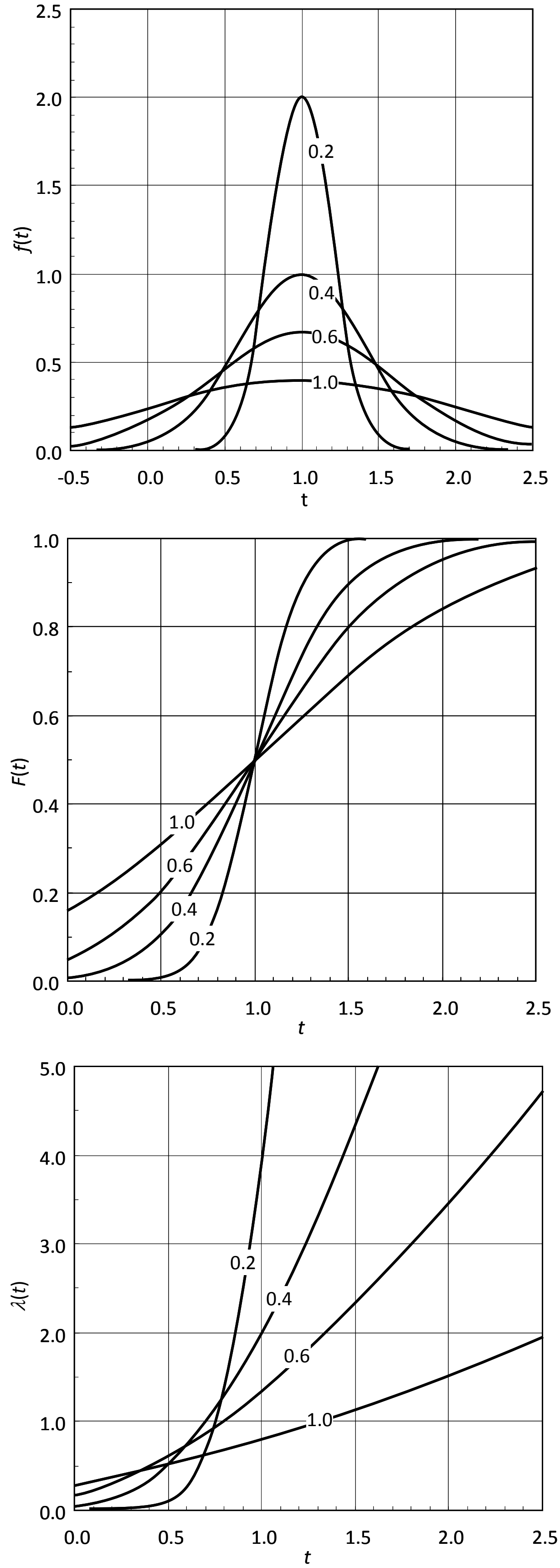

Fig. 50: Graphical representation of the density function (top), distribution function (center) and failure rate of the normal distribution (bottom) as a function of the dispersion parameter σ, when $\mu = 1$.

Normal distribution	
$f(t)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)$	Failure density function
$F(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left(-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right)dv$	Failure probability function
$\lambda(t)=\frac{f(t)}{R(t)}$	Failure rate (cannot be further simplified)
$E(t)$ = MTTF = μ	Expected value (mean) of the distribution
$Var(t) = \sigma^2$	Variance of the distribution
$t_{50} = \mu$	Median of the distribution (50% failure probability)

Parameters of the normal distribution	
t	Lifetime parameter
μ	Scale parameter, $\mu = E(t)$
σ	Shape parameter, $\sigma^2 = Var(t)$

Table 8: Parameters and characteristic values of the normal distribution.

6.2.4. Log-normal distribution

In order to exploit the advantages of normal distribution, such as known characteristics, relatively simply representation, etc. within the context of reliability, the following procedure can be applied: if we express the values of an asymmetrically distributed random variable such as lifetime as logarithms, these will be approximately normally distributed. Using logarithms in this way transforms the range of numbers between 0 and 1 into the range -∞ to 0, so that the left-hand part of the distribution is greatly stretched, the right-hand part greatly compressed.

A continuous random variable X is log-normally distributed if ln(X) is normally distributed. The density function of the log-normal distribution is expressed by

$$f(t)=\frac{1}{t\zeta\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(t)-\eta}{\zeta}\right)^2\right), \quad (6.55)$$

consequently, the distribution function results as

$$F(t)=\int_{0}^{t}\frac{1}{v\zeta\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(v)-\eta}{\zeta}\right)^2\right)dv\ . \quad (6.56)$$

The distribution function cannot be further simplified analytically. Some examples of the density and distribution function of log-normal distribution are contained in Fig. 51.

