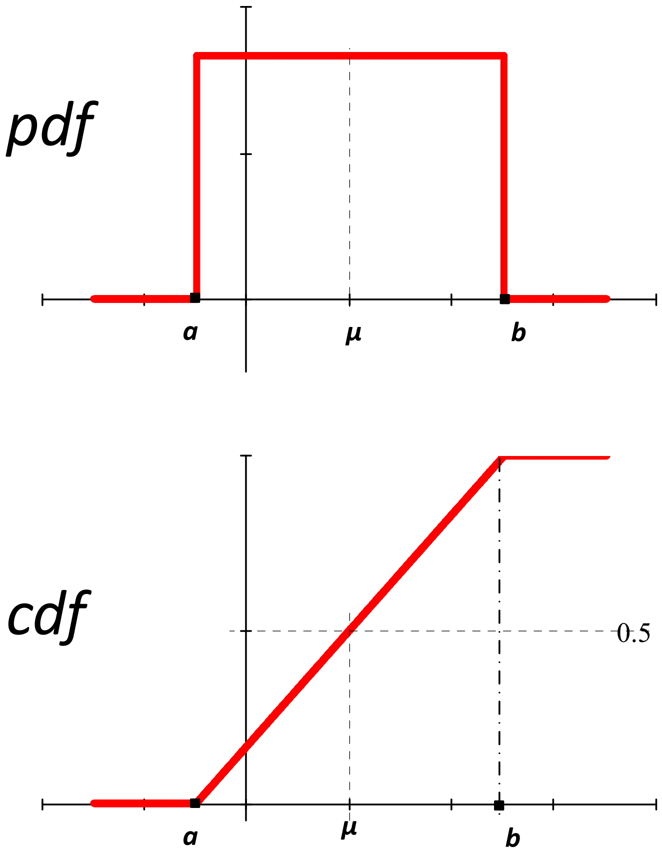
SELECTED CONTINUOUS PROBABILITY DISTRIBUTIONS

In probability theory and statistics, the *continuous uniform distribution* is a family of symmetric probability distributions such that for each member of the family, all intervals of the same length on the distribution's support are equally probable.

The ***support*** is defined by the two parameters, *a* and *b*, which are its minimum and maximum values.

Selected Continuous Probability Distributions



$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & if\ x \in [a,b] \\ 0, & otherwise \end{cases}$$

$$F_X(x) = \begin{cases} 0, & if\ x < a \\ \dfrac{x-a}{b-a}, & if\ x \in [a,b] \\ 1, & if\ x > b \end{cases}$$

$$E[X] = \frac{a+b}{2} \qquad Var(X) = \frac{(b-a)^2}{12}$$

$$Sk(X) = 0 \qquad Ex(X) = -1.2$$

Selected Continuous Probability Distributions

In probability theory, the *normal* (or *Gaussian*) distribution is a very common continuous probability distribution.

Normal distributions are important in statistics and are often used in the natural and social sciences to represent real-valued random variables whose distributions are not known.

A random variable with a Gaussian distribution is said to be normally distributed.

The pdf of the normal distribution is

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

where μ is a location parameter, and σ – shape parameter.

The normal distribution is sometimes informally called the bell curve.

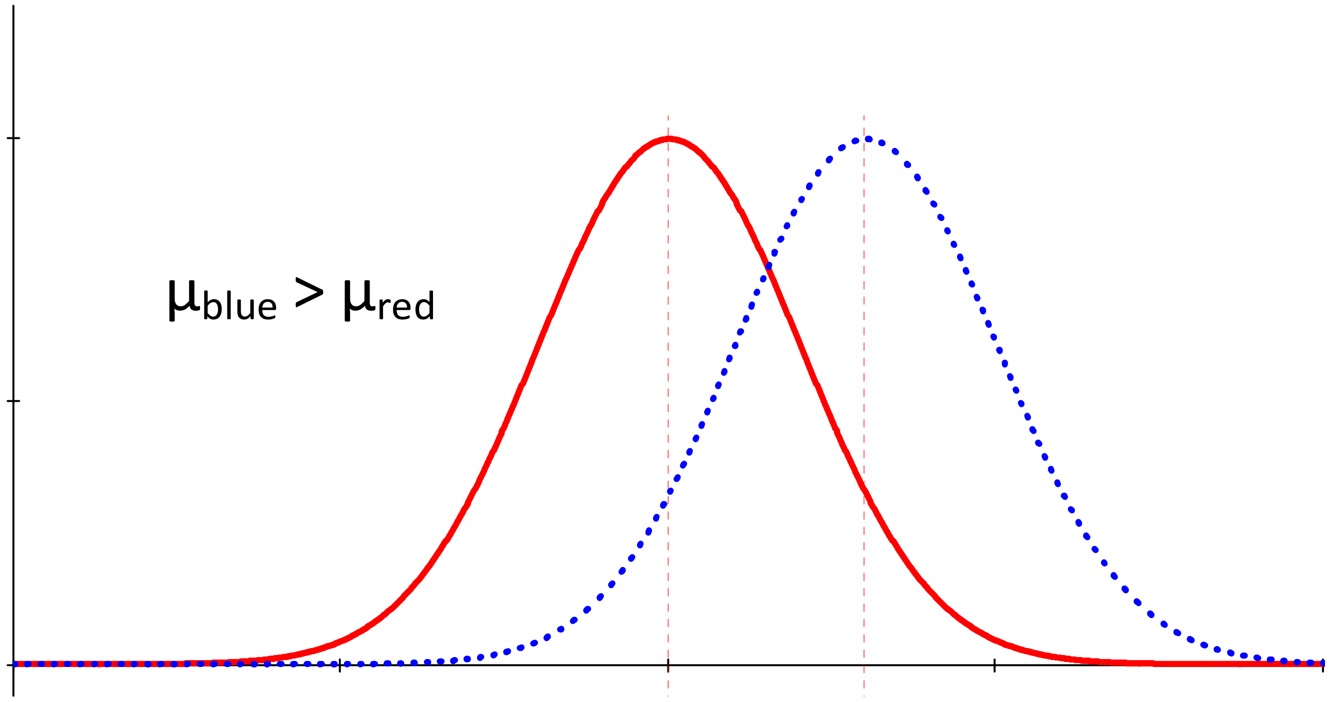
However, many other distributions are bell-shaped (e.g. *Cauchy*, *logistic*, *Student's t*, etc.).

The parameter μ is actually equal to the mean (and median) of X, and σ is the standard deviation of X.

