# Dependability

## 1. Exponential distribution

The exponential probability distribution is used to model the behavior of a random variable $\boldsymbol{T}$ having the following properties

- $\boldsymbol{T}$, with $\boldsymbol{T} \geq 0$, represents the time interval between two events of some kind or between the first event and the origin of the reference frame
- the occurrence of one event does not affect the probability of occurrence of a second one
- the rate of events occurrence is constant
- the occurrence of successive events can not coincide in time.

Under these conditions, the probability density function $p(t)$ is expressed by

$$p(t) = \lambda e^{-\lambda t}, \text{ for } t \in [0, +\infty[ ,$$

where $\lambda$, called the *event rate*, is the number of events per unit of time.

The probability that an event occurs in the time interval $\boldsymbol{T}$ less than $t$, denoted by $P(\boldsymbol{T} < t)$ is determined by

$$P(\boldsymbol{T} < t) = \int_0^t p(u)\, du = \int_0^t \lambda e^{-\lambda u}\, du = 1 - e^{-\lambda t} .$$

It should be noted that

- $P(\boldsymbol{T} \geq t) = 1 - P(\boldsymbol{T} < t) = e^{-\lambda t}$ represents the probability that no event occurs in the time interval $\boldsymbol{T}$ less than $t$
- the exponential distribution is indeed *memoryless*, that is, the conditional probability that some event occurs in the time interval $[t_0, t_0 + \Delta t[$, provided no event has occurred in the time interval $[0, t_0[$, is equal to the probability that some event occurs in the time interval $[0, \Delta t[$

$$P(\boldsymbol{T} < t_0 + \Delta t \mid \boldsymbol{T} \geq t_0) = 1 - P(\boldsymbol{T} \geq t_0 + \Delta t \mid \boldsymbol{T} \geq t_0) =$$
$$= 1 - \frac{e^{-\lambda (t_0 + \Delta t)}}{e^{-\lambda t_0}} = 1 - e^{-\lambda \Delta t} = P(\boldsymbol{T} < \Delta t) .$$

The expected value, or mean, $E(\boldsymbol{T})$ and variance $Var(\boldsymbol{T})$ of a random variable $\boldsymbol{T}$ that satisfies an exponential distribution with event rate $\lambda$, are given by

$$E(\boldsymbol{T}) = \int_0^{+\infty} \lambda t\, e^{-\lambda t}\, dt = \frac{1}{\lambda}$$

$$Var(\boldsymbol{T}) = \int_0^{+\infty} \lambda [t - E(\boldsymbol{T})]^2 e^{-\lambda t}\, dt = \int_0^{+\infty} \lambda t^2 e^{-\lambda t}\, dt - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \ .$$

## 2. Reliability

In reliability theory, the exponential distribution is used to model the failures on the modules of a physical system. In this framework, the random variable $\boldsymbol{T}$ represents the time interval to module failures and the event rate $\lambda$ expresses the module *failure rate* or, when converted to failures per billion hours of module operation, the module *failures in time* (FIR). Its reciprocal, equal to the mean of the random variable $\boldsymbol{T}$, is called *mean time to failure* (MTTF).

In a system consisting of $N$ distinct modules whose individual failure behavior can be modeled by the exponential distribution and if one assumes that module failures are independent of one another, the probability that no failures occur in the time interval $[0, t_0[$, or that the system operates as specified, is expressed by

$$P(\boldsymbol{T}_1 \geq t_0, \boldsymbol{T}_2 \geq t_0, \cdots, \boldsymbol{T}_N \geq t_0) = \prod_{i=1}^{N} P(\boldsymbol{T}_i \geq t_0) =$$

$$= \prod_{i=1}^{N} e^{-\lambda_i t_0} = e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_N) t_0} \ ,$$

which means that under this conditions the *system*, or *overall*, *failure rate* is the sum of the failure rates of the constituent modules or, putting it in terms of the mean time to failure of the modules, that the *system*, or *overall*, *mean time to failure* is given by

$$\mathrm{MTTF}_{\mathrm{sys}} = \frac{1}{\sum_{i=1}^{N} \frac{1}{\mathrm{MTTF}_{\mathrm{n}}}} \ .$$

When a physical system fails to operate according to the specifications. The failed module has to be located and replaced by a properly functioning one. The time evolution of the system operation is thus described by a succession of alternate time intervals: *uptimes*, where there is proper operation, and *downtimes*, where the system is inactive or malfunctioning.


System time evolution

The system time evolution can now be modeled as a random variable $\boldsymbol{T}$ equal to the summation of two independent exponentially distributed random variables: $\boldsymbol{T}_{\text{sys}}$ that represents the time interval to system failures and $\boldsymbol{T}_{\text{rep}}$ that represents the time interval to system repair.