Unlike normal distribution, the log-normal distribution can effectively describe very different failure behavior, as with the Weibull distribution, for example. Furthermore, all findings and methods of normal distribution can easily be transferred to log-normal distribution. There is one major disadvantage, however, in that the failure rate of the log-normal distribution initially rises as lifetime increases, then falls again after a maximum is reached and tends towards zero for very long lifetimes. There are therefore limitations on how correctly the monotonously rising failure rate of fatigue and wear failures can be described.

Whereas a random variable resulting from the additive interaction of numerous influencing factors is normally distributed, in log-normal distribution the random factors are linked through multiplication, i.e. they are proportionally dependent upon one another. If we assume that a fracture occurs in stages and view the growth of the crack at each stage as a random variable that is proportional to the achieved crack length, according to the central limit theorem the log-normal distribution is the model for describing a failure characteristic of this kind, [1], p. 58. This is a theoretical reason for employing the log-normal distribution in questions of reliability.

Table 9 below shows some fundamental relationships. In addition to *e*, 10 can also be used as the basis for the logarithm. This is particularly common when talking about strength. In this case, the expression ln(.) must be replaced by lg(.) and exp(.) by 10^(.).

Log-normal distribution	
$f(t) = \frac{1}{t\zeta\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(t)-\eta}{\zeta}\right)^2\right)$	Failure density function
$F(t) = \int_0^t \frac{1}{v\zeta\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(v)-\eta}{\zeta}\right)^2\right) dv$	Failure probability function
$\lambda(t) = \frac{f(t)}{R(t)}$	Failure rate (cannot be further simplified)
$E(t) = MTTF = \exp\left(\eta + \frac{\zeta^2}{2}\right)$	Expected value (mean) of the distribution
$Var(t) = \exp\left(2\eta + \zeta^2\right)\left(\exp(\zeta^2) - 1\right)$	Variance of the distribution
$t_{50} = \exp(\eta)$	Median of the distribution (50% failure probability)

Parameters of the log-normal distribution	
t	Lifetime parameter
η	Scale parameter, $\eta = \ln(E(t)) - \frac{\zeta^2}{2}$
ζ	Shape parameter, $\zeta^2 = \ln\left(\frac{Var(t)}{E^2(t)} + 1\right)$

Table 9: Parameters and characteristic values of the log-normal distribution.