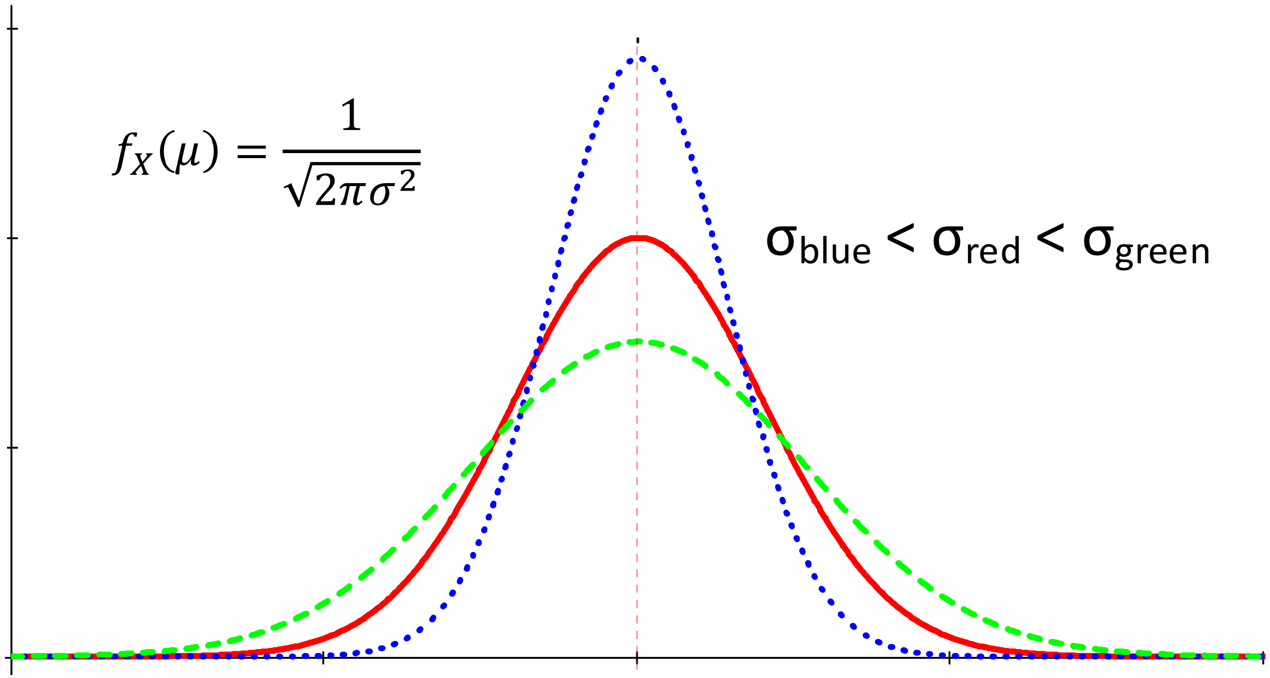
Furthermore, the standard deviation of a r.v. is the square root of the variance of this variable. Therefore, in the case of the normal distribution, σ^2 is its variance.

Since the *pdf* of normal distribution is symmetric, the skewness is equal to 0.

Excess kurtosis is equal to 0 as well.

