Safety and Reliability of Embedded Systems

(Sicherheit und Zuverlässigkeit eingebetteter Systeme)

Stochastic Reliability Analysis

Content

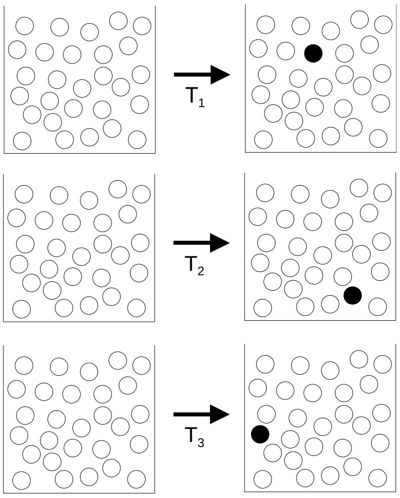
- Definition of Reliability
- Hardware- vs. Software Reliability
- Tool Assisted Reliability Modeling
- Descriptions of Failures over Time
- Reliability Modeling
- Examples of Distribution Functions
  - The exponential distribution
  - The Weibull distribution
  - The Poisson distribution
- Musa's Execution Time Model
- Determination of Model Parameters
- Selection of Models Based on Failure Observation

Definition

- ☐ Reliability
  - Part of the quality with regard to the behavior of an entity during or after given time intervals under given application conditions (translated from DIN 40041)
  - The property of an entity to fulfill its reliability requirements during or after a given time span under given application conditions (translated from DIN ISO 9000 Teil 4)
  - A measure for the capability of an item under consideration to remain functional, expressed by the probability that the demanded function is executed without failure under given conditions during a given time span (Birolini)

Hardware- vs. Software Reliability

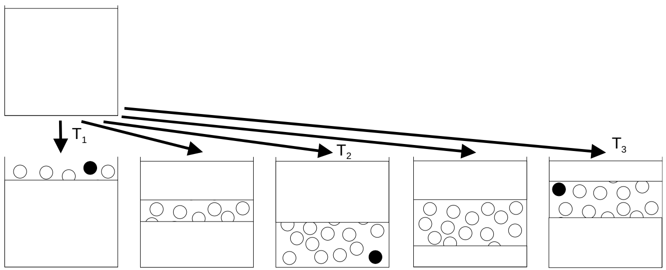
- ☐ Hardware Reliability (typical assumptions)
  - Failures are a result of physical degradation
  - When the faulty component is substituted, the reliability becomes the initial value of this component
  - The reliability of the system does not exceed the initial value of the system reliability through the substitution of components with new components
  - Hardware reliability is determined by fairly constant parameters



Hardware- vs. Software Reliability

☐ Software Reliability (typical assumptions)

- Failures are a result of design errors that are contained in the product from the start and appear accidentally
- After fault correction the system reliability exceeds its initial value (under the assumption that no additional faults are introduced)
- Faults that are introduced during debugging decrease reliability
- Reliability parameters are assumed to vary



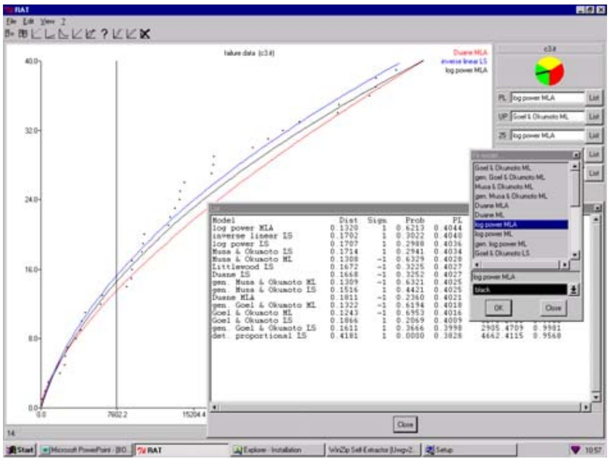
Tool Assisted Reliability Modeling

☐ How reliable is my system now?

☐ How reliable will it be at the planned release date?

☐ How many failures will have occurred by then?

...



PL: log power MLA
UP: Goel & Okumoto ML
25: log power MLA

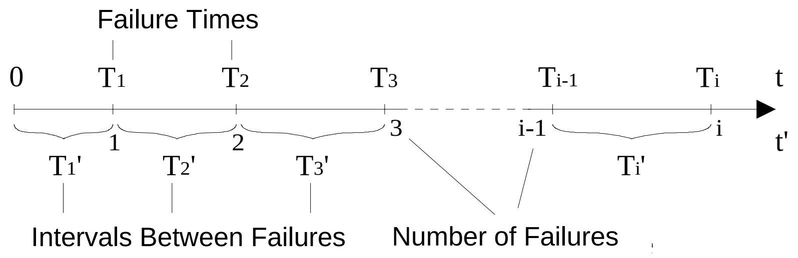
```
Model                        Dist   Sign    Prob      PL
log power MLA              0.1320      1  0.6213  0.4044
inverse linear LS          0.1702      1  0.3022  0.4040
log power LS               0.1707      1  0.2980  0.4036
Musa & Okumoto LS          0.1714      1  0.2941  0.4034
Musa & Okumoto ML          0.1308     -1  0.6329  0.4028
Littlewood LS              0.1672     -1  0.3225  0.4027
Duane LS                   0.1668     -1  0.3252  0.4027
gen. Musa & Okumoto ML     0.1309     -1  0.6321  0.4025
gen. Musa & Okumoto LS     0.1516      1  0.4421  0.4025
Duane MLA                  0.1811     -1  0.2360  0.4021
gen. Goel & Okumoto ML     0.1322     -1  0.6194  0.4018
Goel & Okumoto ML          0.1243     -1  0.6953  0.4016
Goel & Okumoto LS          0.1866      1  0.2069  0.4009
gen. Goel & Okumoto LS     0.1611      1  0.3666  0.3998
det. proportional LS       0.4181      1  0.0000  0.3828
```

Tool Assisted Reliability Modeling

- ☐ Use of models
  - Which models do exist?
  - How can I find out, which model fits my purposes best?
  - How can I define the model parameters in order to get dependable reliability predictions?