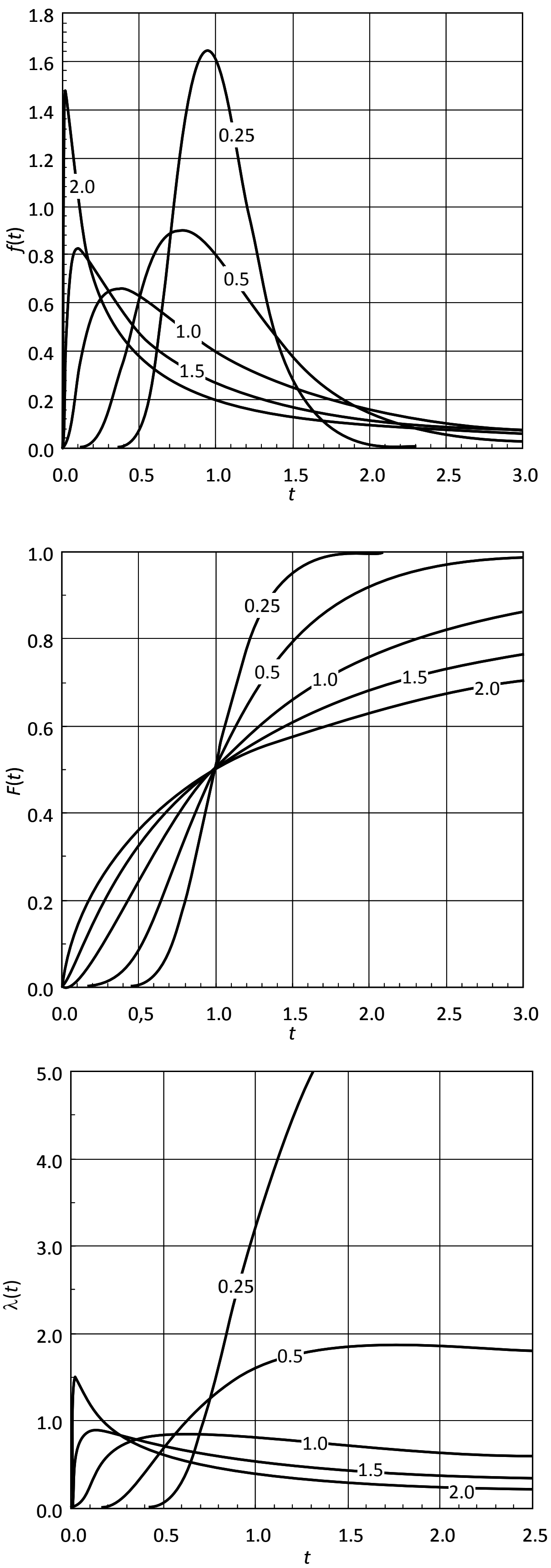

Fig. 51: Graphical representation of the density function (top), distribution function (center) and failure rate of the log-normal distribution (bottom) as a function of the dispersion parameter ζ, when $\eta = 0$.

6.2.5. Mixture of distributions

A mixture of distributions of failure times occurs when products that are subject to different failure modes (e.g. due to different manufacturing conditions or material batches) are grouped in a population. This means, for example, that Group 1 fails only due to cause A, and Group 2 only due to cause B. In the Weibull model, these two failure mechanisms constitute different shape parameters b_1 and b_2 and/or different characteristic lifetimes T_1 and T_2.

Consequently, when data from such a mixture of distributions is entered in the Weibull plot, it is not a straight line that is obtained in most cases, but a curve bending away from the time axis. Inasmuch as the cause of failure can be determined for each part, a separate Weibull analysis can be undertaken after the data has been separated accordingly.

If we mix two populations with a size n_1 or n_2, the distribution functions of which are $F_1(t)$ or $F_2(t)$, the resulting distribution function of the mixture (with $n = n_1 + n_2$) according to the theorem of total probability is:

$$F(t) = \frac{n_1}{n} \cdot F_1(t) + \frac{n_2}{n} \cdot F_2(t) \; . \qquad (6.57)$$

Generally, the following applies:

$$F(t) = \frac{1}{n} \sum_{i=1}^{k} n_i \cdot F_i(t) \text{, with } n = \sum_{i=1}^{k} n_i \; . \qquad (6.58)$$

Based on the density functions (relative frequencies), we can also say that:

$$f(t) = \frac{1}{n} \sum_{i=1}^{k} n_i \cdot f_i(t) \; . \qquad (6.59)$$

If an individual product *can* fail for various reasons without the probability of occurrence of the individual failure causes being known, the situation is different. In this case, we refer to competing failure modes. The survival probability $R(t)$ of a part then equals the product of the survival probabilities in terms of the individual k failure causes:

$$R(t) = R_1(t) \cdot R_2(t) \cdots R_k(t) \; . \qquad (6.60)$$

Weibull analysis of a data set based on indistinguishable, competing failure modes (which as such cannot be distinguished through testing of parts), is only possible with the aid of computers.

6.3. Tables

6.3.1. Standard normal distribution

Distribution function $F(u_0)$ of the standard normal distribution as per equation (3.12)

