



Test and Depentability

Slide set 3: Distributions

Prof. G. Kemnitz

Institute for Computer Science, TU Clausthal (TV_F3_engl.pdf)
April 20, 2023

Contents slide set TV_F3: Distributions

Distribution

- 1.1 Characteristics
- 1.2 Linear transformation
- 1.3 Distribution of counts

CVD Approximations

- 2.1 Binomial distribution
- 2.2 Poisson distribution
- 2.3 Range, Poisson
- 2.4 Normal distribution
- 2.5 Range estimate ND
- 2.6 Varianzerhöhung
- 2.7 Bereichs. mit Abhängigk.

Mischverteilung

- 3.1 Eigenschaften
- 3.2 Anwendungen
- 3.3 Bereichs. Tscheb.

Pareto-Verteilung

- 4.1 Eigenschaften
- 4.2 Fehlernachweislänge
- 4.3 Schaden durch MF

Gamma- und Exp-Vert.

Ausfälle

- 6.1 Kenngrößen
- 6.2 Voralterung
- 6.3 Reserveeinheiten
- 6.4 Wartung

lecture	8	9	10	11
slide	1	50	106	157



Distribution



Random event, experiment, ..., distribution

- Random event: Event that is neither certain nor impossible, but occurs with a certain probability.
- Random experiment: Experiment with several possible results and a random outcome.
- Random variable: Variable that picks up its values depending on the outcome of a random experiment according to a probability distribution.
- Distribution: A distribution assigns probabilities to the possible values of a random variable (i.e. the results of a random experiment).

Random variables of further interest:

- Number of malfunctions (occurred, detected, ...),
- Number of faults (created, eliminated, avoided, ...),
- Lifetime,
- damage costs, ...



Bernoulli experiment

The simplest random experiment. Two-point distribution:

$$\begin{aligned}\mathbb{P}(X = 0) &= 1 - p \\ \mathbb{P}(X = 1) &= p\end{aligned}\tag{1}$$

Random experiments with more than two possible outcomes can be split into one Bernoulli experiment per outcome:

$$A_i = \begin{cases} 0 & \text{event not occurred} \\ 1 & \text{event occurred} \end{cases}$$

X	random variable.
$\mathbb{P}(\dots)$	probability of occurrence of an event.
p	probability of occurrence.



Characteristics



Characteristic parameters of a random variable

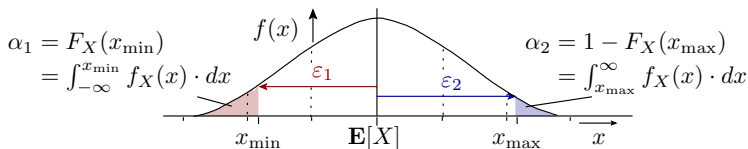
Name	Definition
Distribution function	$F_X(x) = \mathbb{P}(X \leq x)$
Density function	$f_X(x) = \frac{dF_X(x)}{dx}$
Expectation value	$\mathbb{E}[X] = \sum_{i=1}^m p_i \cdot x_i$ or $\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx$
k-th moment	$m_k = \mathbb{E}[X^k]$
k-th centred moment	$\mathbb{E}[(X - \mathbb{E}[X])^k]$
Variance (2nd centr. m.)	$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$
Standard deviation	$\text{sd}[X] = \sqrt{\text{Var}[X]}$
probable range	$[x_{\min}, x_{\max}]$

$F_X(x)$	Distribution of the random variable X .
$f_X(x)$	density function of the random variable X .
$\mathbb{E}[X]$	expectation value of the random variable X .
$\text{Var}[X]$	variance deviation of the random variable X .
$\text{sd}[X]$	standard deviation of the random variable X .



Probable range, $\mathbb{E}[X]$ and $\text{sd}[X]$

Range $[x_{\min}, x_{\max}]$ that the value of the random variable X takes with probability $1 - (\alpha_1 + \alpha_2)$:



$\alpha_1 = F_X(x_{\min})$ Error probability that values are below the estimated range.

$\alpha_2 = 1 - F_X(x_{\max})$ Error probability that values lie above the estimated range.

With $\alpha_1 = 0$ there is only an upper limit and with $\alpha_2 = 0$ only a lower limit.

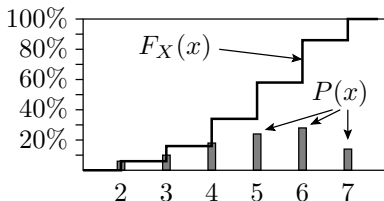
$F_X(x)$	Distribution of the random variable X .
$f_X(x)$	density function of the random variable X .
ϵ_1, ϵ_2	Interval radius, distance of the lower or upper range limit from the expected value.



Discrete distribution

Random variable X can only adopt a (over*-) countable set of discrete values x_i , e.g.:

x_i	2	3	4	5	6	7
$P(x) = \mathbb{P}(X = x_i) = p_i$	6%	10%	18%	24%	28%	14%
$F_X(x) = \mathbb{P}(X \leq x_i)$	6%	16%	34%	58%	86%	100%



$\mathbb{P}(\dots)$ probability of occurrence of an event.

$P(\dots)$ probability of occurrence of a value.

$F_X(x)$ Distribution of the random variable X .

* if the set of possible values is infinite.



Expected value and variance

x_i	2	3	4	5	6	7
$P(x) = \mathbb{P}(X = x_i) = p_i$	6%	10%	18%	24%	28%	14%
$F_X(x) = \mathbb{P}(X \leq x_i)$	6%	16%	34%	58%	86%	100%

Expected value (mean value weighted with the probabilities of occurrence)

$$\mathbb{E}[X] = \sum_{i=1}^m p_i \cdot x_i \quad (2)$$

(m – Number of possible values). Für das Beispiel:

$$6\% \cdot 2 + 10\% \cdot 3 + 18\% \cdot 4 + 24\% \cdot 5 + 28\% \cdot 6 + 14\% \cdot 7 = 5$$

Variance (2nd centred moment):

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{i=1}^m p_i \cdot (x_i - \mathbb{E}[X])^2 \quad (3)$$

For the example:

$$6\% \cdot (2 - 5)^2 + 10\% \cdot (3 - 5)^2 + \dots + 14\% \cdot (7 - 5)^2 = 1.96$$



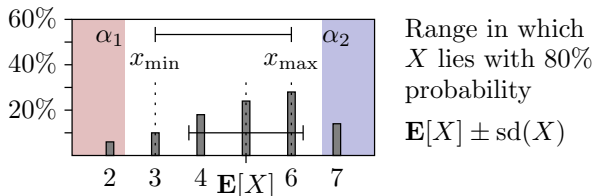
Standard deviation

Square root of the variance, measure of the deviation from the expected value or the width of the most likely range:

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (4)$$

For the example:

$$\text{sd}[X] = \sqrt{1.96} = 1.4$$



Error probabilities for X out of range $[x_{\min}, x_{\max}]$:

$$\alpha_1 = \mathbb{P}[X < x_{\min}] = \sum_{x_i < x_{\min}} \mathbb{P}[X = x_i]$$

$$\alpha_2 = \mathbb{P}[X > x_{\max}] = \sum_{x_i > x_{\max}} \mathbb{P}[X = x_i]$$



Shift rule

The variance is also the difference between the expected value of the squares and the square of the expected value*:

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \quad (5)$$

Derivation:

$$\begin{aligned} \sum_{i=1}^m p_i \cdot (x_i - \mathbb{E}[X])^2 &= \sum_{i=1}^m p_i \cdot (x_i^2 - 2 \cdot x_i \cdot \mathbb{E}[X] + \mathbb{E}[X]^2) \\ &= \underbrace{\sum_{i=1}^m p_i \cdot x_i^2}_{\mathbb{E}[X^2]} - 2 \cdot \mathbb{E}[X] \cdot \underbrace{\sum_{i=1}^m p_i \cdot x_i}_{\mathbb{E}[X]} + \mathbb{E}[X]^2 \cdot \underbrace{\sum_{i=1}^m p_i}_1 \end{aligned}$$

For the example above:

$$\text{Var}[X] = 6\% \cdot 2^2 + 10\% \cdot 3^2 + 18\% \cdot 4^2 + 24\% \cdot 5^2 + 28\% \cdot 6^2 + 14\% \cdot 7^2 - 5^2 = 1,96$$

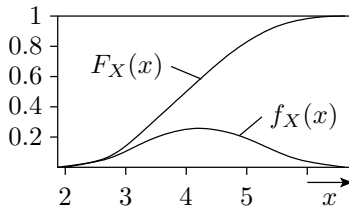
* numerical problems due to *difference of large numbers* possible.

Continuous distributions

Random variable X is continuous and has an infinite number of expressions in the interval $a \leq X \leq b$. Description by its density function:

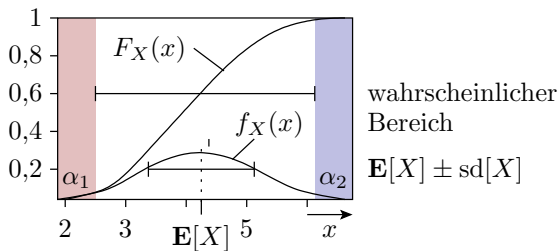
$$f_X(x) = \frac{dF_X(x)}{dx}$$

$$F_X(x) = \int_{-\infty}^x f(u) \cdot du \quad \text{für } X \in \mathfrak{R}$$



$F_X(x)$ Distribution of the random variable X .

$f_X(x)$ density function of the random variable X .



Erwartungswert:

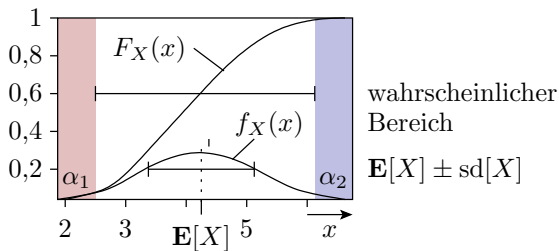
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx \quad (6)$$

Varianz:

$$\text{Var}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot (x - \mathbb{E}[X])^2 \cdot dx \quad (7)$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot x^2 \cdot dx - \mathbb{E}[X]^2$$

$$\text{Var}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x^2 \cdot dx - \mathbb{E}[X]^2$$



Probability that X lies in a range $[x_{\min}, x_{\max}]$:

$$\begin{aligned} \mathbb{P}(x_{\min} \leq x \leq x_{\max}) &= F_X(x_{\max}) - F_X(x_{\min}) \\ &= \int_{x_{\min}}^{x_{\max}} f_X(x) \cdot dx \end{aligned}$$

Range limits:

$$x_{\min} = F^{-1}(\alpha_1)$$

$$x_{\max} = F^{-1}(1 - \alpha_2)$$

α_1, α_2 Probability of error, that the values are below or above the estimated range.



$\mathbb{E}[X]$ and $\text{Var}[X]$ of a data sample

For a data sample of a random variable X

$$w = (w_1, w_2, \dots, w_{\#w})$$

the estimator for the expected value used in the following is the mean value:

$$\hat{\mathbb{E}}[X] = \frac{1}{\#w} \cdot \sum_{i=1}^{\#w} w_i \quad (8)$$

The estimator for the variance is the mean square deviation from the estimated expected value:

$$\hat{\text{Var}}[X] = \frac{1}{\#w - 1} \cdot \sum_{i=1}^{\#w} (w_i - \hat{\mathbb{E}}[X])^2 \quad (9)$$

Divisor is one less than sample size $\#w$, i.e. estimation of variance requires at least sample size $\#w = 2$.

$\#w$ Size of data sample.
. $\hat{\cdot}$. estimated value.



Linear transformation



Linear transformation

Linear transformations are the multiplication and addition of a random variable with real numbers. The expected value multiplies and increases by the same values:

$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \quad (10)$$

With the variance, the shift is omitted and the scaling factor enters in the square*:

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X] \quad (11)$$

In particular, the variance is shift invariant and remains the same when the distribution is mirrored:

$$\text{Var}[-X] = (-1)^2 \cdot \text{Var}[X] = \text{Var}[X]$$

* Proofs by recalculation follow as exercises.

a, b any real numbers.



Check by example

Realisation x from X	1	2	3
Realisation y from $Y = 5 - 2X$	3	1	-1
$\mathbf{P}(Y = y) = \mathbf{P}(X = x)$	0.3	0.5	0.2

$$\mathbb{E}[X] = 0,3 + 1 + 0,6 = 1,9$$

$$\text{Var}[X] = 0,3 + 2 + 1,8 - 1,9^2 = 0,49$$

$$\mathbb{E}[Y] = 0,9 + 0,5 - 0,2 = 1,2$$

$$\text{Var}[Y] = 2,7 + 0,5 + 0,2 - 1,2^2 = 1,96$$

$$\mathbb{E}[Y] = 5 - 2 \cdot \mathbb{E}[X]$$

$$\text{Var}[Y] = (-2)^2 \cdot \text{Var}[X]$$



Sum of random variables

The distribution of the sum of random variables assigns to each of the possible values of the sum the probability that the sum has that value (convolution):

x	1	3	4	y	2	3	4
$f_X(x)$	0.1	0.4	0.5	$f_Y(y)$	0.3	0.6	0.1

$$f_{X+Y} = f_X * f_Y :$$

$$\mathbb{P}(X + Y = 3) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 2)$$

$$\mathbb{P}(X + Y = 4) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 3)$$

$$\mathbb{P}(X + Y = 5) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 4) + \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 2)$$

$$\mathbb{P}(X + Y = 6) = \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 3) + \mathbb{P}(X = 4) \cdot \mathbb{P}(Y = 2)$$

$$\mathbb{P}(X + Y = 7) = \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 4) + \mathbb{P}(X = 4) \cdot \mathbb{P}(Y = 3)$$

$$\mathbb{P}(X + Y = 8) = \mathbb{P}(X = 4) \cdot \mathbb{P}(Y = 4)$$



For the sum of random variables, the expected value is equal to the sum of the expected values:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad (12)$$

The variance is the sum of the variances plus double covariance:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y] \quad (13)$$

with the covariance*:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \quad (14)$$

For independent random variables, the covariance is zero and the variance is the sum of the variances of the summands:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \quad (15)$$

* Proofs by recalculation follow as exercises.



Linearisation around the expectation value

Each random variable X is the sum of the expected value and a random deviation from the expected value with expected value zero:

$$X = \mathbb{E}[X] + X_{\Delta} \quad \text{with } \mathbb{E}[X_{\Delta}] = 0 \quad \text{and} \quad \text{Var}[X_{\Delta}] = \text{Var}[X]$$

Product of two independent random variables:

$$\begin{aligned} X \cdot Y &= (\mathbb{E}[X] + X_{\Delta}) \cdot (\mathbb{E}[Y] + Y_{\Delta}) \\ \frac{X \cdot Y}{\mathbb{E}[X] \cdot \mathbb{E}[Y]} &= 1 + \frac{X_{\Delta}}{\mathbb{E}[X]} + \frac{Y_{\Delta}}{\mathbb{E}[Y]} + \frac{X_{\Delta} \cdot Y_{\Delta}}{\mathbb{E}[X] \cdot \mathbb{E}[Y]} \end{aligned}$$

(~~xxx~~ – for small relative deviations negligible). Expectation value:

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (16)$$

For independent random variables, the variances add up according to eq. 3.15 die Varianzen. According to eq. 3.11 the scaling factors are squared and shifts are omitted:

$$\frac{\text{Var}[X \cdot Y]}{\mathbb{E}[X]^2 \cdot \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (17)$$



Quotient of two random variables:

$$\frac{X}{Y} = \frac{\mathbb{E}[X] + X_{\Delta}}{\mathbb{E}[Y] + Y_{\Delta}} = \frac{(\mathbb{E}[X] + X_{\Delta}) \cdot (\mathbb{E}[Y] - Y_{\Delta})}{(\mathbb{E}[Y] + Y_{\Delta}) \cdot (\mathbb{E}[Y] - Y_{\Delta})}$$

$$\frac{\frac{X}{\mathbb{E}[X]}}{\frac{Y}{\mathbb{E}[Y]}} = \frac{1 + \frac{X_{\Delta}}{\mathbb{E}[X]} - \frac{Y_{\Delta}}{\mathbb{E}[Y]} - \frac{X_{\Delta} \cdot Y_{\Delta}}{\mathbb{E}[X] \cdot \mathbb{E}[Y]}}{1 - \left(\frac{Y_{\Delta}}{\mathbb{E}[Y]}\right)^2} = 1 + \frac{X_{\Delta}}{\mathbb{E}[X]} - \frac{Y_{\Delta}}{\mathbb{E}[Y]}$$

$$\frac{X \cdot Y}{\mathbb{E}[X] \cdot \mathbb{E}[Y]} = 1 + \frac{X_{\Delta}}{\mathbb{E}[X]} + \frac{Y_{\Delta}}{\mathbb{E}[Y]} + \frac{X_{\Delta} \cdot Y_{\Delta}}{\mathbb{E}[X] \cdot \mathbb{E}[Y]}$$

(~~xxx~~ – for small relative deviations negligible). Expectation value:

$$\mathbb{E}[X/Y] = \mathbb{E}[X] / \mathbb{E}[Y] \quad (18)$$

For independent random variables, the relative variances also add up for small deviations when dividing:

$$\frac{\text{Var}[X/Y]}{\mathbb{E}[X/Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (19)$$

For a general function of several random variables $f(X_1, X_2, \dots)$, the expected value is the function of the expected values:

$$\mathbb{E}[f(X_1, X_2, \dots)] = f(\mathbb{E}[X_1], \mathbb{E}[X_2], \dots) \quad (20)$$



The weighting factors with which the individual variances enter into the total variance are the squares of the derivatives of the function according to the random variable at the point of the expected value:

$$\text{Var} [f (X_1, X_2, \dots)] = \sum_{i=1}^n \left(\frac{\partial (f (X, X_2, \dots))}{\partial X_i} \Big|_{\mathbb{VE}} \right)^2 \cdot \text{Var} [X_i] \quad (21)$$

Check on the example of multiplication and division:

$$\begin{aligned} \text{Var} [X \cdot Y] &= \left(\frac{\partial (X \cdot Y)}{\partial X} \Big|_{\mathbb{VE}} \right)^2 \cdot \text{Var} [X] + \left(\frac{\partial (X \cdot Y)}{\partial Y} \Big|_{\mathbb{VE}} \right)^2 \cdot \text{Var} [Y] \\ &= \mathbb{E} [Y]^2 \cdot \text{Var} [X] + \mathbb{E} [X]^2 \cdot \text{Var} [Y] \end{aligned}$$

$$\frac{\text{Var}[X \cdot Y]}{\mathbb{E}[X]^2 \cdot \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \sqrt{\quad}$$

$$\begin{aligned} \text{Var} [X/Y] &= \left(\frac{\partial (X/Y)}{\partial X} \Big|_{\mathbb{VE}} \right)^2 \cdot \text{Var} [X] + \left(\frac{\partial (X/Y)}{\partial Y} \Big|_{\mathbb{VE}} \right)^2 \cdot \text{Var} [Y] \\ &= \left(\frac{1}{\mathbb{E}[Y]} \right)^2 \cdot \text{Var} [X] + \left(-\frac{\mathbb{E}[X]}{\mathbb{E}[Y]^2} \right)^2 \cdot \text{Var} [Y] \end{aligned}$$

$$\frac{\text{Var}[X/Y]}{\mathbb{E}[X]^2 / \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \sqrt{\quad}$$



Measured value and measuring error

In metrology, the following applies to every measured value:

$$X_M = X + X_F \quad (22)$$

All three random variables have each an expectation value and a variance. With the measured value and the measurement error as independent random variables, the following applies to them:

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (23)$$

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (24)$$

- $\mathbb{E}[X_F]$ – Expected value, systematic measurement error
- $\text{sd}[X_F] = \sqrt{\text{Var}[X_F]}$ – Standard deviation, random measurement error.

X_M	random variable measured value.
X	random variable value to measure.
X_F	Random variable measuring error.
$\mathbb{E}[\dots]$	expectation value of the random variable.
$\text{Var}[\dots]$	standard deviation of the random variable.



