Basic Concepts

1.1. Introduction

Definitions of terms and parameters used throughout this book are given in this chapter. They are based¹ mainly on [IEC 15a] and its companion standard [IEC 16a], although not all standards (even IEC ones) are in line with [IEC 15a]. The assessments of the production availability and the reliability parameters are most of the time based on the assumption that the failure rate is a constant, so the meaning of this assumption is given as well as its limits. In addition to this, the characteristics of the so-called "bathtub curve" are provided. The bathtub curve is mainly relevant for the early phase that is crucial for a proper evaluation of the economics of a project.

1.2. Definition of terms

1.2.1. Risk

During the life of a plant, events may occur which could impact human life, environment, equipment or project profitability. These events can then be named unwanted events. Then for each of these four categories, a curve (see Figure 1.1) characterized by (1) its frequency of occurrence (or its probability) and (2) its severity (amount of the consequence to human life, environment, equipment or project profitability) can be determined. This curve is called a risk:

¹ There are three main bodies in charge of producing standards: ISO (International Organization for Standardization), IEC (International Electrotechnical Commission) and IUT (International Telecommunication Union). According to an agreement with ISO, IEC produces standards on dependability (section 1.3.2) for all sectors (not only in the electrical and electronic fields). However, ISO issues oil- and gas-specific standards on dependability.

- The severity axis could be the plant production unavailability and the probability axis the probability of reaching this production unavailability.

- The severity axis could be the potential number of deaths and the probability axis the probability of occurrence of these events.

[IEC 13a] considers the risk as a combination of the probability and the severity without considering the magnitude of the consequences. Then this definition is to be discarded, as risk causing the plant to be in the red is not at all to be considered on the same level as a risk causing a decrease of 10% in the plant profitability.

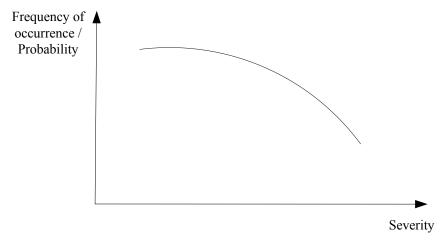


Figure 1.1. Risk in the two-dimension space

For safety (or environmental) risks, a major step is to be carried out before trying to assess the risks: the identification of the hazards (i.e. the potential sources of impact on human life and on the environment²) and their characterization. This theme is not considered in this book.

1.2.2. Time definitions

Figure 1.2 shows the failure-to-repair cycle of a repairable item (TBF = operating Time Between Failures, RT = Repair Time).

² Sometimes damage to equipment is also considered.

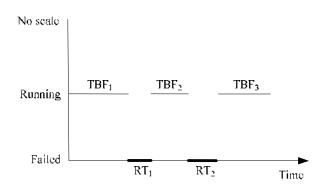


Figure 1.2. Failure-to-repair cycle

The Mean operating Time Between Failures (MTBF) is calculated as follows:

$$MTBF = \frac{\sum_{i} TBF_{i}}{Number of failures} = \frac{\sum Operating times}{Number of failures}$$
[1.1]

The word "operating" was introduced into the definition to prevent the use of $\sum (operating time + repair time)$ as the numerator.

The Mean Repair Time (MRT) is calculated as follows:

$$MRT = \frac{\sum_{i} RT_{i}}{Number of repairs} = \frac{\sum Repair times}{Number of repairs}$$
[1.2]

The acronym MTTR was used in the past instead of MRT. [IEC 15a] considers the MTTR as the Mean Time To Restoration, an acronym no longer used in this book as the restoration time is a mixture of repair times, start-up times, mobilization times, etc., which are different physical events.

For non-repairable items, there is obviously no Repair Time (and then no MRT) and the acronym MTTF (Mean operating Time To Failure) is used instead of MTBF.

The Mean operating Time To Failure is calculated as follows:

 $MTTF = \frac{\sum Operating times}{Number of failed items}$ [1.3]

1.2.3. Failures and repairs

1.2.3.1. Definitions

A failure is the loss of the ability to perform as required. The failure causes are the set of circumstances that lead to the failure. The failure modes are the effects by which the failure is observed. The failure effects are the consequences of the failure. Figure 1.3 shows an example of the relationship of failure causes, failure modes and failure effects for a pressure safety valve.