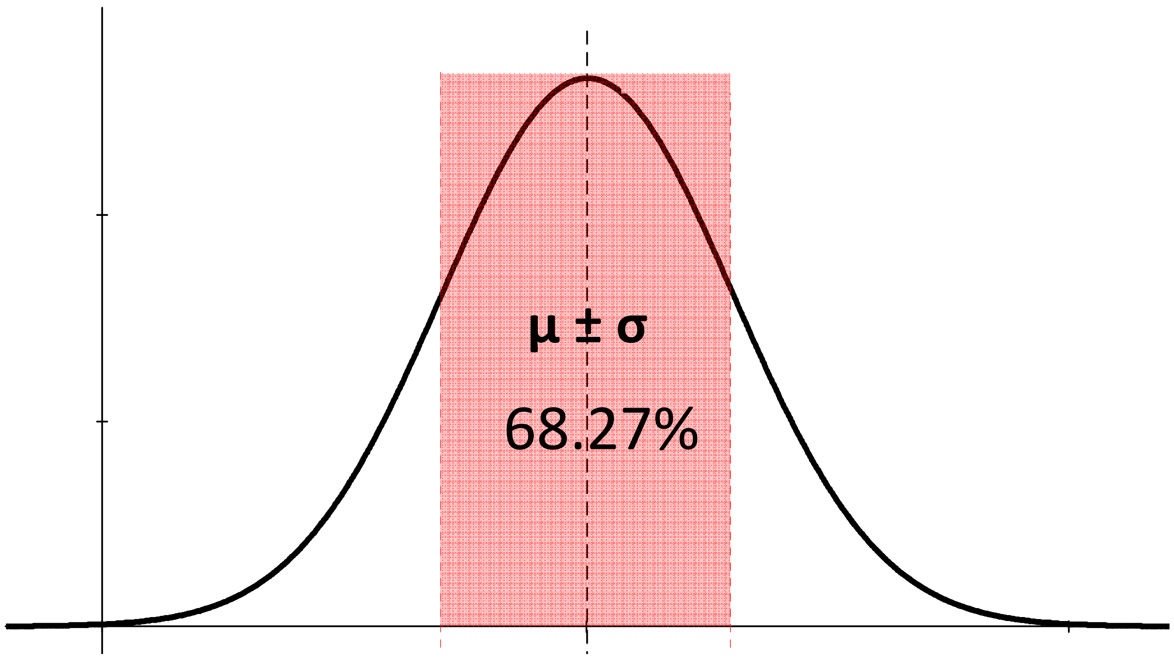


Location parameter μ defines the position (locus) of the distribution's center

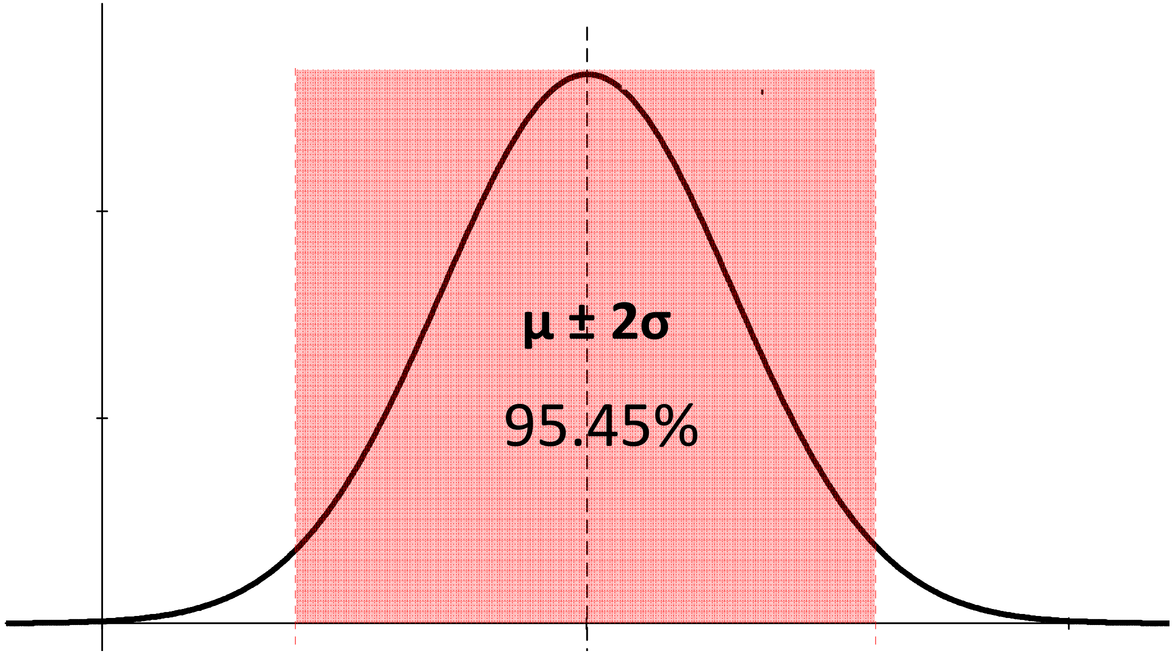


Shape parameter σ affects the spread of the curve, and its height

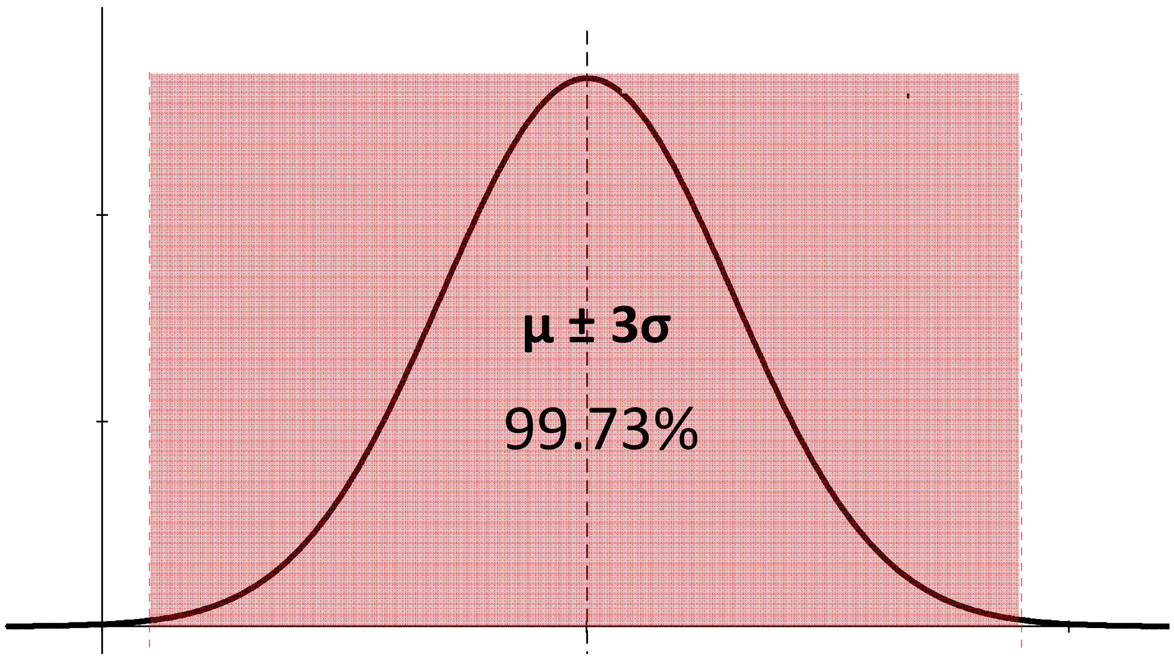
Selected Continuous Probability Distributions

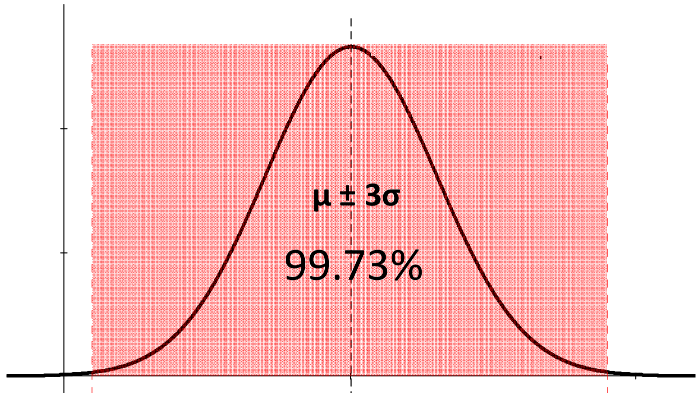


Selected Continuous Probability Distributions



Selected Continuous Probability Distributions



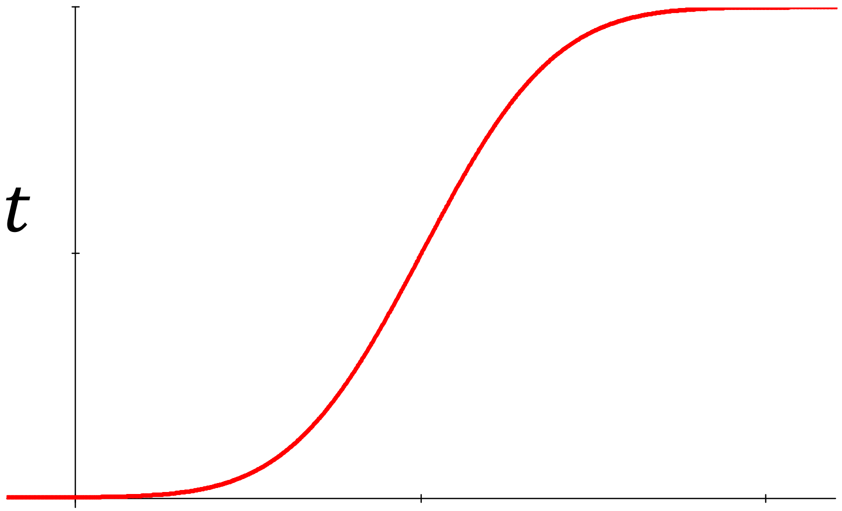


In the empirical sciences the so-called *three-sigma rule of thumb* expresses a conventional heuristic that nearly all values are taken to lie within three standard deviations of the mean, and thus it is empirically useful to treat 99.7% probability as near certainty.

In the social sciences, a result may be considered "significant" if its confidence level is of the order of a *two-sigma effect* (95%), while in particle physics, there is a convention of a *five-sigma effect* (99.99994%) being required to qualify as a discovery.

The *cdf* of a normal distribution cannot be expressed in terms of the elementary functions but can be computed at a given value *x*:

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$



The ***standard normal distribution*** is the special case of a normal distribution. It has $\mu = 0$ and $\sigma = 1$.

cdf: $$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

pdf: $$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

The values of *Φ(x)* are often tabulated.

Consider normally distributed r.v. X with parameters μ and σ. If we define the random variable

$$Z = \frac{X - \mu}{\sigma}$$

then Z is distributed according to the standard normal distribution.

Using tabulated values of *Φ(x)*, we can obtain values for any normal distribution with arbitrary μ and σ.

Nonnegative random variable X has the ***exponential distribution*** with parameter $\lambda > 0$ if it has the *pdf*

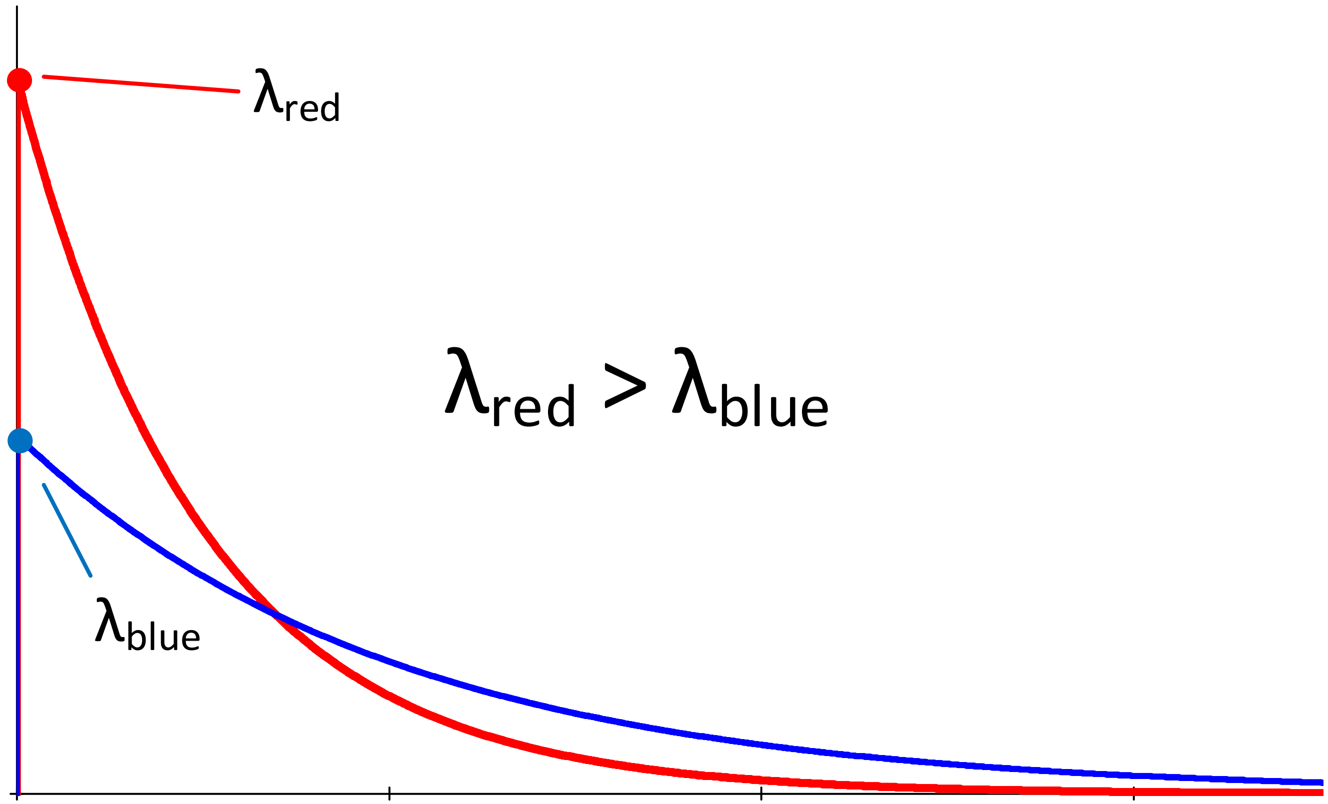
$$f_X(x) = \lambda e^{-\lambda x}, x \geq 0$$