Let $p_{\text{sys}}(t)$ and $p_{\text{rep}}(t)$ denote the probability density functions of the random variables $\boldsymbol{T}_{\text{sys}}$ and $\boldsymbol{T}_{\text{rep}}$, respectively, then the probability density distribution $p(t)$ of the random variable $\boldsymbol{T}$, their sum, is determined by

$$\begin{aligned} p(t) &= \int_0^t p_{\text{sys}}(u)\, p_{\text{rep}}(t-u)\, du = \int_0^t \lambda_{\text{sys}} e^{-\lambda_{\text{sys}} u} \lambda_{\text{rep}} e^{-\lambda_{\text{rep}}(t-u)}\, du = \\ &= \lambda^2 t\, e^{-\lambda t} \qquad \Leftarrow \quad \lambda = \lambda_{\text{sys}} = \lambda_{\text{res}} \\ &= \frac{\lambda_{\text{sys}} \lambda_{\text{res}}}{\lambda_{\text{sys}} - \lambda_{\text{res}}} \cdot \left( e^{-\lambda_{\text{rep}} t} - e^{-\lambda_{\text{sys}} t} \right) \qquad \Leftarrow \quad \lambda_{\text{sys}} \neq \lambda_{\text{res}} \ , \end{aligned}$$

and its expected value $E(\boldsymbol{T})$ by

$$\begin{aligned} E(\boldsymbol{T}) &= \int_0^{+\infty} \lambda^2 t^2 e^{-\lambda t}\, dt = \frac{2}{\lambda} \qquad \Leftarrow \quad \lambda = \lambda_{\text{sys}} = \lambda_{\text{res}} \\ &= \int_0^{+\infty} \frac{\lambda_{\text{sys}} \lambda_{\text{res}}}{\lambda_{\text{sys}} - \lambda_{\text{res}}} \cdot t \left( e^{-\lambda_{\text{rep}} t} - e^{-\lambda_{\text{sys}} t} \right) dt = \frac{1}{\lambda_{\text{sys}}} + \frac{1}{\lambda_{\text{res}}} \qquad \Leftarrow \quad \lambda_{\text{sys}} \neq \lambda_{\text{res}} \ . \end{aligned}$$

Calling the mean of the random variable $\boldsymbol{T}_{\text{rep}}$ *mean time to repair* (MTTR) and the mean of the random variable $\boldsymbol{T}$ *mean time between failures* (MTBF), one gets the following relation that can be applied both to the whole system or to a single module

$$\begin{aligned} \text{MTBF}_{\text{sys}} &= \text{MTTF}_{\text{sys}} + \text{MTTR}_{\text{sys}} \\ \text{MTBF}_{\text{mod}} &= \text{MTTF}_{\text{mod}} + \text{MTTR}_{\text{mod}} \ . \end{aligned}$$

As a measure of how reliable a system, or a module, is, one defines its *availability*, a figure of merit that describes the proportion of its uptime to its downtime during normal operation.

For nonredundant systems with repair, *availability* is expressed by the ratio

$$\begin{aligned} \textit{availability}_{\text{sys}} &= \frac{\text{MTTF}_{\text{sys}}}{\text{MTBF}_{\text{sys}}} = \frac{\text{MTTF}_{\text{sys}}}{\text{MTTF}_{\text{sys}} + \text{MTTR}_{\text{sys}}} \\ \textit{availability}_{\text{mod}} &= \frac{\text{MTTF}_{\text{mod}}}{\text{MTBF}_{\text{mod}}} = \frac{\text{MTTF}_{\text{mod}}}{\text{MTTF}_{\text{mod}} + \text{MTTR}_{\text{mod}}} \ . \end{aligned}$$

Resource redundancy is the most popular method to deal with availability improvements. The rationale is to include in the design of the system extra modules that are prepared to replace modules proved to be defective on-line. In this way, the MTTF of the concerned modules is extended and so is the overall MTTF.

Suppose that in the design of a system a given module is replicated so that its availability is increased. What is the improvement gain that was obtained?

The reasoning rests in the following premises

- in order for the module to fail, both replicas must have a failure
- the module availability, when two modules are operative, is given by

$$availability_{2\,\mathrm{mod}} = \frac{\frac{\mathrm{MTTF}_{\mathrm{mod}}}{2}}{\frac{\mathrm{MTTF}_{\mathrm{mod}}}{2} + \mathrm{MTTR}_{\mathrm{mod}}}$$

- the module availability, when only one module is operative, is given by

$$availability_{1\,\mathrm{mod}} = \frac{\mathrm{MTTF}_{\mathrm{mod}}}{\mathrm{MTTF}_{\mathrm{mod}} + \mathrm{MTTR}_{\mathrm{mod}}}$$