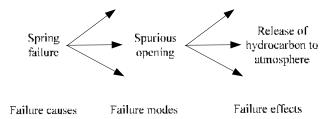


Figure 1.3. Relationship of failure causes, failure modes and failure effects

Failure modes are often classified in several ways depending on the intent of their use, e.g. according to [LEE 12]:

- Condition category (used for maintenance purposes) that emphasizes the causes.
- Performance category that emphasizes the effects.
- Safety category: see section 1.2.4.4.
- Detection category: see section 1.2.4.4.

These categories are used for defining the purpose of Failure Mode and Effects Analysis (section 4.2).

"The failure mechanism is the physical, chemical, thermodynamic or other process or combination that leads to the failure. It is an attribute of the failure event that can be deduced technically" (from [ISO 16]).

The circumstances that induce or activate the processes are termed the root causes of failure.

1.2.3.2. Measures

Several definitions are provided in the standards, which are as follows:

– The instantaneous failure rate $\lambda(t)$ (also named failure rate, hazard function, hazard rate, force of mortality):

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \text{Probability}(\text{Failure occurs within}[t, t + \Delta t]/$$

no failure occurred within [0, t]) [1.4]

also written as [IEC 16a]:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E([N(t + \Delta t) - N(t)] / up \text{ state at } t = 0])$$
[1.5]

- The instantaneous failure intensity z(t) (also named failure intensity, failure frequency, Rate of OCcurrence Of Failures [ROCOF]):

$$z(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E([N(t + \Delta t) - N(t)]) \text{ as good as new at } t = 0)$$
[1.6]

where:

-E(x): expectation of x

-N(t): number of failures in the time interval [0, t].

– The conditional failure intensity $\lambda_v(t)$ (also named Vesely failure rate):

$$\lambda_{v}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E([N(t + \Delta t) - N(t)]/$$

up state at t and as good as new at t = 0) [1.7]

Using [1.7], it can be shown that:

$$\lambda_{\rm v}(t) = z(t) * A(t)$$
[1.8]

The probability of failure upon demand (the probability of starting failure) is the ratio of the total start failures to the number of attempted item starts [IEE 07a]:

$$\gamma = \frac{\text{Total start failures}}{\text{Number of attempted item starts}}.$$
[1.9]

An instantaneous repair rate $\mu(t)$ can also be defined:

$$\mu(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \operatorname{Probability}(\operatorname{Repair completed within} [t, t + \Delta t]/$$

Repair started at t = 0 and not completed at t) [1.10]

Within this book, reliability data³ are data on failure frequencies (and probabilities of failure to start), repair times and failure mode percentages.

1.2.3.3. Phases of use of an item

The three phases of use of any item are:

- OFF, i.e. standby phase or on-guard phase;

- transition OFF - ON, i.e. (nearly instantaneous) switch to running phase;

- ON, i.e. running phase.

These are shown in Figure 1.4.

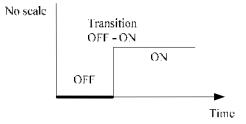


Figure 1.4. The three phases of use of an item

The qualification of the three phases of use of normally running items and on-guard items along their life are as follows:

1) For normally running items (the ones considered in production availability studies): standby phase (or even mothballed phase if the standby period exceeds several weeks), transition standby phase to running phase and running phase. Each of these phases has its own reliability characteristics:

i) Standby failure rate for the standby phase: it is, most of the time, considered as negligible.

ii) Probability of failure upon demand γ for the standby phase to running phase: this parameter is often considered for rotating machines (also named failure to start) only and is given in reliability data books (Chapter 6).

iii) Failure rate λ for the running phase: the failure rate provided within the reliability data books. According to [JAN 15], a major gas turbine manufacturer bases its gas turbine maintenance requirements on independent counts of starts and

^{3 [}ISO 16] uses "reliability and maintainability data" instead of reliability data alone.

running hours, e.g. hot gas path inspection is to be performed every 24,000 hr or 1,200 starts, whichever criteria limit is reached first. They implemented this approach as life limiters are different for starts and running hours: definitively γ and λ are not to be considered in the same way.