u_0	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	0.5039894	0.5079783	0.5119665	0.5159534	0.5199388	0.5239222	0.5279032	0.5318814	0.5358564
0.1	0.5437953	0.5477584	0.5517168	0.5556700	0.5596177	0.5635595	0.5674949	0.5714237	0.5753454
0.2	0.5831662	0.5870644	0.5909541	0.5948349	0.5987063	0.6025681	0.6064199	0.6102612	0.6140919
0.3	0.6217195	0.6255158	0.6293000	0.6330717	0.6368307	0.6405764	0.6443088	0.6480273	0.6517317
0.4	0.6590970	0.6627573	0.6664022	0.6700314	0.6736448	0.6772419	0.6808225	0.6843863	0.6879331
0.5	0.6949743	0.6984682	0.7019440	0.7054015	0.7088403	0.7122603	0.7156612	0.7190427	0.7224047
0.6	0.7290691	0.7323711	0.7356527	0.7389137	0.7421539	0.7453731	0.7485711	0.7517478	0.7549029
0.7	0.7611479	0.7642375	0.7673049	0.7703500	0.7733726	0.7763727	0.7793501	0.7823046	0.7852361
0.8	0.7910299	0.7938919	0.7967306	0.7995458	0.8023375	0.8051055	0.8078498	0.8105703	0.8132671
0.9	0.8185887	0.8212136	0.8238145	0.8263912	0.8289439	0.8314724	0.8339768	0.8364569	0.8389129
1.0	0.8437524	0.8461358	0.8484950	0.8508300	0.8531409	0.8554277	0.8576903	0.8599289	0.8621434
1.1	0.8665005	0.8686431	0.8707619	0.8728568	0.8749281	0.8769756	0.8789995	0.8809999	0.8829768
1.2	0.8868606	0.8887676	0.8906514	0.8925123	0.8943502	0.8961653	0.8979577	0.8997274	0.9014747
1.3	0.9049021	0.9065825	0.9082409	0.9098773	0.9114920	0.9130850	0.9146565	0.9162067	0.9177356
1.4	0.9207302	0.9221962	0.9236415	0.9250663	0.9264707	0.9278550	0.9292191	0.9305634	0.9318879
1.5	0.9344783	0.9357445	0.9369916	0.9382198	0.9394292	0.9406201	0.9417924	0.9429466	0.9440826
1.6	0.9463011	0.9473839	0.9484493	0.9494974	0.9505285	0.9515428	0.9525403	0.9535213	0.9544860
1.7	0.9563671	0.9572838	0.9581849	0.9590705	0.9599408	0.9607961	0.9616364	0.9624620	0.9632730
1.8	0.9648521	0.9656205	0.9663750	0.9671159	0.9678432	0.9685572	0.9692581	0.9699460	0.9706210
1.9	0.9719334	0.9725711	0.9731966	0.9738102	0.9744119	0.9750021	0.9755808	0.9761482	0.9767045
2.0	0.9777844	0.9783083	0.9788217	0.9793248	0.9798178	0.9803007	0.9807738	0.9812372	0.9816911
2.1	0.9825708	0.9829970	0.9834142	0.9838226	0.9842224	0.9846137	0.9849966	0.9853713	0.9857379
2.2	0.9864474	0.9867906	0.9871263	0.9874545	0.9877755	0.9880894	0.9883962	0.9886962	0.9889893
2.3	0.9895559	0.9898296	0.9900969	0.9903581	0.9906133	0.9908625	0.9911060	0.9913437	0.9915758
2.4	0.9920237	0.9922397	0.9924506	0.9926564	0.9928572	0.9930531	0.9932443	0.9934309	0.9936128
2.5	0.9939634	0.9941323	0.9942969	0.9944574	0.9946139	0.9947664	0.9949151	0.9950600	0.9952012
2.6	0.9954729	0.9956035	0.9957308	0.9958547	0.9959754	0.9960930	0.9962074	0.9963189	0.9964274
2.7	0.9966358	0.9967359	0.9968333	0.9969280	0.9970202	0.9971099	0.9971972	0.9972821	0.9973646
2.8	0.9975229	0.9975988	0.9976726	0.9977443	0.9978140	0.9978818	0.9979476	0.9980116	0.9980738
2.9	0.9981929	0.9982498	0.9983052	0.9983589	0.9984111	0.9984618	0.9985110	0.9985588	0.9986051
3.0	0.9986938	0.9987361	0.9987772	0.9988171	0.9988558	0.9988933	0.9989297	0.9989650	0.9989992
3.1	0.9990646	0.9990957	0.9991260	0.9991553	0.9991836	0.9992112	0.9992378	0.9992636	0.9992886
3.2	0.9993363	0.9993590	0.9993810	0.9994024	0.9994230	0.9994429	0.9994623	0.9994810	0.9994991
3.3	0.9995335	0.9995499	0.9995658	0.9995811	0.9995959	0.9996103	0.9996242	0.9996376	0.9996505
3.4	0.9996752	0.9996869	0.9996982	0.9997091	0.9997197	0.9997299	0.9997398	0.9997493	0.9997585
3.5	0.9997759	0.9997842	0.9997922	0.9997999	0.9998074	0.9998146	0.9998215	0.9998282	0.9998347
3.6	0.9998469	0.9998527	0.9998583	0.9998637	0.9998689	0.9998739	0.9998787	0.9998834	0.9998879
3.7	0.9998964	0.9999004	0.9999043	0.9999080	0.9999116	0.9999150	0.9999184	0.9999216	0.9999247
3.8	0.9999305	0.9999333	0.9999359	0.9999385	0.9999409	0.9999433	0.9999456	0.9999478	0.9999499
3.9	0.9999539	0.9999557	0.9999575	0.9999593	0.9999609	0.9999625	0.9999641	0.9999655	0.9999670
4.0	0.9999696	0.9999709	0.9999721	0.9999733	0.9999744	0.9999755	0.9999765	0.9999775	0.9999784
4.1	0.9999802	0.9999811	0.9999819	0.9999826	0.9999834	0.9999841	0.9999848	0.9999854	0.9999861
4.2	0.9999872	0.9999878	0.9999883	0.9999888	0.9999893	0.9999898	0.9999902	0.9999907	0.9999911
4.3	0.9999918	0.9999922	0.9999925	0.9999929	0.9999932	0.9999935	0.9999938	0.9999941	0.9999943
4.4	0.9999948	0.9999951	0.9999953	0.9999955	0.9999957	0.9999959	0.9999961	0.9999963	0.9999964
4.5	0.9999968	0.9999969	0.9999971	0.9999972	0.9999973	0.9999974	0.9999976	0.9999977	0.9999978
4.6	0.9999980	0.9999981	0.9999982	0.9999983	0.9999983	0.9999984	0.9999985	0.9999986	0.9999986
4.7	0.9999988	0.9999988	0.9999989	0.9999989	0.9999990	0.9999990	0.9999991	0.9999991	0.9999992