Example 3.1: Measuring error

The measured value R_M of a resistor batch is on average 1010Ω and has a standard deviation of 11.18Ω . The measurement has a systematic fault R_F of 12Ω and a standard deviation of 5Ω . What are the expectation value and standard deviation of the value to be measured?

$$\mathbb{E}[R_M] = 1010 \Omega, \text{sd}[R_M] = 11.18 \Omega, \mathbb{E}[R_F] = 12 \Omega \text{ und } \text{sd}[R_F] = 5 \Omega.$$

X	Resistance value to be measured.
R_F	measuring error of resistance value.
R_M	measured resistance value.



$$\mathbb{E}[R_M] = 1010 \Omega, \text{sd}[R_M] = 11.18 \Omega, \mathbb{E}[R_F] = 12 \Omega \text{ und } \text{sd}[R_F] = 5 \Omega.$$

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (3.23)$$

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (3.24)$$

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (3.4)$$

The random variable in the exercise is R instead of X and the expected value and standard deviation of the actual value are required:

$$\mathbb{E}[R] = \mathbb{E}[R_M] - \mathbb{E}[R_F] = 1010 \Omega - 12 \Omega = 998 \Omega$$

$$\text{Var}[R] = \text{Var}[R_M] - \text{Var}[R_F] = (11.18 \Omega)^2 - (5 \Omega)^2 = 100 \Omega^2$$

$$\text{sd}[R] = 10 \Omega$$

The (actual) measured value has a smaller standard deviation than the measured value.

X	Resistance value to be measured.
R_F	measuring error of resistance value.
R_M	measured resistance value.



Distribution of counts



Distribution of counts

A random count value X , e.g. the number of correctly performed or incorrectly performed services, can be expressed as a sum of

$$X = \sum_{i=1}^{\#X} X_i$$

»potential count values« X_i with the Bernoulli distribution:

$$\mathbb{P}(X_i = k) = \begin{cases} 1 - p_i & k = 0 \\ p_i & k = 1 \end{cases}$$

Count value X	potential count values $X_i \in \{0, 1\}$
malfunctions	service requests
faults	potential faults
detectable faults	existing faults
...	



Expected value and variance

k	0	1
$\mathbf{P}[X_i = k]$	$1 - p_i$	p_i

Expected values of the single events:

$$\mathbb{E}[X_i] = (1 - p_i) \cdot 0 + p_i \cdot 1 = p_i$$

Variance according to the shift rule:

$$\text{Var}[X_i] = (1 - p_i) \cdot 0^2 + p_i \cdot 1^2 - p_i^2 = p_i \cdot (1 - p_i)$$



The expected value of the sum is the sum of the expected values (eq. 3.10):

$$\mathbb{E}[X] = \sum_{i=1}^{\#X} p_i \quad (25)$$

For the variance, it is often assumed that the events to be counted, such as the occurrence of different malfunctions, do not depend on each other (variance of the sum equals the sum of the variances of the summands, covariance zero, eq. 3.12):

$$\text{Var}[X] = \sum_{i=1}^{\#X} p_i \cdot (1 - p_i) \quad (26)$$



Distribution calculation

When adding the next Bernoulli distributed random variable X_i , the count value increases by one with probability p_i and remains the same with probability $1 - p_i$:

$$\mathbb{P}_i(X = k) = p_i \cdot \mathbb{P}_{i-1}(X = k - 1) + (1 - p_i) \cdot \mathbb{P}_{i-1}(X = k)$$

Distribution calculation:

	p_i	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\mathbb{P}[X_1 = k]$	30%	70%	30%			
$\mathbb{P}[X_1 + X_2 = k]$	50%	35%	50%	15%		
$\mathbb{P}[\dots + X_3 = k]$	40%	21%	44%	29%	6%	
$\mathbb{P}[\dots + X_4 = k]$	10%	18.9%	41.7%	30.5%	8.3%	0.6%

$$\mathbb{P}_1(X = 0) = 1 - p_1$$

$$\mathbb{P}_1(X = 1) = p_1$$

Repeat for $i = 2$ to N

$$\mathbb{P}_i(X = 0) = \mathbb{P}_{i-1}(X = 0) \cdot (1 - p_i)$$

$$\mathbb{P}_i(X = i) = \mathbb{P}_{i-1}(X = i - 1) \cdot p_i$$

Repeat for $k = 1$ to $i - 1$

$$\begin{aligned} \mathbb{P}_i(X = k) &= \mathbb{P}_{i-1}(X = k) \cdot (1 - p_i) \\ &\quad + \mathbb{P}_{i-1}(X = k - 1) \cdot p_i \end{aligned}$$

$\mathbb{P}_i(X = k)$ Probability that the sum of the first i summands is k .



Expected value and variance for the example

p_i	30%	50%	40%	10%	
k	0	1	2	3	4
$P[X = k]$	18.9%	41.7%	30.5%	8.3%	0.6%

Expected value of the sum of all $N = 4$ summands:

$$\mathbb{E}[X] = 18.9\% \cdot 0 + 41.7\% \cdot 1 + 30.5\% \cdot 2 + 8.3\% \cdot 3 + 0.6\% \cdot 4 = 1.3$$

As the sum of all p_i according to eq. 3.25 the calculation is shorter:

$$\mathbb{E}[X] = 30\% + 50\% + 40\% + 10\% = 1.3$$

Variance is according to the shift rule eq. 3.5:

$$18.9\% \cdot 0^2 + 41.7\% \cdot 1^2 + 30.5\% \cdot 2^2 + 8.3\% \cdot 3^2 + 0.6\% \cdot 4^2 - 1.3^2 = 0.79$$

Simplified calculation according to eq. 3.26:

$$\text{Var}[X] = 0.3 \cdot 0.7 + 0.5 \cdot 0.5 + 0.4 \cdot 0.6 + 0.1 \cdot 0.9 = 0.79$$



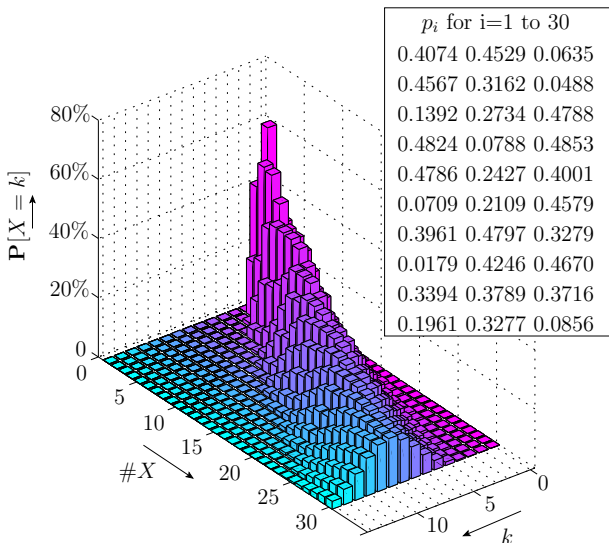
Example of a count distribution

Counting distribution
calculated step-by-
step with Matlab.
The probabilities of
occurrence of the
counting events
are shown in the
box in the picture.
Expected value
and variance for
all 30 summands:

$$\mathbb{E}[X] = 7.05$$

$$\text{Var}[X] = 2.19$$

Likely range about
5 to 15.





Summary



Random variable and distribution

A random variable is a variable for the result of a random experiment.

The distribution assigns probabilities to the possible values of a discrete random variable.

The distribution function $F_X(x)$ indicates the probability that a random variable X has at most the value x :

$$F_X(x) = \mathbb{P}(X \leq x)$$

The density function describes the increase of the distribution function with x for a continuous random variable:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Bernoulli distribution:

$$\begin{aligned}\mathbb{P}(X = 0) &= 1 - p \\ \mathbb{P}(X = 1) &= p\end{aligned}\tag{3.1}$$

Expected value, variance, standard deviation

Expected value of a discrete distribution:

$$\mathbb{E}[X] = \sum_{i=1}^m p_i \cdot x_i \quad (3.2)$$

Variance of a discrete distribution:

$$\text{Var}[X] = \sum_{i=1}^m p_i \cdot (x_i - \mathbb{E}[X])^2 \quad (3.3)$$

Standard deviation:

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (3.4)$$

Shift rule:

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \quad (3.5)$$

Expected value of a continuous distribution:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx \quad (3.6)$$

Variance of a continuous distribution:

$$\text{Var}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x^2 \cdot dx - \mathbb{E}[X]^2 \quad (3.7)$$

Expected value and variance of a data sample

Estimator for the expected value:

$$\hat{\mathbb{E}}[X] = \bar{w} = \frac{1}{\#w} \cdot \sum_{i=1}^{\#w} w_i \quad (3.8)$$

Estimator for the variance:

$$\hat{\text{Var}}[X] = \frac{1}{\#w-1} \cdot \sum_{i=1}^{\#w} \left(w_i - \hat{\mathbb{E}}[X] \right)^2 \quad (3.9)$$



Linear transformations and sums

Expected value of a linear transformation:

$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \quad (3.10)$$

Variance of a linear transformation:

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X] \quad (3.11)$$

Expected value of the sum of two random variables:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad (3.12)$$

Variance of the sum of two random variables:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y] \quad (3.13)$$

Definition of covariance:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \quad (3.14)$$

Variance of a sum of independent random variables:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \quad (3.15)$$

For small deviations from the expected value

Expected value of the product of two independent random variables:

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (3.16)$$

Variance of the product of two independent random variables:

$$\frac{\text{Var}[X \cdot Y]}{\mathbb{E}[X]^2 \cdot \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.17)$$

Expected value of the quotient of two independent random variables:

$$\mathbb{E}[X/Y] = \mathbb{E}[X] / \mathbb{E}[Y] \quad (3.18)$$

Variance of the quotients of two independent random variables:

$$\frac{\text{Var}[X/Y]}{\mathbb{E}[X/Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.19)$$

Measured values and counts

Expected value of measured values:

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (3.23)$$

Variance of measured values:

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (3.24)$$

Expected value of counts:

$$\mathbb{E}[X] = \sum_{i=1}^{\#X} p_i \quad (3.25)$$

Variance of independent counts:

$$\text{Var}[X] = \sum_{i=1}^{\#X} p_i \cdot (1 - p_i) \quad (3.26)$$

Calculation of the distribution of count values:

- Starting with the two-point distribution for the first count value
- Calculate the distribution for n counts by convolving the distribution for $n - 1$ counts with the two-point distribution for count value n .



CVD Approximations



Binomial distribution



Binomial distribution

$$\binom{4}{2} = \frac{4 \cdot 3}{2!}$$

For the special case that equally probable events are counted, the sum of the counted events is binomially distributed

$$X \sim \text{Bin}(n, p)$$

Binomial distribution:

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (27)$$

Expected value of a binomial distribution:

$$\mathbb{E}[X] = n \cdot p \quad (28)$$

Variance and standard deviation of a binomial distribution:

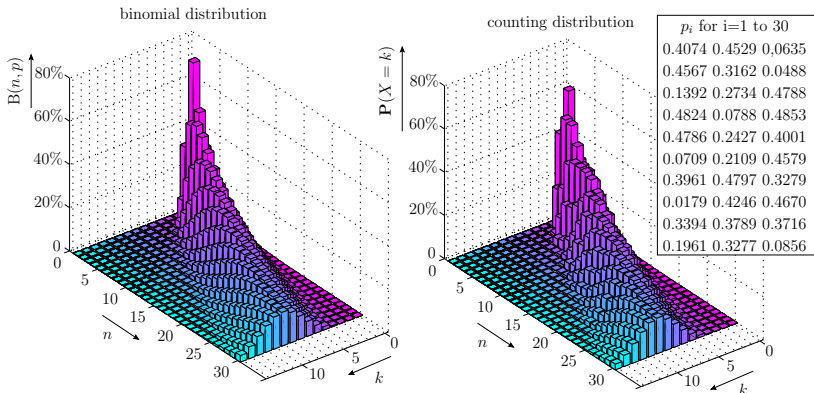
$$\text{Var}[X] = n \cdot p \cdot (1 - p) \quad (29)$$

$$\text{sd}[X] = \sqrt{n \cdot p \cdot (1 - p)} \quad (30)$$

n (# X) Number of counting attempts, or summed random variables.
 p probability of occurrence.



Binomial vs. general counting distribution



A binomial distribution with $p = \frac{1}{n} \cdot \sum_{i=1}^n p_i$ and $n = \#X$ approximates a counting distribution well and is calculated from only the two parameters n and p .

 n

($\#X$) Number of counting attempts, or summed random variables.

**Example 3.2: Range estimation for the number of faults**

The average probability of detecting 10 faults is 30%. What is the probability that at least two out of 10 faults are detected?

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.27)$$

At least two faults:

$$\begin{aligned} \mathbb{P}(X \geq 2) &= \mathbb{P}(X \leq 1) \\ &= 1 - \sum_{k=0}^1 \binom{10}{k} \cdot 0,3^k \cdot (1 - 0,3)^{10-k} \\ &= 1 - (0,7^{10} + 10 \cdot 0,3 \cdot 0,7^9) \\ &\approx 85\% \end{aligned}$$



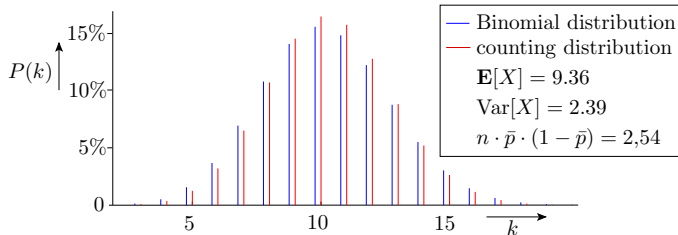
Upper variance limit

Theorem

With the same number of independent count values $n = \#X$ and $\bar{p} = \frac{1}{n} \cdot \sum_{i=1}^n p_i$, the variance of the binomial distribution is an upper bound of the variance of a general counting distribution:

$$\text{Var}[X] = \sum_{i=1}^n p_i \cdot (1 - p_i) \leq n \cdot \bar{p} \cdot (1 - \bar{p}) \quad (31)$$

For the distributions two slides before, for $n = 30$:





Proof

Replace the individual probabilities of occurrence of the events to be counted by the mean probability and a difference that is zero on average:

$$p_i = \bar{p} + \delta_i \text{ with } \sum_{i=1}^n \delta_i = 0$$

Variance of the count distribution:

$$\begin{aligned} \text{Var}[X] &= \sum_{i=1}^n (\bar{p} + \delta_i) \cdot (1 - \bar{p} - \delta_i) \\ &= \sum_{i=1}^n (\bar{p} - \bar{p}^2 - \bar{p} \cdot \delta_i + \delta_i - \bar{p} \cdot \delta_i - \delta_i^2) \\ &= \underbrace{\sum_{i=1}^n (\bar{p} - \bar{p}^2)}_{n \cdot \bar{p} \cdot (1 - \bar{p})} + \underbrace{\sum_{i=1}^n (\delta_i - 2 \cdot \bar{p} \cdot \delta_i)}_{(1 - 2\bar{p}) \cdot \sum_{i=1}^n \delta_i = 0} - \underbrace{\sum_{i=1}^n \delta_i^2}_{\geq 0} \end{aligned}$$

$$\text{Var}[X] \leq n \cdot \bar{p} \cdot (1 - \bar{p}) \text{ (variance binomial distribution)} \sqrt{\quad}$$



Fact

Probable ranges for (independent) counting results (faults, malfunctions, ...) calculated from the mean probability of occurrence

$$\bar{p} = \frac{1}{n} \cdot \sum_{i=1}^n p_i$$

via a binomial distribution are easier to calculate and, if the probabilities of error are the same, are upper limits or, if the range size is the same, the error probabilities are upper limits.

p_i	probability of occurrence of counting event i .
n	(# X) Number of counting attempts, or summed random variables.
\bar{p}	average probability of occurrence.



Poisson distribution



Poisson distribution

When counting many rare events, e.g. the malfunctions in millions of service requests, of which only a few occur, the probability of occurrence of the individual events and the deviation of the variance from the expected value tend towards zero:

$$\begin{aligned} p_i &\rightarrow 0 \\ \text{Var}[X_i] - \mathbb{E}[X_i] &= p_i \cdot (1 - p_i) - p_i = p_i^2 \rightarrow 0 \end{aligned}$$

The variance of the events to be counted and that of the sum tend towards the expected value:

$$\begin{aligned} \text{Var}[X_i] &= \mathbb{E}[X_i] \\ \text{Var}[X] &= \sum_{i=1}^{\#X} \text{Var}[X_i] = \mathbb{E}[X] = \sum_{i=1}^{\#X} \mathbb{E}[X_i] = \lambda \end{aligned}$$

The distribution of the sum tends towards the Poisson distribution:

$$X \sim \text{Pois}(\lambda)$$



The Poisson distribution

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad (32)$$

has only the parameter

$$\lambda = \mathbb{E}[X] = \text{Var}[X] = \sum_{i=1}^n p_i$$

which is the sum of the probabilities of occurrence that a "potential count value" is one, i.e. a "real count value". This parameter is also expectation value and variance.

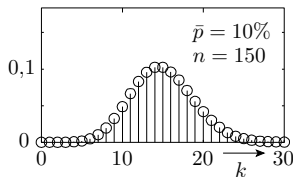
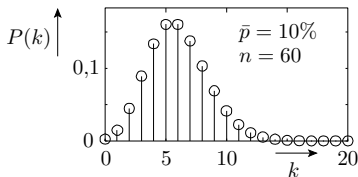
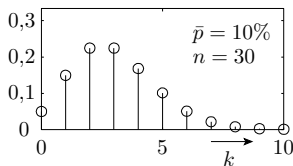
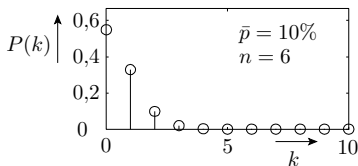
A Poisson distribution with $\lambda = n \cdot p$ approximates a counting distribution well for $p \ll 1$. Estimated range with equal error probabilities guaranteed larger than with actual or binomial distribution.

- λ Parameter, expected value and variance Poisson distribution.
- n ($\#X$) Number of counting attempts, or summed random variables.
- p_i probability of occurrence of counting event i .

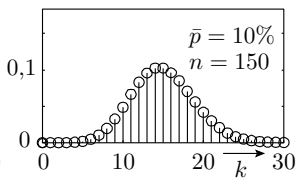
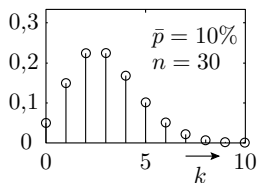
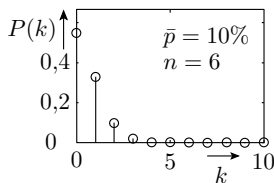


Number of count attempts and distribution

$$\mathbb{P}(X = k) = P(k) = e^{-\bar{p} \cdot n} \cdot \frac{(\bar{p} \cdot n)^k}{k!} \quad (33)$$



- λ Parameter, expected value and variance Poisson distribution.
- n (# X) Number of counting attempts, or summed random variables.
- \bar{p} average probability of occurrence.



Rough estimate of the likely ranges:

- For $\lambda = \mathbb{E}[X] = \text{Var}[X] p \cdot n < 3$ no lower bound $x_{\min} > 0$. Upper bound:

$$k_{\max} > 3 \dots 5 \cdot \lambda$$

- For $\mathbb{E}[X] \approx 3 \dots 10$ also lower bound:

$$k_{\max} < \frac{\lambda}{3 \dots 5}$$

- For $\mathbb{E}[X] > 10$ estimation via normal distribution (see later sec. 3.2.5):

$$[k_{\min}, k_{\max}] \approx \mathbb{E}[X] \cdot \left(1 \mp 2 \dots 4 \cdot \sqrt{\lambda}\right)$$



Circuit yield and fault count

The expected circuit yield is the probability that the number of detectable faults from the originating processes is zero. For a Poisson distributed number of faults according to eq. 3.32 with $\lambda = FC \cdot \mu_{FCP}$:

$$Y = e^{-FC \cdot \mu_{FCP}} \quad (34)$$

For a circuit classified as fault-free

$$\frac{1}{Y} = e^{FC \cdot \mu_{FCP}}$$

Circuits must be manufactured.

Y	yield.
FC	fault coverage, percentage of detectable faults.
μ_{FCP}	expected number of faults from creation process.



Production costs and defect count

Assuming a proportional increase in the expected number of faults and the pure manufacturing costs per circuit with the number of transistors:

$$\mu_{FCP} = \xi_{Tr} \cdot \#Tr$$

$$C_{MIC} = C_{Tr} \cdot \#Tr$$

(see eq. 1.71). Cost per circuit not sorted out as defective:

$$C_{IS} = \frac{C_{ICF}}{Y} = C_{Tr} \cdot \#Tr \cdot e^{FC \cdot \xi_{Tr} \cdot \#Tr} \quad (35)$$

$\#Tr$	number of t ransistors.
ξ_{Tr}	Defect emergence rate in faults per t r.
C_{Tr}	transistor costs in Euro per t ransistor.
C_{MIC}	m anufacturing costs per i ntegrated c ircuit.
C_{IC}	overall costs per i ntegrated c ircuit.