2) For on-guard items (the ones considered in reliability studies): on-guard phase, transition on-guard phase to running phase and running phase. Each of these phases has its own reliability characteristics:

i) Failure rate λ for the on-guard phase (standby failure rate): this parameter is used for calculating PFD_{avg} and PFH (section 1.2.4.3). It is given in reliability data books.

ii) Probability of failure upon demand γ for the on-guard phase to running phase: this parameter is most of the time not considered. As a consequence, PFD_{avg} and PFH considered in reliability calculations are too high, causing often to recommend a too-high proof test frequency.

iii) Failure rate λ for the running phase: the failure rate is not considered as the running phase is short compared to the standby phase.

Based on [NUR 03], it could be considered that standby failure rate is most appropriate for shutoff valves but that standby failure rate and probability of failure upon demand γ are to be considered for on-guard rotating machines such as emergency diesel generators.

A causal analysis coupled with expert judgment [PIE 92] allows us to determine the standby failure rate, the probability of failure upon demand γ and the failure rate λ (for the running phase) from field data.

REMARK 1.1.– The PFD_{avg} (see section 1.2.4.3) is not the probability of failure upon demand γ .

1.2.3.4. Failure severity

According to [IEC 15a], a complete failure is a failure characterized by the loss of all required functions. However, this definition is not used in the oil and gas industry for characterizing the severity of the failure. The ones of [ISO 16] are used instead in production availability studies:

- Critical failure (also called a complete failure): immediate cessation of the ability to perform the required function.

- Degraded failure: the ability to perform the required function is not stopped but other functions are compromised.

Qualifiers, such as catastrophic, critical, major, etc., used in safety/reliability studies make references to the effect of the failure on the system (e.g. catastrophic is, most of the time, used to qualify failures that could affect human life).

1.2.4. IEC 61508 terms

1.2.4.1. Definitions

[IEC 10] has "set out a generic approach for all safety lifecycle activities for systems comprised of electrical and/or electronic and/or programmable electronic (E/E/PE) elements that are used to perform safety functions", which suggests that this standard does not apply to safety elements such as Pressure Safety Valves or dykes.

The functional safety is defined as the part of the overall safety relating to the equipment under control and of its control system that depends on the correct functioning of the E/E/PE safety-related systems and other risk reduction measures, and the safety integrity as the probability of an E/E/PE safety-related system satisfactorily performing the specified safety functions under all of the stated conditions within a stated period.

1.2.4.2. Risk and safety integrity

Safety (related) systems are designed to prevent a process deviation (e.g. a high pressure [HP]) or an external event (e.g. a load drop) from becoming an unwanted event, and ending in a risk. These safety systems are either on-guard systems if these events occur from time to time or active systems if these events are always present.

For on-guard systems, the risk occurs if the unwanted event occurs (at time t) and if the safety system is not available at t. Then the frequency of occurrence of the risk is equal to the product of the frequency of the unwanted event by the mean unavailability of the safety system over the calculation period. This calculation period is the duration T between two full proof tests of all of the items making up the safety system:

Up to the first issue of [IEC 10], the mean unavailability was known as the Fractional Dead Time (FDT) and is calculated as follows:

$$FDT = \frac{1}{T} \int_0^T (1 - A(t)) dt$$
 [1.11]

where A(t) is the availability of the system (section 1.2.3.).

For active systems, the risk occurs if the safety system fails. Then the yearly frequency of occurrence of the risk is equal to the yearly frequency of occurrence of failure of the safety system.

1.2.4.3. Measures

[IEC 10] has defined:

– the Probability of dangerous Failure on Demand (PFD(t)) as the unavailability of an E/E/PE safety-related system to perform the specified safety function when a demand occurs from the equipment under control or its control system, and the average Probability of dangerous Failure on Demand (PFD_{avg}) as the mean unavailability (the FDT above)

- the average frequency of a dangerous Failure per Hour (PFH)⁴ as the average frequency of a dangerous failure of an E/E/PE safety-related system to perform the specified safety function over a given period.

The Safety Integrity Level (SIL) is a discrete level, corresponding to a range of safety integrity values as shown on Table 1.1^5 .

SIL	PFD _{avg}	$PFH (hr^{-1})$
4	$\geq 10^{-5}$ to < 10^{-4}	$\ge 10^{-9}$ to $< 10^{-8}$
3	$\ge 10^{-4}$ to < 10^{-3}	$\ge 10^{-8}$ to $< 10^{-7}$
2	$\geq 10^{-3}$ to < 10^{-2}	$\geq 10^{-7}$ to < 10^{-6}
1	$\geq 10^{-2}$ to < 10^{-1}	$\ge 10^{-6}$ to $< 10^{-5}$

Table 1.1. The four safety integrity levels

⁴ The P of PFH is confusing as the PFH is a frequency not a probability.