Description of failure over time

- ☐ Failure times
- ☐ Time intervals between failures
- ☐ Total number of failures at a point in time
- ☐ Failures within a given time interval



Modeling of Reliability

- Lifetime T
  - Large number of similar systems under consideration
  - Simultaneous start of the systems at time t = 0
  - Observed time of the first failure of each system is the so-called lifetime T of this system
  - Plot of the fraction of failed systems over t is the so-called empirical distribution function of the lifetime (or empirical life distribution)

Modeling of Reliability

- If the number of systems becomes larger (approximates infinity), the empirical life distribution approximates the life distribution F(t)
  - Here, lifetime T is a random variable and F(t) is the probability that an arbitrary system is not operational at t
    $F(t) = P\{T \leq t\}$
  - F(t) is the probability that lifetime T is less or equal to t, meaning that a system has already failed by t.
  - We use the following assumptions:

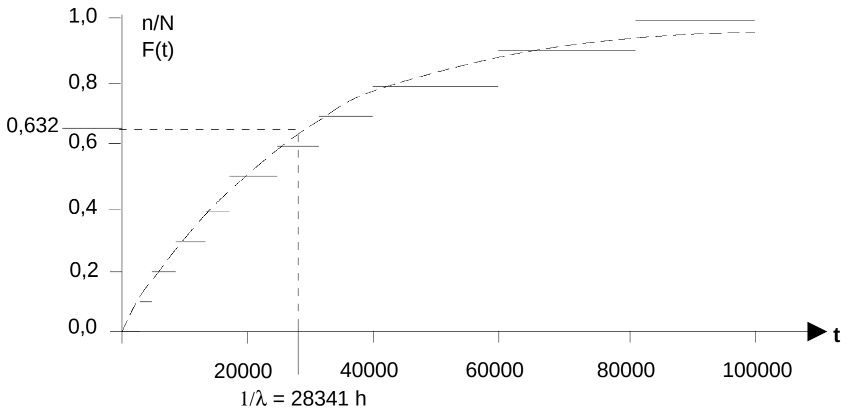
    $F(t = 0) = 0$, i.e. a new system is intact, and

    $\lim\limits_{t \to \infty} F(t) = 1$, i.e. every system fails sometimes
- Failure Times of 10 Systems

T_i(h)	2810	5411	8701	13130	17327	24899	31230	40006	59880	80017
$^n/_N$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
F(t)	0,094	0,174	0,264	0,371	0,457	0,585	0,668	0,765	0,879	0,941

Modeling of Reliability

☐ Life distribution F(t)



Modeling of Reliability

☐ Reliability function R(t)

- F(t) gives the probability that at time t at least one failure has occurred; thus R(t) = 1 - F(t) is the probability that at time t no failure has occurred yet

☐ Probability density f(t)

- The probability density f(t) describes the modification of the probability that a system fails over time

$$f(t) = \frac{d\,F(t)}{dt}$$

Modeling of Reliability

- ☐ MTBF, MTTF
  - A relevant measure for reliability is the Mean Time To Failure (MTTF) or Mean Time Between Failure (MTBF)
  - The MTTF resp. MTBF defines the mean value of the lifetime resp. the mean value for the time interval between two successive failures
  - It is determined by calculating the following integral:

$$\overline{T} = E(T) = \int_{0}^{\infty} t\, f(t)\, dt$$

- ☐ Failure rate
  - The failure rate is the relative boundary value of failed entities at time t in a time interval that approximates zero, referring to the entities still functional at the beginning of the time interval

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

Modeling of Reliability

- ☐ The conditional probability that a system that operated failure free until t also survives the period Δt is

$$\frac{R(t + \Delta t)}{R(t)}$$

- ☐ Thus, the probability that the product fails within Δt is

$$1 - \frac{R(t + \Delta t)}{R(t)} = 1 - \frac{1- F(t + \Delta t)}{1 - F(t)} = \frac{1 - F(t) - (1 - F(t + \Delta t))}{1 - F(t)} = \frac{F(t + \Delta t) - F(t)}{1 - F(t)}$$

Modeling of Reliability

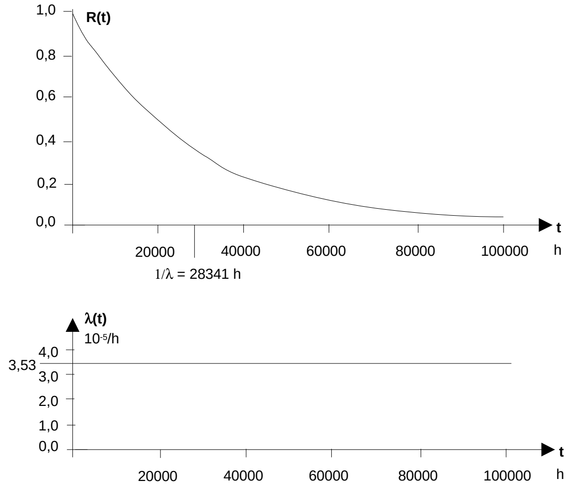
- ☐ As the given probability for short time intervals Δt is proportional to Δt, we divide the term by Δt and determine the boundary value when Δt approximates 0

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

- ☐ Thus the probability that a system, that is operational at time t fails within the (short) time interval Δt, is approximately Δt λ(t)

Modeling of Reliability

- ☐ R(t) and failure rate



Example for the Distribution Function

- Assumption: For the given data (table p. 10) lifetime is exponentially distributed: $F(t) = 1 – e^{-\lambda t}$
- The parameter λ (failure rate) has to be determined based on failure observations in order to achieve an optimal adjustment of the function, according to a predetermined criterion. The Maximum-Likelihood-Method provides the following parameter λ for the exponential distribution:

$$\lambda = \frac{N}{\sum\limits_{i=1}^{N} T_i} = 0{,}0000353 \,/\, h$$

- Reliability: $R(t) = 1 – F(t) = e^{-\lambda t}$

Example for the Distribution Function

- The failure rate λ is constant over time

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{dF(t) \,/\, dt}{R(t)} = \frac{- dR(t) \,/\, dt}{R(t)} = \frac{\lambda\, e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