- the module availability of the redundant configuration is given by

$$availability_{\mathrm{redund}\,2} = 1 - (1 - availability_{2\,\mathrm{mod}}) \cdot (1 - availability_{1\,\mathrm{mod}}) \ .$$

This result can be expressed in terms of the MTTF and MTTR of single modules if the substitutions are made to the expression above, yielding

$$availability_{\mathrm{redund}\,2} = 1 - \left(1 - \frac{\frac{\mathrm{MTTF}_{\mathrm{mod}}}{2}}{\frac{\mathrm{MTTF}_{\mathrm{mod}}}{2} + \mathrm{MTTR}_{\mathrm{mod}}}\right) \cdot \left(1 - \frac{\mathrm{MTTF}_{\mathrm{mod}}}{\mathrm{MTTF}_{\mathrm{mod}} + \mathrm{MTTR}_{\mathrm{mod}}}\right) =$$

$$= \frac{\frac{\mathrm{MTTF}_{\mathrm{mod}}^2}{2} + \frac{3\,\mathrm{MTTF}_{\mathrm{mod}} \cdot \mathrm{MTTR}_{\mathrm{mod}}}{2}}{\frac{\mathrm{MTTF}_{\mathrm{mod}}^2}{2} + \frac{3\,\mathrm{MTTF}_{\mathrm{mod}} \cdot \mathrm{MTTR}_{\mathrm{mod}}}{2} + \mathrm{MTTR}_{\mathrm{mod}}^2} =$$

$$= \frac{\frac{\mathrm{MTTF}_{\mathrm{mod}}^2}{2\,\mathrm{MTTR}_{\mathrm{mod}}} + \frac{3}{2} \cdot \mathrm{MTTF}_{\mathrm{mod}}}{\frac{\mathrm{MTTF}_{\mathrm{mod}}^2}{2\,\mathrm{MTTR}_{\mathrm{mod}}} + \frac{3}{2} \cdot \mathrm{MTTF}_{\mathrm{mod}} + \mathrm{MTTR}_{\mathrm{mod}}} \ ,$$

which means that

$$\mathrm{MTTF}_{\mathrm{mod\text{-}redund\,2}} = \mathrm{MTTF}_{\mathrm{mod}} \cdot \frac{\mathrm{MTTF}_{\mathrm{mod}} + 3\,\mathrm{MTTR}_{\mathrm{mod}}}{2\,\mathrm{MTTR}_{\mathrm{mod}}} \simeq$$

$$\simeq \frac{\frac{\mathrm{MTTF}_{\mathrm{mod}}}{2}}{\frac{\mathrm{MTTR}_{\mathrm{mod}}}{\mathrm{MTTF}_{\mathrm{mod}}}} \ , \text{ if } \mathrm{MTTF}_{\mathrm{mod}} >> 3\,\mathrm{MTTR}_{\mathrm{mod}} \ .$$

In the general case, for a redundancy degree of $K \in \mathbb{N}$, one gets

$$availability_{\mathrm{redund}\ K} = 1 - \prod_{k=1}^{K} (1 - availability_{k\ \mathrm{mod}}) \ ,$$

or in terms of the MTTF and MTTR of single modules

$$\text{MTTF}_{\text{mod - redund K}} = \text{MTTF}_{\text{mod}}^{K-1} \cdot \frac{\text{MTTF}_{\text{mod}} + O\left[K(K+1)/2 \cdot \text{MTTR}_{\text{mod}}\right]}{K! \cdot \text{MTTR}_{\text{mod}}^{K-1}} \simeq$$

$$\simeq \frac{\frac{\text{MTTF}_{\text{mod}}}{K!}}{\left(\frac{\text{MTTR}_{\text{mod}}}{\text{MTTF}_{\text{mod}}}\right)^{K-1}} \quad , \text{ if } \text{MTTF}_{\text{mod}} >> O\left[K(K+1)/2 \cdot \text{MTTR}_{\text{mod}}\right] .$$

For large values of $K$ the above approximation is often not valid. A more accurate way to determine the module MTTF with $K$ redundancy is through the $K$ redundancy availability

$$\text{MTTF}_{\text{mod - redund K}} = \frac{\text{MTTR}_{\text{mod}} \cdot \textit{availability}_{\text{redund K}}}{\left(1 - \textit{availability}_{\text{redund K}}\right)} .$$

Finally, one needs to know what was the reliability improvement gained by a system where the availability of some constituent module was increased in some way. The improvement gain can be computed through the application of Amdahl's Law

$$\text{reliability improvement} = \frac{1}{\text{frac sys non improv failure rate} + \frac{\text{frac sys improv failure rate}}{\left(\frac{\text{MTTF}_{\text{mod - improv}}}{\text{MTTF}_{\text{mod - non improv}}}\right)}} .$$