⁵ The SILs of the PFH were calculated by dividing the SILs of the PFD_{avg} by 10,000 hr and not by 8,760 hr (exactly 1 year) to avoid the use of digits after the comma.

1.2.4.4. Undetected dangerous failures

[IEC 10] also defines6:

- the undetected failures (failures detected by proof tests only) and the detected failures (self-revealed failures, failures detected by diagnostic tests, by operator intervention or actual operation of the system)

- the dangerous failures (failures preventing the safety system from acting, or decreasing its probability of acting properly) and the safe failures (failures resulting in spurious operation or increasing the probability of spurious operation).

The undetected dangerous failures make the highest contribution to safety system unrevealed unavailability (section 3.5.2).

The proof test (periodical test) is a periodic test performed to detect dangerous hidden failures, whereas the diagnostic test is a test performed by automatic online diagnostic test.

EXAMPLE 1.1.– For a solenoid valve, a diagnostic test consists of checking the integrity of the electric line, whereas the proof test consists of checking to see whether the solenoid valve moves.

1.3. Definition of parameters

1.3.1. Reliability

The reliability R of an item I is its ability to perform as required, without failure, for a given time interval under given conditions, i.e.:

R(t) = Probability (I does not fail over [0,t])[1.12]

1.3.1.1. Reliability and failure rate

Using the definition of the conditional probability ([2.3]), and noting Probability as Pr, the instantaneous failure rate can be written as follows [GON 86]:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{\Pr\left(I \text{ fails within } [t, t + \Delta t] \text{ AND } I \text{ not failed on } [0, t]\right)}{\Pr(I \text{ does not fail over } [0, t])}$$

⁶ IEC 61508 gives also the definition of the safe failure fraction (ratio of safe + dangerous detected failures to safe + dangerous failures). This concept of safe failure fraction is no longer in use in the oil and gas industry since the issue of the second edition of [IEC 16b].

Otherwise written:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{(\Pr (I \text{ fails within } [0, t + \Delta t] - \Pr (I \text{ fails within } [0, t]))}{R(t)}$$

or:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{R(t) - R(t + \Delta t)}{R(t)}$$
$$\lambda(t) = \frac{-\frac{dR(t)}{dt}}{R(t)}$$

and then:

$$R(t) = e^{-\int_0^t \lambda(t)dt}$$
[1.13]

1.3.1.2. Reliability and MTTF

Let us consider t the random variable measuring the time of good functioning of an item [GON 86]. The cumulative density function F(t) (section 2.2.3) of t (the probability of failure over [0, t]) is by definition of the reliability:

$$F(t) = 1 - R(t)$$

The probability density function is then:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}$$

The MTTF is then:

$$MTTF = \int_0^\infty t f(t)dt = -\int_0^\infty t \frac{dR(t)}{dt}dt$$

As the MTTF is not infinite it comes:

$$MTTF = \int_0^\infty R(t)dt$$
 [1.14]

If it is considered that after the repair the item is "as good as new", then the same formula is valid for assessing the MTBF.

1.3.2. Maintainability

The maintainability M of an item I is its ability to be retained in, or restored to, a state to perform as required, under given conditions of use and maintenance, i.e.:

M(t) = Probability (I repaired at t / repair started at t = 0) [1.15]

1.3.2.1. Maintainability and repair rate

As for the failure rate and the reliability, it can be shown that:

$$M(t) = 1 - e^{-\int_0^t \mu(t)dt}$$
[1.16]

1.3.2.2. Maintainability and MRT

As for the failure rate and the MTTF, it can be shown that:

$$MRT = \int_{0}^{\infty} (1 - M(t))dt$$
 [1.17]

1.3.3. Availability and production availability

The availability A of an item is its ability to perform a required function at a given instant under given conditions, i.e.:

$$A(t) = Probability (System not failed at t)$$
[1.18]

For non-repairable items, the availability is obviously synonymous with reliability.