6.3.2. t-distribution

α /2 -quantile of the t-distribution with degree of freedom f (two-sided) as per equation (6.31)

	α=0.05	0.01	0.001
f=1	12.71	63.66	636.62
2	4.30	9.92	31.60
3	3.18	5.84	12.92
4	2.78	4.60	8.61
5	2.57	4.03	6.87
6	2.45	3.71	5.96
7	2.36	3.50	5.41
8	2.31	3.36	5.04
9	2.26	3.25	4.78
10	2.23	3.17	4.59
11	2.20	3.11	4.44
12	2.18	3.05	4.32
13	2.16	3.01	4.22
14	2.14	2.98	4.14
15	2.13	2.95	4.07
16	2.12	2.92	4.01
17	2.11	2.90	3.97
18	2.10	2.88	3.92
19	2.09	2.86	3.88
20	2.09	2.85	3.85
25	2.06	2.79	3.73
30	2.04	2.75	3.65
35	2.03	2.72	3.59
40	2.02	2.70	3.55
45	2.01	2.69	3.52
50	2.01	2.68	3.50
100	1.98	2.63	3.39
200	1.97	2.60	3.34
300	1.97	2.59	3.32
400	1.97	2.59	3.32
500	1.96	2.59	3.31
100000	1.96	2.58	3.29

6.3.3. Median values of failure probability

Values for $F_{BB,1-\alpha}$ for α = 50 %, rank *i* and sample size *n* as per equation (6.35)

	n=1	2	3	4	5	6	7	8	9	10
i=1	0.500000	0.292893	0.206299	0.159104	0.129449	0.109101	0.094276	0.082996	0.074125	0.066967
2		0.707107	0.500000	0.385728	0.313810	0.264450	0.228490	0.201131	0.179620	0.162263
3			0.793701	0.614272	0.500000	0.421407	0.364116	0.320519	0.286237	0.258575
4				0.840896	0.686190	0.578593	0.500000	0.440155	0.393085	0.355100
5					0.870551	0.735550	0.635884	0.559845	0.500000	0.451694
6						0.890899	0.771510	0.679481	0.606915	0.548306
7							0.905724	0.798869	0.713763	0.644900
8								0.917004	0.820380	0.741425
9									0.925875	0.837737
10										0.933033

	n=11	12	13	14	15	16	17	18	19	20
i=1	0.061069	0.056126	0.051922	0.048305	0.045158	0.042397	0.039953	0.037776	0.035824	0.034064
2	0.147963	0.135979	0.125791	0.117022	0.109396	0.102703	0.096782	0.091506	0.086775	0.082510
3	0.235786	0.216686	0.200449	0.186474	0.174321	0.163654	0.154218	0.145810	0.138271	0.131474
4	0.323804	0.297576	0.275276	0.256084	0.239393	0.224745	0.211785	0.200238	0.189885	0.180550
5	0.411890	0.378529	0.350163	0.325751	0.304520	0.285886	0.269400	0.254712	0.241543	0.229668
6	0.500000	0.459507	0.425077	0.395443	0.369670	0.347050	0.327038	0.309207	0.293220	0.278805
7	0.588110	0.540493	0.500000	0.465147	0.434833	0.408227	0.384687	0.363714	0.344909	0.327952
8	0.676196	0.621471	0.574923	0.534853	0.500000	0.469408	0.442342	0.418226	0.396603	0.377105
9	0.764214	0.702424	0.649837	0.604557	0.565167	0.530592	0.500000	0.472742	0.448301	0.426262
10	0.852037	0.783314	0.724724	0.674249	0.630330	0.591773	0.557658	0.527258	0.500000	0.475420
11	0.938931	0.864021	0.799551	0.743916	0.695480	0.652950	0.615313	0.581774	0.551699	0.524580
12		0.943874	0.874209	0.813526	0.760607	0.714114	0.672962	0.636286	0.603397	0.573738
13			0.948078	0.882978	0.825679	0.775255	0.730600	0.690793	0.655091	0.622895
14				0.951695	0.890604	0.836346	0.788215	0.745288	0.706780	0.672048
15					0.954842	0.897297	0.845782	0.799762	0.758457	0.721195
16						0.957603	0.903218	0.854190	0.810115	0.770332
17							0.960047	0.908494	0.861729	0.819450
18								0.962224	0.913225	0.868526
19									0.964176	0.917490
20										0.965936

Continued: Values for $F_{BB,1-\alpha}$ for α = 50 %. Rank i and sample size n as per equation (6.35)

	n=21	22	23	24	25	26	27	28	29	30
i=1	0.032468	0.031016	0.029687	0.028468	0.027345	0.026307	0.025345	0.024451	0.023618	0.022840
2	0.078644	0.075124	0.071906	0.068952	0.066231	0.063717	0.061386	0.059221	0.057202	0.055317
3	0.125313	0.119704	0.114576	0.109868	0.105533	0.101526	0.097813	0.094361	0.091145	0.088141
4	0.172090	0.164386	0.157343	0.150879	0.144925	0.139422	0.134323	0.129583	0.125166	0.121041
5	0.218905	0.209107	0.200147	0.191924	0.184350	0.177351	0.170864	0.164834	0.159216	0.153968
6	0.265740	0.253844	0.242968	0.232986	0.223791	0.215294	0.207419	0.200100	0.193279	0.186909
7	0.312584	0.298592	0.285798	0.274056	0.263241	0.253246	0.243983	0.235373	0.227350	0.219857
8	0.359434	0.343345	0.328634	0.315132	0.302695	0.291203	0.280551	0.270651	0.261426	0.252809
9	0.406288	0.388102	0.371473	0.356211	0.342153	0.329163	0.317123	0.305932	0.295504	0.285764
10	0.453143	0.432860	0.414315	0.397292	0.381614	0.367125	0.353696	0.341215	0.329585	0.318721
11	0.500000	0.477620	0.457157	0.438375	0.421075	0.405088	0.390271	0.376500	0.363667	0.351680
12	0.546857	0.522380	0.500000	0.479458	0.460537	0.443053	0.426847	0.411785	0.397749	0.384639
13	0.593712	0.567140	0.542843	0.520542	0.500000	0.481018	0.463423	0.447071	0.431833	0.417599
14	0.640566	0.611898	0.585685	0.561625	0.539463	0.518982	0.500000	0.482357	0.465916	0.450559
15	0.687416	0.656655	0.628527	0.602708	0.578925	0.556947	0.536577	0.517643	0.500000	0.483520
16	0.734260	0.701408	0.671366	0.643789	0.618386	0.594912	0.573153	0.552929	0.534084	0.516480
17	0.781095	0.746156	0.714202	0.684868	0.657847	0.632875	0.609729	0.588215	0.568167	0.549441
18	0.827910	0.790893	0.757032	0.725944	0.697305	0.670837	0.646304	0.623500	0.602251	0.582401
19	0.874687	0.835614	0.799853	0.767014	0.736759	0.708797	0.682877	0.658785	0.636333	0.615361
20	0.921356	0.880296	0.842657	0.808076	0.776209	0.746754	0.719449	0.694068	0.670415	0.648320
21	0.967532	0.924876	0.885424	0.849121	0.815650	0.784706	0.756017	0.729349	0.704496	0.681279
22		0.968984	0.928094	0.890132	0.855075	0.822649	0.792581	0.764627	0.738574	0.714236
23			0.970313	0.931048	0.894467	0.860578	0.829136	0.799900	0.772650	0.747191
24				0.971532	0.933769	0.898474	0.865677	0.835166	0.806721	0.780143
25					0.972655	0.936283	0.902187	0.870417	0.840784	0.813091
26						0.973693	0.938614	0.905639	0.874834	0.846032
27							0.974655	0.940779	0.908855	0.878959
28								0.975549	0.942798	0.911859
29									0.976382	0.944683
30										0.977160

6.3.4. Confidence bounds of failure probability

Values for $F_{BB,1-\alpha}$ for α = 5 % (one-sided), rank *i* and sample size *n* as per equation (6.35)

	n=1	2	3	4	5	6	7	8	9	10
i=1	0.950000	0.776393	0.631597	0.527129	0.450720	0.393038	0.348164	0.312344	0.283129	0.258866
2		0.974679	0.864650	0.751395	0.657408	0.581803	0.520703	0.470679	0.429136	0.394163
3			0.983048	0.902389	0.810745	0.728662	0.658739	0.599689	0.549642	0.506901
4				0.987259	0.923560	0.846839	0.774678	0.710759	0.655059	0.606624
5					0.989794	0.937150	0.871244	0.807097	0.748632	0.696463
6						0.991488	0.946624	0.888887	0.831250	0.777559
7							0.992699	0.953611	0.902253	0.849972
8								0.993609	0.958977	0.912736
9									0.994317	0.963229
10										0.994884

	n=11	12	13	14	15	16	17	18	19	20
i=1	0.238404	0.220922	0.205817	0.192636	0.181036	0.170750	0.161566	0.153318	0.145869	0.139108
2	0.364359	0.338681	0.316340	0.296734	0.279396	0.263957	0.250124	0.237661	0.226374	0.216106
3	0.470087	0.438105	0.410099	0.385390	0.363442	0.343825	0.326193	0.310263	0.295802	0.282619
4	0.564374	0.527327	0.494650	0.465657	0.439784	0.416572	0.395641	0.376679	0.359426	0.343664
5	0.650188	0.609138	0.572619	0.540005	0.510752	0.484396	0.460549	0.438883	0.419120	0.401028
6	0.728750	0.684762	0.645201	0.609585	0.577444	0.548347	0.521918	0.497828	0.475797	0.455582
7	0.800424	0.754700	0.712951	0.674972	0.640435	0.608988	0.580295	0.554046	0.529967	0.507818
8	0.864925	0.818975	0.776045	0.736415	0.700014	0.666626	0.635991	0.607845	0.581936	0.558035
9	0.921180	0.877149	0.834341	0.793927	0.756273	0.721397	0.689170	0.659402	0.631885	0.606415
10	0.966681	0.928130	0.887334	0.847282	0.809135	0.773308	0.739886	0.708799	0.679913	0.653069
11	0.995348	0.969540	0.933950	0.895953	0.858336	0.822234	0.788092	0.756039	0.726054	0.698046
12		0.995735	0.971947	0.938897	0.903342	0.867889	0.833637	0.801047	0.770279	0.741349
13			0.996062	0.974001	0.943153	0.909748	0.876229	0.843656	0.812496	0.782931
14				0.996343	0.975774	0.946854	0.915355	0.883574	0.852530	0.822689
15					0.996586	0.977321	0.950102	0.920305	0.890094	0.860446
16						0.996799	0.978682	0.952975	0.924706	0.895919
17							0.996987	0.979889	0.955535	0.928646
18								0.997154	0.980967	0.957831
19									0.997304	0.981935
20										0.997439

Continued: Values for $F_{BB,1-\alpha}$ for α = 5 %, rank i and sample size n as per equation (6.35)

	n=21	22	23	24	25	26	27	28	29	30
i=1	0.132946	0.127305	0.122123	0.117346	0.112928	0.108830	0.105019	0.101466	0.098145	0.095034
2	0.206725	0.198122	0.190204	0.182892	0.176121	0.169831	0.163974	0.158507	0.153392	0.148596
3	0.270552	0.259467	0.249249	0.239801	0.231040	0.222894	0.215300	0.208205	0.201561	0.195326
4	0.329211	0.315913	0.303638	0.292273	0.281723	0.271902	0.262739	0.254170	0.246139	0.238598
5	0.384408	0.369091	0.354932	0.341807	0.329608	0.318242	0.307626	0.297691	0.288372	0.279615
6	0.436976	0.419800	0.403899	0.389139	0.375405	0.362595	0.350620	0.339402	0.328873	0.318971
7	0.487389	0.468495	0.450975	0.434691	0.419520	0.405354	0.392098	0.379670	0.367996	0.357009
8	0.535936	0.515457	0.496435	0.478728	0.462209	0.446767	0.432302	0.418728	0.405966	0.393947
9	0.582801	0.560868	0.540456	0.521423	0.503642	0.486998	0.471392	0.456731	0.442936	0.429934
10	0.628099	0.604844	0.583155	0.562893	0.543933	0.526162	0.509478	0.493789	0.479012	0.465073
11	0.671891	0.647456	0.624606	0.603215	0.583162	0.564337	0.546640	0.529979	0.514270	0.499439
12	0.714200	0.688736	0.664852	0.642436	0.621378	0.601576	0.582931	0.565355	0.548765	0.533086
13	0.755006	0.728687	0.703907	0.680579	0.658611	0.637911	0.618387	0.599956	0.582536	0.566055
14	0.794250	0.767276	0.741757	0.717644	0.694870	0.673358	0.653028	0.633803	0.615608	0.598371
15	0.831824	0.804437	0.778364	0.753611	0.730147	0.707918	0.686861	0.666910	0.647995	0.630052
16	0.867552	0.840059	0.813656	0.788434	0.764414	0.741576	0.719880	0.699275	0.679704	0.661107
17	0.901156	0.873966	0.847520	0.822039	0.797622	0.774300	0.752066	0.730889	0.710729	0.691536
18	0.932194	0.905891	0.879785	0.854314	0.829696	0.806040	0.783383	0.761729	0.741056	0.721330
19	0.959900	0.935404	0.910191	0.885089	0.860525	0.836718	0.813780	0.791758	0.770660	0.750474
20	0.982809	0.961776	0.938324	0.914115	0.889944	0.866226	0.843181	0.820923	0.799504	0.778941
21	0.997560	0.983602	0.963485	0.940992	0.917709	0.894404	0.871478	0.849149	0.827535	0.806692
22		0.997671	0.984326	0.965047	0.943437	0.921014	0.898515	0.876331	0.854678	0.833674
23			0.997772	0.984988	0.966480	0.945688	0.924064	0.902318	0.880831	0.859815
24				0.997865	0.985597	0.967801	0.947767	0.926886	0.905845	0.885013
25					0.997950	0.986158	0.969022	0.949692	0.929506	0.909126
26						0.998029	0.986677	0.970153	0.951480	0.931944
27							0.998102	0.987159	0.971204	0.953145
28								0.998170	0.987606	0.972184
29									0.998233	0.988024
30										0.998292

7. Literature

[1] B. Bertsche, G. Lechner: Zuverlässigkeit im Fahrzeug- und Maschinenbau, 3rd Edition, Springer, 2004

[2] Bosch Product Engineering System: Committed BES Practice Zuverlässigkeitsgestaltung von Designelementen, 2010

[3] Standard IEC 60050-191: International Electrotechnical Vocabulary, Chapter 191: Dependability and quality of service, 1990

[4] J. C. Laprie (ed.): Dependability: Basic Concepts and Terminology. In: Dependable Computing and Fault-Tolerant Systems, Vol. 5, Springer, 1992

[5] E. Haibach: Betriebsfestigkeit, VDI, 1989

[6] A. Haldar, S. Mahadevan: Probability, Reliability, and Statistical Methods in Engineering Design, John Wiley, 2000

[7] VDA Series on Quality Management in the Automotive Industry, Vol. 4 "Sicherung der Qualität während der Produktrealisierung – Methoden und Verfahren", Chapter 2 "Entwicklungsabläufe"

[8] VDA Series on Quality Management in the Automotive Industry, Vol. 3 "Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten", Chapter 2 "Zuverlässigkeitsmethoden und Hilfsmittel"

[9] Series Quality management in the Bosch Group, No. 14 Failure Modes and Effects Analysis FMEA

[10] MIL-HDBK-217F (Notice 2): Reliability Prediction of Electronic Equipment, U.S. Department of Defense, 1995

[11] Bosch Product Engineering System: PE Guide, 2010

[12] G. Yang: Life Cycle Reliability Engineering, John Wiley, 2007

[13] B. Bertsche, P. Göhner, U. Jensen, W. Schinköthe, H.-J. Wunderlich: Zuverlässigkeit mechatronischer Systeme, Springer, 2009

[14] G. S. Wasserman: Reliability Verification, Testing, and Analysis in Engineering Design, Dekker, 2003

[15] P.D.T. O'Connor: Practical Reliability Engineering, 4th Edition, Wiley, 2002

[16] Series Quality management in the Bosch Group, No. 6 Evaluation of Field Data

[17] A. Krolo: Planung von Zuverlässigkeitstests mit weitreichender Berücksichtigung von Vorkenntnissen, Report 110 of the IMA, University of Stuttgart, 2004

8. Index

2020-04-06 - SOCOS

Robert Bosch GmbH
C/QMM
Postfach 30 02 20
D-70442 Stuttgart
Germany
Phone +49 711 811-4 47 88
Fax +49 711 811-2 31 26
www.bosch.com