and *cdf*

$$F_X(x) = 1 - e^{-\lambda x}.$$

$$E[X] = \frac{1}{\lambda} \qquad Var(X) = \frac{1}{\lambda^2} \qquad Sk(X) = 2 \qquad Ex(X) = 6$$

Selected Continuous Probability Distributions



The ***Weibull distribution*** is one of the most widely used lifetime distributions in reliability engineering. It is a versatile distribution that can take on the characteristics of other types of distributions, based on the value of the shape parameter.

The Weibull distribution is a continuous probability distribution. It is named after Swedish mathematician Waloddi Weibull, who described it in detail in 1951.

Nonnegative random variable *X* has the Weibull distribution if it has the *pdf*

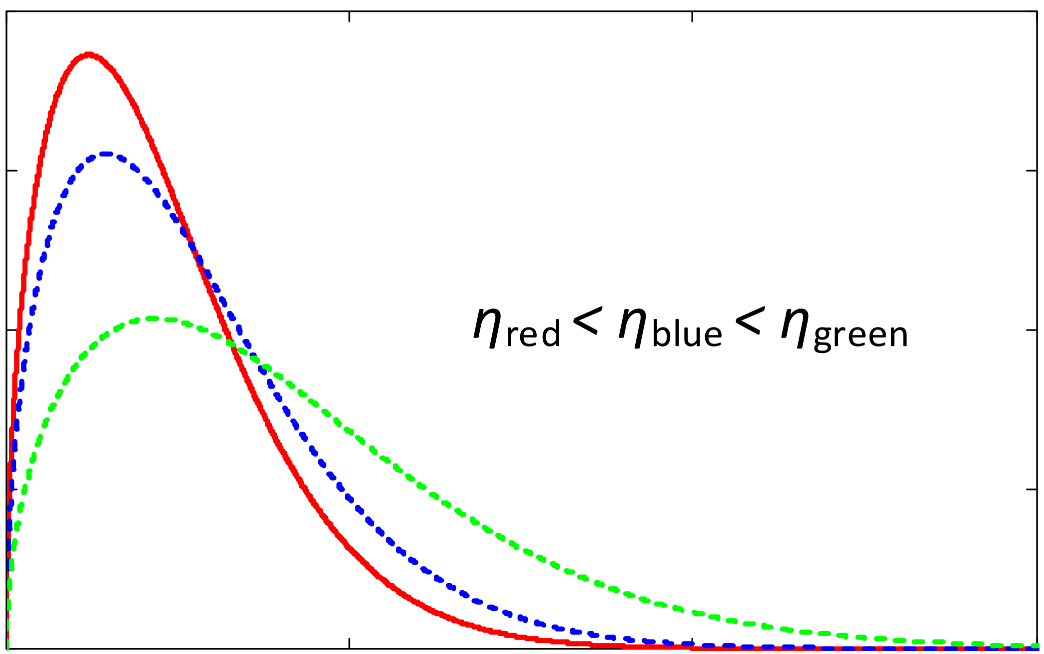
$$f_X(x) = \begin{cases} \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^{\beta}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

and *cdf*

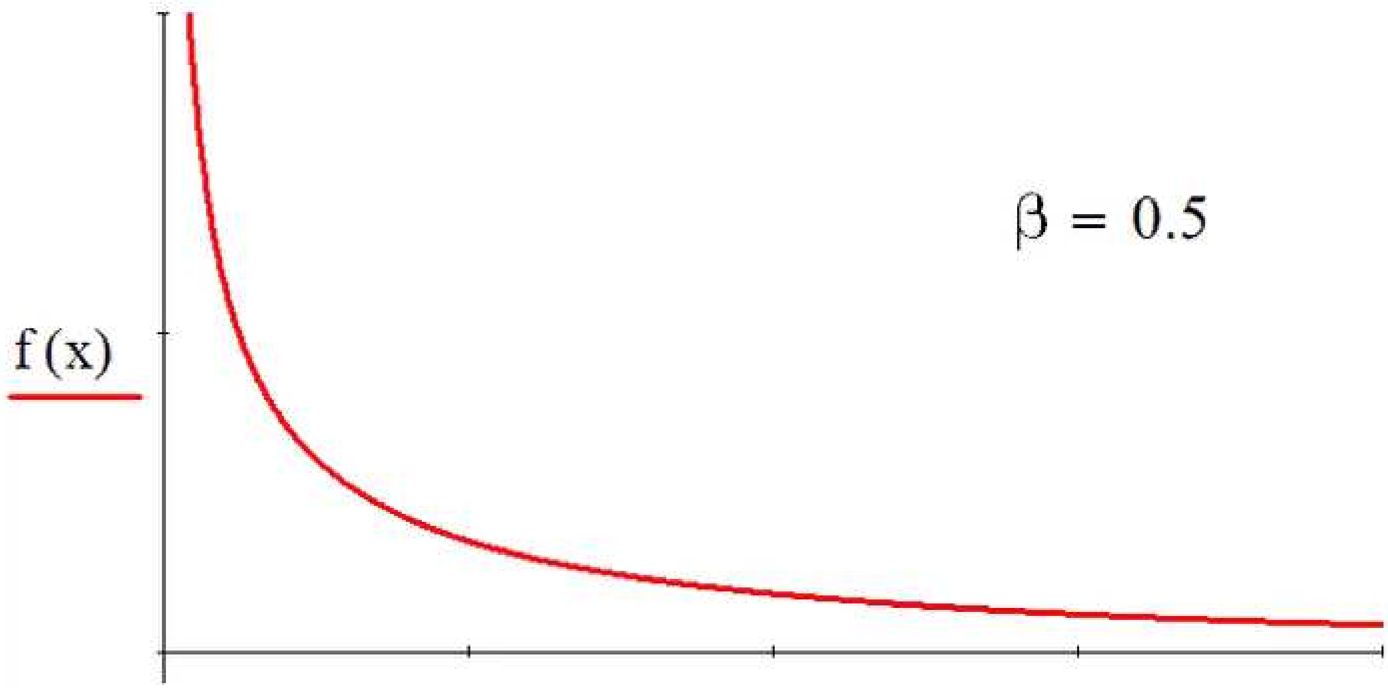
$$F_X(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\eta}\right)^{\beta}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where $\eta > 0$ is a *scale parameter*, and $\beta > 0$ is a *shape parameter*.