- A constant failure rate causes an exponential distribution of the lifetime
- Determination of the MTTF

$$\bar{T} = E(T) = \int_0^\infty t\, f(t)\, dt = \int_0^\infty t\, \lambda\, e^{-\lambda t}\, dt = \lambda \int_0^\infty t\, e^{-\lambda t}\, dt = \lambda \left(\frac{\lambda\, e^{-\lambda t}}{\lambda^2} (-\lambda\, t – 1) \right)_0^\infty = \frac{1}{\lambda}$$

- If lifetime is exponentially distributed, the MTTF is the reciprocal of the failure rate and thus constant

The Exponential Distribution

☐ Life distribution: $F(t) = 1 - e^{-\lambda t}$

☐ Density function: $f(t) = \lambda\, e^{-\lambda t}$

☐ Reliability function: $R(t) = 1 - F(t) = e^{-\lambda t}$

☐ Failure rate: $\lambda(t) = \lambda$

☐ MTTF: $\overline{T} = \frac{1}{\lambda}$

The Weibull Distribution

☐ Life Distribution : $F(t) = 1 – e^{-(\lambda t)^{\beta}}$; $\lambda, \beta > 0$

☐ or:

$$F(t) = 1 – e^{-\frac{1}{\alpha} t^{\beta}};\ \alpha, \beta > 0, \text{ d. h. } \frac{1}{\alpha} = \lambda^{\beta}$$

☐ Density:

$$f(t) = \frac{dF(t)}{dt} = \lambda\, \beta\, (\lambda\, t)^{\beta - 1}\, e^{-(\lambda t)^{\beta}}$$

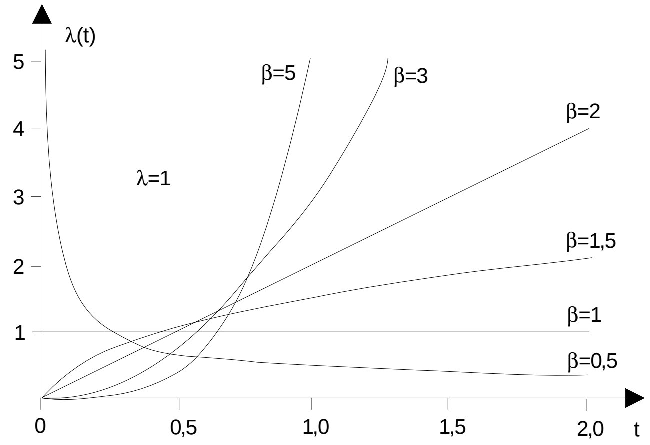
☐ Reliability: $R(t) = e^{(-\lambda t)^{\beta}}$

☐ Failure rate:

$$\lambda(t) = \frac{f(t)}{R(t)} = \lambda\, \beta\, (\lambda\, t)^{\beta - 1}$$

The Weibull Distribution

- ☐ Failure rate of the Weibull distribution depending on the form parameter β



The Poisson Distribution

Assumptions

- ☐ The probability of more than one failure within the (short) time interval Δt can be ignored. Thus, failures occur relatively infrequently
- ☐ The probability of a failure within Δt, respectively within [t, t + Δt], is λ Δt (see definition of failure rate). The probability is proportional to the length of the time interval
- ☐ $P_x(t)$ is the probability, that within time interval [0, t] x failures occur

The Poisson Distribution

☐ No failures

- The probability that within time interval [0, t+Δt] no failures occur is determined by multiplying the probability that until time t no failures have occurred ($P_0(t)$) and the probability that within [t, t+Δt] no failures occur (1-λ Δt):

  $$P_0(t+\Delta t) = P_0(t)(1-\lambda\,\Delta t) \Leftrightarrow \frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = -\lambda\,P_0(t)$$

- For Δ t towards 0 one receives:

  $$\lim_{\Delta t \to 0} \frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = \frac{d\,P_0(t)}{dt} = -\lambda\,P_0(t)$$

- $P_0(0) = 1$, since new systems (t=0) are always operational by definition. For a constant value of λ and $P_0(0) = 1$ the differential equation has the solution:

  $$P_0(t) = e^{-\lambda t}$$

The Poisson Distribution

- The probability that a new system shows no failures until t is R(t)

  $$R(t) = P_0(t) = e^{-\lambda t}$$

- Using the definitions for F(t) and f(t), we get:

  $$F(t) = 1 - R(t) = 1 - e^{-\lambda t} \quad and \quad f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t}$$

The Poisson Distribution

- ☐ Failures
  - The probability that within time interval [0, t+Δt] x failures occur can be determined as follows:

$$
\begin{aligned}
P_x(t+\Delta t) = P_0(t)\ & \left[P\left(x \quad failures \quad between \quad t \quad and \quad t+\Delta t\right)\right] \\
& + \ \ldots \\
& + \ P_{x-2}(t)\ \left[P\left(2 \quad failures \quad between \quad t \quad and \quad t+\Delta t\right)\right] \\
& + \ P_{x-1}(t)\ \left[P\left(1 \quad failure \quad between \quad t \quad and \quad t+\Delta t\right)\right] \\
& + \ P_x(t)\ \left[P\left(no \quad failure \quad between \quad t \quad and \quad t+\Delta t\right)\right]
\end{aligned}
$$

The Poisson Distribution

- Due to the precondition the probability to observe more than one failure in Δt is zero. Therefore we get:

$$
\begin{aligned}
P_x(t+\Delta t) = \ & P_{x-1}(t)\ \left[P\left(1 \quad failure \quad between \quad t \quad and \quad t+\Delta t\right)\right] \\
& + \ P_x(t)\ \left[P\left(no \quad failure \quad between \quad t \quad and \quad t+\Delta t\right)\right] \\
= \ & P_{x-1}(t)(\lambda \quad \Delta t) + P_x(t)(1-\lambda \quad \Delta t) \\
= \ & P_x(t) - (\lambda \quad \Delta t)\left[P_x(t) - P_{x-1}(t)\right] \\
& \Leftrightarrow
\end{aligned}
$$

$$
\frac{P_x(t+\Delta t) - P_x(t)}{\Delta t} = -\lambda\left[P_x(t) - P_{x-1}(t)\right]
$$