**Example 3.3: Circuit costs**

Defect emergence rate $\xi_{Tr} = 10^{-5}$ [fault per transistor], manufacturing cost $C_{Tr} = 10^{-6}$ [Euro per transistor], circuit size $\#Tr \in \{10^5, 10^6, 10^7\}$ [transistors], fault coverage $FC = 1$.

What is the expected cost per sellable circuit?

$\#Tr$	10^4	10^5	10^6
$\mu_{FCP} = \xi_{Tr} \cdot \#Tr$ in faults	0,1	1	10
$C_{MIC} = C_{Tr} \cdot \#Tr$ in euro	1	10	10
$Y = e^{-FC \cdot \mu_{FCP}}$	90.5%	36.8%	$4.54 \cdot 10^{-5}$
$C_{IC} = \frac{C_{MIC}}{Y}$ in euro	0.11	2.72	$2.2 \cdot 10^5$

From $\mu_{FCP} \geq 2 \dots 3$ expected faults per circuit, circuits need repair facilities:

- deactivate / replace defective function blocks,
- Sale e.g. as processors with less cache, ...

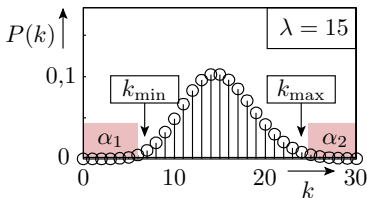


Range, Poisson

Estimate k_{\min}

Input k_{\min} and α_1 . Numerical search of $\lambda(k_{\min}, \alpha_1)$, such that

$$\sum_{k=0}^{k_{\min}-1} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \leq \alpha_1$$



α_1	$k_{\min} = 1$	2	3	4	5	6
0.5%	5.298	7.430	9.273	10.978	12.593	14.150
1%	4.606	6.638	8.406	10.045	11.605	13.109
2%	3.912	5.834	7.516	9.084	10.580	12.027
10%	2.303	3.890	5.323	6.681	7.993	9.275
20%	1.609	2.995	4.279	5.514	6.721	7.906

Example estimates:

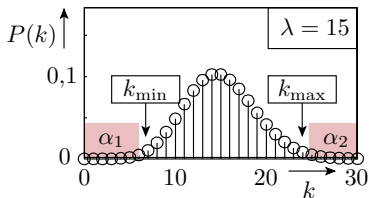
- $\lambda = 7$ und $\alpha_1 \leq 1\% \Rightarrow k_{\min} = 2$
- $k_{\min} = 1$ und $\alpha_1 = 2\% \Rightarrow \lambda \geq 3,912$

α_1 Probability of error, that the values are below the estimated range.
 k_{\min} lower range limit.

Estimate k_{\max}

Input k_{\max} und α_2 . Numerical search of $\lambda(k_{\max}, \alpha_2)$, such that

$$\sum_{k=0}^{k_{\max}} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \geq 1 - \alpha_2$$

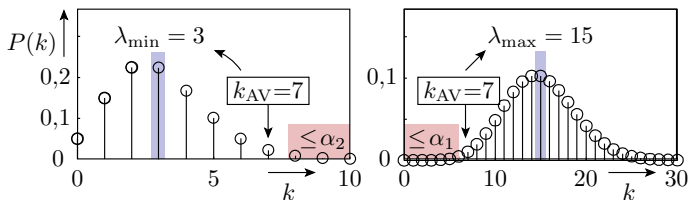


α_2	$k_{\max} = 0$	1	2	3	4	5	6
0.5%	0.005	0.103	0.338	0.672	1.078	1.537	2.037
1%	0.01	0.148	0.436	0.823	1.279	1.785	2.330
2%	0.02	0.215	0.567	1.016	1.529	2.089	2.684
10%	0.105	0.532	1.102	1.744	2.432	3.152	3.894
20%	0.223	0.824	1.534	2.296	3.089	3.903	4.733

Example estimates:

- $\lambda = 2$ und $\alpha_2 \leq 1\% \Rightarrow k_{\max} = 6$
- $k_{\max} = 3$ und $\alpha_2 = 2\% \Rightarrow \lambda \leq 1,016$

α_2 Probability of error, that the values are above the estimated range.
 k_{\min} upper range limit.

Estimate $[\lambda_{\min}, \lambda_{\max}]$ from an actual count value

Values $[\lambda_{\min}, \lambda_{\max}] = f(\alpha, k_{AV})$ taken from the slides before:

$\alpha_1 = \alpha_2$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
0.5%	[0.10, 5.30]	[0.34, 7.43]	[0.67, 9.27]
1%	[0.15, 4.60]	[0.44, 6.64]	[0.82, 8.41]
2%	[0.22, 3.91]	[0.57, 5.83]	[1.02, 7.52]
10%	[0.53, 2.30]	[1.10, 3.89]	[1.74, 5.32]
20%	[0.82, 1.61]	[1.53, 2.99]	[2.30, 4.28]

k_{AV} actual count value.

α_1, α_2 Probability of error, that the values are below or above the estimated range.



$[\lambda_{\min}, \lambda_{\max}] = f(\alpha, k_{AV})$ for actual count values $k_{AV} = 4$ to 6:

$\alpha_1 = \alpha_2$	$k_{ACV} = 4$	$k_{AV} = 5$	$k_{AC} = 6$
0.5%	[1.08, 11.0]	[1.54, 12.6]	[2.04, 14.2]
1%	[1.28, 10.0]	[1.79, 11.6]	[2.33, 13.1]
2%	[1.53, 9.08]	[2.09, 10.6]	[2.68, 12.0]
10%	[2.43, 6.68]	[3.15, 7.99]	[3.89, 9.28]
20%	[3.09, 5.51]	[3.90, 6.73]	[4.73, 7.91]

$\lambda_{\max} = f(\alpha_1, k_{AV})$ for actual count values $k_{AV} = 0$ ($\lambda_{\min} = 0$):

$$\sum_{k=0}^{\infty} e^{-\lambda_{\max}} \cdot \frac{\lambda_{\max}^k}{k!} = e^{-\lambda_{\max}} = \alpha_1$$

$$\lambda_{\max} = -\ln(\alpha_1)$$

α_1	0.5%	1%	2%	10%	20%
λ_{\max}	5.30	4.61	3.91	2.30	1.61%

k_{AV} actual count value.

k_{\min} lower range limit.

k_{\max} upper range limit.

* For $k_{AV} = 0$ ist k_{\min} always zero.



Example 3.4: Poisson distributed number of claims

- a) No damage has occurred in the past 10 years. What is the expected number of claims in the next 10 years with a probability of error of $\alpha_1 = 1\%$?
- b) In one year of use, 5 damage events occurred. What range of the expected number of damage events can be estimated from this information for the next 10 years of use with the error probabilities $\alpha_1 = \alpha_2 = 1\%$?



- a) No damage has occurred in the past 10 years. What is the expected number of claims in the next 10 years with a probability of error of $\alpha_1 = 1\%$?

α_1	0.5%	1%	2%	10%	20%
λ_{\max}	5.30	4.61	3.91	2.30	1.61%

At least $\lambda_{\min} = 0$ and maximum $\lambda_{\max} = 4,61$.

α_1	Probability of error, that the values are below the estimated range.
k_{\min}	upper range limit.
λ_{\max}	maximum expected value for a poisson distributed actual value.



- b) In one year of use, 5 damage events occurred. What range of the expected number of damage events can be estimated from this information for the next 10 years of use with the error probabilities $\alpha_1 = \alpha_2 = 1\%$?

$\alpha_1 = \alpha_2$	$k_{ACV} = 4$	$k_{AV} = 5$	$k_{AC} = 6$
0.5%	[1.08, 11.0]	[1.54, 12.6]	[2.04, 14.2]
1%	[1.28, 10.0]	[1.79, 11.6]	[2.33, 13.1]
2%	[1.53, 9.08]	[2.09, 10.6]	[2.68, 12.0]
10%	[2.43, 6.68]	[3.15, 7.99]	[3.89, 9.28]
20%	[3.09, 5.51]	[3.90, 6.73]	[4.73, 7.91]

At least $10 \cdot \lambda_{\min} = 17,9$ and maximum $10 \cdot \lambda_{\max} = 116$.



Example 3.5: Malfunction rate

With $n = 10^5$ service requests, three malfunctions have been observed. What are the lower and upper bounds for the error rate that can be inferred from the error probabilities $\alpha_1 = \alpha_2 = 1\%$ assuming a Poisson distribution?

$\alpha_1 = \alpha_2 = 1\%$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$[\lambda_{\min}, \lambda_{\max}]$	[0.15, 4.60]	[0.44, 6.64]	[0.82, 8.41]

Estimable range of the malfunction rate:

$$\zeta_{\min} = \frac{\lambda_{\min}}{n} = 0.82 \cdot 10^{-5} \text{ [MF/DS]}$$

$$\zeta_{\max} = \frac{\lambda_{\max}}{n} = 8.41 \cdot 10^{-5} \text{ [MF/DS]}$$

Small counts allow only rough estimates. More accurate estimates require larger counts.

k_{AV} actual count value.
 $\left[\frac{\text{MF}}{\text{DS}}\right]$ value in **m**alfunction per **d**elivered **s**ervice.



Example 3.6: Masking probability

A monitoring unit has failed to detect $k_{AV} = 5$ out of $\#MF = 10,000$ malfunctions. What is the range of the masking probability p_M with an error probability of $\alpha_1 = \alpha_2 = 1\%$?

$\alpha_1 = \alpha_2 = \frac{\alpha}{2} = 0,5\%$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$[\lambda_{\min}, \lambda_{\max}]$	$[1.08, 11.0]$	$[1.54, 12.6]$	$[2.04, 14.2]$

Estimable range of masking probability:

$$p_{M. \min} = \frac{\lambda_{\min}}{\#MF} = 1.54 \cdot 10^{-4}$$

$$p_{M. \max} = \frac{\lambda_{\max}}{\#MF} = 12.6 \cdot 10^{-4}$$

$\#MF$	number of malfunctions .
k_{AV}	actual count value.
α_1, α_2	Probability of error, that the values are below or above the estimated range.
$\lambda_{\min/\max}$	minimum or maximum expected value for a poisson distributed actual value.
p_M	probability of masking.
$[\frac{MF}{DS}]$	value in m alfunction per d elivered s ervice.



Example 3.7: Range of reliability

When testing a system with $\#DS = 10^3$ services, $k_{AV} = 6$ malfunctions were observed. What is the range of reliability resulting from this experiment on an error probability $\alpha_1 = \alpha_2 = 10\%$?

$\alpha_1 = \alpha_2 = \frac{\alpha}{2} = 10\%$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$[\lambda_{\min}, \lambda_{\max}]$	[2.43, 6.68]	[3.15, 7.99]	[3.89, 9.28]

- Estimable range of the MF rate:

$$\zeta_{\min} = 3.89 \cdot 10^{-3} \text{ [MF/DS]}$$

$$\zeta_{\max} = 9.28 \cdot 10^{-3} \text{ [MF/DS]}$$

- Assessable range of reliability:

$$R_{\min} = \frac{1}{\zeta_{\max}} = 108 \text{ [DS/MF]}$$

$$R_{\max} = \frac{1}{\zeta_{\min}} = 257 \text{ [DS/MF]}$$

R reliability.
 $\left[\frac{DS}{MF}\right]$ value in delivered services per malfunction.



Normal distribution

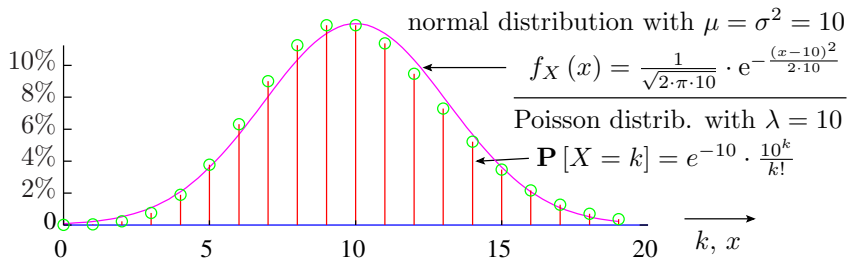
Normal distribution

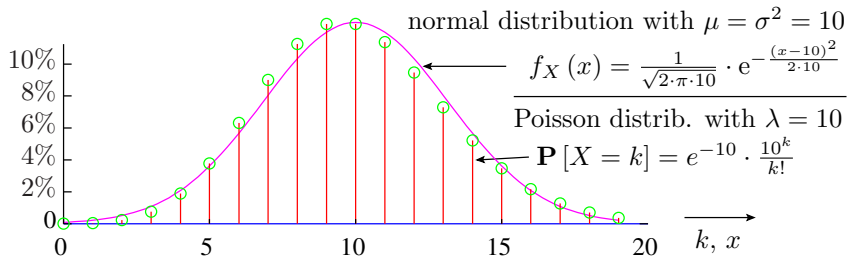
The sum of very many independent random variables tends towards a normal distribution under very general conditions:

- no summand has dominant influence, ...

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \quad \text{mit } \sigma = \text{sd}[X], \mu = \mathbb{E}[X] \quad (36)$$

Example: Poisson and normal distribution with $\mu = \sigma^2 = \lambda = 10$:





For independent count values, the approximation of the count distribution by a normal distribution is usually already sufficient under the condition

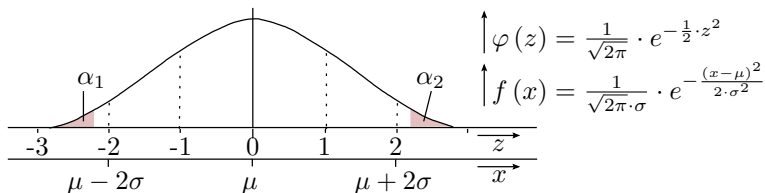
$$10 \leq \mu \leq n - 10 \quad \text{mit} \quad \mu = \sum_{i=1}^n p_i$$

- n (# X) Number of counting attempts, or summed random variables.
- p_i probability of occurrence of counting event i .
- μ expectation value.
- σ standard deviation.



Range estimate ND

Range estimation with normal distribution



- Transformation of a random variable X with expected value μ and standard deviation σ into random variable Z with expected value zero and standard deviation one::

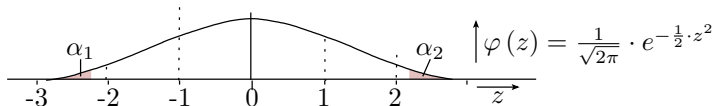
$$Z = \frac{X - \mu}{\sigma} \quad (37)$$

- Transform the values x of X into z of Z so that $F_Z(z) = \Phi(z)$:

$$z = \frac{x - \mu}{\sigma} \quad (38)$$

- Reading the error probabilities from a table of the standardised normal distribution function.

Standardised normal distribution



Distribution function table for $z \geq 0$ in steps of 0,1:

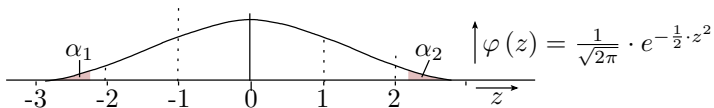
$$\Phi(z) = \int_{-\infty}^z \varphi(z) \cdot dz$$

z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7257	0.7580	0.7881	0.8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0,9772	0,9821	0,9861	0,9893	0,9918	0,9938	0,9953	0,9965	0,9974	0,9981
3,...	0,9987	0,9990	0,9993	0,9995	0,9997	0,9998	0,9998	0,9999	0,9999	1,0000

Due to symmetry for $z < 0$ applies:

$$\Phi(-z) = 1 - \Phi(z)$$

Inverse standardised normal distribution



z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7257	0.7580	0.7881	0.8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0.9772	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981
3,...	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Inverse standardised normal distribution for range estimation::

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10



Probability of range affiliation

z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7257	0.7580	0.7881	0.8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0.9772	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981
3,...	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Starting from a normally distributed random variable X with expected value μ and standard deviation σ and a range with the limits x_{\min} and/or x_{\max} :

- 1 Transformation of the range limit according to (eq. 3.38)

$$z_{\min} = \frac{x_{\min} - \mu}{\sigma} \quad (39)$$

$$z_{\max} = \frac{x_{\max} - \mu}{\sigma} \quad (40)$$

- 2 Take $\Phi(z)$ or, for $z < 0$, $\Phi(-z)$ from the table.
- 3 Determination of the error probabilities:

$$\alpha_1 = \Phi(z_{\min}) = 1 - \Phi(-z_{\min}) = 1 - \Phi\left(\frac{\mu - x_{\min}}{\sigma}\right) \quad (41)$$

$$\alpha_2 = 1 - \Phi(z_{\max}) = 1 - \Phi\left(\frac{x_{\max} - \mu}{\sigma}\right) \quad (42)$$



Likely range

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10

Starting from normally distributed random variable X with exp. value μ and standard deviation σ and permissible error probabilities α_1 and α_2 :

1 Reading from the table:

$$z_{\min} = \Phi^{-1}(\alpha_1) = -\Phi^{-1}(1 - \alpha_1) \quad (43)$$

$$z_{\max} = \Phi^{-1}(1 - \alpha_2) \quad (44)$$

2 Transformation:

$$x_{\min} = \mu + \sigma \cdot z_{\min} = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (45)$$

$$x_{\max} = \mu + \sigma \cdot z_{\max} = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (46)$$

■ Special case symmetrical range:

$$\text{sr}(x) = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (47)$$

sr (...) symmetrical range of probable values.



Example 3.8: Range estimation normal distribution

Random variable X , $\mu = 20$, $\sigma = 5$.

- What is the probability of $X \geq 30$?
- What is the probability of $X \leq 15$?
- Which upper bound x_{\max} is exceeded only with an error probability $\alpha_2 \leq 1\%$?
- Which lower bound x_{\min} is undershot only with an error probability $\alpha_1 \leq 2\%$?

μ expectation value.

σ standard deviation.

α_1, α_2 Probability of error, that the values are below or above the estimated range.



Random variable X , $\mu = 20$, $\sigma = 5$.

a) What is the probability of $X \geq 30$?

z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7257	0.7580	0.7881	0.8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0.9772	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981
3,...	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

$$\alpha_2 = 1 - \Phi\left(\frac{x_{\max} - \mu}{\sigma}\right) \quad (3.42)$$

$$\alpha_2 = 1 - \Phi\left(\frac{30 - 20}{5}\right)$$

$$= 1 - \Phi(2)$$

$$\alpha_2 = 2.27\%$$



Random variable X , $\mu = 20$, $\sigma = 5$.

b) What is the probability of $X \leq 15$?

z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7257	0.7580	0.7881	0.8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0.9772	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981
3,...	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

$$\alpha_1 = 1 - \Phi\left(\frac{\mu - x_{\min}}{\sigma}\right) \quad (3.41)$$

$$\alpha_1 = 1 - \Phi\left(\frac{20 - 15}{5}\right)$$

$$= 1 - \Phi(1)$$

$$\alpha_1 = 15.87\%$$



Random variable X , $\mu = 20$, $\sigma = 5$.

c) Which upper bound x_{\max} is exceeded only with an error probability $\alpha_2 \leq 1\%$?

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10

$$x_{\max} = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.46)$$

$$\begin{aligned} x_{\max} &= 20 + 5 \cdot \Phi^{-1}(1 - 1\%) \\ &= 20 + 5 \cdot 2.33 \\ x_{\max} &= 31.65 \end{aligned}$$

-
- μ expectation value.
 - σ standard deviation.
 - α_1 Probability of error, that the values are below the estimated range.



Random variable X , $\mu = 20$, $\sigma = 5$.

d) Which lower bound x_{\min} is undershot only with an error probability $\alpha_1 \leq 2\%$?

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10

$$x_{\min} = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.45)$$

$$\begin{aligned} x_{\max} &= 20 + 5 \cdot \Phi^{-1}(1 - 2\%) \\ &= 20 + 5 \cdot 2.05 \\ x_{\max} &= 9.75 \end{aligned}$$

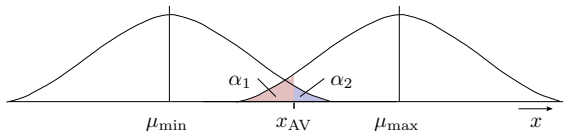
-
- μ expectation value.
 - σ standard deviation.
 - α_2 Probability of error, that the values are above the estimated range.



Range estimate for the expected value

The expected value for an observed event is

- at least so large that the probability of a greater than observed result occurring is smaller than α_2 , and
- at most so great that a smaller result than the observed result occurs, is smaller than α_1 .



Range limits of the expected value:

$$\mu_{\min} = x_{AV} - \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (48)$$

$$\mu_{\max} = x_{AV} + \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (49)$$

$$sr(\mu) = x_{AV} \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (50)$$

x_{AV} actual value.

$sr(\dots)$ symmetrical range of probable expected values.



Example 3.9: Range estimate expected value

$$x_{AV} = 100, \sigma = 10, \alpha_1 = \alpha_2 = 1\%$$

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10

$$\Phi^{-1}(1 - 1\%) = 2,33$$

$$[\mu_{\min}, \mu_{\max}] = 100 \mp 10 \cdot 2,33$$

x_{AV}	actual value.
σ	standard deviation.
μ_{\min}	minimum expectation value.
μ_{\max}	maximum expectation value.



Range estimation of independent counts

For independent counts, the expected value is $\mu = n \cdot p$ and the variance is not greater than

$$\text{Var} [X] \leq n \cdot p \cdot (1 - p) \quad (3.31)$$

Upper bound of variance and standard deviation:

$$\text{Var} [X] \leq \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right) \approx x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right) \quad (51)$$

$$\frac{\text{Var}[X]}{\mathbb{E}[X]^2} \leq \left(\frac{1}{\mu} - \frac{1}{x_{\text{mcv}}}\right) \approx \left(\frac{1}{x_{\text{AV}}} - \frac{1}{x_{\text{mcv}}}\right) \quad (52)$$

$$\sigma = \text{sd} [X] = \sqrt{\text{Var} [X]}$$

$\text{Var} [X]$	variance deviation of the random variable X .
n	(# X) Number of counting attempts, or summed random variables.
p	average probability of occurrence.
σ	standard deviation.
μ	expectation value.
x_{mcv}	m aximum possible c ount v alue.
x_{AV}	a ctual v alue.



Lower and upper bound of the probable range:

$$x_{\min} = \mu - \sqrt{\mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (53)$$

$$x_{\max} = \mu + \sqrt{\mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (54)$$

$$\text{sr}(x) = \mu \mp \sqrt{\mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (55)$$

Lower and upper bound of the expected value:

$$\mu_{\min} = x_{\text{AV}} - \sqrt{x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (56)$$

$$\mu_{\max} = x_{\text{AV}} + \sqrt{x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (57)$$

$$\text{sr}(\mu) = x_{\text{AV}} \mp \sqrt{x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (58)$$

x_{AV}	actual value.
x_{mcv}	maximum possible count value.
$\Phi^{-1}(\cdot)$	Inverse function to the distribution function of the standardised normal distribution.
α_1, α_2	Probability of error, that the values are below or above the estimated range.



Example 3.10: Range estimate MF rate, reliability

During the processing of 20,000 DS, 100 MF were observed. In what range is the MF rate and reliability in 99% of cases in a future observation period of equal length (no dependencies)?

$$\#DS = 20.000 \text{ DS}, x_{AV} = 100 \text{ MF}, x_{mcv} = 20.000 \text{ DS}, \alpha_1 = \alpha_2 = 0,5\%$$

DS	delivered service.
MF	malfunction.
$\#DS$	number of delivered services.
x_{AV}	actual value.
α_1, α_2	Probability of error, that the values are below or above the estimated range.



#DS = 20.000 DS, $x_{AV} = 100$ MF, $x_{m_{cv}} = 20.000$ DS, $\alpha_1 = \alpha_2 = 0,5\%$

Range of the expected number of malfunctions:

α	2.27%	0.13%	0	2%	1%	0.5%	0.2%	0.1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2.05	2.33	2.57	2.88	3.10

$$sr(\mu) = x_{AV} \mp \sqrt{x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{m_{cv}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.58)$$

$$sr(\mu) = 100 \mp \sqrt{100 \cdot \left(1 - \frac{100}{20.000}\right)} \cdot \Phi^{-1}(1 - 0,5\%)$$

$$= 100 \mp 25,6$$

$$\mu_{\min} = 74,4$$

$$\mu_{\max} = 125,6$$



$$\#DS = 20.000 \text{ DS}, x_{AV} = 100 \text{ MF}, x_{mcv} = 20.000 \text{ DS}, \alpha_1 = \alpha_2 = 0,5\%$$

Range of MF rate ($\zeta = \mu/\#DS$):

$$\zeta_{\min} = \frac{74,4}{20.000} = 3,7 \cdot 10^{-3} \left[\frac{\text{MF}}{\text{DS}} \right]; \quad \zeta_{\max} = \frac{125,7}{20.000} = 6,4 \cdot 10^{-3} \left[\frac{\text{MF}}{\text{DS}} \right]$$

Range of reliability ($R = 1/\zeta$):

$$R_{\min} = \frac{1}{\zeta_{\max}} = 159 \left[\frac{\text{DS}}{\text{MF}} \right]; \quad R_{\max} = \frac{1}{\zeta_{\min}} = 269 \left[\frac{\text{DS}}{\text{MF}} \right]$$



Varianzerhöhung

Varianzerhöhung durch Abhängigkeiten

Abhängigkeiten erhöhen Varianz und Standardabweichung und die Breiten der wahrscheinlichen Bereiche.

Wenn z.B. zwei Zählereignisse immer paarweise gleichzeitig eintreten, ist das beschreibbar durch eine Summe von halb so vielen unabhängigen Zufallsvariablen mit den möglichen Werten 0 und 2:

$$X = \sum_{i=1}^{\#X/2} X_i \quad \text{mit} \quad \mathbb{P}(X_i = k) = \begin{cases} 1 - p_i & k = 0 \\ p_i & k = 2 \end{cases}$$

Erwartungswert der Summanden:

$$\mathbb{E}[X_i] = 0 \cdot (1 - p_i) + 2 \cdot p_i = 2 \cdot p_i$$

Varianz der Summanden (nach Shift rule):

$$\begin{aligned} \text{Var}[X_i] &= (1 - p_i) \cdot 0^2 + p_i \cdot 2^2 - (2 \cdot p_i)^2 \\ &= 2^2 \cdot p_i \cdot (1 - p_i) \end{aligned}$$



Der gesamte Erwartungswert ist derselbe wie für $\#X$ unabhängige Zählerereignisse mit paarweise gleichen Eintrittswahrscheinlichkeiten:

$$\mathbb{E}[X] = \sum_{i=1}^{\#X/2} 2 \cdot p_i$$

Die Varianz der Summe verdoppelt sich gegenüber der einer Summe unabhängige Zählerereignisse:

$$\text{Var}[X] = \sum_{i=1}^{\#X/2} 2^2 \cdot p_i \cdot (1 - p_i) = 2 \cdot \left(2 \cdot \sum_{i=1}^{\#X/2} p_i \cdot (1 - p_i) \right)$$

Die Standardabweichung und Breiten wahrscheinlicher Bereiche vergrößern sich um $\sqrt{2}$.



Varianzerhöhung

Die Varianzerhöhung sei definiert als Verhältnis aus Varianz und Erwartungswert:

$$\kappa = \max \left(1, \frac{\text{Var}[X]}{\mathbb{E}[X]} \right) \quad (59)$$

Für kleine $p_i \ll 1$ ist sie im Beispiel $\kappa = 2$. Analog lässt sich zeigen, wenn immer n_{idt} Zählereignisse gleichzeitig eintreten:

$$\kappa = n_{\text{idt}}$$

κ variance increase due to dependencies.

n_{idt} Number of identical counts.

Schätzen der Varianzerhöhung

- Experimentelle Bestimmung von $\#w \geq 2$ Zählwerten w_i .
- Schätzen des Erwartungswerts der Zählwertstichprobe:

$$\hat{\mathbb{E}}[X] = \bar{w} = \frac{1}{\#w} \cdot \sum_{i=1}^{\#w} w_i \quad (3.8)$$

- Schätzen der Varianz der Zählwertstichprobe:

$$\hat{\text{Var}}[X] = \frac{1}{\#w-1} \cdot \sum_{i=1}^{\#w} \left(w_i - \hat{\mathbb{E}}[X] \right)^2 \quad (3.9)$$

- Varianzerhöhung:

$$\kappa = \max \left(1, \frac{\text{Var}[X]}{\mathbb{E}[X]} \right) \quad (3.59)$$



Example 3.11: Varianzerhöhung

$w = 10$ Wiederholungen eines Zählversuchs:

Versuch i	1	2	3	4	5	6	7	8	9	10
Ergebnis w_i	44	87	58	62	59	57	65	57	75	67

DS delivered service.

MF malfunction.

#DS number of delivered services.

x_{AV} actual value.

α_1, α_2 Probability of error, that the values are below or above the estimated range.

$\#w = 10$ Wiederholungen eines Zählversuchs:

Versuch i	1	2	3	4	5	6	7	8	9	10
Ergebnis w_i	44	87	58	62	59	57	65	57	75	67

$$\hat{\mathbb{E}}[X] = \bar{w} = \frac{1}{\#w} \cdot \sum_{i=1}^{\#w} w_i \quad (3.8)$$

$$\hat{\text{Var}}[X] = \frac{1}{\#w-1} \cdot \sum_{i=1}^{\#w} (w_i - \hat{\mathbb{E}}[X])^2 \quad (3.9)$$

$$\kappa = \max\left(1, \frac{\text{Var}[X]}{\mathbb{E}[X]}\right) \quad (3.59)$$

$$\hat{\mathbb{E}}[X] = \frac{1}{10} \cdot \sum_{i=1}^{10} w_i = 63,1$$

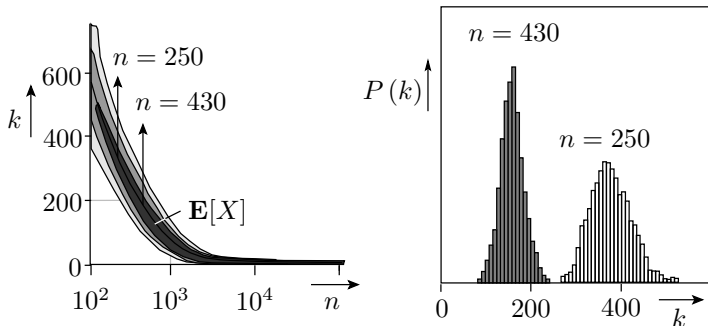
$$\hat{\text{Var}}[X] = \frac{1}{9} \cdot \sum_{i=1}^{10} (w_i - 63,1)^2 = 135$$

$$\kappa = \frac{135}{63,1} \approx 2,14$$

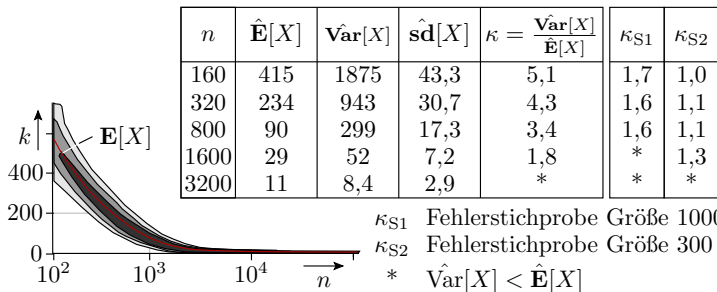
Die Abhängigkeiten erhöhen die Varianz so, als ob mehr als 2 Zählereignisse fast immer gemeinsam eintreten.

Experiment mit Haftfehlern

Kombinatorische Beispielschaltung (Benchmark c3540). 3606 simulierte, unterschiedlich nachweisbare Haftfehler. Zählwert X ist die Anzahl der nicht nachweisbaren Haftfehler. Abschätzung von $\mathbb{P}(X = k) = P(k)$ aus einer Stichprobe von $\#w = 1000$ Zählwerten für verschiedene Zufallstestsätze der Länge n .



Varianzerhöhung im Experiment



Zwischen den nicht nachweisbaren Fehlern gibt es offenbar Abhängigkeiten, die die Varianz so stark erhöhen, als ob 3...5 Modellfehler identisch nachweisbar wären. Identisch nachweisbare Fehler wurden jedoch nicht mitgezählt. Bleiben als Abhängigkeitsursache implizit nachweisbare Fehler sowie geteilte Steuer- und Beobachtungsbedingungen.

Bei weniger nicht nachweisbaren Fehlern oder einer Fehlerstichprobe

Fehlermodellierung und Vorhersagbarkeit

Die Breite des wahrscheinlichen Bereichs der Fehlerüberdeckung:

$$\text{Bereich } FC = \frac{\mu_{\max} - \mu_{\min}}{\#F_M} \sim \frac{\sqrt{\kappa \cdot FC \cdot (1 - FC) \cdot \#F_M}}{\#F_M} \sim \sqrt{\frac{\kappa}{\#F_M}}$$

μ_{\min} , μ_{\max} – Schranken zu erwartenden Anzahl nachweisbare Fehler.

- Für im Verhältnis zur Testobjektgröße kleiner $\#F_M$ ist keine Varianzerhöhung zu erwarten ($\kappa < 1$). Vorhersagbarkeit FC wächst mit der Modellfehleranzahl $\#F_M$.
- Für im Verhältnis zur Testobjektgröße großes $\#F_M$ ist eine Varianzerhöhung $\kappa \sim \#F_M$ zu erwarten. Keine Verbesserung der Vorhersagbarkeit von FC durch mehr Modellfehler.

κ	variance increase due to dependencies.
μ_{\min}	minimum expectation value.
μ_{\max}	maximum expectation value.
$\#F_M$	number of model faults.
FC	fault coverage, percentage of detectable faults.



$$\sqrt{100 \cdot \left(1 - \frac{100}{20.000}\right)} \cdot \Phi^{-1}\left(1 - \frac{1}{5}\right) = 100 + 25 \cdot \frac{1}{6}$$

Thesen:

- Es hat wenig Nutzen, zu viele Modellfehler im Verhältnis zur Testobjektgröße zur Abschätzung von FC zu verwenden.
 - Insbesondere Fehlermodell, bei denen die Anzahl der Modellfehler überproportional mit der Testobjektgröße zunimmt, wie z.B. beim Pfadverzögerungsmodell, sind nicht zielführend.



Bereichs. mit Abhängigk.



Abhängigkeiten in Bereichsschätzungen

α	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2,05	2,33	2,57	2,88	3,10

Wenn Abhängigkeiten zwischen Zählwerten bestehen, der Zählwert aber immer noch normalverteilt ist, lassen sich diese durch eine (max. mögliche) Varianzerhöhung κ berücksichtigen. Der garantierbare wahrscheinliche Bereich verbreitert sich gegenüber Gl. 3.53 bis 3.55 um $\sqrt{\kappa}$:

$$x_{\min} = \mu - \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (60)$$

$$x_{\max} = \mu + \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (61)$$

$$\text{sr}(x) = \mu \mp \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (62)$$

μ expectation value.

κ variance increase due to dependencies.

α_1, α_2 Probability of error, that the values are below or above the estimated range.



Example 3.12: Anzahl Schadensfälle mit Varianzerhöhung

Der zu erwartende Zählwert für die Anzahl von Schadensfällen sei 100. Irrtumswahrscheinlichkeit $\alpha_1 = \alpha_2 = 1\%$, Varianzerhöhung $\kappa = 2$. In welchem Bereich wird bei künftigen Wiederholungen des Versuchs die Anzahl der Schadensfälle liegen?

$$\text{sr}(x) = \mu \mp \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.62)$$

α	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2,05	2,33	2,57	2,88	3,10

Ohne maximale Zählwertgröße:

$$1 - \frac{\mu}{x_{\text{mcv}}} = 1$$

$$(x_{\min}, x_{\max}) = 100 \mp \sqrt{2 \cdot 100} \cdot 2,33$$

$$x_{\min} = 76,0$$

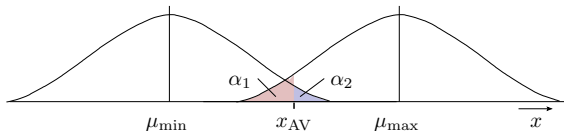
$$x_{\max} = 133,0$$



Bereichsschätzung für den Erwartungswert

Der Erwartungswert zu einem beobachteten Ereignis ist

- mindestens so groß, dass die Wahrscheinlichkeit, dass ein größeres als das beobachtete Ergebnis eintritt, kleiner α_2 , und
- maximal so groß, dass ein kleineres als das beobachtete Ergebnis eintritt, kleiner α_1 , ist.



Der garantierbare wahrscheinliche Bereich für den Erwartungswert verbreitert sich gegenüber Gl. 3.56 und 3.57 gleichfalls um $\sqrt{\kappa}$:

$$\mu_{\min} = x_{AV} - \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (63)$$

$$\mu_{\max} = x_{AV} + \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (64)$$



$$SR(\mu) = x_{AV} \mp \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (65)$$

μ_{\min}	minimum expectation value.
μ_{\max}	maximum expectation value.
$SR(\dots)$	symmetrical range of probable expected values.
x_{AV}	actual value.
κ	variance increase due to dependencies.
α_1, α_2	Probability of error, that the values are below or above the estimated range.
α	error probability that the value is outside the estimated range.



Example 3.13: Anzahl nicht nachweisbarer Modellfehler

Von 1000 Modellfehlern wurden 32 nicht erkannt. Gesucht ist der Bereich für die zu erwartende Anzahl der nicht nachweisbaren Modellfehler mit $\alpha_1 = \alpha_2 = 1\%$ und Varianzerhöhung $\kappa = 2$.

$$sr(\mu) = x_{AV} \mp \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right) \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \quad (3.65)$$

α	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1 - \alpha)$	2	3	4	2,05	2,33	2,57	2,88	3,10

$$(\mu_{\min}, \mu_{\max}) = 32 \mp \sqrt{2 \cdot 32 \cdot \left(1 - \frac{32}{1.000}\right) \cdot 2,33}$$

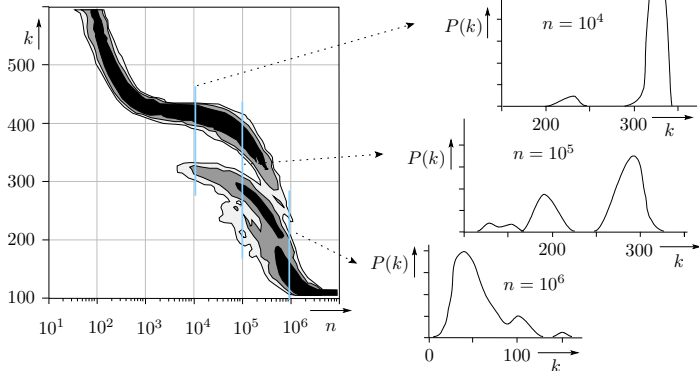
$$\mu_{\min} = 13,7$$

$$\mu_{\max} = 50,3$$

-
- x_{AV} actual value.
 - κ variance increase due to dependencies.
 - α error probability that the value is outside the estimated range.

Nicht normalverteilte Zählwerte

Dasselbe Experiment mit der kleineren Benchmark-Schaltung c2670:



Im Bereich von $n = 10^4$ bis 10^6 mehrere Gipfel. Keine näherungsweise Normalverteilung.

Wie ist das möglich?



Zusammenfassung

Binomialverteilung

Für den Sonderfall, dass gleichwahrscheinliche Ereignisse gezählt werden, ist die Summe der gezählten Ereignisse binomialverteilt:

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.27)$$

Erwartungswert:

$$\mathbb{E}[X] = n \cdot p \quad (3.28)$$

Varianz:

$$\text{Var}[X] = n \cdot p \cdot (1 - p) \quad (3.29)$$

Standardabweichung:

$$\text{sd}[X] = \sqrt{n \cdot p \cdot (1 - p)} \quad (3.30)$$

Bei gleicher Anzahl von unabhängigen Zählwerten und gleicher mittlerer Eintrittswahrscheinlichkeit ist die Varianz nach Gl. 3.29 eine obere Schranke für die Varianz eines Zählwertes nach Gl. 3.26, bei dem sich die Eintrittswahrscheinlichkeiten unterscheiden.

Poissonverteilung

Beim Zählen vieler seltener Ereignisse sind die Zählwerte näherungsweise poisson-verteilt:

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.27)$$

Der Verteilungsparameter λ , gleichzeitig Erwartungswert und Varianz, ist das Produkt aus der Anzahl der Zählversuche und der mittleren Eintrittswahrscheinlichkeit:

$$\mathbb{P}(X = k) e^{-\bar{p} \cdot n} \cdot \frac{(\bar{p} \cdot n)^k}{k!} \quad (3.33)$$

Ab $\lambda \geq 10$ können poissonverteilte Zählwerte als Näherungsweise normalverteilt mit Varianz gleich Erwartungswert angenähert werden.

Für eine poisson-verteilte Fehleranzahl beträgt die Ausbeute

$$Y = e^{-FC \cdot \mu_{FCP}} \quad (3.34)$$

Zu erwartende Schaltkreiskosten in Abhängigkeit von der Transistoranzahl:

$$C_{IS} = C_{Tr} \cdot \#Tr \cdot e^{\xi_{Tr} \cdot \#Tr} \quad (3.35)$$



Bereichsschätzung Poissonverteilung

Geeignet für Zählwerte von vielen seltenen Ereignissen und bekannte Erwartungswerte oder experimentell bestimmte Realisierungen im Bereich von null bis 10. Für praktische Schätzungen sind

- auf slide 3.59 die Funktion $\lambda(k_{\min}, \alpha_1)$
- auf slide 3.59 die Funktion $\lambda(k_{\max}, \alpha_2)$ und
- ab slide 3.61 $[\lambda_{\min}, \lambda_{\max}]$ in Abhängigkeit von der Irrtumswahrscheinlichkeit und einem Istzählwert

tabelliert. Aus diesen Tabellen lassen sich

- für gegebene Irrtumswahrscheinlichkeiten und Erwartungswerte Minima und Maxima für die Zählwerte,
- für gegebene Erwartungswerte und Minima bzw. Maxima für die Zählwerte die Irrtumswahrscheinlichkeiten oder
- für gegebene Irrtumswahrscheinlichkeiten und Zählwerte der Bereich, in dem der Erwartungswert liegt

ablesen.

Normalverteilung, Bereichsschätzungen

Eine Normalverteilung ist eine stetige Verteilung mit der Dichtefunktion:

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \quad (3.36)$$

Für Bereichsschätzungen werde die normalverteilte Zufallsvariable und die Bereichsgrenzen normiert

$$Z = \frac{X-\mu}{\sigma} \quad (3.37)$$

$$z = \frac{x-\mu}{\sigma} \quad (3.38)$$

so dass die Intervallzugehörigkeiten und aus der Tabelle der standardisierten Normalverteilung auf slide 3.74 abgelesen werden können.

Irrumswahrscheinlichkeit, Wert kleiner als untere Bereichsgrenze:

$$\alpha_1 = 1 - \Phi\left(\frac{\mu - x_{\min}}{\sigma}\right) \quad (3.41)$$

Irrumswahrscheinlichkeit, Wert größer als obere Bereichsgrenze:

$$\alpha_2 = 1 - \Phi\left(\frac{x_{\max} - \mu}{\sigma}\right) \quad (3.42)$$



Bereichsgrenzen für gegebene Irrtumswahrscheinlichkeiten:

$$x_{\min} = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.45)$$

$$x_{\max} = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.46)$$

Bereichsschätzung für den Erwartungswert:

$$\mu_{\min} = x_{AV} - \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.48)$$

$$\mu_{\max} = x_{AV} + \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.49)$$

Bereichsschätzung unabhängiger Zählwerte:

$$x_{\min} = \mu - \sqrt{\mu \cdot \left(1 - \frac{\mu}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.53)$$

$$x_{\max} = \mu + \sqrt{\mu \cdot \left(1 - \frac{\mu}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.54)$$

$$\mu_{\min} = x_{AV} - \sqrt{x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.56)$$

$$\mu_{\max} = x_{AV} + \sqrt{x_{AV} \cdot \left(1 - \frac{x_{AV}}{x_{mcv}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.57)$$



Varianzerhöhung

Abhängigkeiten erhöhen Varianz und Standardabweichung und die Breiten der wahrscheinlichen Bereiche. Definition der Varianzerhöhung:

$$\kappa = \max \left(1, \frac{\text{Var}[X]}{\mathbb{E}[X]} \right) \quad (3.59)$$

Die einzusetzenden Erwartungswerte und Varianzen lassen über Gl. 3.8 und 3.9 aus einer Stichprobe mehrerer experimentell bestimmter Zählwerte abschätzen.

Für vollständige Haftfehlermengen ohne identisch nachweisbare Fehler kann sie durch andere Nachweisabhängigkeiten deutlich größer als eins sein. Beschränkung auf Modellfehlerstichproben mindert die Abhängigkeiten.

Es hat wenig Nutzen, zu viele Modellfehler im Verhältnis zur Testobjektgröße zur Abschätzung von Fehlerüberdeckungen zu verwenden.

Abhängigkeiten in Bereichsschätzungen

Bereichsgrenzen in Abhängigkeit vom Erwartungswert, der Varianzerhöhung und den Irrtumswahrscheinlichkeiten:

$$x_{\min} = \mu - \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.60)$$

$$x_{\max} = \mu + \sqrt{\kappa \cdot \mu \cdot \left(1 - \frac{\mu}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.61)$$

Wahrscheinlicher Bereich des Erwartungswertes für einen experiment bestimmten Istwert und zugelassene Irrtumswahrscheinlichkeiten:

$$\mu_{\min} = x_{\text{AV}} - \sqrt{\kappa \cdot x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.63)$$

$$\mu_{\max} = x_{\text{AV}} + \sqrt{\kappa \cdot x_{\text{AV}} \cdot \left(1 - \frac{x_{\text{AV}}}{x_{\text{mcv}}}\right)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.64)$$



Mischverteilung



Mischverteilung

Aus einer Grundgesamtheit von $\#X_{\text{ges}}$ Objekten, von den jeweils $\#X_i$ Objekte eine Verteilung $X_i \sim F_{X_i}$ haben wählt eine Zufallsvariable Y zufällig ein Objekt aus:

$$\mathbb{P}(Y = i) = \frac{\#X_i}{\#X_{\text{ges}}} = h_i$$

$$F_X(x) = \mathbb{P}[(X \leq x)] = \sum_{i=1}^{\#i} h_i \cdot F_{X_i}(x)$$

Für diskrete Verteilungen:

$$\mathbb{P}(X = x) = \sum_{i=1}^{\#i} h_i \cdot \mathbb{P}(X_i = x) \quad (66)$$

Für stetige Verteilungen:

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{i=1}^{\#i} h_i \cdot f_{X_i}(x) \quad (67)$$



Zufallsvariablen mit einer Mischverteilung

- Eigenschaft einer Schraube (z.B. Länge) bei zufälliger Auswahl auf einer Kiste mit Schrauben unterschiedlicher Hersteller.
- Fehleranzahl eines SW-Bausteins bei zufällige Auswahl aus Angeboten unterschiedlicher Programmierer mit unterschiedlichen Fehlerentstehungsraten.
- Schadenshöhe eines zufälligen Schadens auf einer Menge unterschiedlicher Schadensklassen mit unterschiedlicher Kostenverteilung.
- ...



Eigenschaften

Varianzvergrößerung durch »Mischung«

Der Erwartungswert ist der gewichtete Mittelwert:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#i} h_i \cdot \mu_i \quad (68)$$

Varianz mit $\mu = \mu_i + \delta_i$:

$$\begin{aligned} \text{Var}[X] &= \sum_{i=1}^{\#i} h_i \cdot \mathbb{E}[(X_i - \mu_i - \delta_i)^2] \\ &= \sum_{i=1}^{\#i} h_i \left(\underbrace{\mathbb{E}[(X_i - \mu_i)^2]}_{\sigma_i^2} - \underbrace{\mathbb{E}[2 \cdot \delta_i \cdot (X_i - \mu_i)]}_{2 \cdot \delta_i \cdot \mathbb{E}[X_i - \mu_i] = 0} + \underbrace{\mathbb{E}[\delta_i^2]}_{\delta_i^2} \right) \\ \text{Var}[X] &= \sum_{i=1}^{\#i} h_i \cdot \sigma_i^2 + \sum_{i=1}^{\#i} h_i \cdot \delta_i^2 \quad (69) \end{aligned}$$

Mittelwert der Einzelvarianzen plus mittlere quadratische Abweichung der Einzelerwartungswerte vom Gesamterwartungswert.

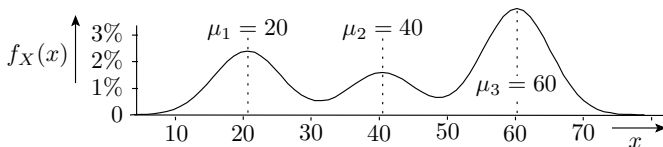
Multimodale Verteilung

Beim Mischen von Grundgesamtheiten mit deutlich abweichenden Erwartungswerten entstehen multimodale Verteilungen. Beispiel: Mischen von 3 normalverteilten Zufallsvariablen X_i (vergl. Gl. 3.67):

h_i	0,3	0,2	0,5
μ_i	20	40	60
σ_i	5	5	5

$$f_X(x) = \frac{dF_X(x)}{dx} = 0,3 \cdot \varphi\left(\frac{x-20}{5}\right) + 0,2 \cdot \varphi\left(\frac{x-40}{5}\right) + 0,5 \cdot \varphi\left(\frac{x-60}{5}\right)$$

$\varphi(z) = \varphi\left(\frac{x-\mu_i}{\sigma_i}\right)$ – Dichte der standardisierten Normalverteilung.





3. Mischverteilung

1. Eigenschaften

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#i} h_i \cdot \mu_i \quad (3.68)$$

$$\text{Var}[X] = \sum_{i=1}^{\#i} h_i \cdot \sigma_i^2 + \sum_{i=1}^{\#i} h_i \cdot \delta_i^2 \quad (3.69)$$

h_i	0,3	0,2	0,5
μ_i	20	40	60
σ_i	5	5	5

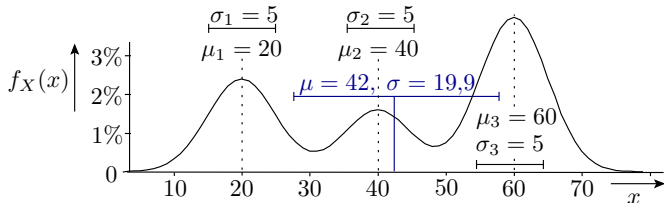
Erwartungswert:

$$\mu = 0,3 \cdot 20 + 0,2 \cdot 40 + 0,5 \cdot 60 = 42$$

Varianz, Standardabweichung:

$$\sigma^2 = 25 + 0,3 \cdot (20 - 42)^2 + 0,2 \cdot (40 - 42)^2 + 0,5 \cdot (60 - 42)^2 = 285$$

$$\sigma = 16,9$$





Anwendungen



Example 3.14: Identisch nachweisbare Fehler

In einer Modellfehlermenge aus $\#F_M = 25$ Fehlern mit einer Nachweiswahrscheinlichkeit $p = 60\%$ seien zehn Fehler identisch und die übrigen Fehler unabhängig voneinander nachweisbar.

$\#F_M = 25$, davon 10 identisch nachweisbar, $p = 60\%$

- Beschreibung als Mischverteilung von zueinander verschobenen Binomialverteilungen.
- Erwartungswert und Varianz.
- Standardabweichung und Varianzerhöhung.



$\#F_M = 25$, davon 10 identisch nachweisbar, $p = 60\%$

- a) Beschreibung als Mischverteilung von zueinander verschobenen Binomialverteilungen.

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.27)$$

$$\mathbb{P}(X = x) = \sum_{i=1}^{\#i} h_i \cdot \mathbb{P}(X_i = x) \quad (3.66)$$

Binomialverteilung ohne die 10 nur gemeinsam nachweisbaren Fehler:

$$\mathbb{P}(X_0 = k) = \begin{cases} \binom{N-10}{k} \cdot p^k \cdot (1 - p)^{N-10-k} & 0 \leq k \leq N - 10 \\ 0 & \text{sonst} \end{cases}$$

Mit den 10 nur gemeinsam nachweisbaren Fehlern verschieben sich alle Wahrscheinlichkeiten zur um 10 größeren Realisierung:

$$\mathbb{P}(X_1 = k) = \begin{cases} \binom{N-10}{k-10} \cdot p^{k-10} \cdot (1 - p)^{N-k} & 10 \leq k \leq N \\ 0 & \text{sonst} \end{cases}$$

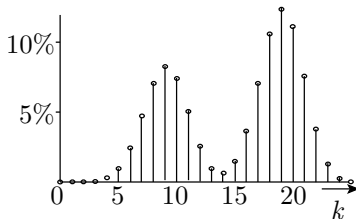


$\#F_M = 25$, davon 10 identisch nachweisbar, $p = 60\%$

- a) Beschreibung als Mischverteilung von zueinander verschobenen Binomialverteilungen.

Mischverteilung:

$$\mathbb{P}(X = k) = (1 - p) \cdot \mathbb{P}(X_0 = k) + p \cdot \mathbb{P}(X_1 = k)$$





$\#F_M = 25$, davon 10 identisch nachweisbar, $p = 60\%$

b) Erwartungswert und Varianz.

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#i} h_i \cdot \mu_i \quad (3.68)$$

$$\text{Var}[X] = \sum_{i=1}^{\#i} h_i \cdot \sigma_i^2 + \sum_{i=1}^{\#i} h_i \cdot \delta_i^2 \quad (3.69)$$

Erwartungswert:

$$\mathbb{E}[X] = 25 \cdot p = 25 \cdot 60\% = 15$$

$$\stackrel{!}{=} (1-p) \cdot \mathbb{E}[X_0] + p \cdot \mathbb{E}[X_1] = (1-0,6) \cdot 9 + 0,6 \cdot 19 = 15\checkmark$$

Varianz als Summe der Varianzen der Summanden:

$$\text{Var}[X] = 15 \cdot p \cdot (1-p) + 10^2 \cdot p \cdot (1-p) = 115 \cdot p \cdot (1-p) = 27,6$$

$$\stackrel{!}{=} \underbrace{n \cdot p \cdot (1-p)}_{3,6} + \underbrace{(1-p) \cdot (\mathbb{E}[X_0] - \mathbb{E}[X])^2}_{+ 0,4 \cdot (9-15)^2} + \underbrace{p \cdot (\mathbb{E}[X_1] - \mathbb{E}[X])^2}_{+ 0,6 \cdot (19-15)^2} = 27,6\checkmark$$



$F_M = 25$, davon 10 identisch nachweisbar, $p = 60\%$

c) Standardabweichung und Varianzerhöhung.

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (3.4)$$

$$\kappa = \max\left(1, \frac{\text{Var}[X]}{\mathbb{E}[X]}\right) \quad (3.59)$$

Standardabweichung:

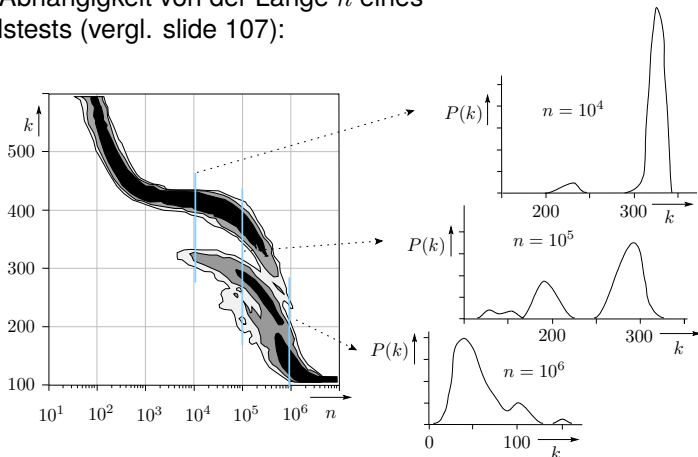
$$\text{sd}[X] = \sqrt{27,6} = 5,25$$

Varianzerhöhung:

$$\kappa = \frac{27,6}{15} = 1,84$$

Dichte nicht nachweisbare Fehler c2670

... in Abhängigkeit von der Länge n eines
Zufallstests (vergl. slide 107):

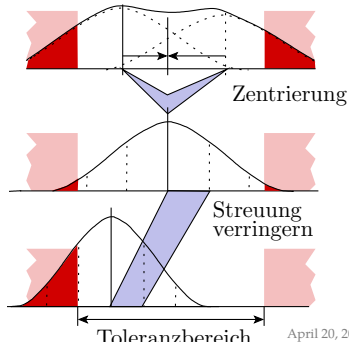


Im Bereich von $n = 10^4$ bis 10^6 multimodale Verteilung. Offenbar ca. 80 sehr ähnlich nachweisbare Fehler mit FF-Rate $\zeta_i \approx 10^{-5} \left[\frac{MF}{DS} \right]$.

Objekten aus unterschiedlichen Prozessen

Bei der mechanischen Fertigung haben die Zielparameter, z.B. bei einer Bohrung Durchmesser und Tiefe, eine Verteilung und einen Toleranzbereich. Entstehungshäufigkeit eines Parameterfehlers ist die Wahrscheinlichkeit, Parameter außerhalb Toleranzbereich. Bei erkennbarer Polarisierung der Messwerte eines Parameters:

- Lokalisierung der Prozesse, deren Ergebnisse gemischt werden.
- Prozesszentrierung: Verschiebung der Verteilung für jeden Einzelprozess mit Hilfe von Einstelloptionen in die Mitte des Toleranzbereichs.
- Prozessverbesserung: Verringerung der Streuung durch technologische Neuerungen neue Maschinen, Verfahren, ...

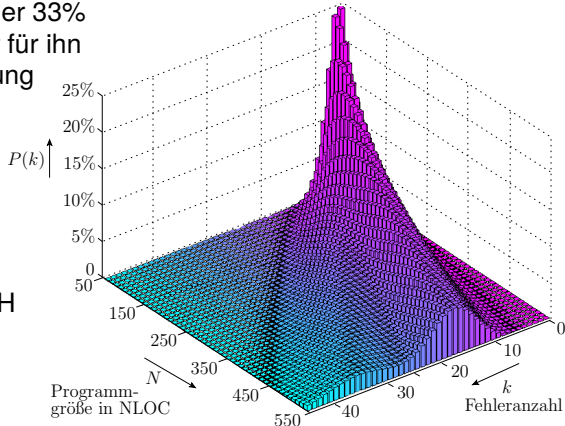


Unterschiedlich gute Programmierer

Ein Anfänger und ein Profi entwickeln Software-Bausteine aus N NLOC (Netto Lines of Code), der Profi 66% mit ca. einem Fehler je 30 NLOC und der Anfänger 33%
 Kunde weiß nicht, wer für ihn programmiert. Verteilung der Fehleranzahl:

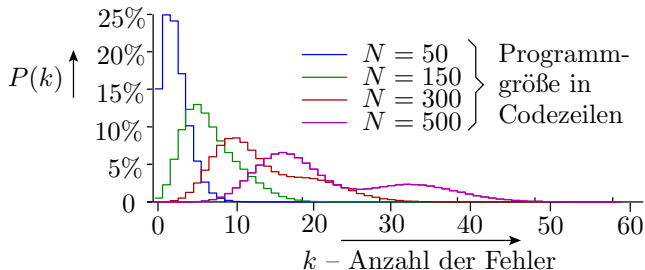
$$\begin{aligned} \mathbb{P}(N, X = k) &= \underbrace{\frac{2}{3} \cdot e^{-\frac{N}{30}} \cdot \frac{\left(\frac{N}{30}\right)^k}{k!}}_{\text{Pois}\left(\frac{N}{30}\right)} \\ &+ \underbrace{\frac{1}{3} \cdot e^{-\frac{N}{15}} \cdot \frac{\left(\frac{N}{15}\right)^k}{k!}}_{\text{Pois}\left(\frac{N}{15}\right)} \end{aligned}$$

GH



Die Wahrscheinlichkeit, dass ein Modul genau k Fehler enthält, ist $2/3$ mal die Wahrscheinlichkeit, dass es k Fehler enthält und vom Profi stammt plus $1/3$ mal die Wahrscheinlichkeit, dass es k Fehler enthält und vom Anfänger stammt (vergl. Gl. 3.66):

$$\mathbb{P}(N, X = k) = \frac{2}{3} \cdot e^{-\frac{N}{30}} \cdot \frac{\left(\frac{N}{30}\right)^k}{k!} + \frac{1}{3} \cdot e^{-\frac{N}{15}} \cdot \frac{\left(\frac{N}{15}\right)^k}{k!}$$



Die Polarisierung nimmt mit der Größe der Software-Bausteine, die vom Profi und vom Anfänger getrennt entwickelt werden, zu.



Bei einem guten Entstehungsprozess streben überwachte Parameter oft gegen eine Normalverteilung. Polarisierungen (mehrere Gipfel) können wichtige Informationen über Schwachstellen und Ansatzmöglichkeiten für Verbesserungen liefern:

- Abhängigkeiten bei der Fehlerentstehung, bei Ausfällen beim Fehlernachweis und beim Versagen von Service-Leistungen,
- Vorliebe oder Neigung befragter Experten, z.B. bei der Einschätzung von Gefährdungen und Risiken,
- Probleme eines Messverfahrens, ...

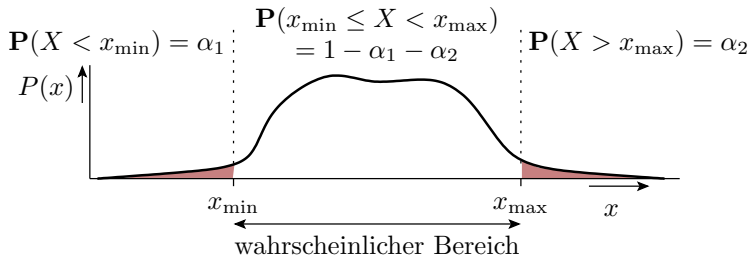
Wenn die Zufallsvariable ein Gütemaß ist, hat man es offenbar mit einer zufälligen Mischung von besser und schlechter funktionierenden Prozessabläufen zu tun. Dann ist es natürlich interessant, warum der Entstehungsprozess mal besser und mal schlechter funktioniert, um das schlechtere Funktionierende zu eliminieren.

Spezielle Form des Lernen aus Fehlern.



Bereichs. Tscheb.

Bereichsschätzung, wenn Verteilung unbekannt

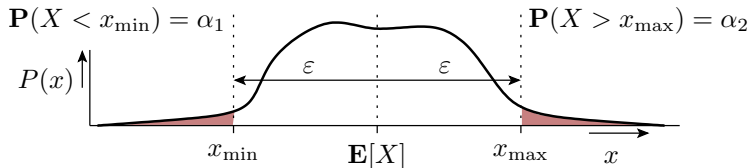


Die Bestimmung eines wahrscheinlichen Intervalls $[x_{\min}, x_{\max}]$

- auch möglich, wenn Verteilung unbekannt, multimodal, ...
- Voraussetzung: eine hinreichend kleine Varianz.

x_{\min}	lower bound of the likely range of the value.
x_{\max}	upper bound of the likely range of the value.
α_1, α_2	Probability of error, that the values are below or above the estimated range.
$\mathbb{P}(\dots)$	probability of occurrence of an event.
$P(\dots)$	probability of occurrence of a value.

Das schwache Gesetz der großen Zahlen



Nach der tschebyscheffschen Ungleichung:

$$\mathbb{P}(|x - \mathbb{E}[X]| \geq \epsilon) \leq \frac{\text{Var}[X]}{\epsilon^2} \quad (70)$$

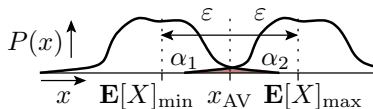
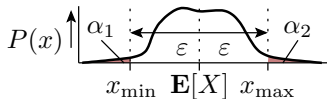
ist die Wahrscheinlichkeit, das der Wert einer Zufallsvariable mehr als ein Intervallradius ϵ von seinem Erwartungswert abweicht, nicht größer als das Verhältnis der Varianz zum Quadrat des Intervallradius ϵ .

Intervallradius für $\alpha = \alpha_1 + \alpha_2$:

$$\epsilon \geq \sqrt{\frac{\text{Var}[X]}{\alpha}} = \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (71)$$

- ϵ Interval radius, distance of the lower und upper range limit from the expected value.
- α error probability that the value is outside the estimated range.

Bereichsschätzung



Wahrscheinlicher Bereich künftiger experimenteller Ergebnisse bei bekanntem Erwartungswert $\mathbb{E}[X]$:

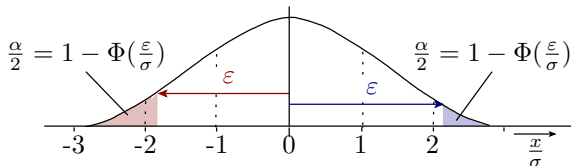
$$[x_{\min}, x_{\max}] = \mathbb{E}[X] \mp \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (72)$$

Wahrscheinlicher Bereich des Erwartungswerts bei einer bekannten Realisierung x_{AV} :

$$[\mathbb{E}[X]_{\min}, \mathbb{E}[X]_{\max}] = x_{AV} \mp \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (73)$$

$x_{\min/\max}$	lower and upper range limit.
$\text{sd}[X]$	standard deviation of the random variable X .
α	error probability that the value is outside the estimated range.
$\mathbb{E}[X]_{\dots}$	minimum and maximum expectation value.
x_{AV}	actual value.

Vergleich mit Intervallradius Normalverteilung



Intervallradius für Normalverteilung und $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$:

$$\varepsilon = \sigma \cdot \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

Intervallradius Tschebyscheffsche Ungleichung:

$$\varepsilon = \sigma \cdot \alpha^{-0,5}$$

α	4,55%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	2	3	4	2,05	2,33	2,57	2,88	3,10
$\alpha^{-0,5}$				5	7,07	10	15,8	22,4



Example 3.15: Bereichsschätzung Erwartungswert

Aus eine Stichprobe gemessener Widerstandswerte in $k\Omega$ soll auf den möglichen Bereich des Erwartungswertes geschlussfolgert werden. Zugelassene Irrtumswahrscheinlichkeit 2%.

$\alpha = 2\%$, Wertestichprobe: $R_i : 10,3, 10,5, 9,7, 8,9, 10,1, 11,0, 10,2, 9,5$

- Ohne Kenntnisse der Verteilung über die Tschebyscheffsche Ungleichung.
- Unter der Annahme, dass die Widerstandswerte normalverteilt sind.



$\alpha = 2\%$, Wertestichprobe: $R_i : 10,3, 10,5, 9,7, 8,9, 10,1, 11,0, 10,2, 9,5$

a) Ohne Kenntnisse der Verteilung über die Tschebyscheffsche Ungleichung.

$$\hat{\mathbb{E}}[X] = \bar{w} = \frac{1}{\#w} \cdot \sum_{i=1}^{\#w} w_i \quad (3.8)$$

$$\hat{\text{Var}}[X] = \frac{1}{\#w-1} \cdot \sum_{i=1}^{\#w} (w_i - \hat{\mathbb{E}}[X])^2 \quad (3.9)$$

$$[x_{\min}, x_{\max}] = \mathbb{E}[X] \mp \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (3.72)$$

Erwartungswert und Standardabweichung der Datenstichprobe:

$$\hat{\mathbb{E}}[R] = \frac{1}{8} (10,3 + \dots) \text{ k}\Omega = 10,025 \text{ k}\Omega$$

$$\hat{\text{sd}}[R] = \sqrt{\frac{1}{7} \left((10,3 - 10,025)^2 + \dots \right) \text{ k}\Omega^2} = 647 \Omega$$

Bereich des Erwartungswerts:

$$10,025 \text{ k}\Omega \mp \frac{647 \Omega}{\sqrt{2\%}} = [5,3 \text{ k}\Omega, 14,8 \text{ k}\Omega]$$



$\alpha = 2\%$, Wertestichprobe: $R_i : 10,3, 10,5, 9,7, 8,9, 10,1, 11,0, 10,2, 9,5$

b) Unter der Annahme, dass die Widerstandswerte normalverteilt sind.

$$x_{\min} = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.45)$$

$$x_{\max} = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.46)$$

α	0,26%	0	4%	2%	1%	0,4%
$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$	3	4	2,05	2,33	2,57	2,88

$$\hat{\mathbb{E}}[R] \mp \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \hat{\text{sd}}[R] = 0,025 \text{ k}\Omega \mp 2,33 \cdot 647 \Omega = [8,5 \text{ k}\Omega, 11,5 \text{ k}\Omega]$$

Weniger als halb so breiter Bereich im Vergleich zu $[5,3 \text{ k}\Omega, 14,8 \text{ k}\Omega]$ aus Aufgabenteil a »ohne Kenntnis der Verteilung«.



Zusammenfassung

Misch- und multimodale Verteilungen

Mischung diskreter Verteilungen:

$$\mathbb{P}(X = x) = \sum_{i=1}^{\#i} h_i \cdot \mathbb{P}(X_i = x) \quad (3.66)$$

Mischung stetiger Verteilungen:

$$f_X(x) = \sum_{i=1}^{\#i} h_i \cdot f_{X_i}(x) \quad (3.67)$$

Erwartungswert:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#i} h_i \cdot \mu_i \quad (3.68)$$

Varianz:

$$\text{Var}[X] = \sum_{i=1}^{\#i} h_i \cdot \sigma_i^2 + \sum_{i=1}^{\#i} h_i \cdot \delta_i^2 \quad (3.69)$$

Bei im Verhältnis zur Standardabweichung großen Abweichungen der Erwartungswert entsteht Multimodalität, z.B.:

- Für die Anzahl der nachweisbaren Fehler bei einem Zufallstest große Fehlerteilmengen, die (fast) gleich nachweisbar sind.
- Mischung von Objekten aus unterschiedlichen, sich ändernden Entstehungsprozessen.



Multimodalität

Multimodalität bedeutet, dass die Verteilung einer diskreten Zufallsvariablen bzw. die Dichtefunktion einer stetigen Zufallsvariablen mehrere Maxima hat. Die Maxima liefern Hinweise auf Verbesserungsmöglichkeiten

- für Entstehungsprozesse zur Fehlervermeidung,
- für die Bewertung erfasster Daten, z.B. das Erkennung unerwarteter Korrelation von Zählwerten oder Vorlieben von Experten bei Befragungen,
- für Messverfahren, ...

Bereichsschätzung

Bereichsschätzung für beliebige incl. multimodale Verteilungen:

- tschebyscheffschen Ungleichung:

$$\mathbb{P}(|x - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\text{Var}[X]}{\varepsilon^2} \quad (3.70)$$

- garantierbarer Intervallradius:

$$\varepsilon \geq \sqrt{\frac{\text{Var}[X]}{\alpha}} = \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (3.71)$$

- garantierbarer Bereiche für Realisierungen und Erwartungswerte:

$$[x_{\min}, x_{\max}] = \mathbb{E}[X] \mp \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (3.72)$$

$$[\mathbb{E}[X]_{\min}, \mathbb{E}[X]_{\max}] = x_{AV} \mp \frac{\text{sd}[X]}{\sqrt{\alpha}} \quad (3.73)$$



Pareto-Verteilung



Pareto-Verteilung

Die Pareto-Verteilung

$$X \sim \text{Par}(k, x_{\min})$$

ist eine stetige Wahrscheinlichkeitsverteilung. Verteilungsfunktion:

$$F_X(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^k & \text{sonst} \end{cases} \quad \text{und } k > 0 \quad (74)$$

Dichtefunktion für $x \geq x_{\min}$:

$$f_X(x) = \frac{k \cdot x_{\min}^k}{x^{k+1}} \quad (75)$$

Typisch für Zufallsprozesse, bei denen ein kleiner Teil der Ursachen für den überwiegenden Teil der Wirkungen verantwortlich ist (Pareto-Prinzip*).

k form factor of the Pareto distribution.

x_{\min} scale parameter of the Pareto distribution.

* In 1906, the Italian economist Vilfredo Pareto studied the distribution of land ownership in Italy and found that about 20% of the population owns about 80% of the land. This has entered the vernacular as the Pareto 20%-80% rule.



Eigenschaften

Eigenschaften der Pareto-Verteilung

- Einen Erwartungswert

$$\mathbb{E}[X] = \int_{x_{\min}}^{\infty} \frac{k \cdot x_{\min}^k}{x^{k+1}} \cdot x \cdot dx = \frac{k \cdot x_{\min}^k}{1-k} \cdot \left(\lim_{x \rightarrow \infty} x^{1-k} - x_{\min}^{1-k} \right)$$

hat eine pareto-verteilte Zufallsvariable nur für $k > 1$:

$$\mathbb{E}[X] = x_{\min} \cdot \frac{k}{k-1} \quad (76)$$

- Eine Varianz existiert nur für $k > 2$:

$$\text{Var}[X] = x_{\min}^2 \cdot \frac{k}{(k-2)(1-k)^2} \quad (77)$$

- Für kleine Exponenten gehört sie zu den endlastigen Verteilungen, bei denen ein erheblicher Teil der Wahrscheinlichkeitsmasse auf große x entfällt.



Formfaktor für die Pareto-20%-80%-Rege

Der Anteil der Ursachen U mit der größten Wirkung:

$$U = \int_{w_{\min}}^{\infty} f(x) \cdot dx = \int_{w_{\min}}^{\infty} \frac{k \cdot x_{\min}^k}{x^{k+1}} \cdot dx = \left(\frac{x_{\min}}{w_{\min}} \right)^k$$

hat mindestens die Wirkung: $w_{\min} = x_{\min} \cdot U^{-\frac{1}{k}}$. Zu erwartende anteilige Wirkung und zu erwartende Gesamtwirkung:

$$\mathbb{E}[X|X \geq w_{\min}] = \int_{w_{\min}}^{\infty} \frac{k \cdot x_{\min}^k}{x^{k+1}} \cdot x \cdot dx = \frac{k}{k-1} \cdot x_{\min} \cdot \left(\frac{x_{\min}}{w_{\min}} \right)^{k-1}$$

$$\mathbb{E}[X] = \int_{x_{\min}}^{\infty} \frac{k \cdot x_{\min}^k}{x^{k+1}} \cdot x \cdot dx = x_{\min} \cdot \frac{k}{k-1}$$

(Voraussetzung $k > 1$). Anteilige Gesamtwirkung: \square

$$W = \frac{\mathbb{E}[X|X \geq w_{\min}]}{\mathbb{E}[X]} = \left(\frac{x_{\min}}{w_{\min}} \right)^{k-1} = U^{\frac{k-1}{k}}$$

$U = 20\%$	$k = 1,1$	$k = 1,2$	$k = 1,15$	$k = 1,175$	$k = 1,16$
$W = U^{\frac{k-1}{k}}$	86,4%	76,5%	81,1%	78,7%	80,1%✓



Fehlernachweislänge



Verteilung der Fehlernachweislänge

Die Nachweislänge N ist die Anzahl der Service-Leitungen bzw. die Zeit bis zum Nachweis eines Fehlers. Die Verteilungsfunktion als Wahrscheinlichkeit, dass die Nachweislänge N nicht größer als eine Testsatzlänge n ist, ist gleich der zu erwartende Fehlerüberdeckung für Testsatzlänge n :

$$F_N(n) = \mathbb{P}(N \leq n) = \mathbb{E}[FC(n)]$$

Grobabschätzung

Bei einem Zufallstest verlangt eine Verringerung des Anteils der nicht nachweisbaren Fehler um eine Dekade in der Regel eine Erhöhung der Testsatzlänge um mehr als eine Dekade.



$$\mathbb{E}[FC(n)] = 1 - \left(\frac{n_{\min}}{n}\right)^k$$

k	1	0,5	0,33	0,25
$\frac{n}{n_{\min}}$ für $1 - FC(n) = 0,1$	10	100	10^3	10^4



Die Nachweislänge N als Wahrscheinlichkeit, dass ein Fehler mit n Tests nachgewiesen wird, ist gleich der zu erwartenden Fehlerüberdeckung und für die Grobabschätzung für ganzzahlige Werte der Testsatzlänge n pareto-verteilt:

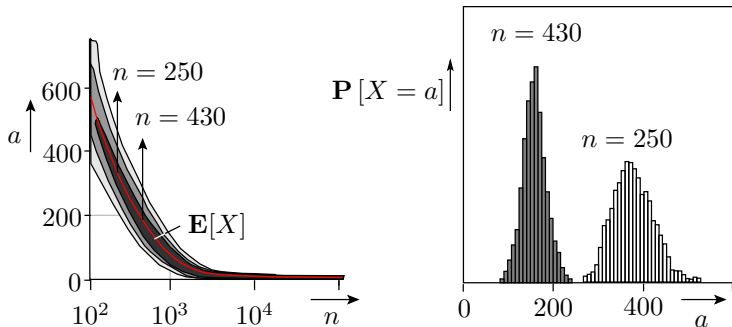
$$F_N(n) = \mathbb{E}[FC(n)] = 1 - \left(\frac{n_{\min}}{n}\right)^k \quad (78)$$

Wegen $k < 1$ hat die Nachweislänge N keinen Erwartungswert. Beim Betrieb von IT-System mit pareto-verteilter Nachweislänge und Beseitigung aller erkennbaren Fehler sind auch nach sehr langer Nutzungsdauer weitere Fehler nichte ausschließbar.

Es folgt ein Experiment zu Untersuchung, wie gut eine Pareto-Verteilung die Verteilung der Nachweislänge von Fehlern in IT-Systemen annähert.

k	form factor of the Pareto distribution.
n_{\min}	scale parameter of the Pareto distribution.
N	detection length (random variable).
n	number of tests.

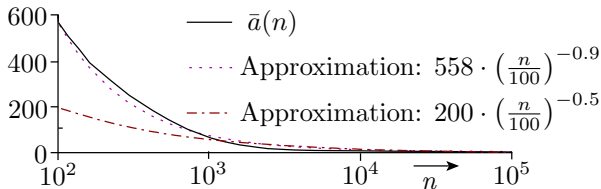
Für das Haftfehlerexperiment



Kombinatorische Beispielschaltung (Benchmark c3540). 3606 simulierte, unterschiedlich nachweisbare Haftfehler. Zu erwartende Anzahl der nicht nachweisbaren Fehler:

$$\bar{a}(n) = 3606 \cdot (1 - \mathbb{E}[FC(n)])$$

Pareto-Näherung

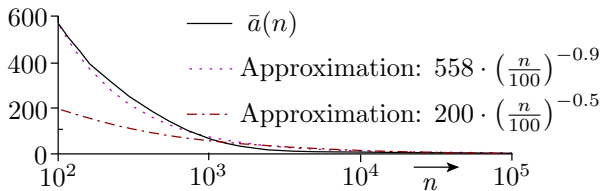


$$\bar{a}(n) = 3606 \cdot (1 - \mathbb{E}[FC(n)])$$

$$\mathbb{E}[FC(n)] = 1 - \frac{\bar{a}(n)}{3606} \approx F_N(t) = 1 - \left(\frac{n}{n_{\min}}\right)^{-k}$$

$$\bar{a}(n) = \frac{3606}{n_{\min}^k} \cdot \left(\frac{n}{n_{\min}}\right)^{-k}$$

- k form factor of the Pareto distribution.
- n_{\min} scale parameter of the Pareto distribution.
- n number of tests.
- \bar{a} mean number of non-detectable faults.



Nicht perfekt. Die Approximation mit $k = 0,9$ nähert Bereich $n < 1000$ und die mit $k = 0,5$ Bereich $n > 1000$ Testschritte besser an.

Die empirische Anfangsannahme, dass eine Erhöhung der Testanzahl n um eine Dekade die Anzahl der nicht nachweisbaren Fehler um den k -ten Teil einer Dekade verringert mit $k \neq f(n)$ ist nur eine grobe Näherung.



Schaden durch MF



Schaden durch Fehlfunktionen

Die möglichen Schäden durch Fehlfunktionen von IT-Systemen sind vom Einsatz abhängig und reichen von »unerheblich« über sehr hoch (Verlust großer Datenmengen) bis zu unbezahlbaren Katastrophen (Krieg mit Atomwaffen (see slide 1.11 *The price of lacking dependability*)).

Anschaulich gilt das Pareto-Prinzip, dass ein kleiner Teil der MF den überwiegenden Teil des Schadens verursacht.

Mangels verfügbaren Schadensstatistik für IT-Fehlfunktionen, betrachten wir die Verteilung von Haftpflichtschäden einer Autoversicherung.



Verteilung von Haftpflichtschäden

Haftpflichtschäden über 100.000 SF (SF – Schweizer Franken) einer Schweizer Autoversicherung¹:

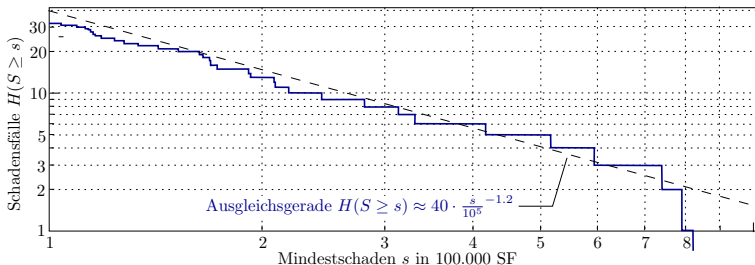
103.765, 109.168, 112.341, 113.800, 114.791,
115.731, 118.264, 123.464, 127.611, 133.504,
142.821, 152.270, 163.491, 164.968, 168.915,
169.346, 172.668, 191.954, 193.102, 208.522,
209.070, 219.111, 243.910, 280.302, 313.898,
330.461, 418.074, 516.218, 595.310, 742.198,
791.874, 822.787, 1.074.499

- Anzahl der Schadensfälle: 33
- Gesamtschadenssumme: 9.458.208 SF

¹Aus Klüppelberg, C. and Villasenor, J. A. (1993) Estimation of distribution tails – A semiparametric approach, Bl. Dtsch. Ges. Versicherungsmath. 21, No.2, 213-235.

Annäherung durch eine Pareto-Verteilung

Schadenshäufigkeit in Abhängigkeit von der Schadenshöhe:



Verteilung der Schadenshöhe S für Schäden $S > 10^5$ ist hier eine Pareto-Verteilung mit Formfaktor $k = 1,2$ und Skalenparameter $s_{\min} = 10^5$:

$$F_S(s) = \mathbb{P}[S \leq s] = 1 - \left(\frac{s}{s_{\min}}\right)^{-k} = 1 - \left(\frac{s}{10^5}\right)^{-1,2}$$

Erwartungswert (eq. 3.76):

$$\mathbb{E}[S] = s_{\min} \cdot \frac{k}{k-1} = \frac{1,2}{1,2-1} \cdot s_{\min} = 600.000 \text{ SF}$$

Eine Pareto-Verteilung hat erst für $k > 2$ eine Varianz. Alle im Kapitel behandelten Bereichsschätzungen incl. über die Tschebyscheffsche Ungleichung nicht anendbar. Abschätzung, wie viel Geld ein Versicherungsunternehmen als Rücklage haben muss, um jeden Schaden erstatten zu können, schwierig. Vermutlich haben Versicherungen deshalb eine max. Deckungssumme.

Der Autor geht davon aus, dass künftig Schäden durch IT z.B. in autonomen Fahrzeugen ähnlich wie heute Haftpflichtschäden durch Personen versichert werden müssen.



Zusammenfassung



Verteilungsfunktion und Dichte der Pareto-Verteilung:

$$F_X(x) = \mathbb{P}(X \leq x) = 1 - \left(\frac{x_{\min}}{x}\right)^k \quad (3.74)$$

$$f_X(x) = \frac{k \cdot x_{\min}^k}{x^{k+1}} \quad (3.75)$$

Erwartungswert für $k > 1$:

$$\mathbb{E}[X] = x_{\min} \cdot \frac{k}{k-1} \quad (3.76)$$

Varianz für $k > 2$:

$$\text{Var}[X] = x_{\min}^2 \cdot \frac{k}{(k-2)(1-k)^2} \quad (3.77)$$

Die Pareto-20-80-Regel beschreibt den Sonderfall $k = 1,16$.

Für grobe Überschläge kann die Fehlernachweislänge als pareto-verteilt mit $0 < k < 1$ betrachtet werden. Kein Erwartungswert.

Die Beispielsschadenszahlen ließen sich durch eine Pareto-Verteilung mit $k = 1,2$ annähern, d.h. Erwartungswert ja, Varianz nein.



Gamma- und Exp-Vert.



Gamma-Verteilung

$\mathcal{G}(\alpha, \beta)$ ist eine stetige Wahrscheinlichkeitsverteilung auf einem rechtsseitig unendlichen Intervall $[0, \infty)$ z.B. zur Modellierung

- von Bedien- und Reparaturzeiten (Warteschlangentheorie),
- kleiner und mittlerer Schäden (Versicherungsmathematik),
- der FF-Raten von Fehlern in IT-Systemen.

Dichte:

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} \quad \text{für } x \geq 0 \quad (79)$$

Gamma-Funktion:

$$\Gamma(\alpha) = \int_0^\infty e^{-z} \cdot z^{\alpha-1} \cdot dz \quad (80)$$

α	shape parameter.
β	scale parameter.
$\Gamma(\dots)$	gamma function.



Gammafunktion

Für $0 < \alpha \leq 1$ beträgt die Gammafunktion

$$\Gamma(\alpha) = \int_0^{\infty} e^{-z} \cdot z^{\alpha-1} \cdot dz \approx \frac{1}{\alpha}$$

α	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
$\Gamma(\alpha)$	9,51	4,59	2,99	2,22	1,77	1,49	1,30	1,16	1,07

Für $\alpha > 1$ gilt:

$$\Gamma(\alpha + 1) = \alpha \cdot \Gamma(\alpha) \quad (81)$$

$\Gamma(\dots)$ gamma function.
 α shape parameter.



Eigenschaften der Gamma-Verteilung

Erwartungswert:

$$\mathbb{E}[X] = \frac{\alpha}{\beta} \quad (82)$$

Varianz:

$$\text{Var}[X] = \frac{\alpha}{\beta^2} \quad (83)$$

Die Summe gamma-verteilter Zufallsvariablen mit gleichem Skalenparameter $X_1 \sim \mathcal{G}(\alpha_1, \beta)$ und $X_2 \sim \mathcal{G}(\alpha_2, \beta)$ ist wieder gamma-verteilt:

$$X_1 + X_2 \sim \mathcal{G}(\alpha_1 + \alpha_2, \beta) \quad (84)$$



Exponentialverteilung-Verteilung

Die Exponentialverteilung $\text{Exp}(\lambda) = \mathcal{G}(1, \lambda)$ ist eine Gamma-Verteilung mit Formparameter $\alpha = 1$ und Skalenparameter $\beta = \lambda$. Dichte:

$$\begin{aligned} f_X(x) &= e^{-\beta \cdot x} \cdot x^{\alpha-1} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \quad \text{für } x \geq 0 \text{ mit } \alpha = 1 \text{ und } \beta = \lambda \\ &= \lambda \cdot e^{-\lambda \cdot x} \quad \text{für } x \geq 0 \end{aligned} \quad (85)$$

λ – Anzahl der zu erwartenden Ereignisse pro Zeitintervall.
Verteilungsfunktion:

$$F_X(x) = \mathbb{P}[X \leq x] = \int_0^x f(z) \cdot dz = 1 - e^{-\lambda \cdot x} \quad \text{für } x \geq 0 \quad (86)$$

Erwartungswert:

$$\mathbb{E}[X] = \frac{1}{\lambda} \quad (87)$$

Varianz:

$$\text{Var}[X] = \frac{1}{\lambda^2} \quad \Rightarrow \quad \text{sd}[X] = \mathbb{E}[X] \quad (88)$$



Korrektur Dichte MF-Rate etc. ab slide 1.129

Die Nachweislänge N als Wahrscheinlichkeit, dass ein Fehler mit n Tests nachgewiesen wird, ist gleich der zu erwartenden Fehlerüberdeckung und für die Grobabschätzungen pareto-verteilt:

$$F_N(n) = \mathbb{E}[FC(n)] = 1 - \left(\frac{n_{\min}}{n}\right)^k$$

Auf Foliensatz 1, slide 1.129 hatten wir unter der vereinfachten Annahme, dass ein Zufallstest alle Fehler mit MF-Rate $\zeta \geq \frac{1}{n}$ nachweist, daraus abgeleitet, Verteilungsfunktion der MF-Rate

$$F(\zeta) = \begin{cases} (n \cdot \zeta)^k & 0 \leq \zeta < \frac{1}{n} \\ 1 & \zeta \geq \frac{1}{n} \end{cases} \quad (1.37)$$

Dichtefunktion der MF-Rate

$$h(\zeta) = \frac{dF(\zeta)}{d\zeta} = \begin{cases} k \cdot n^k \cdot \zeta^{k-1} & 0 \leq \zeta < \frac{1}{n} \\ 0 & \text{sonst} \end{cases} \quad (1.38)$$

-
- n_{\min} reference test number for $FC = 0$.
 - n number of tests for which all detected faults are eliminated including n_{\min} .
 - k form factor of the distribution of the malfunction rate ($0 < k < 1$).



Korrektur p_{NEF} und Dicht

Die Wahrscheinlichkeit, dass n Zufallstests einen Fehler mit einer MF-Rate ζ nicht nachweisen, so das er nicht beseitigt wird, ist nicht die auf slide 1.129 unterstellte Sprungfunktion

$$p_{\text{NEF}}(\zeta, n) = \begin{cases} 1 & \zeta < \frac{1}{n} \\ 0 & \zeta \geq \frac{1}{n} \end{cases}$$

sondern in Anlehnung an Gl. 2.9 eine abklingende Exponentialfunktion:

$$p_{\text{NEF}}(\zeta, n) = e^{-\zeta \cdot n}$$

Die Dichtefunktion ist folglich nicht ab $\zeta \geq \frac{1}{n}$ null, sondern nimmt im Bereich $\zeta \approx 1/n$ stetig ab::

$$h(\zeta) = \frac{1}{\int_0^1 h(\zeta) \cdot d\zeta} \cdot n^k \cdot \zeta^{k-1} \cdot e^{-\zeta \cdot n}$$

p_{FNE}	probability of fault n ot e liminated.
ζ	malfunction rate of a single fault.
n	number of tests for which all detected faults are eliminated including n_{min} .
k	form factor of the distribution of the malfunction rate ($0 < k < 1$).



5. Gamma- und Exp-Vert.

Das Normierungsintegral $\int_0^1 h(\zeta) \cdot d\zeta$ ist für $n \gg 1$ die Gamma-Funktion

$$\Gamma(\alpha) = \int_0^\infty e^{-z} \cdot z^{\alpha-1} \cdot dz \quad (3.80)$$

$$\int_0^1 h(\zeta) \cdot d\zeta = \int_0^1 n^k \cdot \zeta^{k-1} \cdot e^{-\zeta \cdot n} \cdot d\zeta$$

Substitution $\zeta = \frac{z}{n}$, $d\zeta = \frac{dz}{n}$ und $n \gg 1$:

$$\int_0^n n^k \cdot \left(\frac{z}{n}\right)^{k-1} \cdot e^{-z} \cdot \frac{dz}{n} = \int_0^n z^{k-1} \cdot e^{-z} \cdot dz = \Gamma(k)$$

Korrigierte Dichtefunktion der MF-Rate:

$$h(\zeta) = \frac{n^k}{\Gamma(k)} \cdot \zeta^{k-1} \cdot e^{-n \cdot \zeta}$$



5. Gamma- und Exp-Vert.

Die mittlere MF-Rate je Fehler ist nicht wie auf slide 1.129

$$\bar{\zeta} = \int_0^{\frac{1}{n}} \zeta \cdot h(\zeta) \cdot d\zeta = \frac{k}{(k+1) \cdot n} \quad (1.39)$$

sondern der Erwartungswert einer Gammaverteilung:

$$\mathbb{E}[\zeta(n)] = \frac{k}{n}$$

Für die gesamte MF-Rate ist nicht wie auf slide 1.130

$$\zeta_F = \mu_{\text{FNE}}(n) \cdot \bar{\zeta} = \frac{\mu_{\text{FNE}}(n) \cdot k}{(k+1) \cdot n}$$

sondern gehorcht der auf slide 1.135 bereits korrigierten Beziehung

$$\zeta_F(n) = \frac{k \cdot \mu_{\text{FNE}}(n)}{n} \quad (1.43)$$

In allen nachfolgenden Gleichungen auf Foliensatz 1 ist die Korrektur, dass die MF-Rate in Wirklichkeit Gamma-verteilt ist, berücksichtigt.

n	number of tests for which all detected faults are eliminated including n_{\min} .
k	form factor of the distribution of the malfunction rate ($0 < k < 1$).
ζ_F	malfunction rate caused by faults.
μ_{FNE}	expected number of not eliminated faults .



Example 3.16: Fehleranzahl und MF-Rate

Die MF-Rate eines Systems beträgt nach der effektiven Testlänge $n_0 = 10^5$ Tests $\zeta(n_0) = 10^{-4}$ MF/SR. Schätzen Sie für gamma-verteilte MF-Raten mit den Formfaktoren $k \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$

- zu erwartende Anzahl nicht beseitigten Fehler μ_{FNE} Testlänge n_0 ,
- die zu erwartende Anzahl der nicht beseitigten Fehler für die zehnfache effektive Testlänge $n_1 = 10^6$,
- die MF-Rate für die zehnfache effektive Testlänge $n_1 = 10^6$.

Für die Aufgabe werden die folgenden Gl. von Foliensatz 1 benötigt:

$$\mu_{\text{FNE}}(n) = \mu_{\text{FNE}}(n_0) \cdot \left(\frac{n}{n_0}\right)^{-k} \quad (1.42)$$

$$\zeta_{\text{F}}(n) = \frac{k \cdot \mu_{\text{FNE}}(n)}{n} \quad (1.43)$$

$$\zeta_{\text{F}}(n) = \zeta_{\text{F}}(n_0) \cdot \left(\frac{n}{n_0}\right)^{-(k+1)} \quad (1.44)$$

μ_{FNE} expected number of **not eliminated faults**.

ζ_{F} malfunction rate caused by faults.



5. Gamma- und Exp-Vert.

Die MF-Rate eines Systems beträgt nach der effektiven Testlänge $n_0 = 10^5$ Tests $\zeta(n_0) = 10^{-4}$ MF/SR. Schätzen Sie für gamma-verteilte MF-Raten mit den Formfaktoren $k \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$

a) zu erwartende Anzahl nicht beseitigten Fehler μ_{FNE} Testlänge n_0 ,

$$\zeta_{\text{F}}(n) = \frac{k \cdot \mu_{\text{FNE}}(n)}{n} \quad (1.43)$$

$$\mu_{\text{FNE}}(n_0) = \zeta_{\text{F}}(n_0) \cdot \frac{n_0}{k} = 10^4 \cdot \frac{10^5}{k} = \frac{10}{k}$$

k	0,3	0,4	0,5	0,6	0,7
$\mu_{\text{FNE}}(10^5)$	33,3	25	20	16,7	14,3

ζ_{F}	malfunction rate caused by faults.
k	form factor of the distribution of the malfunction rate ($0 < k < 1$).
μ_{FNE}	expected number of not eliminated faults.
n	number of tests.



5. Gamma- und Exp-Vert.

Die MF-Rate eines Systems beträgt nach der effektiven Testlänge $n_0 = 10^5$ Tests $\zeta(n_0) = 10^{-4}$ MF/SR. Schätzen Sie für gamma-verteilte MF-Raten mit den Formfaktoren $k \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$

b) die zu erwartende Anzahl der nicht beseitigten Fehler für die zehnfache effektive Testlänge $n_1 = 10^6$,

$$\mu_{\text{FNE}}(n) = \mu_{\text{FNE}}(n_0) \cdot \left(\frac{n}{n_0}\right)^{-k} \quad (1.42)$$

$$\mu_{\text{FNE}}(10^6) = \mu_{\text{FNE}}(10^5) \cdot \left(\frac{10^6}{10^5}\right)^{-k} = \frac{10}{k} \cdot 10^{-k}$$

k	0,3	0,4	0,5	0,6	0,7
$\mu_{\text{FNE}}(10^6)$	16,7	9,9	6,3	4,2	2,9

μ_{FNE}	expected number of not eliminated faults .
n_0	number of tests with known MF rate or number of faults, respectively.
n	number of tests including n_0 .
k	form factor of the distribution of the malfunction rate ($0 < k < 1$).



5. Gamma- und Exp-Vert.

Die MF-Rate eines Systems beträgt nach der effektiven Testlänge $n_0 = 10^5$ Tests $\zeta(n_0) = 10^{-4}$ MF/SR. Schätzen Sie für gamma-verteilte MF-Raten mit den Formfaktoren $k \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$

c) die MF-Rate für die zehnfache effektive Testlänge $n_1 = 10^6$.

$$\zeta_F(n) = \zeta_F(n_0) \cdot \left(\frac{n}{n_0}\right)^{-(k+1)} \quad (1.44)$$

$$\zeta_F(10^6) = \zeta_F(10^5) \cdot \left(\frac{10^6}{10^5}\right)^{-(k+1)} = 10^{-4} \cdot 10^{-(k+1)}$$

k	0,3	0,4	0,5	0,6	0,7
$\zeta_F(10^6)$	$5,0 \cdot 10^{-6}$	$4,0 \cdot 10^{-6}$	$3,2 \cdot 10^{-6}$	$2,5 \cdot 10^{-6}$	$2,0 \cdot 10^{-6}$

ζ_F	malfunction rate caused by faults.
n_0	number of tests with known MF rate or number of faults, respectively.
n	number of tests including n_0 .
k	form factor of the distribution of the malfunction rate ($0 < k < 1$).



Ausfälle



Ausfälle

Hardware und Mechanik unterliegt einem Verschleiß, der zu Ausfällen führen kann. Bei einem Ausfall entsteht ein Fehler, der oft mehr Fehlfunktionen (MF) als alle vom Test nicht erkannten Fehler zusammen oder ein komplettes Versagen (NS) verursacht.

Kenngößen:

- Lebensdauer, Überlebenswahrscheinlichkeit,
- Ausfallrate, ...

Maßnahmen zum Umgang mit Ausfällen:

- Voralterung,
- Wartung,
- Redundanz (kalte oder heiße Reserve).

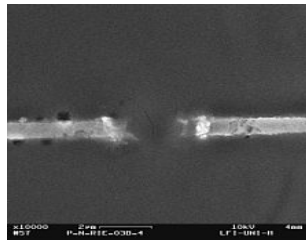
In Software entstehen während des Betriebs keine neuen Fehler, ausgenommen geplante Obsoleszenz (Programm-Features zur Vortäuschung von Ausfällen).

MF	malfunction.
NS	no service.

Verschleiß elektronischer Bauteile

Langsam ablaufende physikalische Vorgänge:

- Korrosion (Stecker, Schalter, Isolationen, Leiterbahnen, ...).
- Elektromigration: strombedingte Wanderung von Metalatomen bei hohen Stromdichten.
- Parameterdrift: Widerstandswerte, Kapazitäten, Schwellspannungen etc.
- Gateoxiddurchschlag: Hochschaukelnde Tunnelströme, Ladungseinlagerung bis zum lokalen Schmelzen des Oxids. Bildung von Kurzschlüssen. Phänomen: Zunahme des Stromverbrauchs über Monate bis zum Ausfall.





Kenngrößen



Kenngrößen des Ausfallverhaltens

- Lebensdauer t_L : Zeit vom Beanspruchungsbeginn bis zum Ausfall. Verteilungsfunktion:

$$F(t) = \mathbb{P}[t_L \leq t]$$

- Überlebenswahrscheinlichkeit:

$$V(t) = \mathbb{P}[t_L > t] = 1 - F(t)$$

- Ausfallrate λ : Relative Abnahme der Überlebenswahrscheinlichkeit mit der Zeit:

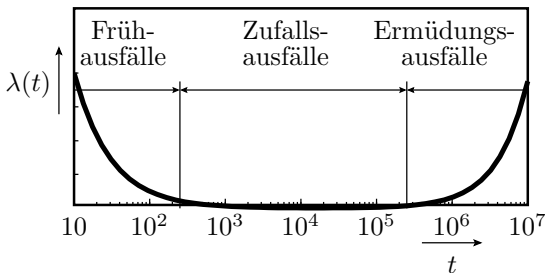
$$\lambda(t) = -\frac{1}{V(t)} \cdot \frac{dV(t)}{dt}$$

- Mittlere Lebensdauer:

$$\mathbb{E}[t_L] = \int_0^{\infty} V(t) \cdot dt$$

t_L	lifetime.
V	Probability of survival.
λ	failure rate.

Ausfallphasen



- Frühausfälle (infant mortalities): Erhöhte Ausfallrate durch Schwachstellen (Materialrisse, lokal stark überhöhte Feldstärke oder Stromdichte, ...).
- Zufallsausfälle: Näherungsweise konstante Ausfallrate.
- Ermüdungsausfälle: Anstieg der Ausfallrate: Materialermüdung, ...

λ failure rate.
 t time.



Zufallsausfälle

In der Hauptnutzungsphase ist die Ausfallrate konstant:

$$\lambda(t) = -\frac{1}{V(t)} \cdot \frac{dV(t)}{dt} = \lambda = \text{konst.}$$

Überlebenswahrscheinlichkeit und Verteilung der Lebensdauer*:

$$V(t) = e^{-\lambda \cdot t} \quad (89)$$

$$F(t) = 1 - e^{-\lambda \cdot t}$$

(Exponentialverteilung vergl. Gl. 3.86). Mittlere Lebensdauer (eq. 3.87):

$$\bar{t}_L = \mathbb{E}[t_L] = \int_0^{\infty} V(t) \cdot dt = \frac{1}{\lambda}$$

Maßeinheit der Ausfallrate: fit (failure in time)

$$1 \text{ fit} = 1 \text{ Ausfall in } 10^9 \text{ Stunden}$$

V	Probability of survival.
fit	failure in time (Number of failures in 10^9 hours).
*	Kontrolle: $\lambda(t) = -\frac{1}{e^{-\lambda \cdot t}} \cdot \frac{d(e^{-\lambda \cdot t})}{dt} = \lambda.$



System aus mehreren Komponenten

Ein System aus mehreren notwendigen Komponenten überlebt, solange alle Komponenten überleben:

$$V(t) = \prod_{i=1}^{\#K} V(t)_i$$

Mit einer konstanten Ausfallrate λ_i für alle Komponenten (eq. 3.89):

$$V(t) = \prod_{i=1}^{\#K} e^{-\lambda_i \cdot t} = e^{-(\sum_{i=1}^{\#K} \lambda_i) \cdot t}$$

ist die Gesamtausfallrate die Summe der Ausfallraten aller Komponenten:

$$\lambda_{\text{Sys}} = \sum_{i=1}^{\#K} \lambda_i \quad (90)$$

V	Probability of survival.
$\#K$	number of components.
λ_{Sys}	failure rate of the system.
λ_i	failure rate of component i .



Ausfallraten in der Hauptnutzungsphase nach [KRG96]

Bauteil	Ausfallrate in fit	Bauteil	Ausfallrate in fit
diskrete HBT	1 bis 100	Widerstände	1 bis 20
digitale IC	50 bis 200	Kondensatoren	1 bis 20
ROM	100 bis 300	Steckverbinder	1 bis 100
RAM	bis 500	Lötstellen	0,1 bis 1
analoge IC	20 bis 300		

- Ausfallrate = Ausfallanzahl / Bauteilanzahl
- Bei mehreren Bauteilen und konstanten Ausfallraten addieren sich die Ausfallraten.

SC Semiconductor components.

IC integrated circuits.

[KRG96] Kärger, R.: Diagnose von Computern, Teubner 1996, S. 68.

Ausfallrate einer Baugruppe

Bauteiltyp	Anzahl n	Ausfallrate λ	$n \cdot \lambda$
Schaltkreise	20	150 fit	3000 fit
diskrete BT	15	30 fit	450 fit
Kondensatoren	15	10 fit	250 fit
Widerstände	30	10 fit	300 fit
Lötstellen	2000	0,5 fit	1000 fit
Baugruppe			5000 fit

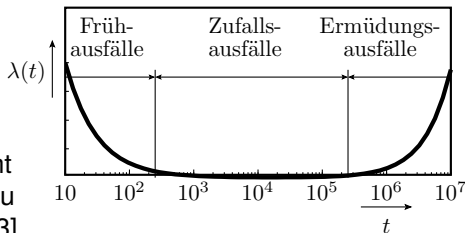
- Im Mittel 1 Ausfall in $2 \cdot 10^5$ Stunden (≈ 23 Jahre) Betriebsdauer.
- Von den heutigen PCs, Handys, ... fallen pro Jahr und hundert Stück nur wenige aus.
- Nach 2 ... 5 Jahren erste Ermüdungsausfälle, z.B. durch Austrocknung von Elektrolytkondensatoren.



Voralterung

Frühausfälle

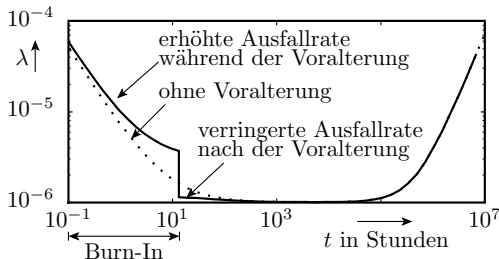
- Auf 100 richtige Fehler kommt etwa ein Beinahefehler, der zu einem Frühausfall führt [BS03].
- Bei 50% fehlerfreien und 50% aussortierten Schaltkreisen $50\%/100 = 0,5\%$ Beinahefehler.
- Die Hälfte wird mit dem Ausschuss aussortiert.
 - $\approx 0,25\%$ (jeder 400ste) Schaltkreis verursacht ein Frühausfall.
 - Bei 20 Schaltkreisen pro Gerät jedes zwanzigste Gerät.
 - Bei großen Systemen fast jedes System.
- Frühausfälle sind Garantiefälle und verursachen Kosten für Reparatur, Ersatz, Auftragsabwicklung, ... Was tun?
- Problem existiert auch bei Mechanik.



λ failure rate.

[BS03] Singh, A. D.: Relating Yield Models to Burn-In Fall-Out in Time. ITC, 12/2003, S.77-84.

Voralterung (Burn-In)



Beschleunigung der Alterung vor dem Einsatz durch »harte Umgebungsbedingungen«:

- überhöhte Spannung,
- überhöhte Temperatur,
- Stress (Burn-In).

Einsatz erst nach der Frühausfallphase, wenn die kränklichen Bauteile »gestorben« und ausgetauscht sind.



Reserveeinheiten



Reserveeinheiten

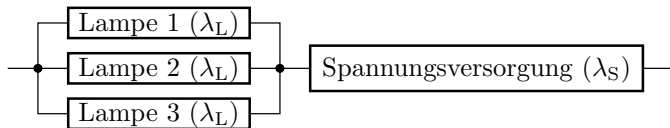
Reserve- oder redundante Einheiten sind Komponenten,

- die für Aufrechterhaltung der Funktion nicht unbedingt erforderlich sind und
- nach einem Ausfall die Funktion der ausgefallenen Komponenten übernehmen.

Ausfallplan mit Reserveeinheiten

Im Ausfallplan werden notwendige Komponenten für die Verfügbarkeit des Services als Reihenschaltung und Reserveeinheiten (Redundanzen) als Parallelschaltung dargestellt*.

Eine Flurbeleuchtung sei verfügbar, wenn mindestens eine von drei Lampen und die Spannungsversorgung funktioniert:



Reserveeinheiten sind erforderlich für

- Systeme ohne Reparaturmöglichkeit, die lange verfügbar sein müssen (z.B. Satelliten),
- hoher Verfügbarkeit (z.B. Serverdienste) zur Überbrückung von Ausfallzeiten der Hauptkomponenten.

* Abgeleitet von Glühlampen. Reihenschaltung fällt mit der ersten Glühlampe aus, bei Parallelschaltung fällt die Beleuchtung erst mit der letzte ganz aus.



Kalte, warme und heiße Reserve

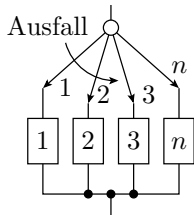
- Heiße Reserve: Reservekomponenten arbeiten parallel (z.B. Mehrversionssystem) und fallen mit derselben Wahrscheinlichkeit wie das aktive System aus.
- Kalte Reserve: Reservekomponenten werden geschont und funktionieren idealerweise noch alle zum Ausfallzeitpunkt der aktiven Komponente.
- Warme Reserve: Reserveeinheiten (z.B. das Reserverad im Auto) altern auch bei Nichtnutzung, nur langsamer.

Die beiden zusätzlichen Lampen auf der Folie zuvor, die für die Verfügbarkeit der Treppenbeleuchtung nicht unbedingt funktionieren müssen, bilden eine heiße Reserve, Ersatzlampen, die erst nach Ausfall der »Hauptlampe« eingeschaltet werden, eine kalte Reserve, ein Ersatzrad im Auto eine warme Reserve, weil der Gummi auch ohne Beanspruchung altert.

Zu erwartende Lebensdauer bei kalte Reserve

Für jede Komponente beginnt die Belastung erst nach Ausfall der vorherigen Komponente.

Phase	mittlere Lebensdauer
1	$\mathbb{E}[t_{L.1}]$
2	$\mathbb{E}[t_{L.2}]$
3	$\mathbb{E}[t_{L.3}]$
...	...



Die zu erwartenden Lebensdauern aller Komponenten addieren sich*:

$$\mathbb{E}[t_{L.Sys}] = \sum_{i=1}^{\#K} \mathbb{E}[t_{L.i}] \quad (91)$$

$t_{L.Sys}$ life time of the system.

$\#K$ number of components.

$t_{L.i}$ life time of component i .

* Annahme, dass Umschalter und ungenutzten Reserveeinheiten Ausfallrate null haben.

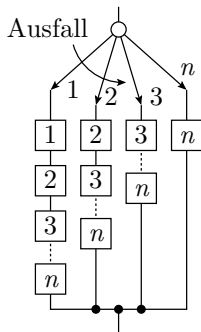
Zu erwartende Lebensdauer bei heiße Reserve

Alle noch lebenden Komponenten können gleichermaßen ausfallen:

$$\mathbb{E}[t_{L.i}] = \frac{1}{\sum_{j=1}^i \lambda_j}$$

Komponenten mit gleicher Ausfallrate λ_K :

Phase	mittlere Lebensdauer
1	$\mathbb{E}[t_{L.K}] / n$
2	$\mathbb{E}[t_{L.K}] / (n - 1)$
...	...

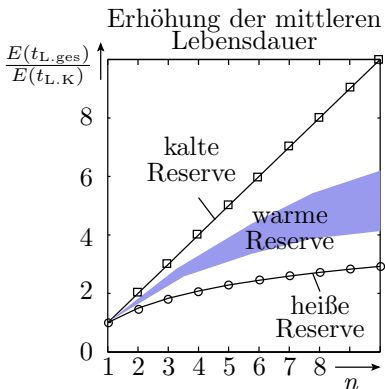
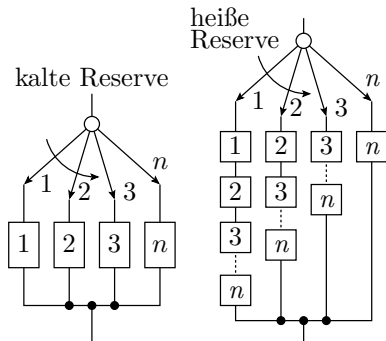


Die zu erwartenden Lebensdauern aller Phasen addieren sich:

$$\mathbb{E}[t_{L.Sys}] = \mathbb{E}[t_{L.K}] \sum_{i=1}^{\#K} \frac{1}{i} \quad (92)$$

$t_{L.Sys}$ life time of the system.
 $t_{L.K}$ life time of a component.

Zu erwartende Lebensdauer warme Reserve



- Die Ausfallrate der »kalten« Ersatzkomponenten ist kleiner als im aktiven Zustand, aber größer null.
- »Warme« Reserveeinheiten verlängert die Lebensdauer mehr als »heiße« und weniger als »kalte«.



Wartung



Wartung

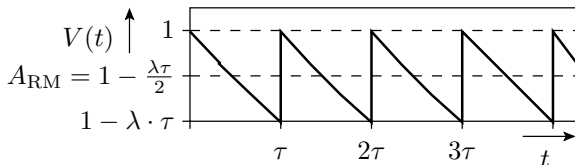
Wartung:

- Test und die Beseitigung aller erkennbaren Fehler, die seit der letzten Wartung entstanden sind.
- Ergänzen und Ersatz von Betriebsstoffen und Verbrauchsmitteln (Getrieben Schmierstoffe, bei Druckern Papier und Toner).
- Planmäßiger Austausch von Verschleißteilen vor der Ermüdungsphase, in der die Ausfallrate stark zunimmt (in PCs die Batterien für den BIOS-RAM, in Servern die Festplatten).

Wartungsintervall τ : Zeit zwischen den Wartungen, z.B. 1 Jahr.

[KFZ-Wartungsintervall: TÜV alle 2 Jahre]

Reparatur bei der nächsten Wartung



Verfügbarkeit als mittlere Überlebenswahrscheinlichkeit, wenn beim Wartungstest alle Ausfälle beseitigt werden für $\lambda \cdot \tau \ll 1$:

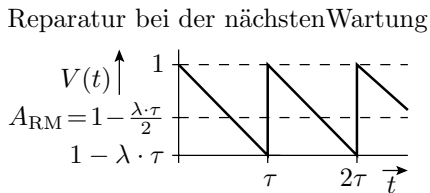
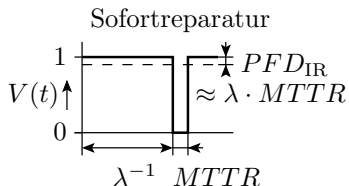
$$A_{RM} = \frac{1}{\tau} \cdot \int_0^{\tau} V(t) \cdot d\tau = \frac{1}{\tau} \cdot \int_0^{\tau} e^{-\lambda \cdot t} \cdot d\tau = \frac{1 - e^{-\lambda \cdot \tau}}{\lambda \cdot \tau}$$

mit

$$e^{-\lambda \cdot \tau} = 1 - \lambda \cdot \tau + \frac{(\lambda \cdot \tau)^2}{2}$$

$$A_{RM} = 1 - \frac{\lambda \cdot \tau}{2}; \quad PFD_{RM} = 1 - A_{RM} = \frac{\lambda \cdot \tau}{2} \quad (93)$$

V	Probability of survival.
A_{RM}	maintenance interval.
τ	maintenance interval.



Wenn erkannte Ausfälle sofort beseitigt werden, ist die PFD der Anteil der Reparaturzeit an der mittleren Zeit zwischen zwei Ausfällen:

$$PFD_{IR} = \frac{MTTR}{MTTF + MTTR} = \lambda \cdot MTTR \quad \text{für } \lambda \cdot MTTR \ll 1$$

Wahrscheinlichkeit der Nichtverfügbarkeit insgesamt:

$$\begin{aligned} PFD &= \eta_{IR} \cdot PFD_{IR} + (1 - \eta_{IR}) \cdot PFD_{RM} \\ &= \eta_{IR} \cdot \lambda \cdot MTTR + (1 - \eta_{IR}) \cdot \frac{\lambda \cdot \tau}{2} \end{aligned} \quad (94)$$

-
- η_{IR} proportion of failures eliminated by immediate repair.
 - PFD_{IR} Probability of failure on demand in case of immediate repair.
 - PFD_{RM} Probability of failure on demand if repair only at next maintenance.
 - $MTTR$ mean time to repair.
 - PFD probability of failure on demand.

Example 3.17: PFD

Ausfallrate $\lambda = 10^{-6} \text{ h}^{-1}$, Anteil der Ausfälle, die sofort beseitigt werden $\eta_D = 75\%$. Wartungsintervall $\tau = 2 \cdot 10^3 \text{ h}$, mittlere Reparaturzeit $MTTR = 4 \text{ h}$. Gesucht PFD .

$$PFD = \eta_{IR} \cdot \lambda \cdot MTTR + (1 - \eta_{IR}) \cdot \frac{\lambda \cdot \tau}{2} \quad (3.95)$$

$$PFD = 10^{-6} \text{ h}^{-1} \cdot \left(\underbrace{0,75 \cdot 4 \text{ h}}_{3 \text{ h}} + \underbrace{\frac{0,25 \cdot 2 \cdot 10^3 \text{ h}}{2}}_{500 \text{ h}} \right) = 2,53 \cdot 10^{-4}$$

Zur Verringerung der PFD bieten sich an:

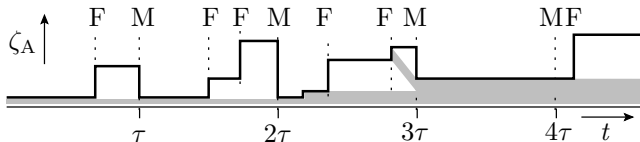
- Verkürzung des Wartungsintervalls τ oder
- die Erhöhung des Anteils η_D der sofort erkannt Ausfälle durch bessere Überwachung und Einschalttests.

Die mittlere Reperaturzeit dürfte sogar größer sein, ohne dass sich die PFD wesentlich erhöht.

PFD probability of failure on demand.

η_{IR} proportion of failures eliminated by immediate repair.

Zuverlässigkeitsverlust durch Ausfälle



- MF-Rate nicht erkennbarer Fehler durch Ausfälle
- M Wartung (maintenance) + Fehlerbeseitigung
- F Ausfall (failure)

Fehler durch Ausfälle mit kleinem ζ_A

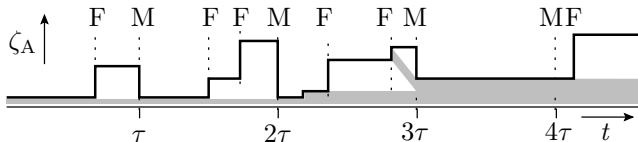
- beeinträchtigen statt der Verfügbarkeit die Zuverlässigkeit und
- werden von Wartungs- und Einschalttests nur mit $FC < 1$ erkannt.

Auch bei regelmäßiger Wartung nimmt die FF-Rate mit der Nutzungsdauer zu und die Zuverlässigkeit ab.

ζ_A malfunction rate of faults due to failures.

τ maintenance interval.

Minderung der Zuverlässigkeitsabnahme



MF-Rate nicht erkennbarer Fehler durch Ausfälle
 M Wartung (maintenance) + Fehlerbeseitigung
 F Ausfall (failure)

- Prophylaktischer Tausch von Hardware, wenn ein Zuverlässigkeitsverlust wahrgenommen wird.
- Gründlicher Wartungstest.
- Ersatz der Gesamtsystem, bevor sich durch Ausfälle verursachte Fehler häufen.

ζ_A malfunction rate of faults due to failures.

τ maintenance interval.



Zusammenfassung

Ausfallverhalten von Hardware und Mechanik

Die Ausfallrate hat typisch einer Badewannenkurve:

- Nutzungsbeginn: Erhöhte Ausfallrate durch Frühausfälle.
- Nach den Frühausfällen näherungsweise konstante Ausfallrate.
- Nach langer Nutzungsdauer Anstieg der Ausfallrate durch Materialermüdung.

Für Zufallsausfälle ist die Lebensdauer exponentialverteilt

$$V(t) = e^{-\lambda \cdot t} \quad (3.89)$$

mit dem Erwartungswert:

$$\mathbb{E}[t_L] = \frac{1}{\lambda}$$

Für Systeme aus mehreren Komponenten addieren sich die Ausfallraten:

$$\lambda_{\text{Sys}} = \sum_{i=1}^{\#K} \lambda_i \quad (3.90)$$

Das Problem der erhöhten Ausfallrate durch Frühausfälle wird durch Voralterung umgangen und das durch Materialermüdung durch Wartungsmaßnahmen, insbesondere rechtzeitigen Ersatz.

Reserveeinheiten und Wartung

Bei kalter Reserve addieren sich die Lebensdauern und ihre Erwartungswerte:

$$\mathbb{E}[t_{L.Sys}] = \sum_{i=1}^{\#K} \mathbb{E}[t_{L.i}] \quad (3.91)$$

Bei heißer Reserve vergrößert die erste Reserveeinheit die zu erwartende Lebensdauer um 50%, die zweite um ein Drittel etc.:

$$\mathbb{E}[t_{L.Sys}] = \mathbb{E}[t_{L.K}] \sum_{i=1}^{\#K} \frac{1}{i} \quad (3.92)$$

Regelmäßige Wartung und Fehlerbeseitigung nur zu den Wartungszeitpunkten:

$$PFD_{RM} = 1 - A_{RM} = \frac{\lambda \cdot \tau}{2} \quad (3.93)$$

Wenn zusätzlich zur regelmäßigen Wartung erkannte Fehler durch Ausfälle sofort beseitigt werden:

$$PFD = \eta_{IR} \cdot \lambda \cdot MTTR + (1 - \eta_{IR}) \cdot \frac{\lambda \cdot \tau}{2} \quad (3.95)$$

Fehler durch Ausfälle, die von den Wartungstests nicht erkannt werden, verringern die Zuverlässigkeit mit der Nutzungsdauer.



Kontrolle Seitennummerreferenzen

(see slide 3.59 *Estimation of k_{\min}*) on page 59

(see slide 3.60 *Estimation of k_{\max}*) on page 60

(see slide 3.61 *Estimation of $[\lambda_{\min}, \lambda_{\max}]$ from an actual count value*)
on page 61

(see slide 3.74 *Standardized normal distribution*) on page 74

(see slide 3.167 *Early failures*) 168