As the scale parameter η increases, Weibull *pdf* stretches along the *x*-axis, while its height decreases to maintain the area under the curve constant:



The shape parameter β affects Weibull pdf in a more dramatic way:



The mean and variance of the Weibull distribution are obtained by the following equations:

$$E[X] = \eta\Gamma\left(1+\frac{1}{\beta}\right) \qquad Var[X] = \eta^2\left[\Gamma\left(1+\frac{2}{\beta}\right)-\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2\right]$$

where *Γ(x)* – is the gamma function:

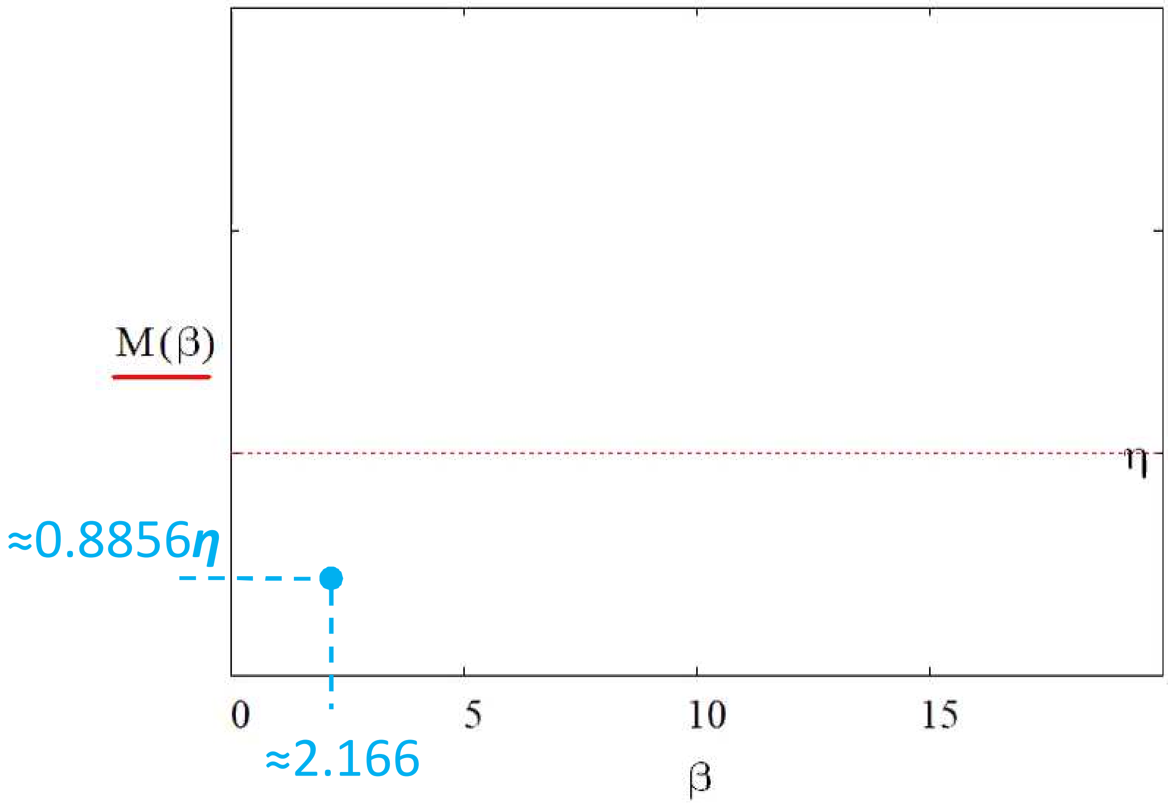
$$\Gamma(x) = \int_0^{\infty} s^{x-1}e^{-s}\,ds$$

$$E[X] = \eta\Gamma\left(1 + \frac{1}{\beta}\right) \qquad Var[X] = \eta^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2\right]$$

As you can see, the mean is directly proportional to the value of η, and the variance – to its square:

$$E[X] \propto \eta \qquad Var[X] \propto \eta^2$$

The effect of β on the mean:



We can specify two special cases of the Weibull distribution:

- if $\beta = 1$, Weibull(η, 1) → Exponential;
- if $\beta = 2$, Weibull(η, 2) → *Rayleigh*.

$$E\left[X_{Rayleigh}\right] = \eta \frac{\sqrt{\pi}}{2}$$

$$Var\left[X_{Rayleigh}\right] = \eta^2 \left(1 - \frac{\pi}{4}\right)$$

BASIC CONCEPTS
OF
RELIABILITY THEORY

Reliability is a characteristic of the item, expressed by the *probability* that it will perform its *required function* under *given conditions* for a *stated time interval*. It is generally designated by *R*.

From a *qualitative* point of view, reliability can be defined as the *ability of the item to remain functional*.

Quantitatively, reliability specifies the *probability that no operational interruptions* will occur during a stated time interval.

To make sense, a numerical statement of reliability (e. g. $R = 0.9$) must be accompanied by the definition of the *required function*, *the operating conditions*, and *the mission duration*.

In general, it is also important to know whether or not the item can be considered new when the mission starts.

An *item* is a functional or structural *unit* of arbitrary complexity (e.g. component, assembly, equipment, subsystem, system) that can be considered as an entity for investigations.

It may consist of *hardware*, *software*, or *both*, and may also include *human resources*.

Often, ideal human aspects and logistic support are assumed, even if (for simplicity) the term *system* is used instead of *technical system*.

Often the mission duration is considered as a parameter *t*, the *reliability function* is then denoted by *R(t)*.

R(t) is the probability that no failure at item level will occur in the interval (0, *t*].

For the systems of *physical nature* reliability function is a *decreasing (non-increasing) function*.

Reliability function is sometimes referred to as a *survival function*, *S(t)*.

A distinction between *predicted* and *estimated* reliability is important:

- the first is calculated on the basis of the item's reliability structure and the reliability characteristics of its components;
- the second is obtained from a statistical evaluation of *reliability tests* or from *field data*.

The concept of reliability can be extended to processes and services, although human aspects can lead to modeling difficulties.

A *failure* occurs when the item stops performing its required function. As simple as this definition is, it can become difficult to apply it to complex items.

The *failure-free time* (hereafter used as a synonym for *failure-free operating time*) is generally a random variable.

A general assumption in investigating failure-free times is that at $t = 0$ the item is free of *defects* and *systematic failures*.

Failures can be classified as *sudden* or *gradual*.

We will consider time at which failure occurs (*failure time*) as a nonnegative continuous r.v. with particular *cdf F(t)* and *pdf f(t)*.

In reliability theory *F(t)* is referred to as probability of failure or ***failure probability function***.

Similarly, *f(t)* is often called *failure density function* or *instantaneous failure rate*.

The failure probability function is sometimes referred to as a *life distribution*.

Basic Concepts of Reliability Theory

We can infer the following formulae:

	$R(t)$	$F(t)$	$f(t)$
$R(t) =$		$1 - F(t)$	$\int_t^\infty f(\tau)\,d\tau$
$F(t) =$	$1 - R(t)$		$\int_0^t f(\tau)\,d\tau$
$f(t) =$	$-\frac{dR(t)}{dt}$	$\frac{dF(t)}{dt}$	

A concept closely-related but different to instantaneous failure rate $f(t)$ is the hazard rate (or hazard function), $h(t)$.

It is defined as the proportional failure rate of the items still functioning at time t (as opposed to $f(t)$, which is the expressed as a proportion of the initial number of systems).