The Poisson Distribution

- With Δt approximating zero:

$$\lim_{\Delta t \to 0} \frac{P_x(t+\Delta t) - P_x(t)}{\Delta t} = \frac{dP_x(t)}{dt} = -\lambda\left[P_x(t) - P_{x-1}(t)\right]$$

- The following term for $P_x(t)$ is a solution for this differential equation

$$P_x(t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

(Poisson Distribution)

The Poisson Distribution

☐ This can be shown very easily

$$\frac{dP_X(t)}{dt} = \frac{d\,\dfrac{(\lambda t)^x e^{-\lambda t}}{x!}}{dt} = \frac{x(\lambda t)^{x-1}\lambda e^{-\lambda t} + (\lambda t)^x(-\lambda)e^{-\lambda t}}{x!}$$

$$= -\lambda\left[\frac{(\lambda t)^x e^{-\lambda t}}{x!} - \frac{(\lambda t)^{x-1} e^{-\lambda t}}{(x-1)!}\right] = -\lambda\left[P_x(t) - P_{x-1}(t)\right]$$

☐ The probability $P_X(t)$ provides the correct value $P_0(t)$ also for the case that we treated separately before

$$\frac{dP_X(t)}{dt} \underset{x=0}{=} \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t} = P_0(t)$$

The Poisson Distribution

☐ $P_X(t)$ fulfills the boundary conditions for t = 0, i.e. $P_0(0) = 1$ and $P_X(0) = 0$, for $x \geq 1$. Furthermore the sum of the probabilities of all $x \geq 0$ for every $t \geq 0$ must be 1, i.e.

$$\sum_{x=0}^{\infty} P_x(t) = \sum_{x=0}^{\infty} \frac{(\lambda t)^x e^{-\lambda t}}{x!} = e^{-\lambda t} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = 1 \Rightarrow \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} \overset{!}{=} e^{-\lambda t}$$

☐ The specified sum on the left hand side of the equation is the power series of the exponential function on the right hand side. The Poisson Distribution thus fulfills the preconditions. If λ is constant, the mean value is $\mu(t)=\lambda t$. This is called a homogeneous Poisson Process. If λ is a function of time, the mean value is

$$\mu(t) = \int_0^t \lambda(\tau)\, d\tau \quad and \quad P_x(t) = \frac{\mu(t)^x\, e^{-\mu(t)}}{x!}$$

This is called a non-homogeneous Poisson Process (NHPP)

Failure Times and Times between Failures

☐ The time of failure i is T_i

☐ The time interval between failure (i - 1) and failure i is $T_i‘$

☐ $T_i = \sum_{j=1}^{i} T_j‘,\ T_0 = 0$

☐ M(t) is the number of failures at t

$$[M(t) \geq i] \Leftrightarrow [T_i \ \leq t]$$

Failure Times and Times between Failures

- ☐ The probability for j failures until time t is

$$P_j(t) = P[M(t) = j] = \frac{(\mu(t))^j e^{-\mu(t)}}{j!}$$

- ☐ The probability for at least i failures at t is

$$P[M(t) \geq i] = \sum_{j=i}^{\infty} \frac{(\mu(t))^j e^{-\mu(t)}}{j!} = P[T_i \leq t]$$

Musa's Execution Time Model

- ☐ A software system fails due to errors in the software randomly at t1, t2, ... (t here refers to execution time, i. e. CPU-seconds)
  - It is assumed that the number of failures observed in Δt is linearly proportional to the number of faults contained in the software at this time
  - $\mu(t)$ is the total number of failures for times $t \geq 0$
  - $\mu(t)$ is a limited function of t
  - The number of failures is a monotonic increasing function of t
  - At t=0 no failures have been observed yet: $\mu(0)=0$
  - After very long execution time ($t \rightarrow \infty$) the value $\mu(t)$ is equal to a. a is the total number of failures in infinite time. (There are also models where infinite numbers of failures are assumed to happen)

Musa's Execution Time Model

☐ Model development

- The number of failures observed in a time interval Δt is proportional to Δt and to the number of errors not yet detected

$$\mu(t+\Delta t)-\mu(t)=b[a-\mu(t)]\Delta t \Rightarrow \frac{\mu(t+\Delta t)-\mu(t)}{\Delta t}=ba-b\mu(t)$$

☐ With Δt $\rightarrow$ 0 we get:

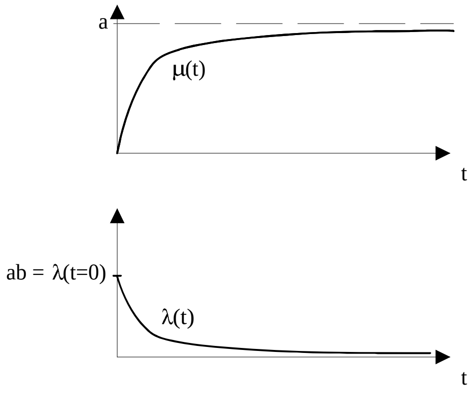
$$\frac{d\mu(t)}{dt}=ba-b\mu(t)=\mu'(t)$$

☐ With $\mu(0)=0$ and $\mu(\infty)=a$ we get: $\mu(t)=a\left(1-e^{-bt}\right)$

☐ The failure rate is: $\lambda(t)=\mu'(t)=abe^{-bt}$

Musa's Elementary Execution Model

☐ The curve for the accumulated number of failures μ(t) approximates asymptotically the expected total number of failures a



Musa's Elementary Execution Model

- ☐ The curve for the failure rate λ(t) for t = 0 starts at the initial failure rate λ_0 = ab and approximates asymptotically the value 0. The initial failure rate is proportional to the expected number of failures a, with the constant of proportionality b

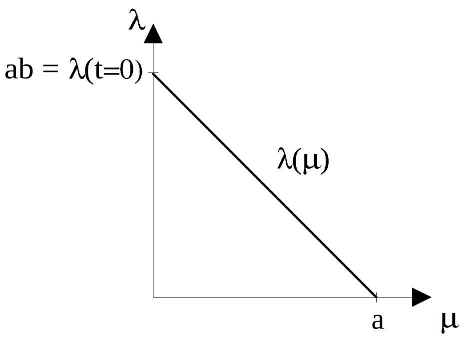
$$\mu(t) = a\left(1-e^{-bt}\right) = a\left(1-e^{-\frac{\lambda_0}{a}t}\right)$$

$$\lambda(t) = abe^{-bt} = \lambda_0 e^{-\frac{\lambda_0}{a}t}$$

$$\mu(t) = a\left(1-e^{-bt}\right) = a\left(1-e^{-\frac{\lambda_0}{a}t}\right) \quad and \quad \lambda(t) = abe^{-bt} = \lambda_0 e^{-\frac{\lambda_0}{a}t} \;\Rightarrow\; e^{-bt} = \frac{\lambda(t)}{ab}$$

Musa's Execution Time Model

$$\mu(t) = a\left(1-\frac{\lambda(t)}{ab}\right) \Rightarrow \lambda(\mu) = b(a-\mu) = ab\left(1-\frac{\mu}{a}\right) = \lambda_0\left(1-\frac{\mu}{a}\right)$$



Musa's Execution Time Model

- ☐ If λ is the present failure rate and a target λ_z is defined, Δμ additional failures will occur until this target is reached

$$\Delta\mu = \mu_z - \mu = a\left(1 - \frac{\lambda_z}{\lambda_0}\right) - a\left(1 - \frac{\lambda}{\lambda_0}\right) = a\left(\frac{\lambda - \lambda_z}{\lambda_0}\right)$$

- ☐ The additional time Δt until this target is reached is

$$\Delta t = t_z - t = -\frac{a}{\lambda_0}\left[\ln\left(\frac{\lambda_z}{\lambda_0}\right) - \ln\left(\frac{\lambda}{\lambda_0}\right)\right] = \frac{a}{\lambda_0}\left[\ln\left(\frac{\lambda}{\lambda_0}\right) - \ln\left(\frac{\lambda_z}{\lambda_0}\right)\right] = \frac{a}{\lambda_0}\ln\left(\frac{\lambda}{\lambda_z}\right)$$

Musa's Execution Time Model

- ☐ If

$$\mu(t) = a\left(1 - e^{-\frac{\lambda_0}{a}t}\right)$$

is inserted into the general equation of the Poisson distribution, we get:

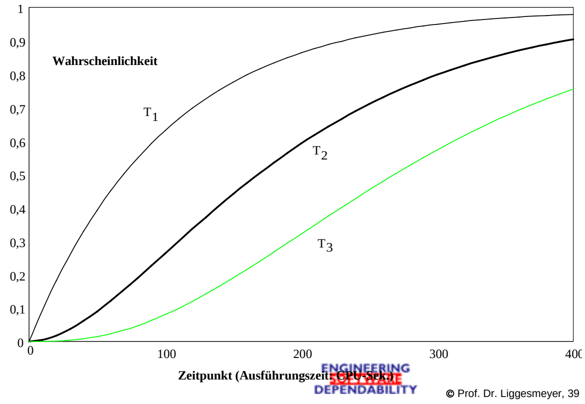
$$P[T_i \le t] = \sum_{j=i}^{\infty} \frac{(\mu(t))^j e^{-\mu(t)}}{j!} = \sum_{j=i}^{\infty} \frac{\left[a\left(1 - e^{-\frac{\lambda_0}{a}t}\right)\right]^j e^{-a\left[1 - e^{-\frac{\lambda_0}{a}t}\right]}}{j!}$$

$$= e^{-a\left[1 - e^{-\frac{\lambda_0}{a}t}\right]} \sum_{j=i}^{\infty} \frac{\left[a\left(1 - e^{-\frac{\lambda_0}{a}t}\right)\right]^j}{j!}$$

Examples of Modeling

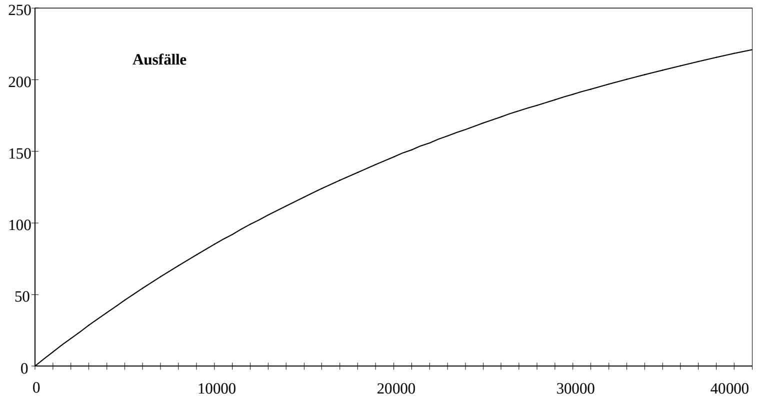
- ☐ For a program with an expected total number of 300 failures with an initial failure rate of 0,01/CPU-second, models are to be generated
- ☐ What is the probability that at a particular execution time at least a certain number of failures will have occurred?

Formula for $P[T_i \leq t]$
for 1, 2 and 3 failures



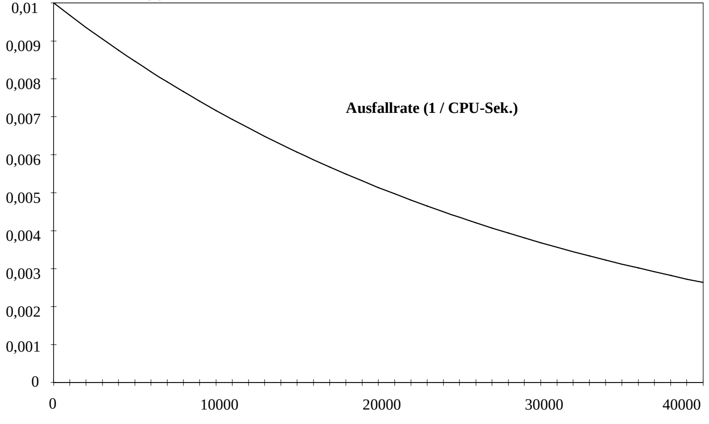
Examples of Modeling

- ☐ What will be the number of failures w.r.t. execution time?
- ☐ Formula for $\mu(t)$



Examples of Modeling

- ☐ How will the failure rate develop depending on the execution time?
- ☐ Formula for $\lambda(t)$



Determination of Model Parameters

- ☐ Least squares
  - Target: Define parameters in such a way that the sum of the squares of the deviations between the calculated and the observed values becomes minimal. If F_i refers to the value of the empirical distribution function at point t_i, the following term is to be minimized:

$$\Delta = \sum_{i=1}^{n} \Delta_i^{\,2} = \sum_{i=1}^{n} \left(F(t_i) - F_i \right)^2$$

- ☐ Maximum-Likelihood-Method
  - Target: Choose parameters in such a way that the probability is maximized to produce a "similar" observation to the present observation. The probability density has to be known

Determination of Model Parameters
Least squares

- ☐ Target: Define parameters in such a way that the sum of the squares of the deviations between the calculated and the observed values becomes minimal. If F_i refers to the value of the empirical distribution function at point t_i, the following term is to be minimized:

$$\Delta = \sum_{i=1}^{n} \Delta_i^2 = \sum_{i=1}^{n} \left(F(t_i) - F_i\right)^2$$

- ☐ For the exponential distribution we get:

$$\Delta_{\exp} = \sum_{i=1}^{n} \Delta_{i_{\exp}}^2 = \sum_{i=1}^{n} \left(F(t_i) - F_i\right)^2 = \sum_{i=1}^{n} \left(1 - e^{-\lambda t_i} - F_i\right)^2$$

Determination of Model Parameters
Least squares

- ☐ The value λ that minimizes this term is to be determined

$$\frac{d\Delta_{\exp}}{d\lambda} = \frac{d\left(\sum_{i=1}^{n} \Delta_{i_{\exp}}^2\right)}{d\lambda} = \sum_{i=1}^{n} 2\left(1 - e^{-\lambda t_i} - F_i\right) t_i e^{-\lambda t_i}$$

- ☐ The value $\hat{\lambda}$ is calculated by determining the root

$$\sum 2\left(1 - e^{-\hat{\lambda} t_i} - F_i\right) t_i e^{-\hat{\lambda} t_i} \overset{!}{=} 0$$

Determination of Model Parameters
Least squares

- ☐ Sometimes numerical method must be used for this task. A Newtonian iteration provides the following results for the Exponential Distribution

$$\lambda_{n+1} = \lambda_n - \frac{f(\lambda_n)}{\frac{df(\lambda_n)}{d\lambda}}$$

- ☐ with: $f_{\exp}(\lambda) = \sum_{i=1}^{n} 2\left(1 - e^{-\lambda t_i} - F_i\right) t_i e^{-\lambda t_i}$
- ☐ and: $\frac{f_{\exp}(\lambda)}{d\lambda} = \sum_{i=1}^{n} 2 t_i^2 \left(\left(F_i - 1\right) e^{-\lambda t_i} + 2 e^{-2\lambda t_i}\right)$
- ☐ For the failure times of the table on page 10 the search for zero points according to the Newtonian iteration provides a value $\hat{\lambda} \approx$ 3,9326702 * 10^{-5}/h for the exponential distribution

---

Determination of Model Parameters
Maximum-Likelihood-Method

- ☐ Target: Choose parameters in such a way that the probability is maximized to produce a "similar" observation to the present observation
- ☐ Precondition: Probability density has to be known
- ☐ Likelihood function
  - Product of the densities at the observed failure times
  - The value is proportional to the probability to observe failure times that do not exceed the deviation Δt w.r.t. the present observation
  - It is a function of the distribution function's parameters that are to be determined
  - Example:
    The parameter λ of the exponential distribution is to be determined with the Maximum-Likelihood-Method

$$F(t) = 1 - e^{-\lambda t}, \; f(t) = \lambda e^{-\lambda t}$$

Determination of Model Parameters
Maximum-Likelihood-Method

With n observed failure times $t_1, \dots, t_n$ we get the Likelihood Function:

$$L(\lambda,t_1,\dots,t_n) = f(\lambda,t_1)\,f(\lambda,t_2)\dots f(\lambda,t_n) = \lambda e^{-\lambda t_1}\lambda e^{-\lambda t_2}\dots\lambda e^{-\lambda t_n}$$

$$= \lambda^n e^{-\lambda(t_1+t_2+\dots+t_n)} = \lambda^n e^{-\lambda\sum_{i=1}^{n} t_i}$$

Due to the monotonicity of the logarithmic function, L und ln L have identical maxima

$$\ln L(\lambda,t_1,\dots,t_n) = n\ln\lambda - \lambda\sum_{i=1}^{n} t_i$$

In order to calculate the value $\hat{\lambda}$ that maximizes the Likelihood Function, the derivation according to λ must be determined

$$\frac{d\left(\ln L(\lambda,t_1,\dots,t_n)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i$$

Determination of Model Parameters
Maximum-Likelihood-Method

☐ $\hat{\lambda}$ is the root. For the exponential distribution we get:

$$\frac{n}{\lambda} - \sum_{i=1}^{n} t_i \overset{!}{=} 0 \Leftrightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i}$$

Model Selection based on Failure Observations

- ☐ U-Plot-Method
- ☐ Prequential-Likelihood-Method
- ☐ Holdout-Evaluation

Model Selection based on Failure Observations
U-Plot

- ☐ U-Plot
  - Graphic method that tests whether a distribution function can be accepted with regard to the present observation
  - Additionally, statistical tests (e.g. Kolmogoroff-Smirnov) might be used
  - If a random variable T is described by the distribution F(t), the $F(t_i)$ of the random variable are equally distributed over the interval [0,1]

Model Selection based on Failure Observations
U-Plot

- The n values U_i are charted in a U-Plot as follows
  - The values U_i are used as y-values in such a way that the value U_i with the position j is attributed to the x-value $^j/_n$
  - If the values U_i are approximately equally distributed, the applied points are located "near by" the function y = x, for $0 \leq x \leq 1$

Model Selection based on Failure Observations
U-Plot

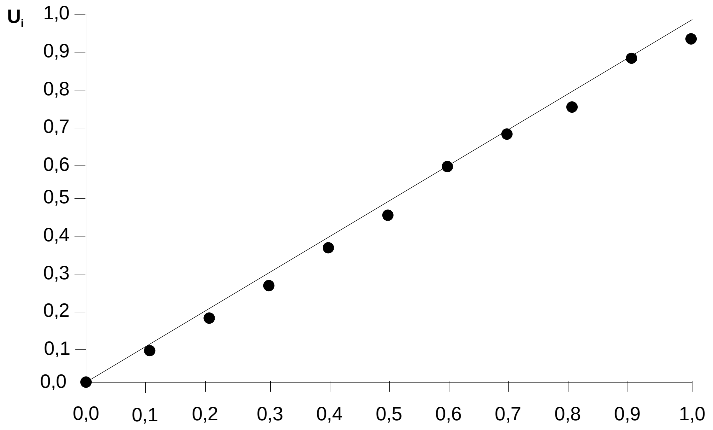
☐ Example

T_i(h)	2810	5411	8701	13130	17327	24899	31230	40006	59880	80017
$^n/_N$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
F(t)	0,094	0,174	0,264	0,371	0,457	0,585	0,668	0,765	0,879	0,941

The values presented in the table for F(t) are the U_i according to the definition stated above

Model Selection based on Failure Observations
U-Plot

☐ U-Plot of the data



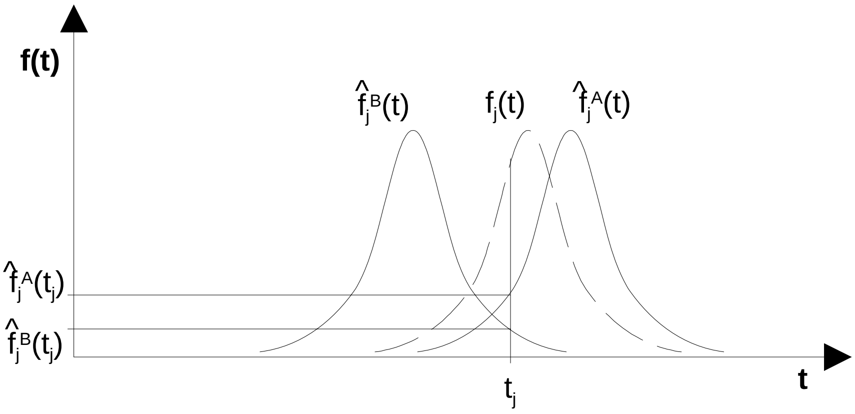
Model Selection based on Failure Observations
Prequential-Likelihood-Method

☐ The Prequential-Likelihood-Method compares the suitability of two distribution functions under consideration with regard to a given failure observation

☐ It is based on the following approach

- The failure interval t_j is a realization of a random variable with the distribution $F_j(t)$ and the density $f_j(t)$
- $F_j(t)$ and $f_j(t)$ are unknown
- The densities of the distribution functions A and B ($\hat{f}_j^A(t)$ resp. $\hat{f}_j^B(t)$) can be determined based on the failure intervals t_1, ... , t_{j-1}
- If the distribution A is more suitable than the distribution B, it can be expected that the value $\hat{f}_j^A(t_j)$ is greater than the value $\hat{f}_j^B(t_j)$
- The quotient $\frac{\hat{f}_j^A(t_j)}{\hat{f}_j^B(t_j)}$ will be greater than 1

Model Selection based on Failure Observations
Prequential-Likelihood-Method



Model Selection based on Failure Observations
Prequential-Likelihood-Method

- If this analysis is done for every observed failure time interval t_j we get the so-called Prequential-Likelihood-Ratio concerning the distributions A and B

$$PLR_i^{AB} = \prod_{j=s}^{j=i} \frac{\hat{f}_j^{A}\left(t_j\right)}{\hat{f}_j^{B}\left(t_j\right)}$$

- If A is more appropriate than B with regard to the present failure data, the PLR shows a rising tendency
- Example
  - We compare the exponential distribution and the normal distribution based on the data from the table on page 10 using the Prequential-Likelihood-Method

Model Selection based on Failure Observations
Prequential-Likelihood-Method

- The parameters of the distributions are determined using a Maximum-Likelihood-Approach. For the exponential distribution, we get:

$$\hat{\lambda}_j = \frac{j-1}{\sum_{k=1}^{j-1} t_k}$$

- For the normal distribution we get: $f\left(\tau,\sigma^2,t\right) = \frac{1}{\sqrt{\left(2\pi\sigma^2\right)}} e^{-(t-\tau)^2/2\sigma^2}$

- The parameters according to the Maximum-Likelihood-Method are:

$$\hat{\tau}_j = \frac{\sum_{k=1}^{j-1} t_k}{j-1}; \hat{\sigma}_j^2 = \frac{1}{j-1}\sum_{k=1}^{j-1}\left(t_k - \tau\right)^2$$