In logistic engineering [BLA 03], the following definitions are often used:

$$Inherent \ availability = \frac{MTBF}{MTBF + MRT}$$

Achieved availability

Mean operating time between maintenance tasks

 $^-$ Mean operating time between maintenance tasks + Mean maintenance task duration

Operational availability

Mean operating time between maintenance tasks

 $\begin{tabular}{ll} Mean operating time between maintenance tasks + Mean maintenance downtime \\ \end{tabular}$

The mean maintenance downtime is the sum the MRT, the logistic delay time and the administrative delay time. However, these definitions of the availability do not specify how to consider degraded production states (i.e. production at 75% of the nominal production). [ISO 08] has therefore defined the production availability⁷ as the ratio of the actual production to the planned production over a specified period.

EXAMPLE 1.2.– The annual production availability of a plant producing at nominal capacity (100%) for T (100%) hours, at 75% for T (75%) hours and at 55% for T (55%) hours is (1 year = 8,760 hr):

Annual production availability = $\frac{T(100\%)*100\% + T(75\%)*75\% + T(55\%)x*55\%}{8,760*100\%}$

1.3.4. Dependability

According to [IEC 15a], the dependability of an item is its ability to perform as and when required.

Then dependability includes at least availability, reliability, maintainability, maintenance support performance and, in some cases, other characteristics such as safety. This is an umbrella word whose value cannot be assessed.

1.3.5. Definitions used by maintenance engineers

Often maintenance engineers use specific definitions for the reliability and the availability of gas turbines (close to definitions of [ISO 99] of availability factor and reliability factor):

- Availability: ability of an item to be in a state so that it performs a required function, at a given instant or over within a given lapse of time, assuming the availability of all necessary means.

$$Availability = \frac{RUN + ABNR}{TT}$$

- Reliability: ability of an item to perform a required function within a given lapse of time with given conditions.

$$Reliability = \frac{RUN + ABNR + SCH}{TT}$$

⁷ Production availability does not only mean availability of the production; it is also used to calculate e.g. the water injection production availability (Chapter 11).

where:

- ABNR: Available But Not Required.
- RUN: RUNning period.
- SCH: SCHeduled downtime for maintenance.
- TT: Total Time = RUN + ABNR + SCH + UNSCH.
- UNSCH: UNSCHeduled downtime for maintenance.

Definitions of [IEE 07a] to be used for electric generating units are similar.

1.3.6. Definitions used in the refinery industry

For making comparisons between units of several plants, the following definitions are commonly used:

- Mechanical availability: the percentage of time available for manufacturing after subtracting maintenance downtimes.

- Operational availability: the percentage of time available for manufacturing after subtracting maintenance and regulatory/process downtimes.

- On-stream factor: the percentage of time available for manufacturing after subtracting all downtimes.

1.4. The exponential law/the constant failure rate

1.4.1. Reliability

Assuming $\lambda(t) = \text{Constant}$, it comes:

$$R(t) = e^{-\lambda t}$$
[1.19]

The reliability curve is then as shown in Figure 1.5.

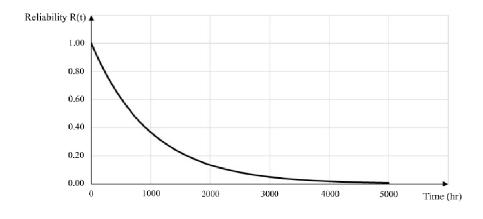


Figure 1.5. The exponential law

And:
$$MTTF = \int_0^\infty R(t)dt = \int_0^\infty e^{-\lambda t}dt = 1/\lambda$$

 $MTTF = 1/\lambda$
[1.20]

If the failure rate is a constant, the MTTF is the reciprocal of the failure rate. As the MTTF is always a constant, the failure rate cannot be calculated in any case as being the reciprocal of the MTTF.

Making t = MTTF, it comes: $R(t = MTTF) = e^{-\lambda * MTTF} = e^{-1} = 0.368$

There is (roughly) one chance out of three for an item to have no failure over its MTTF.

If a repairable item is as good as new after a repair:

$$MTBF = 1/\lambda$$
[1.21]

1.4.2. Validity

It is nearly always assumed that the assumption "the failure rate is a constant" is valid. According to [ISO 16], in Appendix C: "if early failures are considered separately and units are taken out of service before they arrive at wear-out, the

assumption of constant failure rate⁸ can be reasonable". Then the stress is put on the identification of the beginning of the wear-out period.

THEOREM 1.1. DRENICK'S THEOREM.– Under certain constraints, systems that are composed of a large quantity of non-exponentially distributed sub-components tend themselves toward being exponentially distributed.

Thus, for major equipment such as compressors, the failure rate should be constant.

The failure rate of electronic items is always considered as constant in reliability data books (section 9.4).

1.4.3. Oil and gas industry

The OREDA handbooks [ORE 15] assume that the failure rate is constant during the useful life period.

1.5. The bathtub curve

1.5.1. Meaning

The experience shows that often the failure rate has a time-profile like a bathtub cross-section as shown in Figure 1.6:

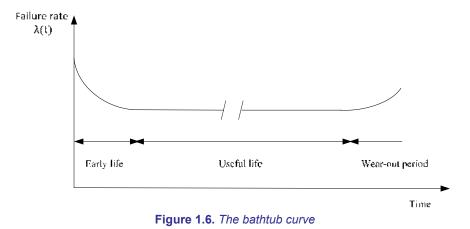
- An early life period (burn-in period, infant mortality period, wear-in period) where the number of failures is high but decreasing with time. This period is caused by:

- the end of the debugging of the items;
- the full cleaning of the piping (e.g. instrument air piping);
- the learning period of the maintenance crews and of the operators;
- the final adjustment of the items all together;
- etc.

- A so-called useful life period (constant failure rate period) where the failure rate is nearly constant.

⁸ Hazard rate in [ISO 16].

- A wear-out period (aging period) where the failure rate steadily increases however high the number of maintenance activities.



For non-repairable items (e.g. electronic items), the bathtub curve shows the behavior of thousands of items but for repairable items, the bathtub curve shows the behavior of a single item (or of several such items).

Formulae are available (section 9.3) for calculating the failure rate of small mechanical items (e.g. seals, bearings). Preventive maintenance tasks are performed (when the anticipated failure rate becomes too high) on a regular basis to prevent the failure from occcuring. Then the failure rate varies with time (saw tooth curve), as shown in Figure 1.7. On average, the failure rate can be considered as constant.

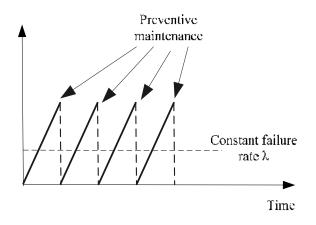


Figure 1.7. Effect of preventive maintenance on the failure rate

Overall, for repairable items, the "true" bathtub curve is likely to be as shown in Figure 1.6. Indeed, material fatigue, corrosion, contact wear, etc. are present right from the beginning and are "kept under control" by measures such as inspection, preventive maintenance, repairs, etc.

1.5.2. Useful life and mission life

There is no correlation between mission/service life and useful life. The typical MTTF of a sensor is of approximately 100 years but its service life is in the order of 20 years.

1.5.3. Validity

[MUN 83] provided the distribution of insured losses on gas turbines caused by errors in design and construction over a period of 35,000 hr. The early life lasts 5,000 hr and the frequency of losses is nearly constant after this period.

1.5.4. Oil and gas industry

According to the author of the book:

- For most of these components (valves, sensors, electric motors, etc.), the burn-in period lasts for approximately six months and the failure rate is multiplied by:

- a factor of approximately six for the first two months;
- a factor of approximately two for the remaining four months.

- For large rotating machines such as compressors and gas turbines, the burn-in period lasts for approximately 24 months and the failure rate is multiplied by:

- a factor of approximately four for the first two months;
- a factor of approximately three for months three to six;
- a factor of approximately two for months seven to 24.

Such high multiplying factors may not be applied to capital spare parts⁹.

⁹ Capital spare parts (insurance spare) are spare parts of high value and, most of the time, have a low failure rate and a long delivery time (e.g. the gas generator is a capital spare part for a gas turbine).

The wear-out period may be initiated after 15 years should the maintenance activities not be performed properly or if the units were not operated as they should be. However, the useful life of e.g. large electric motors can be as high as 30 years. Usually items such as vessels, separators can be kept running for a longer period.

Some experts consider that the high number of failures in the burn-in period is mainly caused by spurious failures.

EXAMPLE 1.3.– Oil pipelines installed in France after World War II are still in use, giving a useful life of at least 75 years.