By definition

$$h(t) = \frac{f(t)}{R(t)}.$$

Omitting the derivation, it is possible to define hazard rate in terms of the other reliability measures:

$$h(t) = -\frac{R'(t)}{R(t)} = \frac{F'(t)}{1 - F(t)} = \frac{f(t)}{\int_t^{\infty} f(\tau)d\tau}$$

The reverse formulae could also be of use:

$$R(t) = e^{-\int_0^t h(\tau)d\tau}$$

$$F(t) = 1 - e^{-\int_0^t h(\tau)d\tau}$$

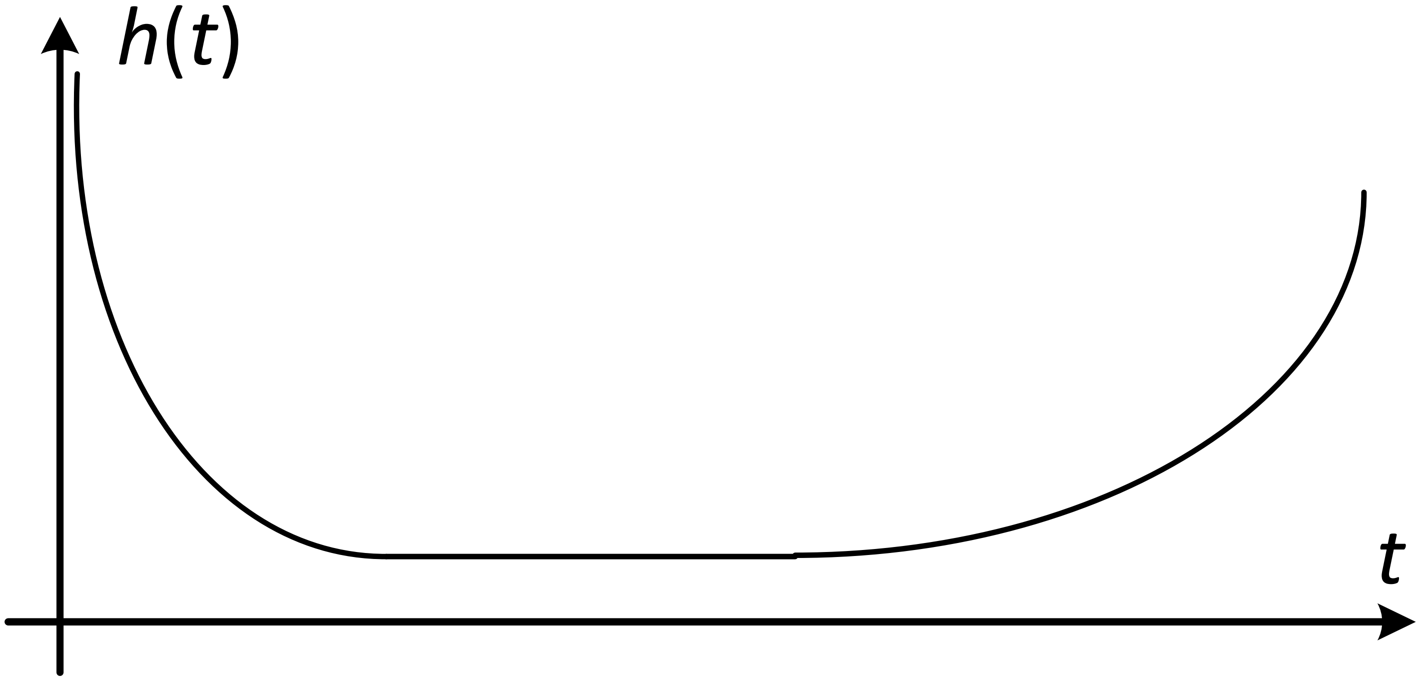
$$f(t) = -\left[e^{-\int_0^t h(\tau)d\tau}\right]'$$

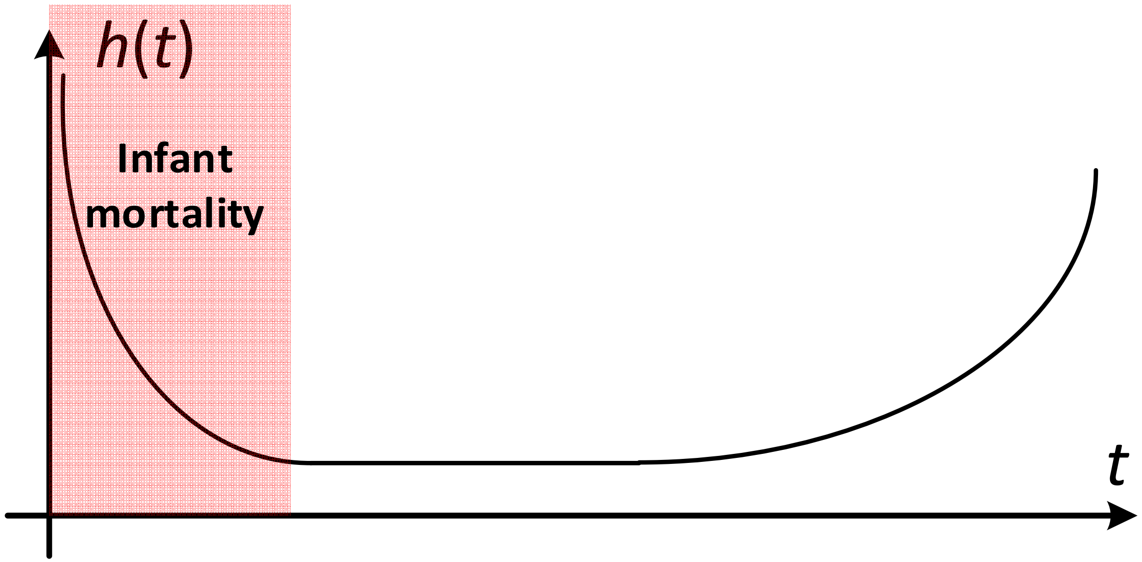
Basic Concepts of Reliability Theory

Now we can update the table:

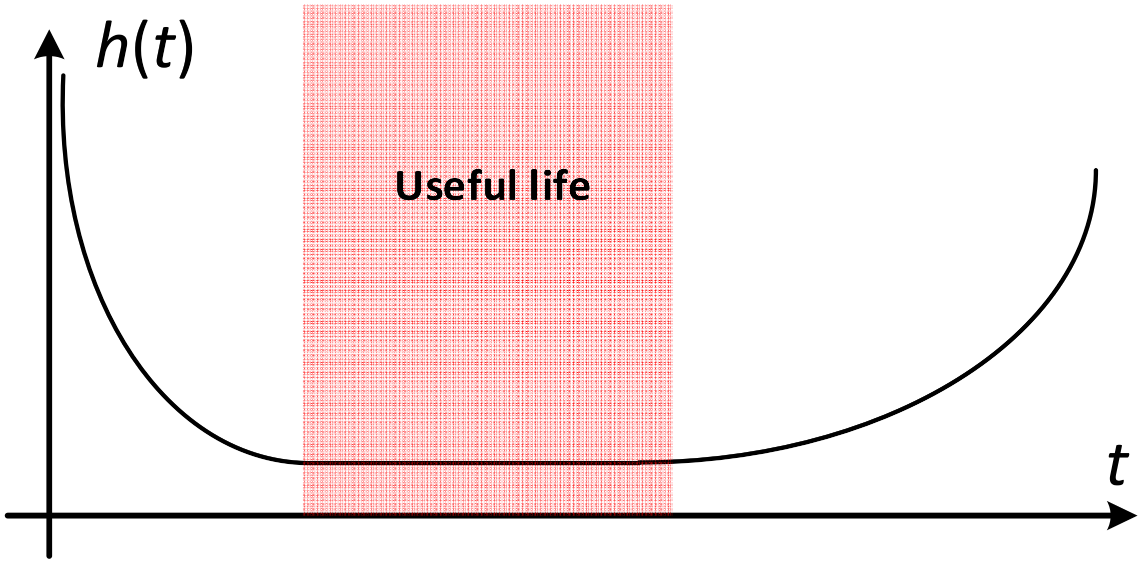
	$R(t)$	$F(t)$	$f(t)$	$h(t)$
$R(t) =$		$1 - F(t)$	$\int_t^{\infty} f(\tau)\, d\tau$	$e^{-\int_0^t h(\tau)d\tau}$
$F(t) =$	$1 - R(t)$		$\int_0^{t} f(\tau)\, d\tau$	$1 - e^{-\int_0^t h(\tau)d\tau}$
$f(t) =$	$-\frac{dR(t)}{dt}$	$\frac{dF(t)}{dt}$		$-\left[e^{-\int_0^t h(\tau)d\tau}\right]'$
$h(t) =$	$-\frac{R'(t)}{R(t)}$	$\frac{F'(t)}{1 - F(t)}$	$\frac{f(t)}{\int_t^{\infty} f(\tau)d\tau}$	