Model Selection based on Failure Observations
Prequential-Likelihood-Method

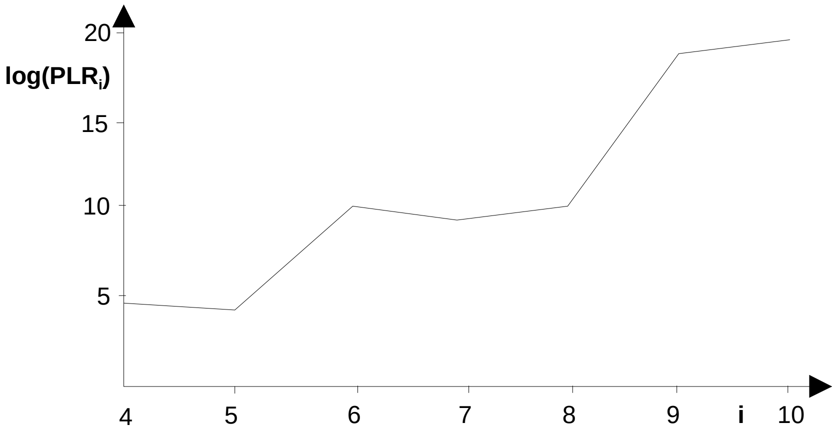
- The following table shows the densities of the exponential distribution and the normal distribution for the arrival time intervals t_i based on the failure times T_1 to T_{i-1} from the table on page 10
- The calculation starts with i = 4. In addition the logarithm of the quotient of the densities and the logarithm of the PLR_i is contained in the table
- The rising of the PLR underlines that the assumption of exponentially distributed arrival times for the present data makes more sense than the assumption of normally distributed arrival times

Model Selection based on Failure Observations
Prequential-Likelihood-Method

i	1	2	3	4	5	6	7	8	9	10
T_i(h)	2810	5411	8701	13130	17327	24899	31230	40006	59880	80017
t_i(h)	2810	2601	3290	4429	4197	7572	6331	8776	19874	20137
$f_i^{Exp}/ 10^{-6}$				74,9	84,8	32,5	52,4	31,3	3,8	7,3
$f_i^{Norm}/ 10^{-9}$				1,1	244558	0,076	101787	10096	0,000008	2362
log (f_i^{Exp}/ f_i^{Norm})				4,83	-0,46	5,63	-0,29	0,49	8,67	0,49
log (PLR_i)				4,83	4,37	10,00	9,71	10,20	18,87	19,36

Model Selection based on Failure Observations
Prequential-Likelihood-Method

- ☐ PLR of the data



Model Selection based on Failure Observations
Holdout Evaluation

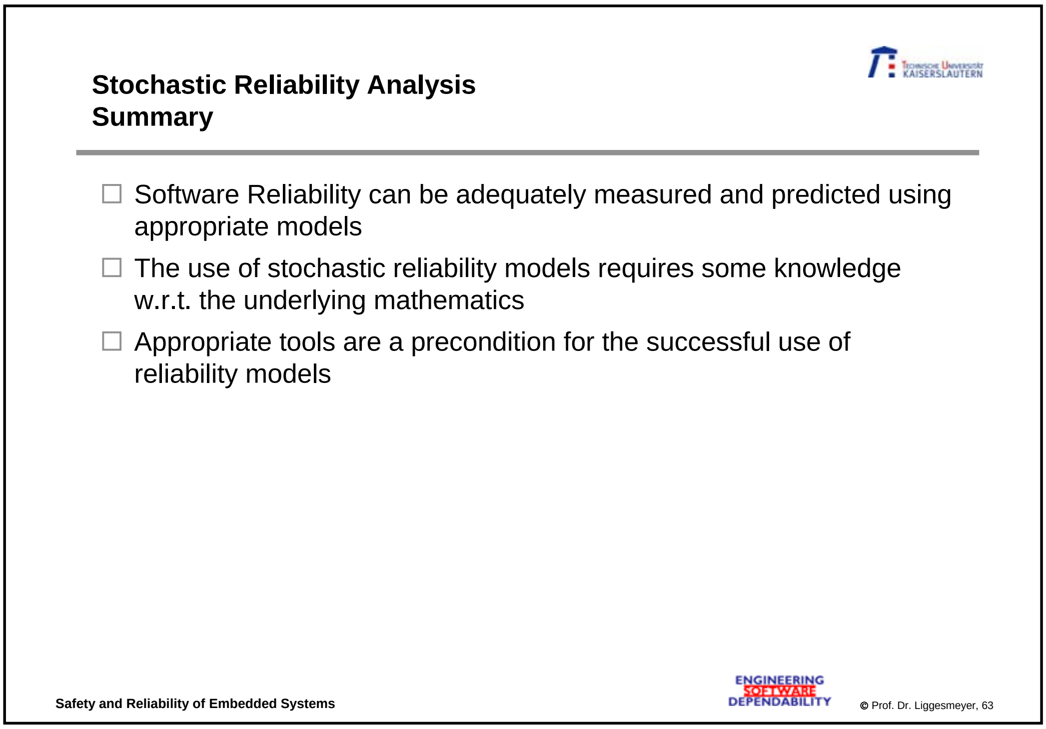
☐ Approach

- Only parts of the failure data are used for model calibration. The remaining data are used to judge the prediction quality of the calibrated model
- If an exponential distribution and a Weibull distribution are calibrated to the first 6 failure times (table p. 10) using a Least-Squares-Algorithm, we get the following results:
- Exponential distribution: $F(t)_{\exp} = 1 - e^{-3,89292*10^{-5} t_i / h}$
- Weibull distribution: $F(t)_{\exp} = 1 - e^{-\frac{1}{1,552005*10^4} t_i^{0,94750} / h}$

Model Selection based on Failure Observations
Holdout Evaluation

☐ The Weibull distribution has – as expected - a better adjustment to the failure times T_1 to T_6. The sum of the deviation squares for the first 6 failure times is 0,000459 compared to 0,000790 in the exponential distribution

☐ The prediction quality of the Weibull distribution is however worse than that of the exponential distribution. The sum of the deviation squares for the failure times T_7 to T_{10} is 0,00446 for the Weibull distribution; for the exponential distribution is only 0,00210. We might prefer to use the exponential distribution in order to avoid „over-calibration“



Stochastic Reliability Analysis
Summary

☐ Software Reliability can be adequately measured and predicted using appropriate models

☐ The use of stochastic reliability models requires some knowledge w.r.t. the underlying mathematics

☐ Appropriate tools are a precondition for the successful use of reliability models