The hazard rate of a large population of statistically identical and independent items often exhibits a typical ***bathtub curve***:

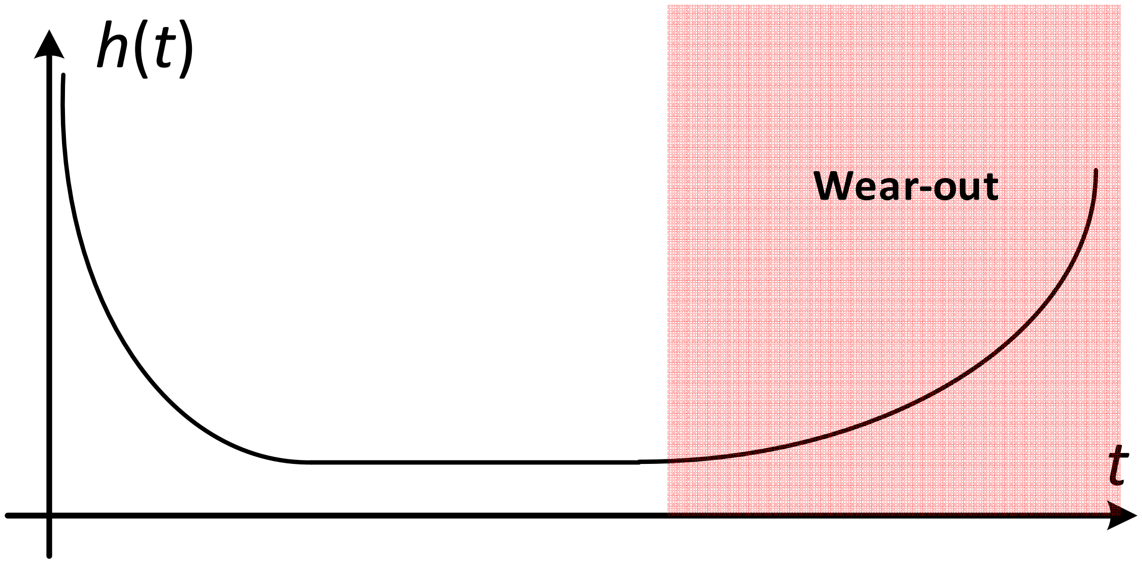




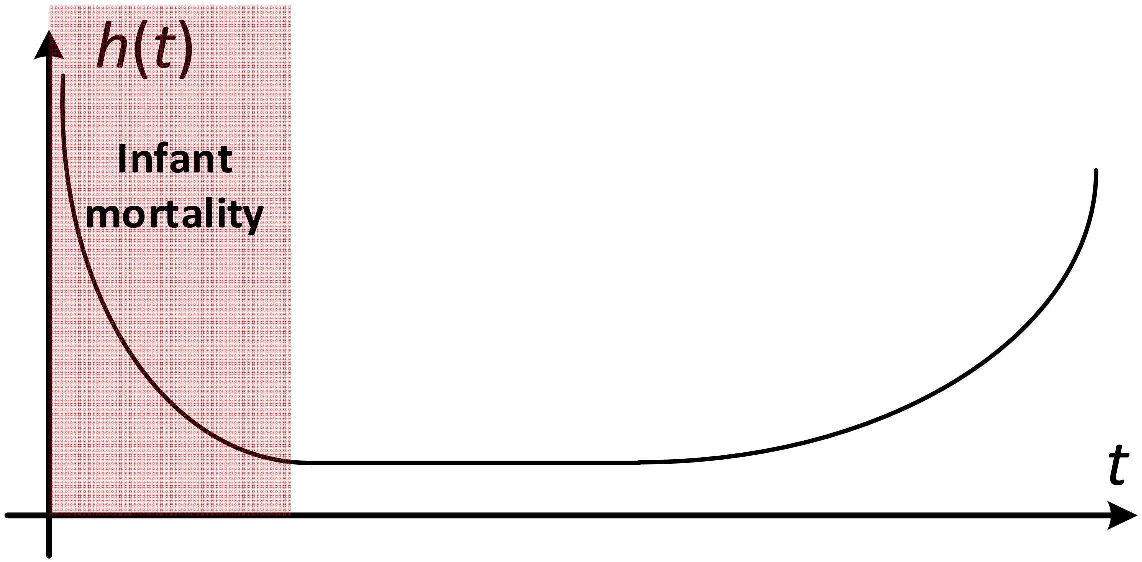
As can be seen from this plot, many products will begin their lives with a higher hazard rate (which can be due to manufacturing defects, poor workmanship, poor quality control of incoming parts, etc.) and exhibit a *decreasing hazard rate*.



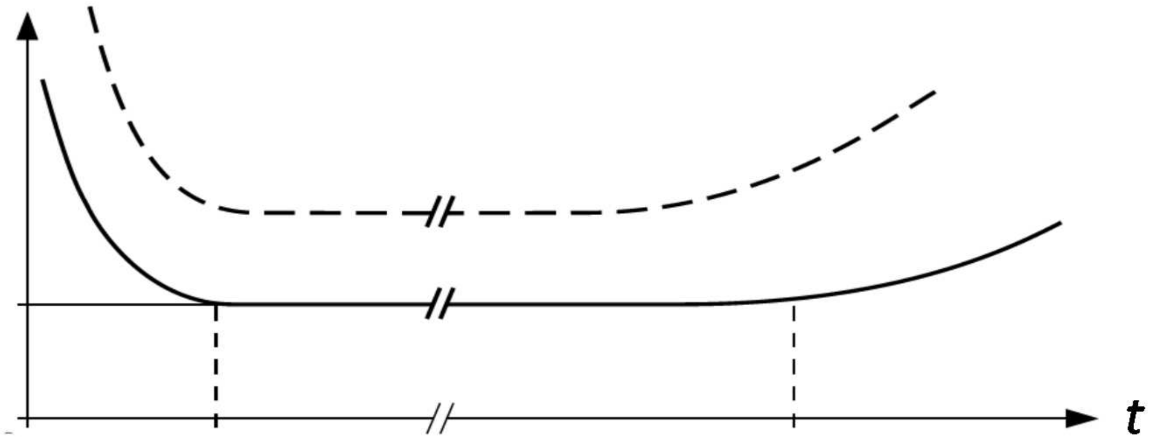
During the *useful life* period the item's hazard rate remains *nearly constant* with respect to time. Some of the main reasons for the occurrence of failures during this period are undetectable defects, higher random stress than expected, abuse, low safety factors, and human error.



As the products experience more use and wear, the *hazard rate begins to rise* as the population begins to experience failures related to aging, wear-out, fatigue, etc.



It should be obvious that it would be best to ship a product at the beginning of the useful life period, rather than right off the production line; thus preventing the customer from experiencing early failures. This practice is what is commonly referred to as "*burn-in*", and is frequently performed for electronic components.



The hazard rate strongly depends upon the item's operating conditions.

Typical figures for *h(t)* are 10^{-10} to 10^{-7} h^{-1} for electronic components at 40°C, doubling for a temperature increase of 10 to 20°C.

Since time to failure, *T*, is a random variable, its mean is an important and the most obvious reliability measure.

In reliability theory it is called *"mean time to failure" (MTTF)*, and, by definition, it can be computed as

$$E[T] = MTTF = \int_0^\infty t f(t) dt .$$

Another useful formula for the MTTF is as follows:

$$MTTF = \int_0^\infty R(t)dt \, .$$

Note that, although

$$MTTF = \int_0^\infty R(t)dt = \int_0^\infty tf(t)dt \, ,$$

$$\boldsymbol{R(t) \neq tf(t) \, !}$$

RELIABILITY MODELS (LIFE DISTRIBUTIONS)

Reliability Models (Life Distributions)

When specifying certain probability distribution as a reliability model for an element, a system or a subsystem, we assume that failure time of such items is a random variable distributed according to this particular distribution.

Generally, *any distribution* can be regarded as a life distribution.

Often, it is assumed that failure probability at $t = 0$ is equal to 0, so, failure time is considered to be a *non-negative continuous r.v.*

However, even a distribution *F(t)* with an infinite support (-∞; +∞) can be employed as a reliability model in cases when

- $F(t = 0)$ is a negligible quantity;
- $F(t = 0) > 0$ a priori, i.e. the assumption on $F(t = 0) = 0$ is not valid.

When failure occurs after random number of on-off cycles, *F(t)* could be discrete.

Reliability Models (Life Distributions)

The most popular reliability models are exponential and Weibull distributions.

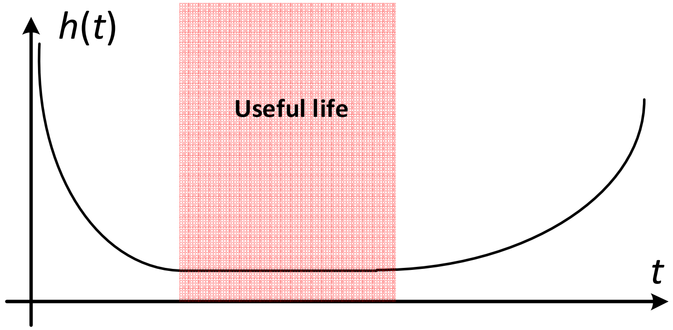
As most of characteristics of these distributions were addressed in preceding sections, we will consider only hazard rate according to these two models.

Exponential Reliability Model (ERM)

Given that $h(t) = \dfrac{f(t)}{R(t)}$ and
$$f(t) = \lambda e^{-\lambda t}$$
$$R(t) = 1 - F(t) = e^{-\lambda t}$$

we obtain *h(t)* = *λ* = *const*.

Therefore, ERM is valid for items in their useful life period.



By using ERM we assume that the manufacturer carried out complete burn-in and that the item will be in use for a time interval not extending into wear-out period.

Weibull Reliability Model (WRM)

Given that $h(t) = \frac{f(t)}{R(t)}$ and

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad R(t) = 1 - F(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

we get

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}$$

The obtained equation offers various distinct shapes of the hazard rate.

First, if $\beta = 1$, Weibull distribution reduces to the exponential:

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\Bigg|_{\beta=1} = \frac{1}{\eta} = const.$$

Second, if $\beta < 1$, the hazard rate of Weibull distribution is a convex decreasing function.

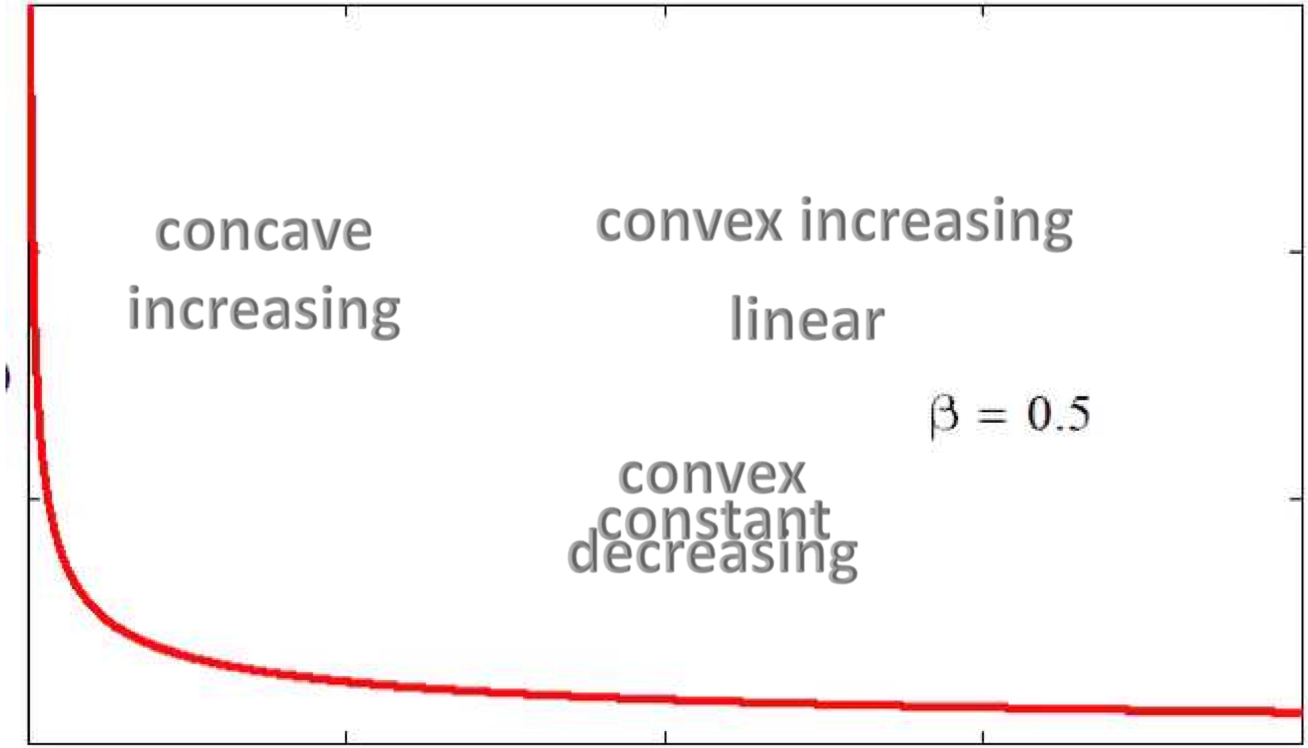
Third, if $1 < \beta < 2$, the hazard rate of Weibull distribution is a concave increasing function.

Next, if $\beta = 2$, Weibull distribution reduces to the Rayleigh distribution, and $h(t)$ becomes linear:

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\Bigg|_{\beta = 2} = \frac{2}{\eta^2}t.$$

Finally, if $\beta > 2$, the hazard rate of Weibull distribution is a convex increasing function.

The following clip illustrates the shapes of the Weibull hazard rate:



Reliability Models (Life Distributions)

Though Weibull distribution can’t provide bathtub shape for the hazard rate, it is widely applied in reliability analysis due to its relative simplicity and flexibility.

There were suggested various compound distributions that demonstrate this particular shape of the hazard rate.

Some of them are based on either exponential or Weibull distributions.

Consider so called Bi-Weibull distribution with a reliability function defined as

$$R(t) = e^{-\left(\frac{t}{\eta 1}\right)^{\beta 1} - \left(\frac{t}{\eta 2}\right)^{\beta 2}}.$$

Adjusting its parameters we can obtain the following shape of the hazard rate function: