



# Test und Verlässlichkeit

## Foliensatz 3: Verteilungen

Prof. G. Kemnitz

Institut für Informatik, TU Clausthal (TV\_F3)  
16. Januar 2024



## Inhalt Foliensatz 3

### Verteilung

- 1.1 Charakteristische Größen
- 1.2 Lineare Transformation
- 1.3 Verteilung von Zählwerten

### Näherungen

- 2.1 Binomialverteilung
- 2.2 Poisson-Verteilung
- 2.3 Bereichsschätz. Poissonvert.
- 2.4 Fehleranzahl und Defektanteil
- 2.5 Normalverteilung
- 2.6 Bereichsschätz. Normalvert.
- 2.7 Bereichsschätzung Zählwerte
- 2.8 Varianzerhöhung

### 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.**
- 5.1 Eigenschaften
- 5.2 MF-Rate  
**Ausfälle**
- 6.1 Kenngrößen
- 6.2 Voralterung
- 6.3 Reserveeinheiten
- 6.4 Wartung

Vorlesung	8	9	10	11
ca. ab Folie	5	54	141	203





# Verteilung



## Zufallsexperiment, Zufallsvariable, Verteilung

- Zufälliges Ereignis: Ereignis, das weder sicher noch unmöglich ist, sondern mit einer gewissen Wahrscheinlichkeit eintritt.
- Zufallsexperiment: Experiment mit mehreren möglichen Ergebnissen und zufälligem Ausgang.
- Zufallsvariable: Veränderliche, die ihre Werte in Abhängigkeit vom Ergebnis eines Zufallsexperiments nach einer Wahrscheinlichkeitsverteilung annimmt.
- Verteilung: Eine Verteilung ordnet den möglichen Werten einer Zufallsvariablen (also den Ergebnissen eines Zufallsexperiments) Wahrscheinlichkeiten zu.

Im weiteren interessierende Zufallsvariablen:

- Anzahl der Fehlfunktionen (aufgetretene, erkannte, ...),
- Anzahl der Fehler (entstandene, beseitigte, vermiedene, ...),
- Lebensdauer,
- Schadenskosten, ...



## Bernoulli-Versuch

Das einfachste Zufallsexperiment. Zweipunktverteilung:

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

Zufallsexperimente mit mehr als zwei möglichen Ergebnissen lassen sich in je einen Bernoulli-Versuch je Ergebnis aufspalten:

$$A_i = \begin{cases} 0 & \text{Ereignis nicht eingetreten} \\ 1 & \text{Ereignis eingetreten} \end{cases}$$

---

$X$	Zufallsvariable.
$\mathbb{P}[\dots]$	Eintrittswahrscheinlichkeit des Ereignisses ...
$p$	Eintrittswahrscheinlichkeit.
$A_i$	Ereignis $A_i$ .



## Charakteristische Größen



## Charakteristische Größen einer Zufallsvariablen

Verteilungsfunktion:

$$F_X(x) = \mathbb{P}[X \leq x] \quad (3.2)$$

Dichtefunktion für stetige Zufallsvariablen:

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (3.3)$$

Verteilung für diskrete Zufallsvariablen:

$$\mathbb{P}[X = x_i]$$

Erwartungswert einer diskreten Zufallsvariablen:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i \quad (3.4)$$

---

$X$	Zufallsvariable.
$x_i$	Realisierung (möglicher Wert) $i$ der Zufallsvariablen $X$ .
$\#x$	Anzahl der möglichen Realisierungen der Zufallsvariablen $X$ .
$p_i$	Eintrittswahrscheinlichkeit der Realisierung $x_i$ .



Erwartungswert einer stetigen Zufallsvariablen:

$$\mathbb{E}[X] = \mu = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx \quad (3.5)$$

Varianz, mittlere quadratische Abweichung vom Erwartungswert:

$$\text{Var}[X] = \sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \quad (3.6)$$

Standardabweichung, Wurzel aus der Varianz:

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (\sigma = \sqrt{\sigma^2}) \quad (3.7)$$

Speziell für Zählergebnisse von  $n$  Zählversuchen:

- Mittlere Eintrittswahrscheinlichkeit:

$$p = \frac{\mu}{n} \quad (3.8)$$

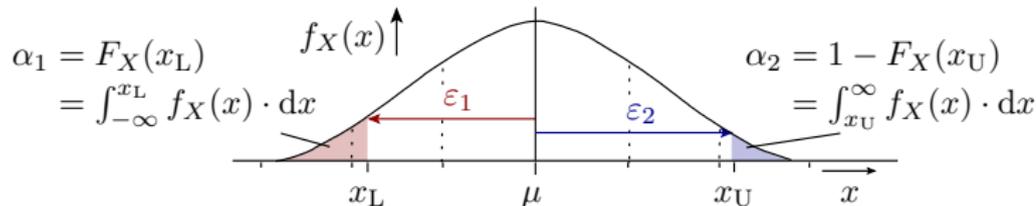
- Varianzkoeffizient des Eintretens bzw. Nichteintretens als Maße der Schätzgenauigkeit:

$$\begin{aligned} \sigma_r &= \frac{\sigma}{\mu} && \text{für } p \leq 50\% \\ \sigma_{\bar{r}} &= \frac{\sigma}{n-\mu} && \text{für } p > 50\% \end{aligned} \quad (3.9)$$

$p$	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
$\sigma_r$	Relative Standardabweichung zum erwarteten Eintrittszahlwert.
$\sigma_{\bar{r}}$	Relative Standardabweichung zum erwarteten Nichteintrittszahlwert.

## Wahrscheinlicher Bereich

Bereich  $[x_L, x_U]$ , die der Wert der Zufallsvariable  $X$  mit einer Wahrscheinlichkeit  $1 - (\alpha_1 + \alpha_2)$  annimmt:



$$\alpha_1 = F_X(x_L)$$

Irrtumswahrscheinlichkeit, dass Werte unterhalb des geschätzten Bereichs liegen.

$$\alpha_2 = 1 - F_X(x_U)$$

Irrtumswahrscheinlichkeit, dass Werte oberhalb des geschätzten Bereichs liegen.

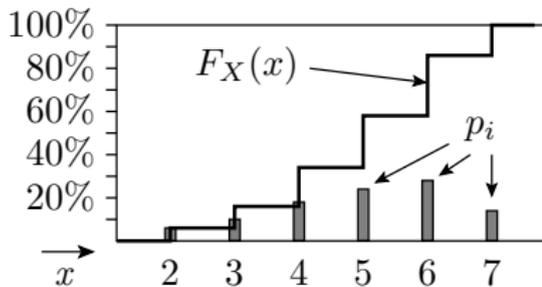
Bei  $\alpha_1 = 0$  gibt es nur eine Ober- und bei  $\alpha_2 = 0$  nur eine Untergrenze.

$\varepsilon_1, \varepsilon_2$	Intervallradius, Abstand der unteren bzw. oberen Bereichsgrenze vom Erwartungswert.
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$\mu, x_L, x_U$	Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.
$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ .
$F_X(x)$	Verteilungsfunktion der Zufallsvariablen $X$ .

## Diskrete Verteilung

Zufallsvariable  $X$  kann nur eine abzählbare Menge von diskreten Werten  $x_i$  annehmen, z.B.:

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



- $\mathbb{P}[X = x_i]$  Verteilung der diskreten Zufallsvariablen  $X$ .
- $F_X(x)$  Verteilungsfunktion der Zufallsvariablen  $X$ .
- $p_i$  Eintrittswahrscheinlichkeit der Realisierung  $x_i$ .



## Erwartungswert und Varianz

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

Erwartungswert, mit den Auftretswahrsch.  $p_i$  gewichteter Mittelwert:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i \quad (3.4)$$

Für das Beispiel:

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

Varianz, mittlere quadratische Abweichung vom Erwartungswert:

$$\text{Var}[X] = \sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \quad (3.6)$$

$$\text{Var}[X] = \sigma^2 = \sum_{i=1}^{\#x} p_i \cdot (x_i - \mathbb{E}[X])^2 \quad (3.10)$$

Für das Beispiel:

$$\sigma^2 = 6\% \cdot (2 - 5)^2 + 10\% \cdot (3 - 5)^2 + \dots + 14\% \cdot (7 - 5)^2 = 1,96$$

$\#x$  Anzahl der möglichen Realisierungen der Zufallsvariablen  $X$ .

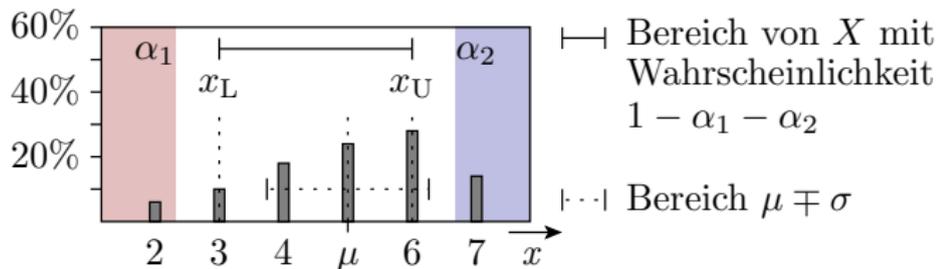
## Standardabweichung

Quadratwurzel aus der Varianz, Maß für die Abweichung vom Erwartungswert bzw. die Breite des wahrscheinlichen Bereichs:

$$\sigma = \sqrt{\sigma^2} \quad (3.7)$$

Für das Beispiel:

$$\sigma = \sqrt{1,96} = 1,4$$



Irrtumswahrscheinlichkeiten für  $X$  außerhalb  $[x_L, x_U]$ :

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

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



## Steinerscher Verschiebungssatz

Die Varianz ist gleichfalls die Differenz aus dem Erwartungswert der Quadrate und dem Quadrat des Erwartungswertes\*:

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \quad (3.11)$$

Herleitung:

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

Für das Beispiel zuvor:

$$\sigma^2 = 6\% \cdot 2^2 + 10\% \cdot 3^2 + 18\% \cdot 4^2 + 24\% \cdot 5^2 + 28\% \cdot 6^2 + 14\% \cdot 7^2 - 5^2 = 1,96$$

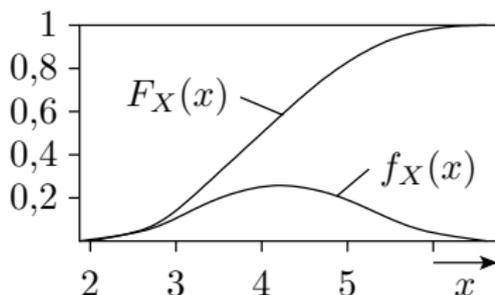
- 
- #x Anzahl der möglichen Realisierungen der Zufallsvariablen  $X$ .
  - $x_i$  Realisierung (möglichen Wert)  $i$  der Zufallsvariablen  $X$ .
  - $p_i$  Eintrittswahrscheinlichkeit der Realisierung  $x_i$ .
  - \* numerische Probleme, wenn  $\mathbb{E}[X^2]$  und  $\mathbb{E}[X]^2$  fast gleich sind.

## Stetige Verteilungen

Zufallsvariable  $X$  ist stetig und hat im Intervall  $a \leq X \leq b$  unendlich viele Ausprägungen. Beschreibung durch die Dichte:

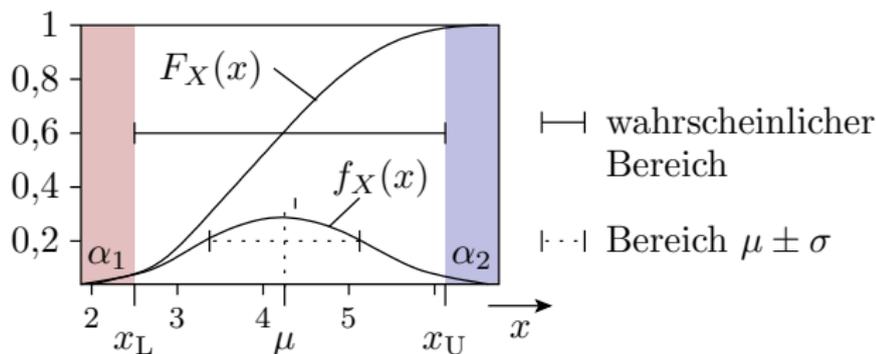
$$f_X(x) = \frac{dF_X(x)}{dx} \quad (3.3)$$

$$F_X(x) = \int_{-\infty}^x f_X(u) \cdot du \quad \text{für } X \in \mathfrak{R} \quad (3.12)$$



---

$F_X(x)$  Verteilungsfunktion der Zufallsvariablen  $X$ .  
 $f_X(x)$  Dichtefunktion der Zufallsvariablen  $X$ .



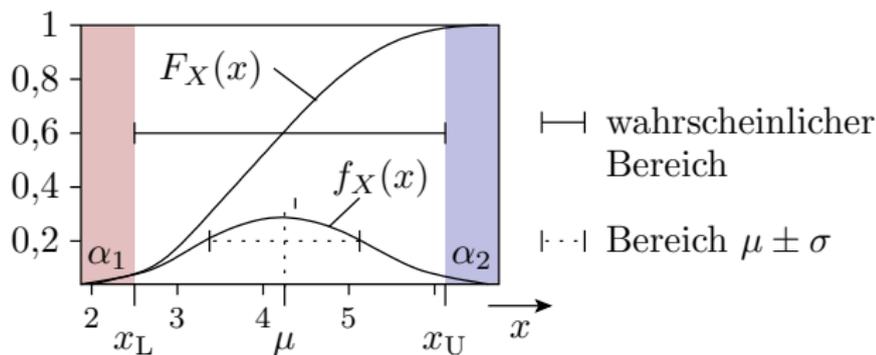
Erwartungswert:

$$\mathbb{E}[X] = \mu = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx \quad (3.5)$$

Varianz:

$$\text{Var}[X] = \sigma^2 = \int_{-\infty}^{\infty} f_X(x) \cdot (x - \mathbb{E}[X])^2 \cdot dx \quad (3.13)$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot x^2 \cdot dx - \mathbb{E}[X]^2 \quad (3.14)$$



Wahrscheinlichkeit, dass  $X$  in einem Bereich  $[x_L, x_U]$  liegt:

$$\begin{aligned} \mathbb{P}[x_L \leq x \leq x_U] &= F_X(x_U) - F_X(x_L) \\ &= \int_{x_L}^{x_U} f_X(x) \cdot dx \end{aligned}$$

Bereichsgrenzen:

$$\begin{aligned} x_L &= F_X^{-1}(\alpha_1) \\ x_U &= F_X^{-1}(1 - \alpha_2) \end{aligned}$$

$\alpha_1, \alpha_2$  Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.  
 $\mu, x_L, x_U$  Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.

## Erwartungswert und Varianz Datenstichprobe

Für eine Datenstichprobe einer Zufallsvariable  $X$

$$v = (v_1, v_2, \dots, v_{\#v})$$

ist der im weiteren verwendete Schätzer für den Erwartungswert:

$$\hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (3.15)$$

Der Schätzer für die Varianz ist die mittlere quadratische Abweichung vom geschätzten Erwartungswert:

$$\hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v - 1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (3.16)$$

$$\hat{\text{sd}}[X] = \hat{\sigma} = \sqrt{\hat{\sigma}^2} \quad (3.17)$$

Divisor um eins kleiner als  $\#v$ , d.h Abschätzungen von Varianz und Standardabweichung erfordern Stichprobengröße  $\#v \geq 2$ .

---

$\#v$	Größe der Datenstichprobe.
$v_i$	Wert $i$ der Datenstichprobe.
$\hat{\cdot}$	Schätzwert.



## Lineare Transformation



## Lineare Transformation

Lineare Transformationen sind Multiplikationen und Additionen einer Zufallsvariablen mit reellen Zahlen. Der Erwartungswert vervielfacht und erhöht sich um dieselben Werte:

$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \quad (3.18)$$

Bei der Varianz entfällt die Verschiebung und der Skalierungsfaktor geht im Quadrat ein:

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X] \quad (3.19)$$

Die Varianz ist insbesondere verschiebungsinvariant und bleibt bei einer Spiegelung der Verteilung gleich:

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

Beweise durch Nachrechnen folgen später als Übungsaufgaben.

---

$a, b$

Beliebige reelle Zahlen.



## Kontrolle am Beispiel

Realisierung $x$ von $X$	1	2	3
Realisierung $y$ von $Y = 5 - 2X$	3	1	-1
$\mathbb{P}[Y = y] = \mathbb{P}[X = x]$	0,3	0,5	0,2

$$\mathbb{E}[X] = 0,3 \cdot 1 + 0,5 \cdot 2 + 0,3 \cdot 3 = 1,9$$

$$\text{Var}[X] = 0,3 \cdot 1^2 + 0,5 \cdot 2^2 + 0,3 \cdot 3^2 - 1,9^2 = 0,49$$

$$\mathbb{E}[Y] = 0,3 \cdot 3 + 0,5 \cdot 1 + 0,2 \cdot (-1) = 1,2$$

$$\text{Var}[Y] = 0,3 \cdot 3^2 + 0,5 \cdot 1^2 + 0,2 \cdot (-1)^2 - 1,2^2 = 1,96$$

$$\mathbb{E}[Y] = 5 - 2 \cdot \mathbb{E}[X] = 5 - 2 \cdot 1,9 = 1,2\checkmark$$

$$\text{Var}[Y] = (-2)^2 \cdot \text{Var}[X] = 4 \cdot 0,49 = 4 \cdot 0,5 - 4 \cdot 0,1 = 1,96\checkmark$$

## Summe von Zufallsvariablen

Die Verteilung der Summe von Zufallsvariablen ordnet jedem der möglichen Werte der Summe die Wahrscheinlichkeit zu, dass die Summe diesen Wert hat (Faltung):

$x_i$	1	3	4
$\mathbb{P}[X = x_i]$	0,1	0,4	0,5

$y_i$	2	3	4
$\mathbb{P}[Y = y_i]$	0,3	0,6	0,1

$$\mathbb{P}[X + Y = 3] = \mathbb{P}[X = 1] \cdot \mathbb{P}[Y = 2]$$

$$\mathbb{P}[X + Y = 4] = \mathbb{P}[X = 1] \cdot \mathbb{P}[Y = 3]$$

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

$$\mathbb{P}[X + Y = 6] = \mathbb{P}[X = 3] \cdot \mathbb{P}[Y = 3] + \mathbb{P}[X = 4] \cdot \mathbb{P}[Y = 2]$$

$$\mathbb{P}[X + Y = 7] = \mathbb{P}[X = 3] \cdot \mathbb{P}[Y = 4] + \mathbb{P}[X = 4] \cdot \mathbb{P}[Y = 3]$$

$$\mathbb{P}[X + Y = 8] = \mathbb{P}[X = 4] \cdot \mathbb{P}[Y = 4]$$



Für die Summe von Zufallsvariablen ist der Erwartungswert gleich der Summe der Erwartungswerte:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad (3.20)$$

Die Varianz ist die Summe der Varianzen plus doppelte Kovarianz:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y] \quad (3.21)$$

mit der Kovarianz:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \quad (3.22)$$

Für unabhängige Zufallsvariablen ist die Kovarianz null und die Varianz die Summe der Varianzen der Summanden:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \quad (3.23)$$

Die Beweise für diskrete Zufallsvariablen durch Nachrechnen folgen als Übungsaufgaben.

---

$X, Y$	Zufallsvariablen.
$\mathbb{E}[\dots]$	Erwartungswert von ...
$\text{Var}[\dots]$	Varianz von ...
$\text{Cov}[X, Y]$	Kovarianz der beiden Zufallsvariablen.



## Produkte und Quotienten von Zufallsvariablen

Jede Zufallsvariable  $X$  ist die Summe von Erwartungswert und einer zufälligen Abweichung von Erwartungswert mit Erwartungswert null:

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

Produkt zweier unabhängiger Zufallsvariablen:

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

Der Erwartungswert ist das Produkt der Erwartungswerte:

$$\frac{\mathbb{E}[X \cdot Y]}{\mathbb{E}[X] \cdot \mathbb{E}[Y]} = 1 + \frac{\mathbb{E}[X_{\Delta}]}{\mathbb{E}[X]} + \frac{\mathbb{E}[Y_{\Delta}]}{\mathbb{E}[Y]} \quad \text{mit } \mathbb{E}[X_{\Delta}] = 0; \mathbb{E}[Y_{\Delta}] = 0$$

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (3.24)$$

Für unabhängige Zufallsvariablen addieren sich nach Gl. 3.23 die Varianzen:

$$\text{Var} \left[ \frac{X \cdot Y}{\mathbb{E}[X] \cdot \mathbb{E}[Y]} \right] = \underbrace{\text{Var}[1]}_0 + \text{Var} \left[ \frac{X_{\Delta}}{\mathbb{E}[X]} \right] + \text{Var} \left[ \frac{Y_{\Delta}}{\mathbb{E}[Y]} \right]$$

~~3.23~~

Vernachlässigbar für  $X_{\Delta} \ll \mathbb{E}[X]$  und  $Y_{\Delta} \ll \mathbb{E}[Y]$ .



Nach Gl. 3.19 Skalierungen quadrieren, Verschiebungen entfallen:

$$X = \mathbb{E}[X] + X_{\Delta} \quad \rightarrow \quad \text{Var} \left[ \frac{X_{\Delta}}{\mathbb{E}[X]} \right] = \text{Var} \left[ \frac{X}{\mathbb{E}[X]} - 1 \right] = \frac{\text{Var}[X]}{\mathbb{E}[X]^2}$$

$$\frac{\text{Var}[X \cdot Y]}{\mathbb{E}[X]^2 \cdot \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.25)$$

Quotient zweier Zufallsvariablen:

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

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

Erwartungswert:

$$\mathbb{E}[X/Y] = \mathbb{E}[X] / \mathbb{E}[Y] \quad (3.26)$$

Varianzen:

$$\frac{\text{Var}[X/Y]}{\mathbb{E}[X/Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.27)$$

Bei Multiplikation und Division addieren sich die quadrierten Varianzkoeffizienten, vorausgesetzt sie sind klein.



## Gemessener Wert und Messfehler

In der Messtechnik gilt für jeden gemessenen Wert:

$$X_M = X + X_F \quad (3.28)$$

Alle drei Größen haben einen Erwartungswert und eine Varianz. Mit dem Messwert und dem Messfehler als unabhängige Zufallsvariablen, gilt für diese:

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (3.29)$$

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (3.30)$$

- $\mathbb{E}[X_F]$  – Erwartungswert, systematischer Messfehler
- $\text{sd}[X_F] = \sqrt{\text{Var}[X_F]}$  – Standardabweichung, zufälliger Messfehler.

---

$X_M$	Zufallsvariable gemessener Wert.
$X$	Zufallsvariable zu messender Wert.
$X_F$	Zufallsvariable Messfehler.
$\mathbb{E}[\dots]$	Erwartungswert von ...
$\text{Var}[\dots]$	Varianz von ...



## Beispiel 3.1: Messfehler

Der gemessene Wert  $R_M$  einer Widerstands-Charge ist im Mittel  $1010 \Omega$  und hat eine Standardabweichung von  $11,18 \Omega$ . Die Messung habe einen systematischen Fehler  $R_F$  von  $12 \Omega$  und eine Standardabweichung von  $5 \Omega$ .

---

$$\mathbb{E}[R_M] = 1010 \Omega, \text{sd}[R_M] = 11,18 \Omega, \mathbb{E}[R_F] = 12 \Omega \text{ und } \text{sd}[R_F] = 5 \Omega.$$

*Welchen Erwartungswert und welche Standardabweichung hat der zu messende Wert?*

---

$R_M$	Gemessener Widerstandswert.
$R_F$	Messfehler des Widerstandswertes.
$\mathbb{E}[\dots]$	Erwartungswert von ...
$\text{Var}[\dots]$	Varianz von ...
$\text{sd}[\dots]$	Standardabweichung von ...



$\mathbb{E}[R_M] = 1010 \Omega$ ,  $\text{sd}[R_M] = 11,18 \Omega$ ,  $\mathbb{E}[R_F] = 12 \Omega$  und  $\text{sd}[R_F] = 5 \Omega$ .

*Welchen Erwartungswert und welche Standardabweichung hat der zu messende Wert?*

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (3.29)$$

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (3.30)$$

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (3.7)$$

Die Zufallsvariable in der Aufgabe ist  $R$  statt  $X$  und gesucht sind Erwartungswert und Standardabweichung des tatsächlichen Wertes:

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

$$\text{Var}[R] = \text{Var}[R_M] - \text{Var}[R_F] = (11,18 \Omega)^2 - (5 \Omega)^2 = 100 \Omega^2$$

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

Der (tatsächliche) Messwert hat eine kleinere Standardabweichung als der gemessene Wert.



## Verteilung von Zählwerten



## Verteilung von Zählwerten

Ein zufälliger Zählwert  $X$ , z.B. die Anzahl der korrekt ausgeführten oder fehlerhaft ausgeführten Service-Leistungen lässt sich als Summe

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

»potentieller Zählwerte«  $X_i$  mit der Bernoulli-Verteilung (Gl. 3.1):

$$\begin{aligned} \mathbb{P}[X_i = 0] &= 1 - p_i \\ \mathbb{P}[X_i = 1] &= p_i \end{aligned}$$

beschreiben.

Zählwert $X$	potentielle Zählwerte $X_i \in \{0, 1\}$
Fehlfunktionen (MF)	geleistete Service-Leistungen (DS)
Fehler	potentielle Fehler
nachweisbare Fehler	vorhandene Fehler
...	

$p_i$       Wahrscheinlichkeit, dass das Ergebnis von Zählversuch  $i$  eins ist.

$n$       Anzahl der Zählversuche, maximaler Zählwert.



## Erwartungswert und Varianz der Summanden

$$\begin{aligned}\mathbb{P}[X_i = 0] &= 1 - p_i \\ \mathbb{P}[X_i = 1] &= p_i\end{aligned}$$

Erwartungswerte der Einzelereignisse:

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

Varianzen nach dem Verschiebungssatz:

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

---

$X_i$       Potentielle Zählwerte, Zufallsvariablen mit Wertebereich  $X_i \in \{0, 1\}$ .  
 $p_i$       Wahrscheinlichkeit, dass das Ergebnis von Zählversuch  $i$  eins ist.



Der Erwartungswert der Summe ist die Summe der Erwartungswerte (Gl. 3.18):

$$\mathbb{E}[X] = \sum_{i=1}^n p_i \quad (3.31)$$

Für die Varianz wird oft unterstellt, dass die zu zählenden Ereignisse, wie das Auftreten unterschiedlicher Fehlfunktionen, nicht voneinander abhängen (Varianz der Summe gleich der Summe der Varianzen der Summanden, Kovarianz null, Gl. 3.20):

$$\text{Var}[X] = \sum_{i=1}^n p_i \cdot (1 - p_i) \quad (3.32)$$

$p_i$   
 $n$

Wahrscheinlichkeit, dass das Ergebnis von Zählversuch  $i$  eins ist.  
Anzahl der Zählversuche, maximaler Zählwert.



## Berechnung der Verteilung

Bei Addition der nächsten bernoulli-verteilten Zufallsvariable  $X_i$  erhöht sich der Zählwert mit Wahrscheinlichkeit  $p_i$  um eins und bleibt mit Wahrscheinlichkeit  $1 - p_i$  gleich:

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

Berechnung der  
Verteilung:

$$\mathbb{P}_1 [X = 0] = 1 - p_1$$

$$\mathbb{P}_1 [X = 1] = p_1$$

	$p_i$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\mathbf{P}_1[X = k]$	30%	70%	30%			
$\mathbf{P}_2[X = k]$	50%	35%	50%	15%		
$\mathbf{P}_3[X = k]$	40%	21%	44%	29%	6%	
$\mathbf{P}_4[X = k]$	10%	18,9%	41,7%	30,5%	8,3%	0,6%

Wiederhole für  $i = 2$  bis  $n$

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

$$\mathbb{P}_i [X = i] = \mathbb{P}_{i-1} [X = i - 1] \cdot p_i$$

Wiederhole für  $k = 1$  bis  $i - 1$

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

$$+ \mathbb{P}_{i-1} [X = k - 1] \cdot p_i$$

---


$$\mathbb{P}_i [X = k] \quad \text{Verteilung der Summe ersten } i \text{ Zählerwerte } X = \sum_{j=1}^i X_j.$$



## Erwartungswert und Varianz für das Beispiel

$p_i$	30%	50%	40%	10%	
$k$	0	1	2	3	4
$\mathbf{P}_4[X = k]$	18,9%	41,7%	30,5%	8,3%	0,6%

Erwartungswert der Summe aller  $n = 4$  Summanden:

$$\mathbb{E}[X] = 18,9\% \cdot 0 + 41,7\% \cdot 1 + 30,5\% \cdot 2 + 8,3\% \cdot 3 + 0,6\% \cdot 4 = 1,3$$

Als Summe aller  $p_i$  nach Gl. 3.31 ist die Berechnung kürzer:

$$\mathbb{E}[X] = 30\% + 50\% + 40\% + 10\% = 1,3$$

Die Varianz beträgt nach dem Verschiebungssatz Gl. 3.11:

$$18,9\% \cdot 0^2 + 41,7\% \cdot 1^2 + 30,5\% \cdot 2^2 + 8,3\% \cdot 3^2 + 0,6\% \cdot 4^2 - 1,3^2 = 0,79$$

Vereinfachte Berechnung nach Gl. 3.32:

$$\text{Var}[X] = 0,3 \cdot 0,7 + 0,5 \cdot 0,5 + 0,4 \cdot 0,6 + 0,1 \cdot 0,9 = 0,79$$



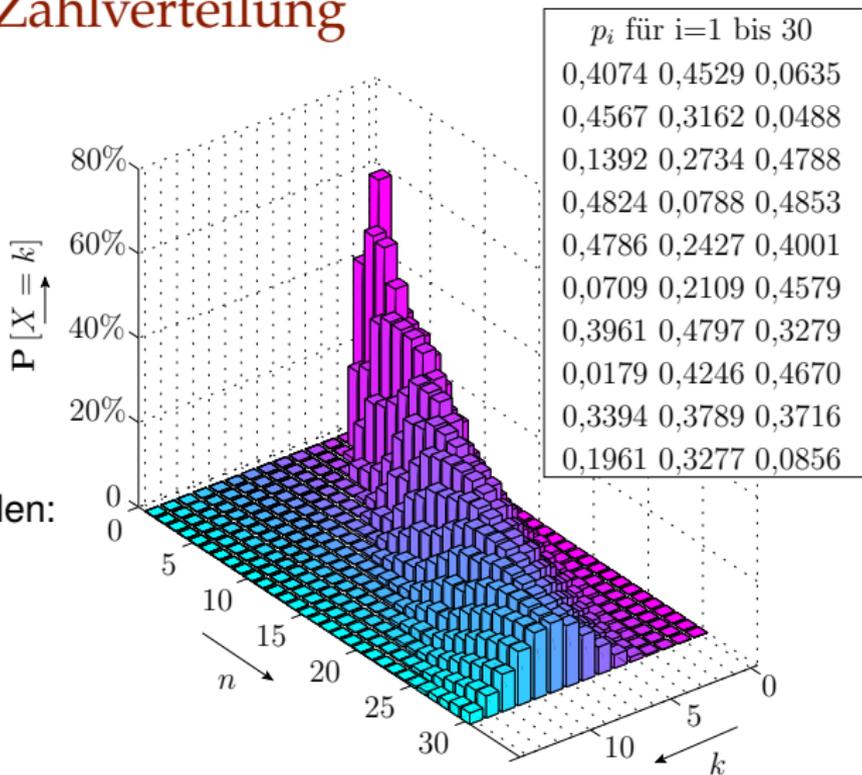
## Beispiel einer Zählverteilung

Mit Matlab schrittweise berechnete Zählverteilung. Die Eintrittswahrscheinlichkeiten der Zählereignisse siehe Kasten im Bild. Erwartungswert und Varianz für alle 30 Summanden:

$$\mathbb{E}[X] = 7,05$$

$$\text{Var}[X] = 2,19$$

Wahrscheinlicher Bereich ca. 5 bis 15.





# Zusammenfassung

## Zufallsvariable und Verteilung

Eine Zufallsvariable ist eine Veränderliche für das Ergebnis eines Zufallsexperiments.

Die Verteilung ordnet den möglichen Werten  $x_i$  einer diskreten Zufallsvariablen  $X$  Wahrscheinlichkeiten  $\mathbb{P}[X = x_i]$  zu.

Die Verteilungsfunktion  $F_X(x)$  gibt die Wahrscheinlichkeit an, dass eine Zufallsvariable  $X$  maximal den Wert  $x$  hat :

$$F_X(x) = \mathbb{P}[X \leq x] \quad (3.2)$$

Die Dichtefunktion beschreibt für eine stetige Zufallsvariable die Zunahme der Verteilungsfunktion mit  $x$ :

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (3.3)$$

Bernoulli-Verteilung:

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

## Erwartungswert, Varianz, Standardabweichung

Erwartungswert einer diskreten Verteilung:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i \quad (3.4)$$

Varianz einer diskreten Verteilung:

$$\text{Var}[X] = \sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \quad (3.6)$$

Standardabweichung:

$$\text{sd}[X] = \sigma = \sqrt{\text{Var}[X]} \quad (3.7)$$

Verschiebesatz:

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \quad (3.11)$$

Erwartungswert einer stetigen Verteilung:

$$\mathbb{E}[X] = \mu = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx \quad (3.5)$$

Varianz einer stetigen Verteilung:

$$\text{Var}[X] = \sigma^2 = \int_{-\infty}^{\infty} f_X(x) \cdot (x - \mathbb{E}[X])^2 \cdot dx \quad (3.13)$$

## Erwartungswert und Varianz einer Datenstichprobe

Schätzer für den Erwartungswert:

$$\hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (3.15)$$

Schätzer für die Varianz:

$$\hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (3.16)$$

## Lineare Transformationen und Summen

Erwartungswert einer linearen Transformation:

$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \quad (3.18)$$

Varianz einer linearen Transformation:

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X] \quad (3.19)$$

Erwartungswert der Summe zweier Zufallsvariablen:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad (3.20)$$

Varianz der Summe zweier Zufallsvariablen:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y] \quad (3.21)$$

Definition der Kovarianz:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \quad (3.22)$$

Varianz einer Summe unabhängiger Zufallsvariablen:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \quad (3.23)$$

## Für kleine Abweichungen vom Erwartungswert

Erwartungswert des Produkts zweier unabhängiger Zufallsvariablen:

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (3.24)$$

Varianz des Produkts zweier unabhängiger Zufallsvariablen:

$$\frac{\text{Var}[X \cdot Y]}{\mathbb{E}[X]^2 \cdot \mathbb{E}[Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.25)$$

Erwartungswert des Quotienten zweier unabhängiger Zufallsvariablen:

$$\mathbb{E}[X/Y] = \mathbb{E}[X] / \mathbb{E}[Y] \quad (3.26)$$

Varianz des Quotienten zweier unabhängiger Zufallsvariablen:

$$\frac{\text{Var}[X/Y]}{\mathbb{E}[X/Y]^2} = \frac{\text{Var}[X]}{\mathbb{E}[X]^2} + \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \quad (3.27)$$



## Messwerte

Erwartungswert gemessener Werte:

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (3.29)$$

Varianz gemessener Werte:

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (3.30)$$

## Verteilung von Zählwerten

Erwartungswert und Varianz:

$$\mathbb{E}[X] = \sum_{i=1}^n p_i \quad (3.31)$$

$$\text{Var}[X] = \sum_{i=1}^n p_i \cdot (1 - p_i) \quad (3.32)$$

Berechnung der Verteilung:

$$\mathbb{P}_1[X = 0] = 1 - p_1$$

$$\mathbb{P}_1[X = 1] = p_1$$

Wiederhole für  $i = 2$  bis  $n$

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

$$\mathbb{P}_i[X = i] = \mathbb{P}_{i-1}[X = i - 1] \cdot p_i$$

Wiederhole für  $k = 1$  bis  $i - 1$

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



# Näherungen



# Binomialverteilung



## Binomialverteilung

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

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

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

Binomialverteilung:

$$\mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (3.33)$$

Erwartungswert einer Binomialverteilung:

$$\mathbb{E}[X] = \mu = n \cdot p \quad (3.34)$$

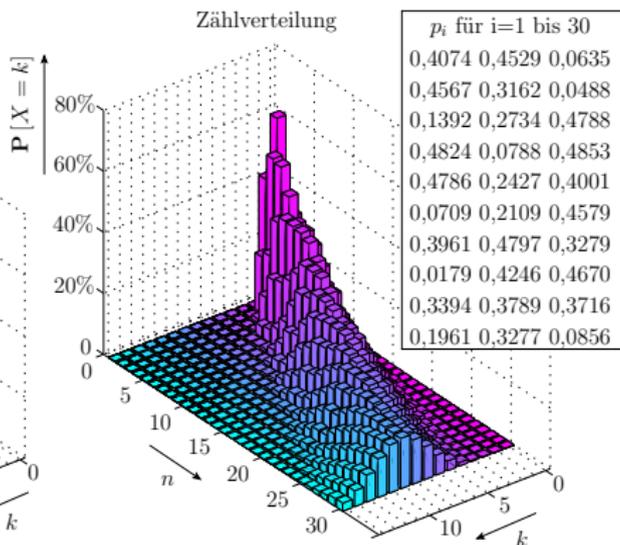
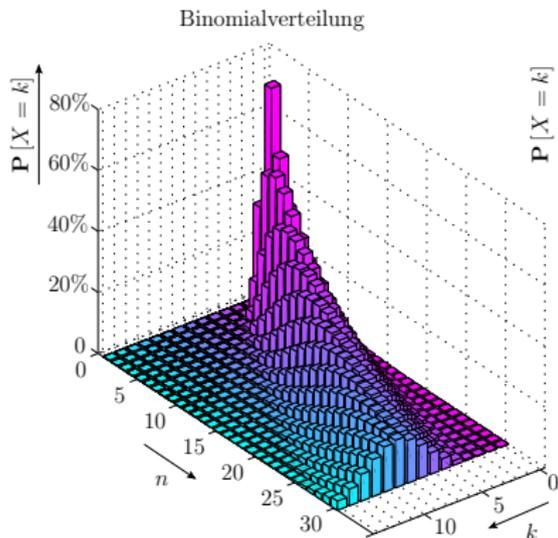
Varianz und Standardabweichung einer Binomialverteilung:

$$\text{Var}[X] = \sigma^2 = n \cdot p \cdot (1-p) \quad (3.35)$$

$$\text{sd}[X] = \sigma = \sqrt{n \cdot p \cdot (1-p)} \quad (3.36)$$

$n$  Anzahl der Zählversuche, maximaler Zählwert.  
 $p$  Eintrittswahrscheinlichkeit.

## Binomialverteilung vs. allgemeine Zählverteilung



Eine Binomialverteilung mit  $p = \frac{1}{n} \cdot \sum_{i=1}^n p_i$  nähert eine Zählverteilung gut an und berechnet sich aus nur zwei Parametern.

$n$  Anzahl der Zählversuche, maximaler Zählwert.

$p$  Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

$p_i$  Eintrittswahrscheinlichkeit Zählereignis  $i$ .

## Bereichsschätzung für eine Fehleranzahl

Die mittlere Nachweiswahrscheinlichkeit von 10 Fehlern sei 30%.

*Wie groß ist die Wahrscheinlichkeit, dass mindestens zwei von 10 Fehlern nachgewiesen werden?*

$$\mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.33)$$

Mindestens zwei Fehler:

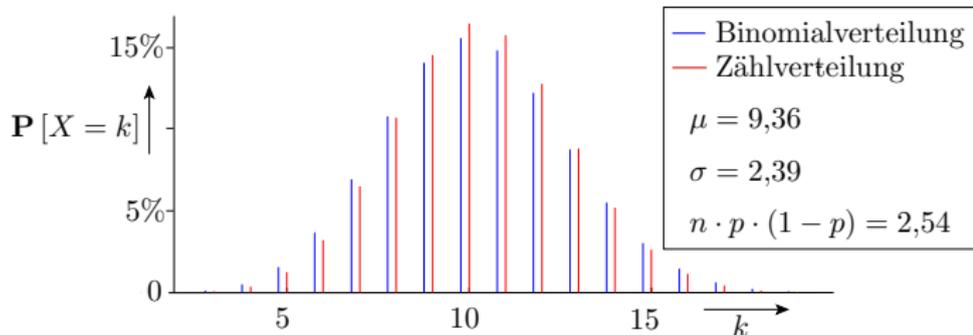
$$\begin{aligned} \mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X \leq 1] \\ &= 1 - \sum_{k=0}^1 \binom{10}{k} \cdot 0,3^k \cdot (1 - 0,3)^{10-k} \\ &= 1 - (0,7^{10} + 10 \cdot 0,3 \cdot 0,7^9) \\ &\approx 85\% \end{aligned}$$

## Varianzobergrenze für unabhängige Zählwerte

Bei gleicher Anzahl von unabhängigen Zählwerten  $n$  und  $p = \frac{1}{n} \cdot \sum_{i=1}^n p_i$  ist die Varianz der Binomialverteilung eine obere Schranke der Varianz einer Zählverteilung:

$$\sigma^2 = \sum_{i=1}^n p_i \cdot (1 - p_i) \leq n \cdot p \cdot (1 - p) \quad (3.37)$$

Damit sind bei Annäherung einer Zählverteilung durch eine Binomialverteilung die geschätzten Bereiche bzw. Irrtumswahrscheinlichkeiten pessimistische Abschätzungen. Für die beiden Verteilungen der Folie zuvor gilt für  $n = 30$ :



## Beweis Ungleichung 3.37 durch Nachrechnen

Ersatz der individuellen Auftrittswahrscheinlichkeiten der zu zählenden Ereignisse durch die mittlere Wahrscheinlichkeit und eine Differenz, die im Mittel null ist:

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

Varianz der Zählverteilung:

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^n (p + \delta_i) \cdot (1 - p - \delta_i) \\ &= \sum_{i=1}^n (p - p^2 - p \cdot \delta_i + \delta_i - p \cdot \delta_i - \delta_i^2) \\ &= \underbrace{\sum_{i=1}^n (p - p^2)}_{n \cdot p \cdot (1-p)} + \underbrace{\sum_{i=1}^n (\delta_i - 2 \cdot p \cdot \delta_i)}_{(1-2p) \cdot \sum_{i=1}^n \delta_i = 0} - \underbrace{\sum_{i=1}^n \delta_i^2}_{\geq 0} \\ \sigma^2 &\leq n \cdot p \cdot (1 - p) \text{ (Varianz Binomialverteilung)} \sqrt{\quad} \end{aligned}$$

## Schätzer für Zählwerte

$$\mu = n \cdot p \quad (3.34)$$

$$\sigma^2 \leq n \cdot p \cdot (1 - p) \quad (3.37)$$

Für praktische Abschätzungen wird die Ungleichheit in (Gl. 3.37) im Weiteren in einen Korrekturfaktor  $\kappa \leq 1$  versteckt, der bei Abhängigkeiten zwischen den Zählwerten deutlich größer eins sein kann (siehe Abschn. 3.2.8):

$$\sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \quad (3.38)$$

Schätzer für Erwartungswert, Standardabweichung und Eintrittswahrscheinlichkeit zu einem experimentell bestimmten Zählwert  $x_{AV}$ :

$$\hat{\mu} = x_{AV} \quad (3.39)$$

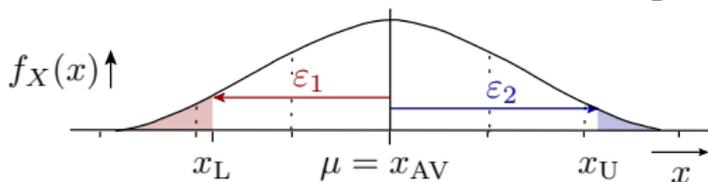
$$\hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{AV}}{n} \quad (3.40)$$

$$\hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)} \quad (3.41)$$

---

$\hat{\mu}, \hat{\sigma}$	Schätzwerte für Erwartungswert und Standardabweichung.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n$	Anzahl der Zählversuche, maximaler Zählwert.

Intervallradius wahr-  $2 \dots 3 \cdot \mu \cdot \sigma_{\bar{r}}$  für  $\mu \leq \frac{n}{2}$   
 scheinlicher Werte:  $2 \dots 3 \cdot (n - \mu) \cdot \sigma_{\bar{r}}$  für  $\mu > \frac{n}{2}$



Ein Maß für die Schätzgenauigkeit ist der Varianzkoeffizient des Eintritts- bzw. Nichteintrittszählwerts (Gl. 3.9). Mit Gl. 3.39 und Gl. 3.41 ergibt sich abschätzungsweise:

$$\begin{aligned} \sigma_{\bar{r}} &= \frac{\sigma}{\mu} = \sqrt{\kappa \cdot \left( \frac{1}{x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} && \text{für } \hat{p} \leq 50\% \\ \sigma_{\bar{r}} &= \frac{\sigma}{n - \mu} = \sqrt{\kappa \cdot \left( \frac{1}{n - x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} && \text{für } \hat{p} > 50\% \end{aligned} \quad (3.42)$$

Später wird gezeigt, dass größere Zählwerte oft normalverteilt und die Intervallradien 2 bis 3-mal so breit wie die Standardabweichung sind.

$\sigma_{\bar{r}}, \sigma_{\bar{r}}$	Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszählwert.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\left[ -\frac{1}{n} \right]^*$	Term für kleine und große $\hat{p}$ und grobe Überschläge vernachlässigbar.



# Poisson-Verteilung



### Poisson-Verteilung

Beim Zählen vieler seltener Ereignisse, z.B. der Fehlfunktionen bei Millionen von Service-Anforderungen, von denen nur wenige eintreten

$$n \rightarrow \infty$$

$$p_i \rightarrow 0$$

$$\text{Var} [X_i] = p_i \cdot (1 - p_i) \rightarrow p_i$$

Strebt die Varianz der zu zählenden Ereignisse und deren Summe gegen den Erwartungswert:

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

Die Verteilung der Summe strebt gegen die Poisson-Verteilung:

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

---

$n$  Anzahl der Zählversuche, maximaler Zählwert.

$p_i$  Wahrscheinlichkeit, dass das Ergebnis von Zählversuch  $i$  eins ist.

$\lambda$  Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).

$$\mathbb{P}[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad (3.43)$$

$$= e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!} \quad (3.44)$$

$$\mathbb{P}[X = k] = e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!}$$

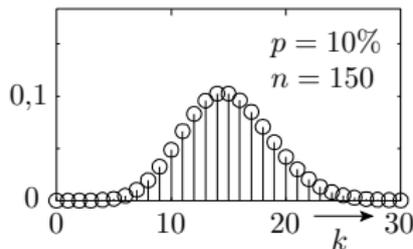
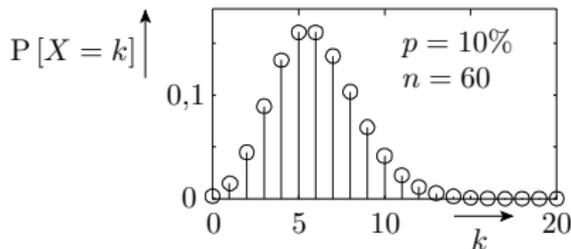
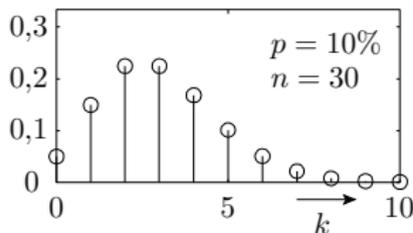
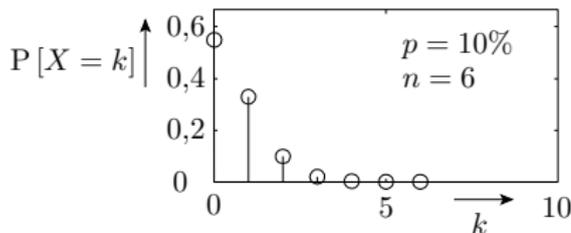
Die Poisson-Verteilung hat nur den einen Parameter  $\lambda$ , der gleichzeitig Erwartungswert und Varianz ist.

---

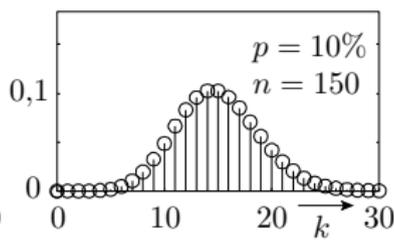
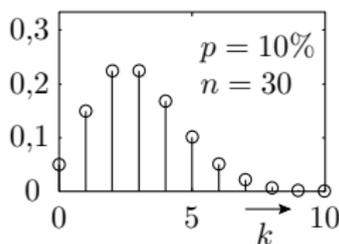
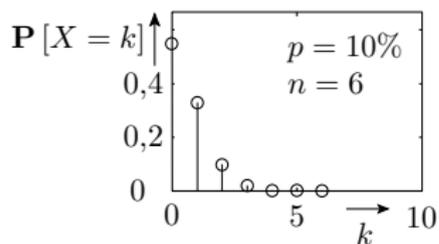
$\lambda$	Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$p$	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

## Anzahl der Zählversuche und Verteilung

$$\mathbb{P}[X = k] = e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!} \quad (3.44)$$



- $\mathbb{P}[X = k]$  Verteilung der diskreten Zufallsvariablen  $X$ .
- $n$  Anzahl der Zählversuche, maximaler Zählwert.
- $p$  Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.



Grobabschätzung der wahrscheinlichen Bereiche:

- Für  $\lambda = p \cdot n < 3$  keine untere Schranke  $k_L > 0$ . Ober Schranke:

$$k_U \approx 3 \dots 5 \cdot p \cdot n$$

- Für  $\lambda = p \cdot n \approx 3 \dots 10$  zusätzlich unter Schranke:

$$k_L \approx \frac{p \cdot n}{3 \dots 5}$$

- Für  $\lambda > 10$  Abschätzung über Normalverteilung (Gl. 3.61):

$$sr(k) \approx p \cdot n \cdot \left( 1 \mp 3 \dots 5 \cdot \sqrt{\frac{1}{n \cdot p}} \right)$$

$k_L, k_U$

Untere und obere Bereichsgrenze.

$sr(x)$

Symmetrischer Bereich der wahrscheinlichen Werte der Zufallsvariablen.

$n$

Anzahl der Zählversuche, maximaler Zählwert.

$p$

Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

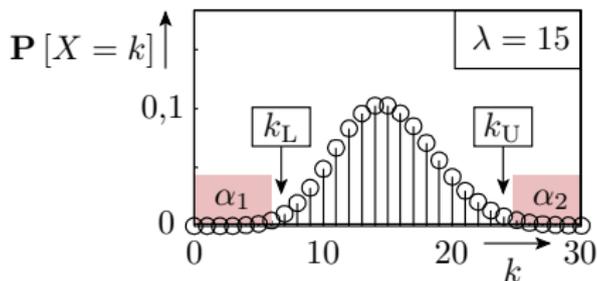


# Bereichschätz. Poissonvert.

Schätzen von  $k_L$ 

Vorgabe  $k_L$  und  $\alpha_1$ . Numerische  
Suche  $\lambda(k_L, \alpha_1)$ , so dass

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



$\alpha_1$	$k_L = 1$	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

Beispielabschätzungen:

- $\lambda = 7$  und  $\alpha_1 \leq 1\% \Rightarrow k_L = 2$
- $k_L = 1$  und  $\alpha_1 = 2\% \Rightarrow \lambda \geq 3,912$

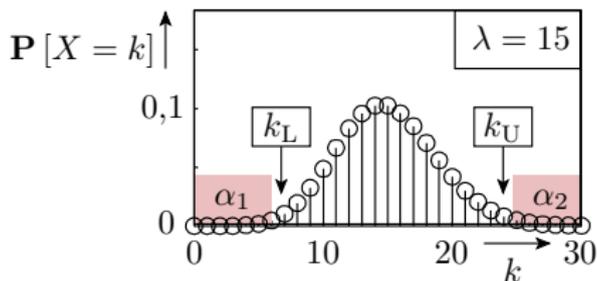
$\alpha_1, \alpha_2$   
 $k_L, k_U$

Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.  
Untere und obere Bereichsgrenze.

Schätzen von  $k_{U \max}$ 

Vorgabe  $k_U$  und  $\alpha_2$ . Numerische  
Suche  $\lambda(k_U, \alpha_2)$ , so dass

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



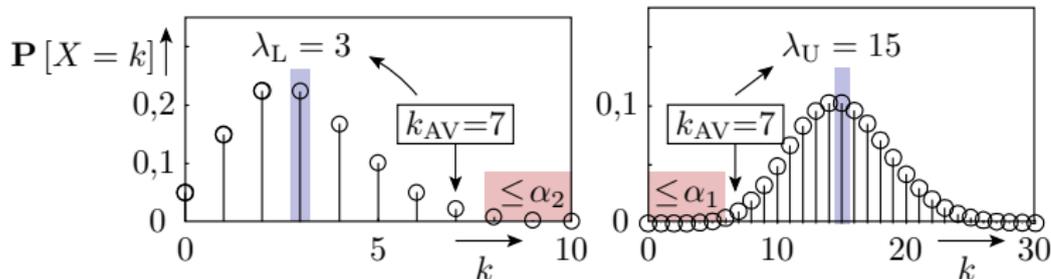
$\alpha_2$	$k_U = 0$	1	2	3	4	5	6
0,5%	0,005	0,103	0,338	0,672	1,078	1,537	2,037
1%	0,01	0,148	0,436	0,823	1,279	1,785	2,330
2%	0,02	0,215	0,567	1,016	1,529	2,089	2,684
10%	0,105	0,532	1,102	1,744	2,432	3,152	3,894
20%	0,223	0,824	1,534	2,296	3,089	3,903	4,733

Beispielabschätzungen:

- $\lambda = 2$  und  $\alpha_2 \leq 1\% \Rightarrow k_U = 6$
- $k_U = 3$  und  $\alpha_2 = 2\% \Rightarrow \lambda \leq 1,016$

$\alpha_2$  Irrtumswahrscheinlichkeit, Werte oberhalb des geschätzten Bereichs.  
 $k_U$  Obere Bereichsgrenze.

## Schätzen von $[\lambda_L, \lambda_U]$ aus einem Ist-Zählwert



Von den Folien zuvor übernommene Werte  $[\lambda_L, \lambda_U] = f(\alpha, k_{AV})$ :

$\alpha_1 = \alpha_2$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
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]

$k_{AV}$

Ist-Zählwert (Actual count value).

$\alpha_1, \alpha_2$

Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

$\lambda_L, \lambda_U$

Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.



$[\lambda_L, \lambda_U] = f(\alpha, k_{AV})$  für die Ist-Zählwerte  $k_{AV} = 4$  bis 6:

$\alpha_1 = \alpha_2$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
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]

$\lambda_U = f(\alpha_1, k_{AV})$  für den Ist-Zählwerte  $k_{AV} = 0$  ( $\lambda_L = 0$ ):

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

$$\lambda_U = -\ln(\alpha_1)$$

$\alpha_1$	0,5%	1%	2%	10%	20%
$\lambda_U$ für $k_{AV} = 0$	5,30	4,61	3,91	2,30	1,61%

$k_{AV}$  Ist-Zählwert (Actual count value).

$\alpha_1, \alpha_2$  Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

$\lambda_L, \lambda_U$  Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

## Beispiel 3.2: Poissonverteilte Anzahl Schadenfälle

- a) *In den vergangenen 10 Jahren ist kein Schaden eingetreten. Wie groß ist die zu erwartende Anzahl der Schadensfälle in den nächsten 10 Jahren mit einer Irrtumswahrscheinlichkeit  $\alpha_1 \leq 1\%$ ?*
- b) *In einem Nutzungsjahr sind 5 Schadensfälle eingetreten. Auf welchen Bereich der zu erwartenden Anzahl der Schadensfälle lässt sich aus dieser Angabe für die nächsten 10 Nutzungsjahre mit den Irrtumswahrscheinlichkeiten  $\alpha_1 = \alpha_2 = 1\%$  schließen?*

---

$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$k_{AV}$	Ist-Zählwert (Actual count value).
$\lambda_L, \lambda_U$	Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.



- a) In den vergangenen 10 Jahren ist kein Schaden eingetreten. Wie groß ist die zu erwartende Anzahl der Schadensfälle in den nächsten 10 Jahren mit einer Irrtumswahrscheinlichkeit  $\alpha_1 \leq 1\%$ ?

$\alpha_1$	0,5%	1%	2%	10%	20%
$\lambda_U$ für $k_{AV} = 0$	5,30	4,61	3,91	2,30	1,61%

Bereich der zu erwartenden Anzahl der Schadensfälle:  $\lambda_L = 0$  bis  $\lambda_U = 4,61$ . Die Anzahl der tatsächlichen Schadensfälle kann natürlich die geschätzte Obergrenze des Erwartungswert noch übersteigen.

---

$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$k_{AV}$	Ist-Zählwert (Actual count value).
$\lambda_L, \lambda_U$	Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.



- b) In einem Nutzungsjahr sind 5 Schadensfälle eingetreten. Auf welchen Bereich der zu erwartenden Anzahl der Schadensfälle lässt sich aus dieser Angabe für die nächsten 10 Nutzungsjahre mit den Irrtumswahrscheinlichkeiten  $\alpha_1 = \alpha_2 = 1\%$  schließen?

$\alpha_1 = \alpha_2$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AF} = 6$
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]

Mindestens  $10 \cdot \lambda_L = 17,9$  und maximal  $10 \cdot \lambda_U = 116$  Schadensfälle.

- $\alpha_1, \alpha_2$  Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.  
 $k_{AV}$  Ist-Zählwert (Actual count value).  
 $\lambda_L, \lambda_U$  Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

### Beispiel 3.3: Fehlfunktionsrate

Bei  $\#DS = 10^5$  Service-Leistungen wurden  $k_{AV} = 3$  Fehlfunktionen beobachtet.

*Auf welche Unter- und Obergrenze für die Fehlfunktionsrate lässt sich mit den Irrtumswahrscheinlichkeit  $\alpha_1 = \alpha_2 = 1\%$  unter Annahme einer Poissonverteilung schließen?*

---

$\#DS$	Anzahl der erbrachten Service-Leistungen (Number of delivered services).
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$k_{AV}$	Ist-Zählwert (Actual count value).
$\left[\frac{MF}{DS}\right]$	Zählwertverhältnis in Fehlfunktionen je erbrachte Service-Leistung.



Bei  $\#DS = 10^5$  Service-Leistungen wurden  $k_{AV} = 3$  Fehlfunktionen beobachtet.

*Auf welche Unter- und Obergrenze für die Fehlfunktionsrate lässt sich mit den Irrtumswahrscheinlichkeit  $\alpha_1 = \alpha_2 = 1\%$  unter Annahme einer Poissonverteilung schließen?*

$\alpha_1 = \alpha_2 = 1\%$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$[\lambda_L, \lambda_U]$	[0,15, 4,60]	[0,44, 6,64]	[0,82, 8,41]

Abschätzbarer Bereich der Fehlfunktionsrate:«

$$\zeta_L = \frac{\lambda_L}{\#DS} = 0,82 \cdot 10^{-5} \text{ [MF/DS]}$$

$$\zeta_U = \frac{\lambda_U}{\#DS} = 8,41 \cdot 10^{-5} \text{ [MF/DS]}$$

Kleine Zählwerte erlauben nur grobe Abschätzungen. Genauere Abschätzungen verlangen größere Zählwerte.

$\lambda_L, \lambda_U$

Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

$\zeta_L, \zeta_U$

Untere und obere Bereichsgrenze der geschätzten Fehlfunktionsrate.

## Beispiel 3.4: Maskierungswahrscheinlichkeit

Eine Überwachungseinheit hat von  $\#MF = 10.000$  Fehlfunktionen  $k_{AV} = 5$  Fehlfunktionen nicht erkannt.

*Welchen Bereich der Maskierungswahrscheinlichkeit ergibt sich aus diesem Versuchsergebnis für eine Irrtumswahrscheinlichkeiten*

$\alpha_1 = \alpha_2 = 10\%$ ?

---

$\#MF$	Anzahl der Fehlfunktionen (Number of malfunctions).
$k_{AV}$	Ist-Zählwert (Actual count value).
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.



Eine Überwachungseinheit hat von  $\#MF = 10.000$  Fehlfunktionen  $k_{AV} = 5$  Fehlfunktionen nicht erkannt.

*Welchen Bereich der Maskierungswahrscheinlichkeit ergibt sich aus diesem Versuchsergebnis für eine Irrtumswahrscheinlichkeiten*

$\alpha_1 = \alpha_2 = 10\%$ ?

$\alpha_1 = \alpha_2 = \frac{\alpha}{2} = 10\%$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$[\lambda_L, \lambda_U]$	$[2,43, 6,68]$	$[3,15, 7,99]$	$[3,89, 9,28]$

Abschätzbarer Bereich der Maskierungswahrscheinlichkeit:

$$p_{ML} = \frac{\lambda_L}{\#MF} = 3,15 \cdot 10^{-4}$$

$$p_{MU} = \frac{\lambda_U}{\#MF} = 7,99 \cdot 10^{-4}$$

$\lambda_L, \lambda_U$

Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

$p_{MK}$

Untere Bereichsgrenze der geschätzten Maskierungswahrscheinlichkeit.

$p_{MU}$

Obere Bereichsgrenze der geschätzten Maskierungswahrscheinlichkeit.

## Beispiel 3.5: Zuverlässigkeitsbereich

Beim Test eines Systems mit  $\#DS = 10^3$  Service-Leistungen wurden  $k_{AV} = 6$  Fehlfunktionen beobachtet.

*Auf welchen Bereich der Zuverlässigkeit kann nach diesem Versuchsergebnis mit den Irrtumswahrscheinlichkeiten  $\alpha_1 = \alpha_2 = 10\%$  geschlossen werden?*

---

$\#DS$	Anzahl der erbrachten Service-Leistungen (Number of delivered services).
$k_{AV}$	Ist-Zählwert (Actual count value).
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.



Beim Test eines Systems mit  $\#DS = 10^3$  Service-Leistungen wurden  $k_{AV} = 6$  Fehlfunktionen beobachtet.

*Auf welchen Bereich der Zuverlässigkeit kann nach diesem Versuchsergebnis mit den Irrtumswahrscheinlichkeiten  $\alpha_1 = \alpha_2 = 10\%$  geschlossen werden?*

$\alpha_1 = \alpha_2 = \frac{\alpha}{2} = 10\%$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$[\lambda_L, \lambda_U]$	$[2,43, 6,68]$	$[3,15, 7,99]$	$[3,89, 9,28]$

- Abschätzbarer Bereich der MF-Rate:

$$\zeta_L = 3,89 \cdot 10^{-3} \text{ [MF/DS]}$$

$$\zeta_U = 9,28 \cdot 10^{-3} \text{ [MF/DS]}$$

- Daraus folgender Bereich der Zuverlässigkeit:

$$R_L = \frac{1}{\zeta_L} = 108 \text{ [DS/MF]}$$

$$R_U = \frac{1}{\zeta_U} = 257 \text{ [DS/MF]}$$

$\zeta_L, \zeta_U$

Untere und obere Bereichsgrenze der geschätzten Fehlfunktionsrate.

$R_L, R_U$

Untere und obere Bereichsgrenze der geschätzten Zuverlässigkeit.



## Fehleranzahl und Defektanteil



## Zu erwartender Defektanteil und Fehleranzahl

Der zu erwartende Defektanteil ist die Wahrscheinlichkeit, dass ein Erzeugnis mindestens einen Fehler hat:

$$\mu_{DL} = 1 - \mathbb{P}[X = 0]$$

Bei konstanter Fehlerentstehungsrate ist die zu erwartende Fehleranzahl für gleiche Produkte aus demselben Entstehungsprozess gleich und die Fehleranzahl ist poissonverteilt

$$\mathbb{P}[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad (3.43)$$

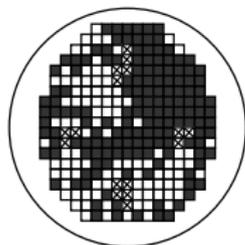
mit  $\lambda = \mu_F$ :

$$\mu_{DL} = 1 - e^{-\mu_F} \quad (3.45)$$

---

$\lambda$	Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).
$\mu_F$	Zu erwartende Fehleranzahl.
$\mu_{DL}$	Zu erwartender Defektanteil.

## Prozessgüteschwankungen und Fehlercluster



Fehlercluster auf einem Schaltkreiswafer [FSB87]

- funktionsfähige Schaltkreise
- erkannte fehlerhafte Schaltkreise
- Strukturen für Test und Diagnose

Örtlichen und zeitlichen Schwankungen von Parametern, die die Fehlerentstehung und Beseitigung beeinflussen:

- Fehlerentstehungsraten,
- Fehlerüberdeckungen, ...

verursachen Fehlercluster (lokale Fehlerhäufungen). Beispiele:

- Cluster von Schreibfehler in Texten,
- die qualitativ niederwertigen »Montagsprodukte«, ...

Fehlercluster liefern Hinweise auf Prozessverbesserungsmöglichkeiten.

## Schaltkreisausbeute und Fehleranzahl

Die zu erwartende Schaltkreisausbeute ist die Wahrscheinlichkeit, dass die Anzahl der nachweisbaren Fehler aus den Entstehungsprozessen null ist. Für eine poisson-verteilte Fehleranzahl nach Gl. 3.45 mit  $\mu_F = FC \cdot \mu_{CF}$ :

$$\mu_Y = e^{-FC \cdot \mu_{CF}} \quad (3.46)$$

Für jeden als fehlerfrei befundenen Schaltkreis müssen im Mittel

$$\frac{1}{\mu_Y} = e^{FC \cdot \mu_{CF}}$$

Schaltkreise gefertigt werden.

---

$\mu_Y$	Zu erwartende Ausbeute.
$FC$	Fehlerüberdeckung (fault coverage), Anteil der nachweisbaren Fehler.
$\mu_{CF}$	Zu erwartende Anzahl der Fehler aus den Entstehungsprozessen.



## Fertigungskosten und Fehleranzahl

Annahme, dass die zu erwartende Fehleranzahl und die reinen Fertigungskosten je Schaltkreis proportional zur Transistoranzahl zunehmen:

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

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

(vergl. Gl. 1.91). Kosten je nicht als defekt aussortierter Schaltkreis:

$$C_{IC} = \frac{C_{MIC}}{\mu_Y} = C_{Tr} \cdot \#Tr \cdot e^{FC \cdot \xi_{Tr} \cdot \#Tr} \quad (3.47)$$

---

$\mu_{CF}$	Zu erwartende Anzahl der Fehler aus den Entstehungsprozessen.
$\xi_{Tr}$	Defektentstehungsrate in Fehlern je Transistor.
$\#Tr$	Anzahl der Transistoren.
$C_{Tr}$	Transistorkosten in Euro pro Transistor.
$C_{MIC}$	Zu erwartende Kosten je gefertigter Schaltkreis.
$C_{IC}$	Kosten je als gut befundener (verkaufbarer) Schaltkreis.

### Beispiel 3.6: Schaltkreiskosten

Fehlerentstehungsrate  $\xi_{Tr} = 10^{-6}$  [Fehler je Transistor], Fertigungskosten  $C_{Tr} = 10^{-6}$  [Euro je Transistor], Schaltkreisgröße  $\#Tr \in \{10^5, 10^6, 10^7\}$  [Transistoren], Fehlerüberdeckung  $FC = 1$ .

*Welche Kosten entfallen auf jeden als gut befundene (verkaufbaren) Schaltkreis?*

---

$\xi_{Tr}$	Defektentstehungsrate in Fehlern je Transistor.
$C_{Tr}$	Transistorkosten in Euro pro Transistor.
$\#Tr$	Anzahl der Transistoren.
$FC$	Fehlerüberdeckung (fault coverage), Anteil der nachweisbaren Fehler.



Fehlerentstehungsrate  $\xi_{Tr} = 10^{-6}$  [Fehler je Transistor], Fertigungskosten  $C_{Tr} = 10^{-6}$  [Euro je Transistor], Schaltkreisgröße  $\#Tr \in \{10^5, 10^6, 10^7\}$  [Transistoren], Fehlerüberdeckung  $FC = 1$ .

*Welche Kosten entfallen auf jeden als gut befundene (verkaufbaren) Schaltkreis?*

$\#Tr$	$10^5$	$10^6$	$10^7$
$\mu_{CF} = \xi_{Tr} \cdot \#Tr$ in Fehlern	0,1	1	10
$C_{MIC} = C_{Tr} \cdot \#Tr$ in Euro	0,1	1	10
$\mu_Y = e^{-FC \cdot \mu_{CF}}$	90,5	36,8%	$4,54 \cdot 10^{-5}$
$C_{IC} = \frac{C_{MIC}}{Y}$ in Euro	0,11	2,72	$2,2 \cdot 10^5$

Ab  $\mu_{CF} \geq 2 \dots 3$  zu erwartenden Fehlern pro Schaltkreis benötigen Schaltkreise Reparaturmöglichkeiten:

- deaktivieren / Ersatz defekter Funktionsblöcke,
- Verkauf z.B. als Prozessoren mit weniger Cache, ...

$\mu_{CF}$  Zu erwartende Anzahl der Fehler je gefertigter Schaltkreis.

$C_{MIC}$  Zu erwartende Kosten je gefertigter Schaltkreis.

$\mu_Y$  Zu erwartende Ausbeute.

$C_{IC}$  Kosten je als gut befundener (verkaufbarer) Schaltkreis.



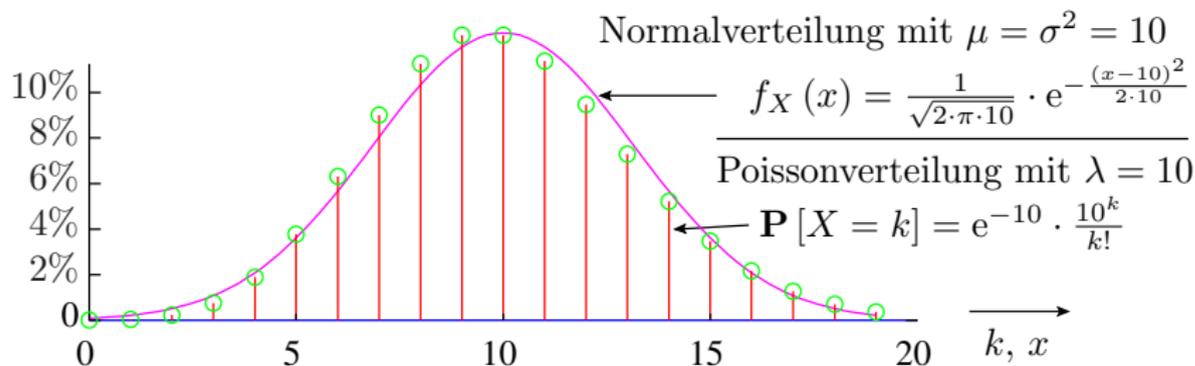
# Normalverteilung

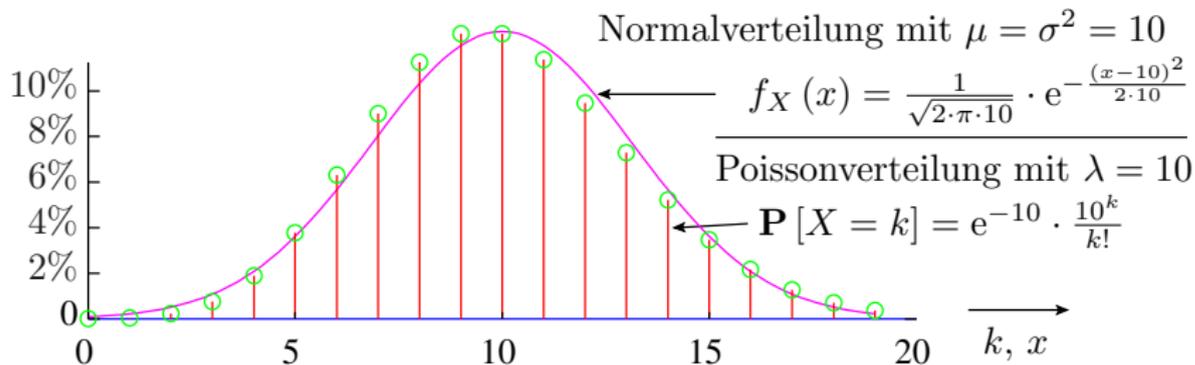
## Normalverteilung

Die Summe sehr vieler unabhängiger Zufallsvariablen strebt unter sehr allgemeinen Bedingungen (kein dominanter Summanden, ...) gegen eine Normalverteilung:

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \quad \text{mit } \sigma = \text{sd}[X], \mu = \mathbb{E}[X] \quad (3.48)$$

Vergleich Poisson- und Normalverteilung mit  $\mu = \sigma^2 = \lambda = 10$ :





Für Zählwerte genügt die Annäherung der Zähl- durch eine Normalverteilung in der Regel bereits unter der Bedingung

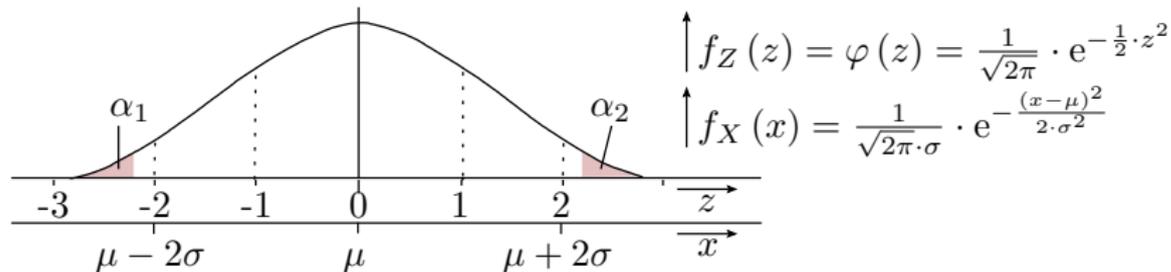
$$10 \cdot \kappa \leq \mu \leq n - 10 \cdot \kappa \quad \text{mit} \quad \mu = \sum_{i=1}^n p_i \quad (3.49)$$

- $n$  Anzahl der Zählversuche, maximaler Zählwert.
- $p_i$  Wahrscheinlichkeit, dass das Ergebnis von Zählversuch  $i$  eins ist.
- $\mu$  Erwartungswert.
- $\sigma$  Standardabweichung.



# Bereichsschätz. Normalvert.

## Bereichsschätzung mit Normalverteilung



- Transformation einer Zufallsvariablen  $X$  mit Erwartungswert  $\mu$  und Standardabweichung  $\sigma$  in eine Zufallsvariable  $Z$  mit Erwartungswert null und Standardabweichung eins:

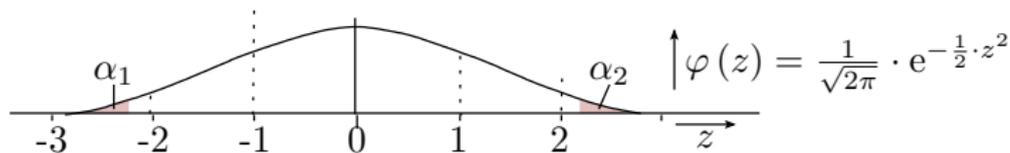
$$Z = \frac{X - \mu}{\sigma} \quad (3.50)$$

- Transformation der Werte  $x$  von  $X$  in  $z$  von  $Z$ :

$$z = \frac{x - \mu}{\sigma} \quad (3.51)$$

- Ablesen der Irrtumswahrscheinlichkeiten aus einer Tabelle der standardisierten Normalverteilungsfunktion  $F_Z(z) = \Phi(z)$ .

## Standardisierte Normalverteilung



Verteilungsfunktion tabelliert für  $z \geq 0$  in Schritten von 0,1:

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

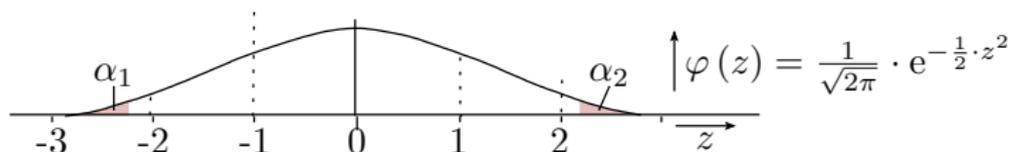
$z$	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0,5000	0,5398	0,5793	0,6179	0,6554	0,6915	0,7257	0,7580	0,7881	0,8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0,9772	0,9821	0,9861	0,9893	0,9918	0,9938	0,9953	0,9965	0,9974	0,9981
3,...	0,9987	0,9990	0,9993	0,9995	0,9997	0,9998	0,9998	0,9999	0,9999	1,0000

Wegen Symmetrie gilt für  $z < 0$ :

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



## Inverse standardisierte Normalverteilung



$z$	...,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 standardisierte Normalverteilung zur Bereichsschätzung:

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

$$\Phi^{-1}(\alpha_1) = -\Phi^{-1}(1 - \alpha_1)$$

$X$  Normalverteilte Zufallsvariable.

$Z$  Zufallsvariable mit Erwartungswert  $\mu = 0$  und Standardabweichung  $\sigma = 1$ .

Werte der Zufallsvariablen  $X$  und  $Z$ .

## Wahrscheinlichkeit der Bereichszugehörigkeit

$z$	...,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

Ausgehend von einer normalverteilten Zufallsvariablen  $X$  mit Erwartungswert  $\mu$  und Standardabweichung  $\sigma$  und einem Bereich mit den Grenzen  $x_L$  und/oder  $x_U$ :

- 1 Transformation der Bereichsgrenze nach (Gl. 3.51)

$$z_L = \frac{x_L - \mu}{\sigma}$$

$$z_U = \frac{x_U - \mu}{\sigma}$$

- 2 Ablesen von  $\Phi(z)$  bzw. für  $z < 0$  von  $\Phi(-z)$  aus der Tabelle.
- 3 Bestimmung der Irrtumswahrscheinlichkeiten:

$$\alpha_1 = \Phi(z_L) = 1 - \Phi(-z_L) = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right) \quad (3.52)$$

$$\alpha_2 = 1 - \Phi(z_U) = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right) \quad (3.53)$$



## Wahrscheinlicher Bereich

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

Ausgehend von einer normalverteilten Zufallsvariablen  $X$  mit Erwartungswert  $\mu$  und Standardabweichung  $\sigma$  und zulässigen Irrtumswahrscheinlichkeiten  $\alpha_1$  oder/und  $\alpha_2$ :

**1** Ablesen aus der Tabelle:

$$z_L = \Phi^{-1}(\alpha_1) = -\Phi^{-1}(1 - \alpha_1)$$

$$z_U = \Phi^{-1}(1 - \alpha_2)$$

**2** Transformation:

$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.54)$$

$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.55)$$

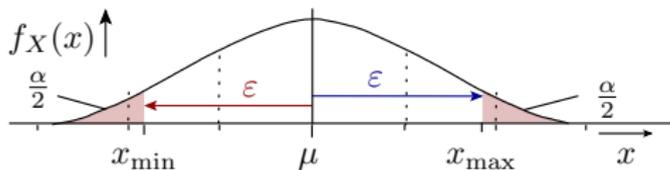
$x_L, x_U$  Untere und obere Schranke des wahrscheinlichen Bereichs von  $X$ .

$z_L, z_U$  Transformierte unteren und obere Schranke.

$\mu, \sigma$  Erwartungswert, Standardabweichung.

$\alpha_1, \alpha_2$  Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

## Symmetrischer Bereich



$\alpha$	4,54%	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

$$\epsilon = \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.56)$$

$$\text{sr}(x) = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.57)$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ mit den möglichen Zählwerten $x$ .
$\mu, x_L, x_U$	Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\epsilon, \mu, \sigma$	Intervallradius, Erwartungswert, Standardabweichung.
$\text{sr}(x)$	Symmetrischer Bereich der wahrscheinlichen Werte der Zufallsvariablen.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.



### Beispiel 3.7: Bereichsschätzung Normalverteilung

Zufallsvariable  $X$ ,  $\mu = 20$ ,  $\sigma = 5$ .

- Mit welcher Wahrscheinlichkeit ist  $X \geq 30$ ?*
- Mit welcher Wahrscheinlichkeit ist  $X \leq 15$ ?*
- Welche obere Schranke  $x_U$  wird nur mit einer Irrtumswahrscheinlichkeit  $\alpha_2 \leq 1\%$  überschritten?*
- Welche untere Schranke  $x_L$  wird nur mit einer Irrtumswahrscheinlichkeit  $\alpha_1 \leq 2\%$  unterschritten?*

---

$\mu, \sigma$	Erwartungswert, Standardabweichung.
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$x_L, x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von $X$ .



Zufallsvariable  $X$ ,  $\mu = 20$ ,  $\sigma = 5$ .

a) Mit welcher Wahrscheinlichkeit ist  $X \geq 30$ ?

$$\alpha_2 = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right) \quad (3.53)$$

$z$	...,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

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

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

$$\alpha_2 = 2,27\%$$



Zufallsvariable  $X$ ,  $\mu = 20$ ,  $\sigma = 5$ .

b) *Mit welcher Wahrscheinlichkeit ist  $X \leq 15$ ?*

$$\alpha_1 = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right) \quad (3.52)$$

$z$	...,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

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

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

$$\alpha_1 = 15,87\%$$



Zufallsvariable  $X$ ,  $\mu = 20$ ,  $\sigma = 5$ .

c) Welche obere Schranke  $x_U$  wird nur mit einer Irrtumswahrscheinlichkeit  $\alpha_2 \leq 1\%$  überschritten?

$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.55)$$

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

$$\begin{aligned} x_U &= 20 + 5 \cdot \Phi^{-1}(1 - 1\%) \\ &= 20 + 5 \cdot 2,33 \\ x_U &= 31,65 \end{aligned}$$

$\Phi^{-1}(\cdot)$  Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.



Zufallsvariable  $X$ ,  $\mu = 20$ ,  $\sigma = 5$ .

d) Welche untere Schranke  $x_L$  wird nur mit einer Irrtumswahrscheinlichkeit  $\alpha_1 \leq 2\%$  unterschritten?

$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.54)$$

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

$$x_L = 20 - 5 \cdot \Phi^{-1}(1 - 2\%)$$

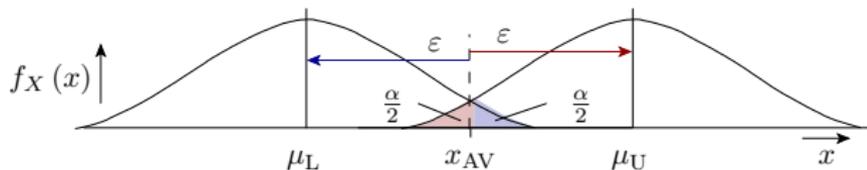
$$= 20 - 5 \cdot 2,05$$

$$x_L = 9,75$$

## Bereichsschätzung für den Erwartungswert

Der Erwartungswert  $\mu$  zu einem beobachteten Zählwert  $x_{AV}$  ist

- mindestens so groß, dass  $\mathbb{P}[X > x_{AV}] < \frac{\alpha}{2}$  und
- maximal so groß, dass  $\mathbb{P}[X < x_{AV}] < \frac{\alpha}{2}$ :



Intervallradius zwischen Erwartungswert und wahrscheinlichen Werten:

$$\varepsilon = \sigma \cdot \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \quad (3.56)$$

Symmetrischer Bereich mit  $x_{AV}$  als Schätzwert:

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

$\mu_L, \mu_U$	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
$\varepsilon$	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$sr(\mu)$	Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.

## Beispiel 3.8: Bereichsschätzung Erwartungswert

$$x_{AV} = 100, \sigma = 10.$$

*In welchem symmetrischen Bereich liegt der Erwartungswert mit Irrtumswahrscheinlichkeit  $\alpha = 2\%$ .*

---

$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$\sigma$	Standardabweichung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.



$$x_{AV} = 100, \sigma = 10.$$

In welchem symmetrischen Bereich liegt der Erwartungswert mit Irrtumswahrscheinlichkeit  $\alpha = 2\%$ .

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

$\alpha$	4,54%	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

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

$$sr(\mu) = 100 \mp 10 \cdot 2,33 = 100 \mp 23,3$$

- $sr(\mu)$       Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.  
 $\Phi^{-1}(\cdot)$     Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.



## Bereichsschätzung Zählwerte

## Schätzen von Zähl- und Erwartungswerten

Mit dem Erwartungswert und der Standardabweichung für Zählwerte:

$$\mu = n \cdot p \quad (3.34)$$

$$\sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \quad (3.38)$$

(Binomialverteilungsnäherung) und

$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.54)$$

$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.55)$$

$$\text{sr}(x) = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.57)$$

ergeben sich folgende Bereichsgrenzen:

$$x_L = p \cdot n - \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.59)$$

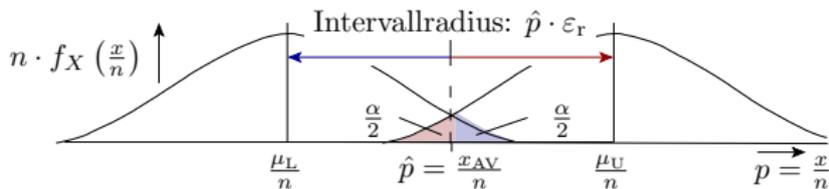
$$x_U = p \cdot n + \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.60)$$

$$\text{sr}(x) = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.61)$$

---

$p$	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\mu, x_L, x_U$	Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.

## Schätzen der Eintrittswahrscheinlichkeit



Schätzer der Eintrittswahrscheinlichkeit für  $n$  Zählversuche:

$$\hat{p} = \frac{x_{AV}}{n} \quad (3.40)$$

Varianzkoeffizient des Eintritts- bzw. Nichteintrittszählwerts:

$$\sigma_{\Gamma} = \frac{\sigma}{\mu} = \sqrt{\kappa \cdot \left( \frac{1}{x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} \quad \text{für } \hat{p} \leq 50\% \quad (3.42)$$

$$\sigma_{\bar{\Gamma}} = \frac{\sigma}{n-\mu} = \sqrt{\kappa \cdot \left( \frac{1}{n-x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} \quad \text{für } \hat{p} > 50\%$$

$f_X(x)$  Dichtefunktion der Zufallsvariablen  $X$ .

$\mu_L, \mu_U$  Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.

$n$  Anzahl der Zählversuche, maximaler Zählwert.

$x_{AV}$  Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.

$\sigma_{\Gamma}, \sigma_{\bar{\Gamma}}$  Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszählwert.

$\kappa$  Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte  $\kappa \leq 1$ .



Relative Intervallradien:

$$\begin{aligned}\varepsilon_r &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n}\right]^*\right)} && \text{für } \hat{p} \leq 50\% \\ \varepsilon_{\bar{r}} &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n}\right]^*\right)} && \text{für } \hat{p} > 50\%\end{aligned}\quad (3.62)$$

Symmetrischer Bereich:

$$sr(p) = \begin{cases} \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) & \text{für } \hat{p} \leq 50\% \\ 1 - \left(1 - \frac{x_{AV}}{n}\right) \cdot (1 \mp \varepsilon_{\bar{r}}) & \text{für } \hat{p} > 50\% \end{cases} \quad (3.63)$$

$\varepsilon_r, \varepsilon_{\bar{r}}$	Intervallradius relativ zum erwarteten Eintritts- bzw. Nichteintritts-Zählwert.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.
$sr(p)$	Geschätzter symmetrischer Bereich der Eintrittswahrscheinlichkeit.
$\left[-\frac{1}{n}\right]^*$	Term für kleine und große $\hat{p}$ und grobe Überschläge vernachlässigbar.

## Geeignete Zählwertgrößen (ACR)

$$\begin{aligned} \varepsilon_r &= \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left( \frac{1}{x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} && \text{für } \hat{p} \leq 50\% \\ \varepsilon_{\bar{r}} &= \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left( \frac{1}{n - x_{AV}} \left[ -\frac{1}{n} \right]^* \right)} && \text{für } \hat{p} > 50\% \end{aligned} \quad (3.62)$$

Geeignete Zählwertgrößen ergeben sich durch Auflösung nach der Anzahl der (nicht) eingetretenen Zählereignisse:

$$\begin{aligned} x_{AV} &\geq \frac{\kappa \cdot \left( \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \right)^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) && \text{für } \hat{p} \leq 50\% \\ n - x_{AV} &\geq \frac{\kappa \cdot \left( \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \right)^2}{\varepsilon_{\bar{r}}^2} \cdot \hat{p} && \text{für } \hat{p} > 50\% \end{aligned} \quad (3.64)$$

Die Terme  $(1 - \hat{p})$  bzw.  $\hat{p}$  sind mindestens 50% und max. eins.

$\varepsilon_r, \varepsilon_{\bar{r}}$	Intervallradius relativ zum erwarteten Eintritts- bzw. Nichteintritts-Zählwert.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.

## Beispiel 3.9: Geeignete Zählwertgröße

Erforderliche Zählwerte für relative Intervallradien  $\varepsilon_r$  bzw.  $\varepsilon_{\bar{r}}$  von 20% und 2% für zu schätzende Eintrittswahrscheinlichkeiten  $\hat{p} \in \{10\%, 50\%, 90\%\}$ . Irrtumswahrscheinlichkeit  $\alpha$  so groß, dass  $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2$ , d.h.  $\alpha = 4,52\%$ ,  $\kappa = 1$ .

---

$\varepsilon_r$	Intervallradius relativ zum erwarteten Eintritts-Zählwert.
$\varepsilon_{\bar{p}}$	Intervallradius relativ zum erwarteten Nichteintritts-Zählwerts.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .



Erforderliche Zählwerte für relative Intervallradien  $\varepsilon_r$  bzw.  $\varepsilon_{\tilde{r}}$  von 20% und 2% für zu schätzende Eintrittswahrscheinlichkeiten  $\hat{p} \in \{10\%, 50\%, 90\%\}$ . Irrtumswahrscheinlichkeit  $\alpha$  so groß, dass  $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2$ , d.h.  $\alpha = 4,52\%$ ,  $\kappa = 1$ .

$$\begin{aligned} x_{AV} &\geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) && \text{für } \hat{p} \leq 50\% \\ n - x_{AV} &\geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot \hat{p} && \text{für } \hat{p} > 50\% \end{aligned} \quad (3.64)$$

$\varepsilon_r$	$\hat{p} = 10\%$		$\hat{p} = 50\%$		$\varepsilon_{\tilde{r}}$	$p = 90\%$	
	$x_{AV.min}$	$n_{min}$	$x_{AV.min}$	$n_{min}$		$x_{AV.max}$	$n_{min}$
20%	90	900	50	100	20%	810	900
2%	9.000	90.000	5.000	50.000	2%	81.000	90.000

$x_{AV}$  Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.

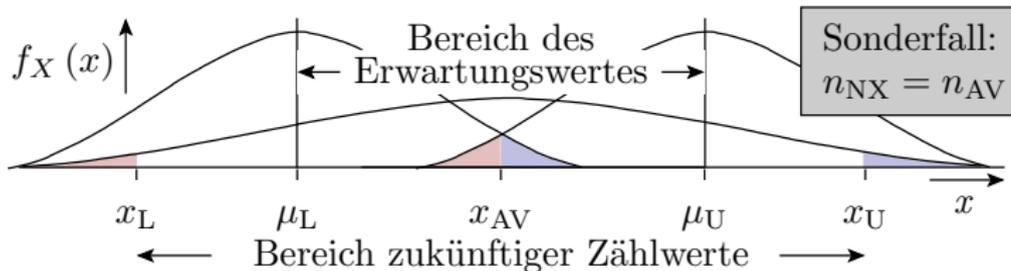
$n$  Anzahl der Zählversuche, maximaler Zählwert.

$n_{min}$  Mindestanzahl der Zählversuche.

$x_{AV.min}$  Minimal erforderliches Zählergebnis.

$x_{AV.max}$  Maximal zulässiges Zählergebnis.

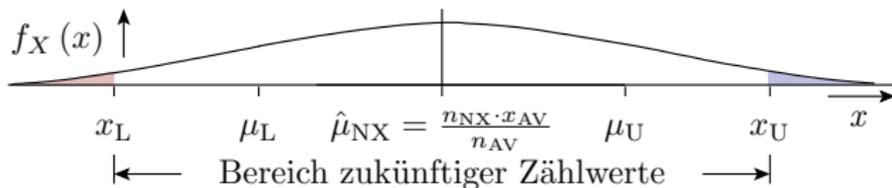
## Schätzung künftiger aus bekannten Zählwerten



Für  $n_{NX} \neq n_{AV}$  betragen die übereinstimmende Eintrittswahrscheinlichkeit und der Erwartungswert für künftige Zählwerte:

$$\hat{p} = \frac{x_{AV}}{n_{AV}} = \frac{\hat{\mu}_{NX}}{n_{NX}} \quad \Rightarrow \quad \hat{\mu}_{NX} = \frac{n_{NX} \cdot x_{AV}}{n_{AV}}$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ .
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$\mu_L, \mu_U$	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswertes.
$x_L, x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von $X$ .
$n_{AV}$	Anzahl der Zählversuche zur Schätzung der Eintrittswahrscheinlichkeit.
$n_{NX}$	Anzahl der Zählversuche zur Bestimmung eines künftigen Zählwerts.
$\hat{\mu}_{NX}$	Erwartungswert des zu schätzenden Zählwerts.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.



	Erwartungswert	Varianz
$X_{AV}$	$p \cdot n_{AV}$	$\kappa \cdot n_{AV} \cdot p \cdot (1 - p)$
$X_{NX}$	$p \cdot n_{NX}$	$\kappa \cdot n_{NX} \cdot p \cdot (1 - p)$

Die Abweichung der künftigen Werte von  $\mu_{NX}$ :

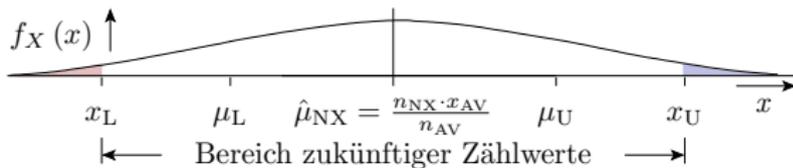
$$X_{\Delta} = X_{NX} - \frac{n_{NX} \cdot X_{AV}}{n_{AV}}$$

Varianz der Differenz nach (Gl. 3.19):

$$\begin{aligned} \text{Var}[X_{\Delta}] &= \text{Var}\left[X_{NX} - \frac{n_{NX} \cdot X_{AV}}{n_{AV}}\right] + \left(\frac{n_{NX}}{n_{AV}}\right)^2 \cdot \text{Var}[X_{AV}] \\ &= \kappa \cdot \left(n_{NX} + \left(\frac{n_{NX}}{n_{AV}}\right)^2 \cdot n_{AV}\right) \cdot p \cdot (1 - p) \\ &= \kappa \cdot n_{NX} \cdot \left(1 + \frac{n_{NX}}{n_{AV}}\right) \cdot p \cdot (1 - p) \end{aligned}$$

$X_{AV}$  Zufallsvariable für die  $x_{AV}$  mit  $n_{AV}$  Zählversuchen bestimmt wird.

$X_{NX}$  Zufallsvariable für die  $x_{NX}$  mit  $n_{NX}$  Zählversuchen bestimmt wird.



$\text{Var}[X_\Delta]$  ist um den Faktor  $\left(1 + \frac{n_{\text{NX}}}{n_{\text{AV}}}\right)$  größer, als die Varianz bei der Bereichsschätzung mit bekanntem Erwartungswert

$$\text{sr}(x) = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.61)$$

Symmetrischer Bereich für die wahrscheinlichen künftigen Werte:

$$\text{sr}(x_{\text{NX}}) = \hat{\mu}_{\text{NX}} \mp \sqrt{\kappa \cdot n_{\text{NX}} \cdot \left(\frac{n_{\text{NX}}}{n_{\text{AV}}} + 1\right) \cdot \hat{p} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.65)$$

mit  $\hat{p} = \frac{x_{\text{AV}}}{n_{\text{AV}}}$  und  $\hat{\mu}_{\text{NX}} = \frac{n_{\text{NX}} \cdot x_{\text{AV}}}{n_{\text{AV}}}$

$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ .
$x_{\text{AV}}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$\mu_L, \mu_U$	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
$x_L, x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von $X$ .
$n_{\text{AV}}$	Anzahl der Zählversuche zur Schätzung der Eintrittswahrscheinlichkeit.
$n_{\text{NX}}$	Anzahl der Zählversuche zur Bestimmung eines künftigen Zählwerts.
$\text{sr}(x_{\text{NX}})$	Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.
$\hat{\mu}_{\text{NX}}$	Erwartungswert des zu schätzenden Zählwerts.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.



### Beispiel 3.10: Bereich MF-Rate und künftige MF-Anzahl

Bei der Abarbeitung von 2.000 Service-Anforderungen wurden 100 Fehlerfunktionen beobachtet. Keine Abhängigkeiten. Zugelassene Irrtumswahrscheinlichkeit 2%.

$$n_{[AV]} = 20.000 \text{ [DS]}, x_{AV} = 100 \text{ [MF]}, \alpha = 2\%, \kappa = 1.$$

- Schätzwert, Intervallradius, symmetrischer Bereich der MF-Rate?
- Erforderliche Anzahl DS zur Verringerung des relativen Intervallradius auf  $\varepsilon_r \leq 10\%$ ?
- Symmetrischer Bereich der Anzahl der Fehlerfunktionen für  $n_{NX} = 10.000 \text{ [DS]}$  mit der geschätzten MF-Rate aus Aufgabenteil a)?

$n_{AV}$	Anzahl der Zählversuche, mit denen $x_{AV}$ bestimmt wurde.
[DS]	Zählwert in erbrachten Service-Leistungen.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
[MF]	Zählwert in Fehlerfunktionen.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$\varepsilon_r$	Relativer Intervallradius der Fehlerfunktionsrate.
$n_{NX}$	Anzahl der Zählversuche zur Bestimmung eines künftigen Zählwerts.



$$n_{[AV]} = 20.000 \text{ [DS]}, x_{AV} = 100 \text{ [MF]}, \alpha = 2\%, \kappa = 1.$$

a) *Schätzwert, Intervallradius, symmetrischer Bereich der MF-Rate?*

Die MF-Rate ist eine sehr kleine Eintrittswahrscheinlichkeit  $\hat{\zeta} \ll 1$ :

$$\varepsilon_r = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{x_{AV}}} \quad \text{für } \hat{p} \ll 50\% \quad (3.62)$$

$$\text{sr}(p) = \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) \quad \text{für } \hat{p} \leq 50\% \quad (3.63)$$

$\alpha$	4,54%	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

$$\hat{\zeta} = \frac{100 \text{ [MF]}}{20.000 \text{ [DS]}} = 0,5\% \left[ \frac{\text{MF}}{\text{DS}} \right]$$

$$\varepsilon_r = \Phi^{-1} (1 - 1\%) \cdot \sqrt{\frac{1}{100}} = 0,233$$

$$\text{sr}(\zeta) = \hat{\zeta} \cdot (1 \mp \varepsilon_r) = [0,38\%, 0,62\%] \left[ \frac{\text{MF}}{\text{DS}} \right]$$

$\Phi^{-1}(\cdot)$  Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.

$\text{sr}(p)$  Geschätzter symmetrischer Bereich der Eintrittswahrscheinlichkeit.

$\hat{\zeta}$  Schätzwert Fehlfunktionsrate.

$\text{sr}(\zeta)$  Symmetrischer Bereich der geschätzten Fehlfunktionsrate.



$$n_{[AV]} = 20.000 \text{ [DS]}, x_{AV} = 100 \text{ [MF]}, \alpha = 2\%, \kappa = 1.$$

b) *Erforderliche Anzahl DS zur Verringerung des relativen Intervallradius auf  $\varepsilon_r \leq 10\%$ ?*

$$x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \quad \text{für } \hat{p} \ll 50\% \quad (3.64)$$

Mindestanzahl der gezählten Fehlfunktionen:

$$x_{AV} \geq \frac{2,33^2}{0,1^2} = 543 \text{ [MF]}$$

Diese Zählergebnis wird erreicht nach etwa

$$n_{AV} = \frac{x_{AV}}{\zeta} = \frac{543 \text{ [MF]}}{0,5\% \left[ \frac{\text{[MF]}}{\text{[DS]}} \right]} = 108.000 \text{ [DS]}$$

[MF] Zählwert in Fehlfunktionen.

[DS] Zählwert in erbrachten Service-Leistungen.



$$n_{[AV]} = 20.000 \text{ [DS]}, x_{AV} = 100 \text{ [MF]}, \alpha = 2\%, \kappa = 1.$$

c) *Symmetrischer Bereich der Anzahl der Fehlfunktionen für  $n_{NX} = 10.000$  [DS] mit der geschätzten MF-Rate aus Aufgabenteil a?*

$$\text{sr}(x_{NX}) = \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1 - \hat{p}) \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \quad (3.65)$$

mit  $\hat{p} = \frac{x_{AV}}{n_{AV}}$  und  $\hat{\mu}_{NX} = \frac{n_{NX} \cdot x_{AV}}{n_{AV}}$

$$\hat{\mu}_{NX} = 0,5\% \cdot 10.000 = 50 \text{ [MF]}$$

$$\begin{aligned} \text{sr}(x_{NX}) &= 50 \mp \sqrt{\left(\frac{10^4}{2 \cdot 10^4} + 1\right) \cdot 50 \cdot (1 - 0,5\%) \cdot 2,33 \text{ [MF]}} \\ &= 50 \mp 20,1 \text{ [MF]} \end{aligned}$$

Relativer Intervallradius:

$$\varepsilon_r = \frac{20,1 \text{ [MF]}}{50 \text{ [MF]}} \approx 40\%$$

$\hat{\mu}_{NX}$  Erwartungswert des zu schätzenden Zählwerts.  
 $\text{sr}(x_{NX})$  Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.



# Varianzerhöhung



## Varianzerhöhung durch Abhängigkeiten

Abhängigkeiten erhöhen Varianz und Standardabweichung und den Intervallradius 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 Verschiebungssatz):

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



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

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

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

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

Die Standardabweichung und Intervallradius vergrößern sich um  $\sqrt{2}$ .

---

$\#X$	Anzahl der Zufallsvariablen.
$p_i$	Wahrscheinlichkeit, dass das Ergebnis von Zählversuch $i$ eins ist.
$p$	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.



## Varianzerhöhung

Die Varianzerhöhung ist das Verhältnis aus Varianz und Erwartungswert:

$$\kappa = \frac{\text{Var}[X]}{\mathbb{E}[X]} \quad (3.66)$$

Wenn

- die Eintrittswahrscheinlichkeit  $p_i$  alle Zählversuche gleich,
- die Zählwerte paarweise gleich und
- für alle Paare unabhängig sind,

ist die Varianzerhöhung, wie auf Folie zuvor gezeigt,  $\kappa = 2$ .

Analog lässt sich zeigen, dass unter sonst gleichen Bedingungen, wenn immer  $\#IDC$  Zählereignisse gleichzeitig eintreten, die Varianzerhöhung

$$\kappa = \#IDC$$

beträgt.

---

$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$\text{Var}[X]$	Varianz der Zufallsvariablen $X$ .
$\mathbb{E}[X]$	Erwartungswert der Zufallsvariablen $X$ .
$\#IDC$	Anzahl der identischen Zählwerte.

## Schätzen der Varianzerhöhung

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

$$\hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (3.15)$$

- Schätzen der Varianz der Zählwertstichprobe:

$$\hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (3.16)$$

- geschätzte Varianzerhöhung nach Gl. 3.66:

$$\hat{\kappa} = \frac{\hat{\text{Var}}[X]}{\hat{\mathbb{E}}[X]} = \frac{\hat{\sigma}^2}{\hat{\mu}} \quad (3.67)$$

---

$\hat{\mathbb{E}}[X]$	Geschätzter Erwartungswert der Zufallsvariablen $X$ .
$\#v$	Größe der Datenstichprobe.
$v_i$	Wert $i$ der Datenstichprobe.
$\hat{\text{Var}}[X]$	Geschätzte Varianz der Zufallsvariablen $X$ .
$\hat{\kappa}$	Geschätzte Varianzerhöhung durch Abhängigkeiten.



### Beispiel 3.11: Varianzerhöhung

# $v = 10$  Wiederholungen eines Zählversuchs:

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

*Abschätzung der Varianzerhöhung  $\hat{\kappa}$ ?*

---

- # $v$  Größe der Datenstichprobe.
- $v_i$  Wert  $i$  der Datenstichprobe.
- $\hat{\kappa}$  Geschätzte Varianzerhöhung durch Abhängigkeiten.



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

*Abschätzung der Varianzerhöhung  $\hat{\kappa}$ ?*

$$\hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (3.15)$$

$$\hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (3.16)$$

$$\hat{\kappa} = \frac{\hat{\sigma}^2}{\hat{\mu}} \quad (3.67)$$

$$\hat{\mu} = \frac{1}{10} \cdot \sum_{i=1}^{10} v_i = 63,1$$

$$\hat{\sigma}^2 = \frac{1}{9} \cdot \sum_{i=1}^{10} (v_i - 63,1)^2 = 135$$

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

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



### Beispiel 3.12: Anzahl Schadensfälle mit Varianzerhöhung

Der zu erwartende Zählwert für die Anzahl von Schadensfällen sei 100. Irrtumswahrscheinlichkeit 2%, Varianzerhöhung 2.

$$x_{AV} = 100 \text{ [D]}, n_{AV} \gg x_{AV}, \alpha = 2\%, \kappa = 2.$$

*In welchem symmetrischen Bereich wird bei künftigen Wiederholungen unter denselben Versuchsbedingungen ( $n_{NX} = n_{AV}$ ) die Anzahl der Schadensfälle liegen?*

---

$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n_{AV}$	Anzahl der Zählversuche, mit denen $x_{AV}$ bestimmt wurde.
$n_{NX}$	Anzahl der Zählversuche, mit denen $x_{NX}$ bestimmt werden soll.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
[D]	Zählwert in Schadensfällen (damages).



$x_{AV} = 100$  [D],  $n_{AV} \gg x_{AV}$ ,  $\alpha = 2\%$ ,  $\kappa = 2$ .

*In welchem symmetrischen Bereich wird bei künftigen Wiederholungen unter denselben Versuchsbedingungen ( $n_{NX} = n_{AV}$ ) die Anzahl der Schadensfälle liegen?*

$$sr(x_{NX}) = \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1 - \hat{p}) \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \quad (3.65)$$

mit  $\hat{p} = \frac{x_{AV}}{n_{AV}}$  und  $\hat{\mu}_{NX} = \frac{n_{NX} \cdot x_{AV}}{n_{AV}}$

$\alpha$	4,54%	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

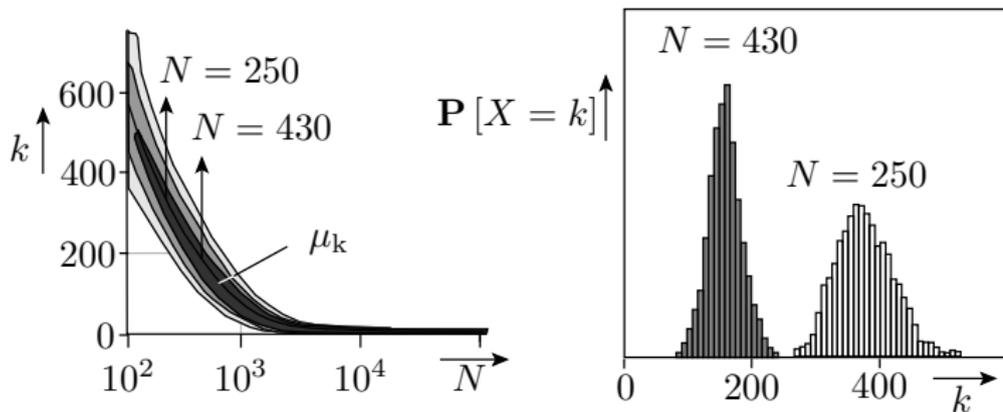
Für  $n_{NX} = n_{AV}$  und  $n_{AV} \gg x_{AV}$  ist  $\mu_{NX} = x_{AV}$ :

$$\begin{aligned} sr(x_{NX}) &= x_{AV} \mp \sqrt{\kappa \cdot 2 \cdot x_{AV} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \\ &= 100 \mp 20 \cdot 2,33 = [53,4, 146,6] \text{ [D]} \end{aligned}$$

$sr(x_{NX})$     Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.

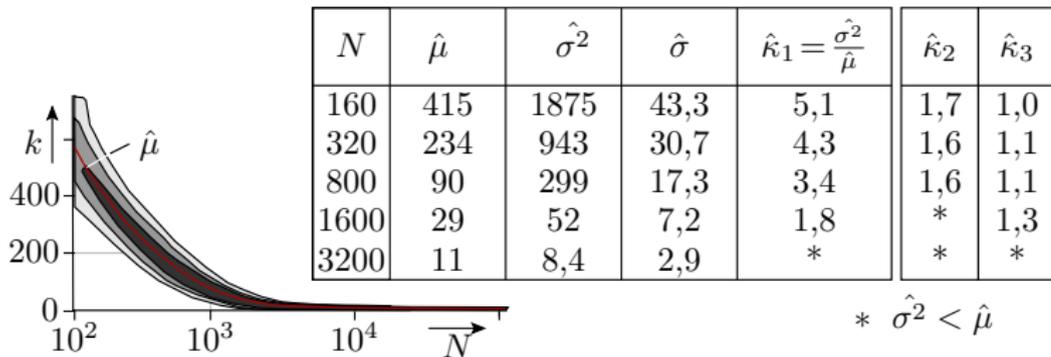
## Experiment mit Haftfehlern

Kombinatorische Beispielschaltung (Benchmark c3540).  $n_{[MF]} = 3606$  simulierte, unterschiedlich nachweisbare Haftfehler. Zählwert  $k$  ist die Anzahl der nicht nachweisbaren Haftfehler. Abschätzung der Verteilung  $\mathbb{P}[X = k]$  mit einer Stichprobe von  $\#v = 1000$  Zählwerten für verschiedene Zufallstestsätze der Länge  $N$ .



- $k$  Anzahl der nicht nachweisbaren Modellfehler.
- $n_{MF}$  Anzahl der Modellfehler.
- $N$  Anzahl der Tests.

## Varianzerhöhung im Experiment



$\hat{\kappa}_1$ : Fehlersimulation mit allen  $n_1 = 3606$  Haftfehlern. Abhängigkeiten bis, als ob 3...5 Modellfehler identisch nachweisbar wären. Erkennbare Identitäten waren aber beseitigt. Bleiben als Abhängigkeiten impliziter Nachweis und geteilte Steuer- und Beobachtungsbedingungen.

$\hat{\kappa}_2, \hat{\kappa}_3$ : Simulation mit Fehlerstichproben  $n_2 = 1.000$  bzw.  $n_3 = 300$ . Abnahme der Abhängigkeiten mit der Verkleinerung der Fehlerstichprobe.

$\hat{\kappa}_{1/2/3}$  Geschätzte Varianzerhöhung für 1 – alle,  $n_1 = 3606$ , 2 – eine Stichprobe von  $n_2 = 1000$  und 3 – eine Stichprobe von  $n_3 = 300$  Modellfehlern.

## Fehlermodellierung und Vorhersagbarkeit

Relativer Intervallradius für große Eintrittswahrscheinlichkeiten:

$$\varepsilon_{\bar{r}} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{n - x_{AV}}} \quad \text{für } \hat{p} \gg 50\% \quad (3.62)$$

Das Experiment zuvor hat gezeigt, dass für eine große Anzahl von Modellfehlern bezogen auf die Testobjektgröße die Varianzerhöhung  $\kappa$  mit der Modellfehleranzahl  $n$  zunimmt. Damit lässt sich der relative Intervallradius  $\varepsilon_{\bar{r},MF}$  als Maß der Schätzgenauigkeit nicht unbegrenzt durch mehr Modellfehler verringern.

---

$\varepsilon_{\bar{p}}$	Intervallradius relativ zum erwarteten Nichteintritts-Zählwerts.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$p$	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\varepsilon_{\bar{r},MF}$	Relativer Intervallradius der Anzahl der nichtnachweisbaren Modellfehler.



# Schlussfolgerungen

- Bei zufälliger Testauswahl hilft eine zu große Modellfehleranzahl im Verhältnis zur Testobjektgröße nicht die Schätzgenauigkeit für die Fehlerüberdeckung zu verbessern.
- Fehlermodelle, bei denen die Anzahl der Modellfehler überproportional mit der Testobjektgröße zunimmt, z.B. Kurzschlüsse und Pfadverzögerungsfehler (siehe Folie 5.30), sind nicht zielführend.

## Abschätzung mit mehreren Zufallstests

Alternative zur Verbesserung der Schätzgenauigkeit: Abschätzung der zu erwartenden Fehlerüberdeckung mit einer Stichprobe von  $n_{MF}$  Modellfehlern und einer Stichprobe von  $\#v$  Zufallstestsätzen wie im Experiment ab Folie 3.101. Vergrößerung der Anzahl der Zählversuche  $n$  auf  $n_{MF} \cdot \#v$  und damit die Anzahl der nicht nachweisbaren Fehler in

$$\varepsilon_{\bar{r}} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{n - x_{AV}}} \quad \text{für } \hat{p} \gg 50\% \quad (3.62)$$

auf etwa  $\#v \cdot n_{MF} \cdot (1 - FC_M)$ :

$$\varepsilon_{\bar{r}, MF} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{\#v \cdot n_{MF} \cdot (1 - FC_M)}} \quad (3.68)$$

$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.
$n_{MF}$	Anzahl der Modellfehler.
$\varepsilon_{\bar{p}}$	Intervallradius reaktiv zum erwarteten Nichteintritts-Zählwerts.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$n$	Anzahl der Zählversuche, maximaler Zählwert.
$\hat{FC}_M$	Schätzwert der Modellfehlerüberbedeckung.

### Beispiel 3.13: Bereich der Modellfehlerüberdeckung

Von 1000 Modellfehlern wurden 32 nicht erkannt. Varianzerhöhung maximal 2. Zulässige Irrtumswahrscheinlichkeit 2%.

$$n_{\text{MF}} = 1000 \text{ [F]}, x_{\text{AV}} = 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

- a) *In welchem Bereich liegt die Modellfehlerüberdeckung und wie groß ist der relative Intervallradius des Anteils der nicht nachweisbaren Fehler bei Abschätzung mit einem Zufallstestsatz ( $\#v = 1$ )?*
- b) *Der relative Intervallradius soll max. 10% betragen. Für wie viele unterschiedliche Zufallstestsätze ist die Anzahl der nicht nachweisbaren Fehler dafür zu mitteln?*

---

$n_{\text{MF}}$	Anzahl der Modellfehler.
[F]	Zählwert in Modellfehlern.
$x_{\text{AV}}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$\kappa$	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$ .
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.



$$n_{MF} = 1000 \text{ [F]}, x_{AV} = 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

a) *In welchem Bereich liegt die Modellfehlerüberdeckung und wie groß ist der relative Intervallradius des Anteils der nicht nachweisbaren Fehler bei Abschätzung mit einem Zufallstestsatz ( $\#v = 1$ )?*

$$\varepsilon_{\bar{r}.MF} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{\#v \cdot n_{MF} \cdot (1 - FC_M)}} \quad (3.68)$$

$$sr(p) = 1 - \left( 1 - \frac{x_{AV}}{n} \right) \cdot (1 \mp \varepsilon_{\bar{r}}) \quad \text{für } \hat{p} > 50\% \quad (3.63)$$

$\alpha$	4,54%	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

$$\hat{FC}_M = 1 - \frac{x_{AV}}{n_{MF}} = 1 - \frac{32}{1.000} = 96,8\%$$

$$\varepsilon_{\bar{r}.MF} \leq \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{x_{AV}}} = 2,33 \cdot \sqrt{\frac{2}{32}} = 58\%$$

$$sr(FC_M) = 1 - \left( 1 - \hat{FC}_M \right) \cdot (1 \mp 58\%) = [94,9\%, 98,7\%]$$

 $\hat{FC}_M$ 

Schätzwert der Modellfehlerüberdeckung.

 $\varepsilon_{\bar{r}.MF}$ 

Relativer Intervallradius der Anzahl der nichtnachweisbaren Modellfehler.

$$n_{MF} = 1000 \text{ [F]}, x_{AV} = 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

- b) *Der relative Intervallradius soll max. 10% betragen. Für wie viele unterschiedliche Zufallstestsätze ist die Anzahl der nicht nachweisbaren Fehler dafür zu mitteln?*

$$\varepsilon_{\bar{r},MF} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{\#v \cdot n_{MF} \cdot (1 - FC_M)}} \quad (3.68)$$

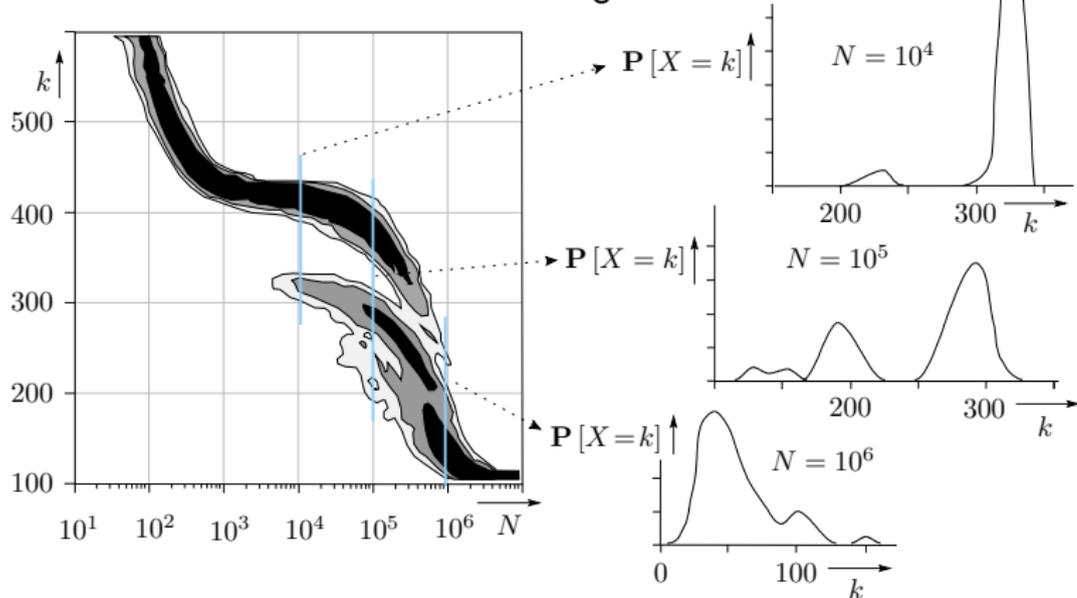
Umstellung nach  $\#v$ :

$$\#v \geq \frac{\kappa \cdot \left( \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \right)^2}{\left( 1 - \hat{F}C_M \right) \cdot n_{MF} \cdot \varepsilon_{\bar{r},MF}} = \frac{2 \cdot 2,33^2}{32 \cdot 0,1^2} = 34$$

Die Modellfehlerüberdeckung muss für mindestens 34 unterschiedliche Zufallstestsätze bestimmt und gemittelt werden.

## Nicht normalverteilte Zählwerte

Verteilung der Anzahl der nicht nachweisbaren Fehler für die kleineren ISCAS-Benchmark-Schaltung c2670:



Im Bereich der Testsatzlänge  $N = 10^4$  bis  $10^6$  mehrere Gipfel.

Wie ist das möglich?



# Zusammenfassung

## Binomialverteilung

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

$$\mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.33)$$

Erwartungswert:

$$\mathbb{E}[X] = \mu = n \cdot p \quad (3.34)$$

Varianz:

$$\text{Var}[X] = \sigma^2 = n \cdot p \cdot (1 - p) \quad (3.35)$$

Die Standardabweichung:

$$\text{sd}[X] = \sigma = \sqrt{n \cdot p \cdot (1 - p)} \quad (3.36)$$

Obergrenze für unabhängige Zählwerte:

$$\sigma^2 \leq n \cdot p \cdot (1 - p) \quad (3.37)$$

$$\sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \quad (3.38)$$

## Schätzer für Zählwerte

Erwartungswert, Standardabweichung, Eintrittswahrscheinlichkeit:

$$\hat{\mu} = x_{AV} \quad (3.39)$$

$$\hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)} \quad (3.41)$$

$$\hat{p} = \frac{x_{AV}}{n} \quad (3.40)$$

Relative Standardabweichung bezogen auf den zu erwartende (Nicht-) Eintrittszählwert als Maß der Schätzgenauigkeit:

$$\begin{aligned} \sigma_{\hat{\Gamma}} &= \frac{\sigma}{\mu} = \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n}\right]^*\right)} && \text{für } \hat{p} \leq 50\% \\ \sigma_{\hat{\bar{\Gamma}}} &= \frac{\sigma}{n-\mu} = \sqrt{\kappa \cdot \left(\frac{1}{n-x_{AV}} \left[-\frac{1}{n}\right]^*\right)} && \text{für } \hat{p} > 50\% \end{aligned} \quad (3.42)$$

## Poissonverteilung

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

$$\mathbb{P}[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad (3.43)$$

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

$$\mathbb{P}[X = k] = e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!} \quad (3.44)$$

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

## Fehleranzahl, Defektanteil, Schaltkreiskosten

Bei konstanter Fehlerentstehungsrate ist die Fehleranzahl gleicher Produkte aus demselben Entstehungsprozess poissonverteilt. Zu erwartender Fehleranteil:

$$\mu_{DL} = 1 - e^{-\mu_F} \quad (3.45)$$

Örtliche und zeitliche Schwankungen von Parametern, die die Fehlerentstehung und Beseitigung beeinflussen, verursachen Fehlercluster.

Fehlercluster liefern Hinweise auf Möglichkeiten zur Verbesserung der Entstehungsprozesse.

Zu erwartende Schaltkreisausbeute für poissonverteilte Fehleranzahl:

$$\mu_Y = e^{-FC \cdot \mu_{CF}} \quad (3.46)$$

Kosten je verkaufbarer (nicht als defekt aussortierter) Schaltkreis:

$$C_{IC} = \frac{C_{MIC}}{\mu_Y} = C_{Tr} \cdot \#Tr \cdot e^{FC \cdot \xi_{Tr} \cdot \#Tr} \quad (3.47)$$

## Bereichsschätzung Poissonverteilung

Geeignet für Zählwerte von vielen seltenen Ereignissen und bekanntem Erwartungswerte oder experimentell bestimmter Realisierungsanzahl im Bereich von null bis 10. Für praktische Schätzungen sind

- auf Folie 3.57 die Funktion  $\lambda(k_L, \alpha_1)$
- auf Folie 3.57 die Funktion  $\lambda(k_U, \alpha_2)$  und
- ab Folie 3.59  $[k_L, k_U]$  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

Die Normalverteilung ist eine stetige Verteilung mit der Dichtefunktion:

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \quad (3.48)$$

Zählwerte meist näherungsweise normalverteilt ab:

$$10 \cdot \kappa \leq \mu \leq n - 10 \cdot \kappa \quad \text{mit} \quad \mu = \sum_{i=1}^n p_i \quad (3.49)$$

Für Bereichsschätzungen Transformation normalverteilter Zufallsvariablen und Bereichsgrenzen in die der standardisierten Normalverteilung:

$$Z = \frac{X - \mu}{\sigma} \quad (3.50)$$

$$z = \frac{x - \mu}{\sigma} \quad (3.51)$$

Bestimmung der Irrtumswahrscheinlichkeiten aus der Tabelle der standardisierten Normalverteilung auf Folie 3.76:

- Wert kleiner als untere Bereichsgrenze:

$$\alpha_1 = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right) \quad (3.52)$$

- Wert größer als obere Bereichsgrenze:

$$\alpha_2 = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right) \quad (3.53)$$

Bereichsgrenzen für gegebene Irrtumswahrscheinlichkeiten:

$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.54)$$

$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.55)$$

Intervallradius und symmetrischer Bereich:

$$\varepsilon = \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.56)$$

$$\text{sr}(x) = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.57)$$

Symmetrischer Bereich des Erwartungswerts um einen experimentell bestimmten Zählwert:

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

## Bereichsschätzung für Zählwerte

Ober Schranke, untere Schranke und symmetrischer Bereich:

$$x_L = p \cdot n - \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (3.59)$$

$$x_U = p \cdot n + \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (3.60)$$

$$\text{sr}(x) = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.61)$$

Schätzen der Eintrittswahrscheinlichkeit:

$$\hat{p} = \frac{x_{AV}}{n} \quad (3.40)$$

$$\varepsilon_r = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n}\right]^*\right)} \quad \text{für } \hat{p} \leq 50\% \quad (3.62)$$

$$\varepsilon_{\bar{r}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n-x_{AV}} \left[-\frac{1}{n}\right]^*\right)} \quad \text{für } \hat{p} > 50\%$$

$$\text{sr}(p) = \begin{cases} \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) & \text{für } \hat{p} \leq 50\% \\ 1 - \left(1 - \frac{x_{AV}}{n}\right) \cdot (1 \mp \varepsilon_{\bar{r}}) & \text{für } \hat{p} > 50\% \end{cases} \quad (3.63)$$

Geeignete Zählwertgröße:

$$\begin{aligned}x_{AV} &\geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) && \text{für } \hat{p} \leq 50\% \\n - x_{AV} &\geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot \hat{p} && \text{für } \hat{p} > 50\%\end{aligned} \quad (3.64)$$

Intervallradius künftiger Zählwerte zu einem bekannten Zählwert:

$$\begin{aligned}\text{sr}(x_{NX}) &= \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1 - \hat{p}) \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \\ \text{mit } \hat{p} &= \frac{x_{AV}}{n_{AV}} \quad \text{und} \quad \hat{\mu}_{NX} = \frac{n_{NX} \cdot x_{AV}}{n_{AV}}\end{aligned} \quad (3.65)$$

## Varianzerhöhung

Varianzerhöhung und ihre Abschätzung:

$$\kappa = \frac{\text{Var}[X]}{\mathbb{E}[X]} \quad (3.66)$$

$$\hat{\kappa} = \frac{\hat{\text{Var}}[X]}{\hat{\mathbb{E}}[X]} \quad (3.67)$$

Zwischen Modellfehlern gibt es auch, wenn identisch nachweisbare Fehler zusammengefasst sind, erhebliche Varianzerhöhungen durch gemeinsame Anregungs- und Beobachtungsbedingungen.

Verbesserung der Schätzung durch Mittelung der Fehlerüberdeckung mehrerer Zufallstestsätze:

$$\varepsilon_{\bar{r},MF} = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \cdot \sqrt{\frac{\kappa}{\#v \cdot n_{MF} \cdot (1 - FC_M)}} \quad (3.68)$$

Können Abhängigkeiten dazu führen, dass Zählwerte wie auf Folie Folie 3.107 eine Verteilung mit mehreren Maxima haben?



# Mischverteilung



### Mischverteilung

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

$$\mathbb{P}[Y = j] = \frac{\#X_j}{\#X_{\text{ges}}} = h_j$$
$$F_X(x) = \mathbb{P}[X \leq x] = \sum_{j=1}^{\#Y} h_j \cdot F_{X_j}(x) \quad (3.69)$$

Für diskrete Verteilungen:

$$\mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i) \quad (3.70)$$

Für stetige Verteilungen:

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{i=1}^{\#Y} h_j \cdot f_{X_j}(x) \quad (3.71)$$



# 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:

$$\mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \quad (3.72)$$

Varianz mit  $\mu = \mu_j + \delta_j$ :

$$\begin{aligned} \text{Var}[X] &= \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{E}[(X_j - \mu_j + \delta_j)^2] \\ &= \sum_{j=1}^{\#Y} h_j \left( \underbrace{\mathbb{E}[(X_j - \mu_j)^2]}_{\sigma_j^2} - 2 \cdot \delta_j \cdot \underbrace{\mathbb{E}[(X_j - \mu_j)]}_0 + \underbrace{\mathbb{E}[\delta_j^2]}_{\delta_j^2} \right) \\ \text{Var}[X] = \sigma^2 &= \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \quad (3.73) \end{aligned}$$

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

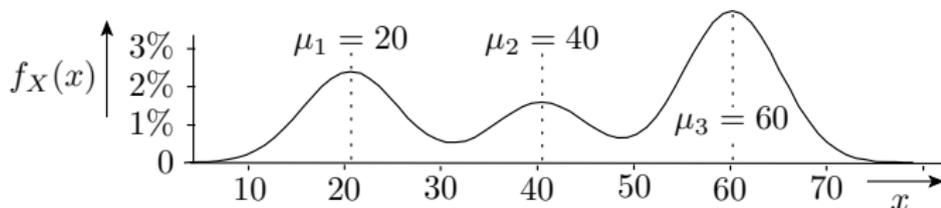
### Multimodale Verteilung

Beim Mischen von Grundgesamtheiten mit deutlich abweichenden Erwartungswerten entstehen multimodale Verteilungen. Beispiel: Mischen von 3 normalverteilten Zufallsvariablen  $X_i$  (vergl. Gl. 3.71):

$h_i$	0,3	0,2	0,5
$\mu_i$	20	40	60
$\sigma_i$	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.





$$\mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \quad (3.72)$$

$$\text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \quad (3.73)$$

$h_i$	0,3	0,2	0,5
$\mu_i$	20	40	60
$\sigma_i$	5	5	5

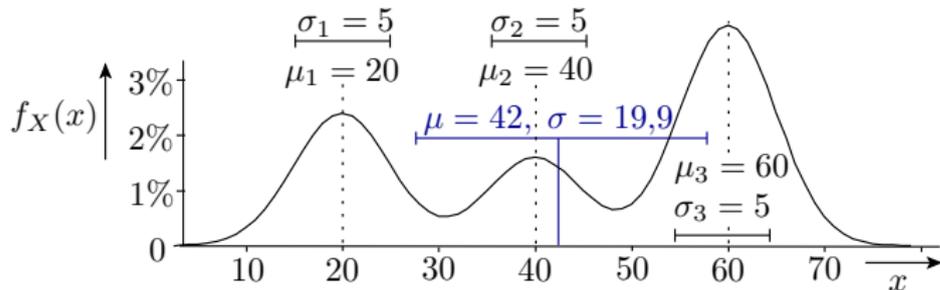
Erwartungswert:

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

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



### Beispiel 3.14: Identisch nachweisbare Fehler

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

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

- Mischverteilung als Zusammensetzung aus zueinander verschobenen Binomialverteilungen?*
- Erwartungswert und Varianz?*
- Standardabweichung und Varianzerhöhung?*

---

$n$  Anzahl der Zählversuche, maximaler Zählwert.  
 $p$  Eintrittswahrscheinlichkeit.



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

a) *Mischverteilung als Zusammensetzung aus zueinander verschobenen Binomialverteilungen?*

$$\mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (3.33)$$

$$\mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i) \quad (3.70)$$

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 um 10 Realisierung:

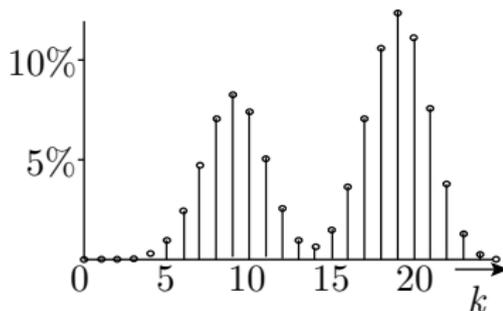
$$\mathbb{P}[X_1 = k + 10] = \begin{cases} \binom{n-10}{k} \cdot p^k \cdot (1-p)^{n-10-k} & 0 \leq k \leq n-10 \\ 0 & \text{sonst} \end{cases}$$



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

Mischverteilung:

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





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

b) Erwartungswert und Varianz?

$$\mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \quad (3.72)$$

$$\text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \quad (3.73)$$

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:

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

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



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

c) *Standardabweichung und Varianzerhöhung?*

$$\text{sd}[X] = \sqrt{\text{Var}[X]} \quad (3.7)$$

$$\kappa = \frac{\text{Var}[X]}{\mathbb{E}[X]} \quad (3.66)$$

Standardabweichung:

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

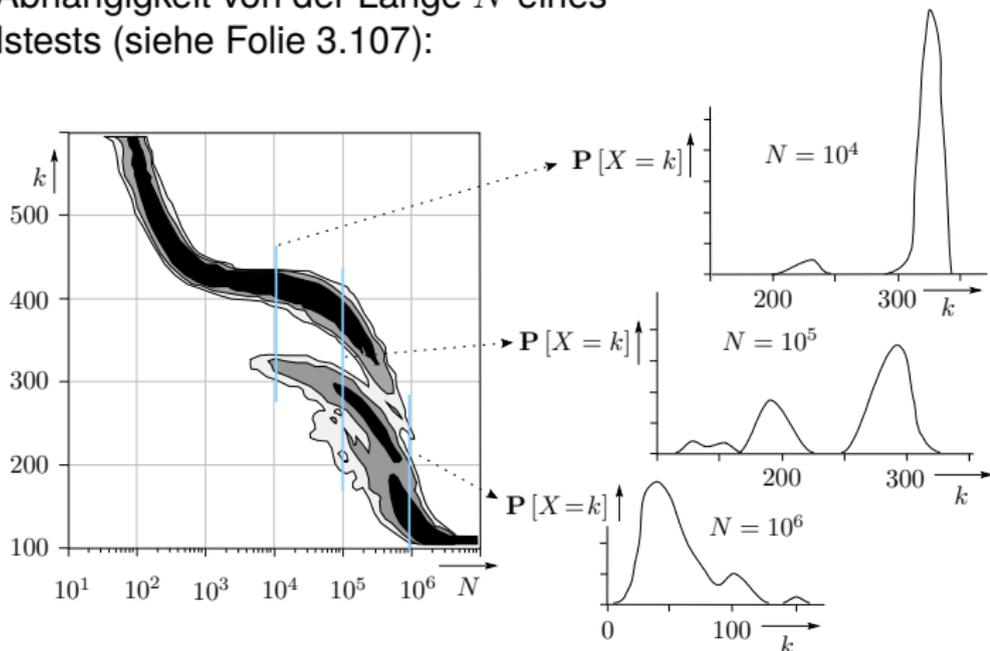
Varianzerhöhung:

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

Etwas kleinere Varianzerhöhung als bei einer Modellfehlermenge aus paarweise identisch nachweisbaren Fehlern.

## Dichte nicht nachweisbare Fehler c2670

... in Abhängigkeit von der Länge  $N$  eines  
Zufallstests (siehe Folie 3.107):

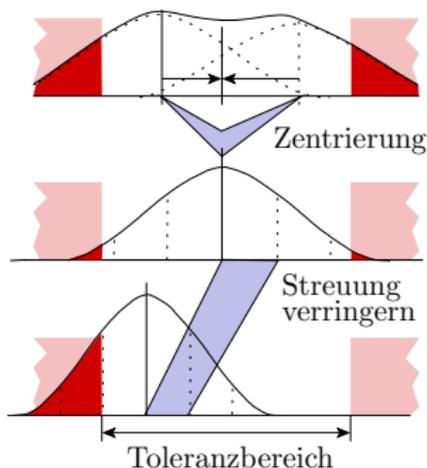


Im Bereich von  $N = 10^4$  bis  $10^6$  multimodale Verteilung. Offenbar ca. 80 sehr ähnlich nachweisbare Fehler mit MF-Rate  $\zeta_i \approx 10^{-5} \left[ \frac{MF}{DS} \right]$ .

## 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, ...

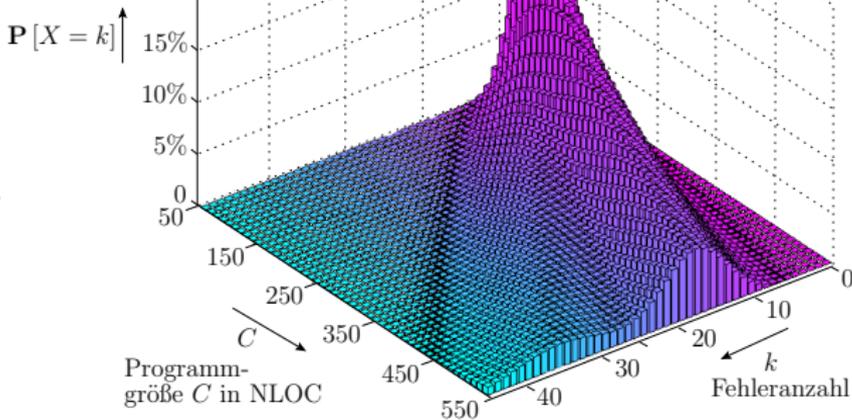


(siehe Folie 1.205 *Prozesszentrierung und -verbesserung*)

## Unterschiedlich gute Programmierer

Ein Anfänger und ein Profi entwickeln Software-Bausteine aus  $C$  Netto Lines of Code (NLOC), der Profi 66% mit ca. einem Fehler je 30 NLOC und der Anfänger 33% mit einem Fehler je 15 NLOC. Der Kunde weiß nicht, wer für ihn programmiert. Verteilung der Fehleranzahl:

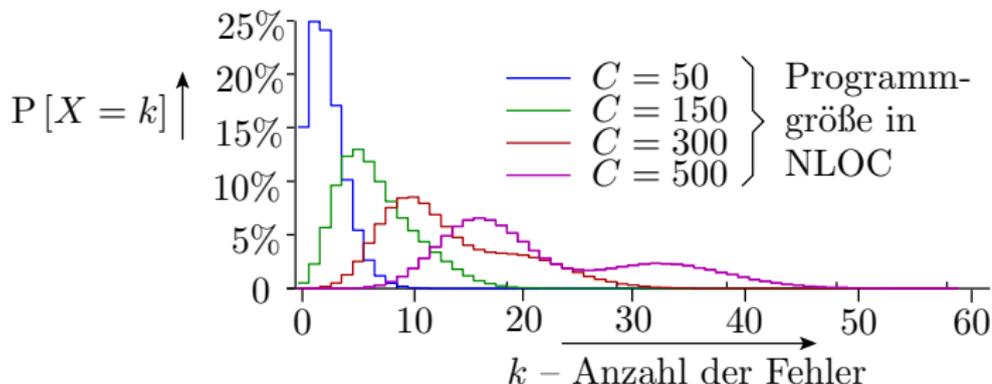
$$\begin{aligned}
 \mathbb{P}[X = k] &= \underbrace{\frac{2}{3} \cdot e^{-\frac{C}{30}} \cdot \frac{\left(\frac{C}{30}\right)^k}{k!}}_{\text{Pois}\left(\lambda = \frac{C}{30}\right)} \\
 &+ \underbrace{\frac{1}{3} \cdot e^{-\frac{C}{15}} \cdot \frac{\left(\frac{C}{15}\right)^k}{k!}}_{\text{Pois}\left(\lambda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





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. 3.70):

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



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

C

Metrik für den Entstehungsaufwand, hier in NLOC (netto lines of code).

NLOC

Netto Lines of Code, Anzahl der Code-Zeilen ohne Kommentar und Leerzeilen.



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, ...

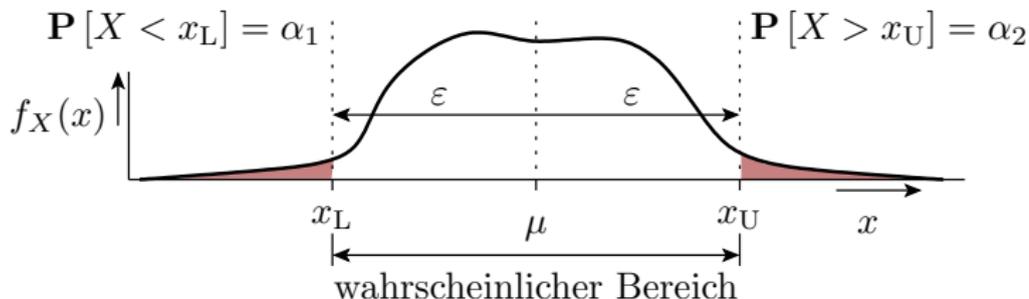
Wenn die Zufallsvariable ein Gütemaß ist, hat man es offenbar mit einer zufälligen Mischung von besser und schlechter funktionierenden Prozessabläufen zu tun. Dann ist es natürlich interessant, warum der Entstehungsprozess mal besser und mal schlechter funktioniert, um das schlechter Funktionierende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

Spezielle Form des Lernen aus Fehlern.



Bereichs. Tscheb.

## Bereichsschätzung, wenn Verteilung unbekannt

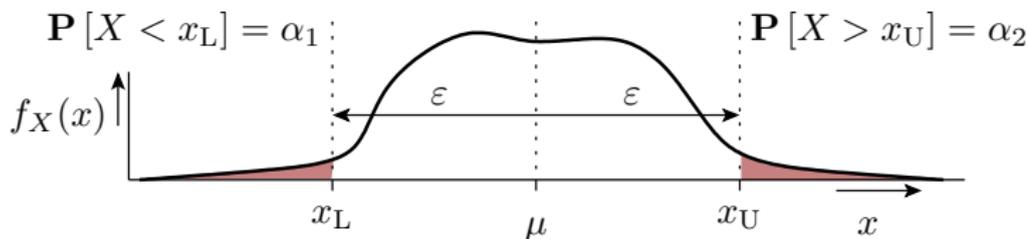


Die Bestimmung eines wahrscheinlichen Intervalls  $[x_{\min}, x_{\max}]$

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

$x_L, x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von $X$ .
$\alpha_1, \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$\mathbb{P}[\dots]$	Eintrittswahrscheinlichkeit des Ereignisses ...
$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ .
$\mu$	Erwartungswert.
$\varepsilon$	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.

## Das schwache Gesetz der großen Zahlen



Nach der tschebyscheffschen Ungleichung ist die Irrtumswahrscheinlichkeit  $\alpha$ , das der Wert einer Zufallsvariablen mehr als ein Intervallradius  $\epsilon$  von seinem Erwartungswert abweicht, nicht größer als das Verhältnis von Varianz und Intervallradius  $\epsilon$ :

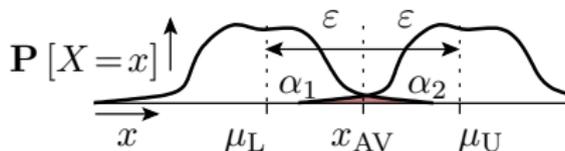
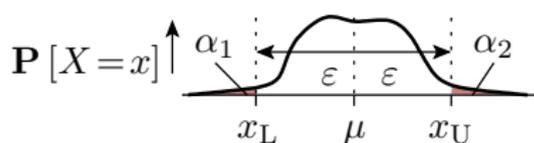
$$\mathbb{P}[|X - \mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2} \quad (3.74)$$

Intervallradius zur Irrtumswahrscheinlichkeit  $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ :

$$\epsilon \leq \sqrt{\frac{\sigma^2}{\alpha}} = \frac{\sigma}{\sqrt{\alpha}} \quad (3.75)$$

$X$	Zufallsvariable.
$\epsilon, \mu, \sigma$	Intervallradius, Erwartungswert, Standardabweichung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.

## Bereichsschätzung



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

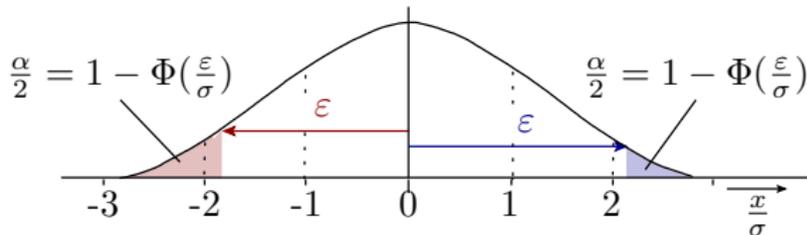
$$\text{sr}[x] = \mu \mp \frac{\sigma}{\sqrt{\alpha}} \quad (3.76)$$

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

$$\text{sr}[\mu] = x_{AV} \mp \frac{\sigma}{\sqrt{\alpha}} \quad (3.77)$$

$\varepsilon, \mu, \sigma$	Intervallradius, Erwartungswert, Standardabweichung.
$\alpha$	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\mu_L, \mu_U$	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
$x_L, x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von $X$ .
$\text{sr}(x)$	Symmetrischer Bereich der wahrscheinlichen Werte der Zufallsvariablen.
$\text{sr}(\mu)$	Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.
$x_{AV}$	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.

## Vergleich mit Intervallradius Normalverteilung



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

$$\epsilon = \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.56)$$

Intervallradius Tschebyscheffsche Ungleichung:

$$\epsilon \leq \frac{\sigma}{\sqrt{\alpha}} \quad (3.75)$$

$\alpha$	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
$\alpha^{-0,5}$	4,68	6,20		5	7,07	10	15,8	22,4

### Beispiel 3.15: Tschebytscheffschen Ungleichung

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 Tschebytscheffsche Ungleichung?*
- Unter der Annahme, dass die Widerstandswerte normalverteilt sind?*

---

$\alpha$  Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.  
 $R_i$  Widerstandswerte in  $k\Omega$ .



$\alpha = 2\%$ , 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?*

$$\hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (3.15)$$

$$\hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (3.16)$$

$$\text{sr}[x] = \mu \mp \frac{\sigma}{\sqrt{\alpha}} \quad (3.76)$$

$$\hat{\mu} = \frac{1}{8} (10,3 + \dots) \text{ k}\Omega = 10,025 \text{ k}\Omega$$

$$\hat{\sigma} = \sqrt{\frac{1}{7} ((10,3 - 10,025)^2 + \dots)} \text{ k}\Omega^2 = 647 \Omega$$

Bereich des Erwartungswerts:

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



$\alpha = 2\%$ , 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?*

$$\text{sr}(x) = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.57)$$

$\alpha$	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

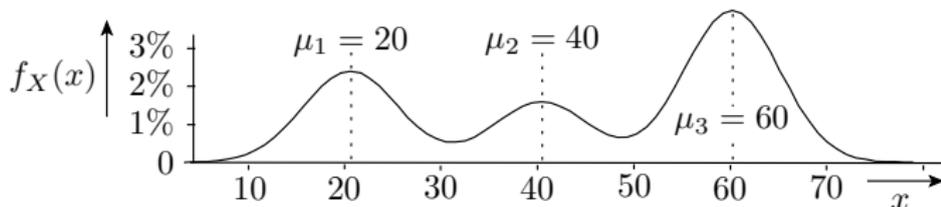
$$\begin{aligned} \text{sr}(R) &= \hat{\mu} \mp \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \hat{\sigma} \\ &= 10,025 \text{ k}\Omega \mp 2,33 \cdot 647 \Omega \\ &= [8,5 \text{ k}\Omega, 11,5 \text{ k}\Omega] \end{aligned}$$

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:

$$\mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i) \quad (3.70)$$

Mischung stetiger Verteilungen:

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{i=1}^{\#Y} h_j \cdot f_{X_j}(x) \quad (3.71)$$

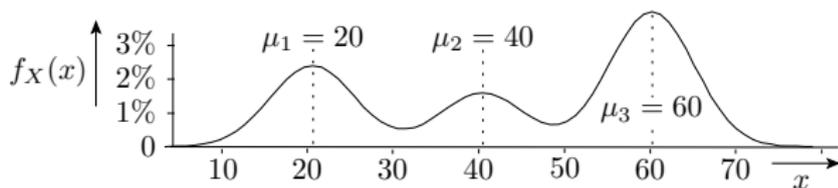
Erwartungswert:

$$\mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \quad (3.72)$$

Varianz:

$$\text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \quad (3.73)$$

## Multimodalität



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

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

Die Maxima liefern Hinweise auf Verbesserungsmöglichkeiten

- für Entstehungsprozesse zur Fehlervermeidung,
- für die Bewertung erfasster Daten, z.B. Erkennen unerwartete Abhängigkeiten zwischen Zählwerten und Vorlieben von Experten bei Befragungen,
- für Messverfahren, ...

## Bereichsschätzung

Bereichsschätzung für beliebige incl. multimodale Verteilungen:

- tschebyscheffschen Ungleichung:

$$\mathbb{P} [ |X - \mu| \geq \varepsilon ] \leq \frac{\sigma^2}{\varepsilon^2} \quad (3.74)$$

- garantierbarer Intervallradius:

$$\varepsilon \leq \frac{\sigma}{\sqrt{\alpha}} \quad (3.75)$$

- garantierbarer Bereiche für Realisierungen und Erwartungswerte:

$$\text{sr} [x] = \mu \mp \frac{\sigma}{\sqrt{\alpha}} \quad (3.76)$$

$$\text{sr} [\mu] = x_{AV} \mp \frac{\sigma}{\sqrt{\alpha}} \quad (3.77)$$



# Pareto-Verteilung



## Pareto-Verteilung

Die Pareto-Verteilung

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

ist eine stetige Wahrscheinlichkeitsverteilung. Verteilungsfunktion:

$$F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^K & \text{sonst} \end{cases} \quad (3.78)$$

Dichtefunktion für  $x \geq x_{\min}$ :

$$f_X(x) = \frac{K \cdot x_{\min}^K}{x^{K+1}} \quad (3.79)$$

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

$K > 0$       Formfaktor der Pareto-Verteilung.

$x_{\min} > 0$       Skalenparameter der Pareto-Verteilung.

\* Der italienische Ökonom Vilfredo Pareto untersuchte 1906 die Verteilung des Grundbesitzes in Italien und fand heraus, dass ca. 20% der Bevölkerung ca. 80% des Bodens besitzen. Das ist in den Sprachgebrauch als Pareto-20%-80%-Regel eingegangen.



### 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  $K > 1$ :

$$\mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K - 1} \quad (3.80)$$

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

$$\text{Var}[X] = \sigma^2 = x_{\min}^2 \cdot \frac{K}{(K - 2)(1 - K)^2} \quad (3.81)$$

- 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%-Regel

Der Anteil der Ursachen  $U$  mit der größten Wirkung:

$$U = \int_{w_{\min}}^{\infty} f(x) \cdot dx = \int_{w_{\min}}^{\infty} \frac{K \cdot x_{\min}^K}{x^{K+1}} \cdot dx = \left( \frac{x_{\min}}{w_{\min}} \right)^K \quad (3.82)$$

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:

$$W = \frac{\mathbb{E}[X|X \geq w_{\min}]}{\mathbb{E}[X]} = \left( \frac{x_{\min}}{w_{\min}} \right)^{K-1} = U^{\frac{K-1}{K}} \quad (3.83)$$

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



# Fehlernachweislänge



## Verteilung der Fehlernachweislänge

Die Zufallsvariable  $X$  sei die Nachweislänge eines Fehlers, d.h. die Anzahl der Service-Leistungen oder die Zeit bis zum Nachweis. Die Verteilungsfunktion als Wahrscheinlichkeit, Nachweislänge  $X$  nicht größer als Testssatzlänge  $N$ , ist die zu erwartende Fehlerüberdeckung:

$$F_X(N) = \mathbb{P}[X \leq N] = \mu_{FC}(N)$$

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 (siehe Abschn. 1.4.5 *Zuverlässigkeit nach Fehlerbeseitigung*):

$$1 - \mu_{FC}(N) = \left(\frac{N}{N_0}\right)^{-K} \quad (1.57)$$

$K$	1	0,5	0,33	0,25
$\frac{N}{N_0}$ für $1 - \mu_{FC}(N) = 0,1$	10	100	$10^3$	$10^4$

---

$\mu_{FC}(N)$	Zu erwartende Fehlerüberdeckung in Abhängigkeit von der Testanzahl.
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .



Daraus resultiert, dass die Nachweislänge  $X$  bei Abnahme des Anteils der nicht nachweisbaren Fehler nach Gl. 1.57 pareto-verteilt ist mit  $N_0$  als Skalenparameter und  $K$  als Formfaktor:

$$X \sim \text{Par}(K, N_0)$$
$$F_X(N) = \mu_{FC}(N) = 1 - \left(\frac{N_0}{N}\right)^K \quad \text{für } N \geq N_0 \quad (3.84)$$

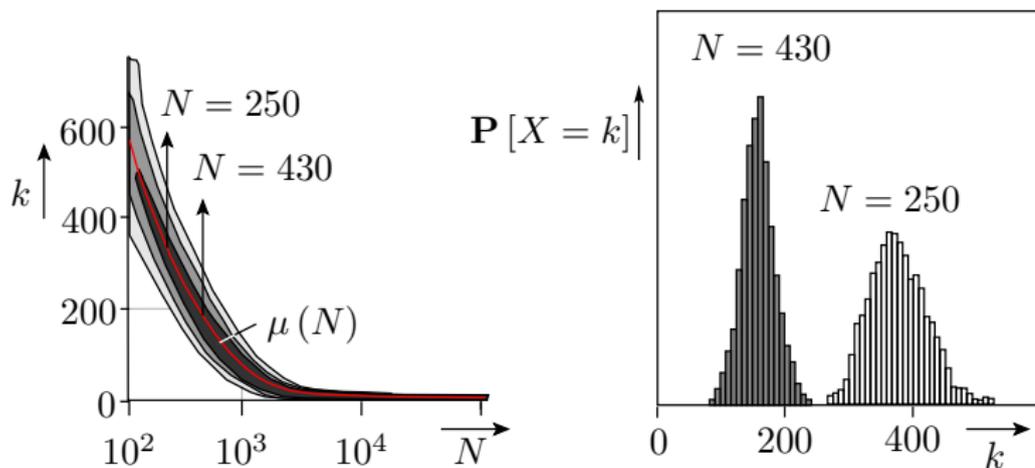
Wegen  $K < 1$  hat die Nachweislänge keinen Erwartungswert. Beim Betrieb von IT-System mit pareto-verteilter Nachweislänge und Beseitigung aller erkennbaren Fehler sind auch nach sehr langer Nutzungsdauer weitere Fehler nicht ausschließbar.

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

---

$F_X(N)$	Verteilungsfunktion der Nachweislänge.
$\mu_{FC}(N)$	Zu erwartende Fehlerüberdeckung in Abhängigkeit von der Testanzahl.
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .
$K$	Formfaktor der Verteilung der Fehlfunktionsrate ( $0 < K < 1$ ).

## Für das Haftfehlerexperiment

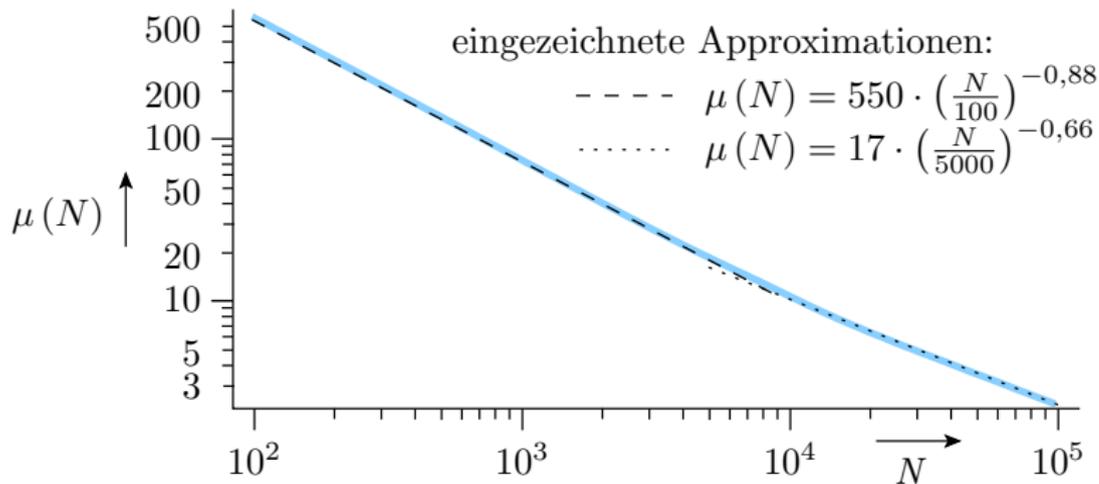


Kombinatorische Beispielschaltung (Benchmark c3540). 3606 simulierte, unterschiedlich nachweisbare Haftfehler. Zu erwartende Anzahl der nicht nachweisbaren Fehler:

$$\mu(N) = 3606 \cdot (1 - \mu_{FC}(N))$$

$\mu(N)$  Zu erwartende Anzahl nicht nachweisbaren Modellfehler als Funktion der Testanzahl  $N$ .  
 $k$  Anzahl nicht nachweisbaren Modellfehler.

## Pareto-Näherung



Formfaktor  $K = 0,88$  nähert den Bereich  $N \in [10^2, 10^4]$  und  $K = 0,66$  den Bereich bis  $N \in [10^4, 10^5]$  besser an. Aber Achtung, für  $N \geq 10^4$  sind die Zählwerte sehr klein und der Schätzfehler für den Erwartungswert  $\mu(N)$  und damit auch für den Formfaktor  $K$  groß.

---

$\mu(N)$       Zu erwartende Anzahl nicht nachweisbaren Modellfehler als Funktion der Testanzahl  $N$ .  
 $K > 0$       Formfaktor der Pareto-Verteilung.



# 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 (siehe Folie 1.12 *Der Preis fehlender Verlässlichkeit*)).

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 einer Schweizer  
Autoversicherung\*:

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

---

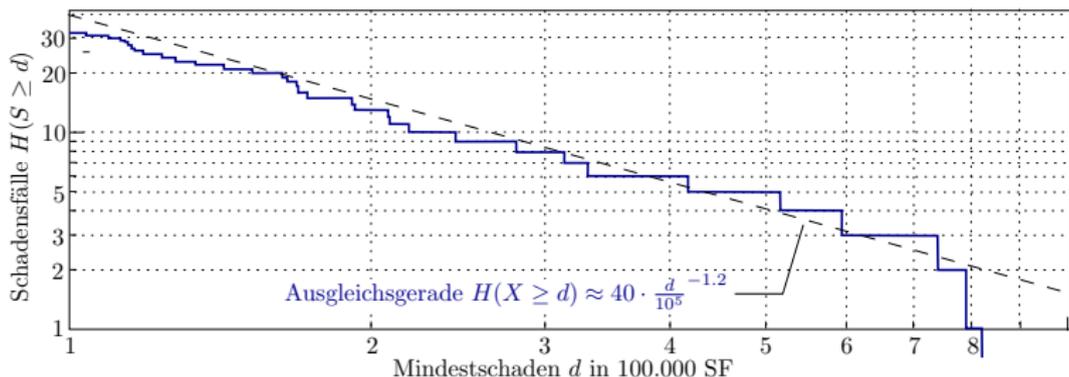
SF                      Schadenskosten in Schweizer Franken.

\*

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

## Annäherung durch eine Pareto-Verteilung

Anzahl der Versicherungsfälle mit einem Schaden größer  $s$ :



Pareto-Verteilung der Schadenshöhe ab  $d_{\min} = 10.000 \text{ SF}$ :

$$F_X(d) = \mathbb{P}[X \leq d] = 1 - \left(\frac{d}{d_{\min}}\right)^{-K} = 1 - \left(\frac{d}{10^5}\right)^{-1,2}$$

$H(X \geq d)$	Anzahl der Schadensfälle mit einem Schadenskosten größer $d$ .
$d$	Schadenskosten in Schweizer Franken.
$d_{\min}$	Betrachteter Mindestschaden, Skalenparameter der Pareto-Verteilung.



Erwartungswert (Gl. 3.80):

$$\mu_d = d_{\min} \cdot \frac{K}{K-1} = \frac{1,2}{1,2-1} \cdot d_{\min} = 600.000 \text{ SF}$$

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.

$\mu_d$	Zu erwartende Schadenskosten in Schweizer Franken.
$d_{\min}$	Betrachteter Mindestschaden, Skalenparameter der Pareto-Verteilung.
$K > 0$	Formfaktor der Pareto-Verteilung.
SF	Schadenskosten in Schweizer Franken.



# Zusammenfassung



Verteilungsfunktion und Dichte der Pareto-Verteilung:

$$F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^K & \text{sonst} \end{cases} \quad (3.78)$$

$$f_X(x) = \frac{K \cdot x_{\min}^K}{x^{K+1}} \quad (3.79)$$

Erwartungswert für  $k > 1$ :

$$\mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K-1} \quad (3.80)$$

Varianz für  $k > 2$ :

$$\text{Var}[X] = \sigma^2 = x_{\min}^2 \cdot \frac{K}{(K-2)(1-K)^2} \quad (3.81)$$

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

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

$$F_X(N) = \mu_{\text{FC}}(N) = 1 - \left(\frac{N_0}{N}\right)^K \quad \text{für } N \geq N_0 \quad (3.84)$$

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



# Gamma- und Exp-Vert.



### Eigenschaften



## 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 MF-Raten von Fehlern in IT-Systemen.

Dichte:

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} \quad \text{für } x \geq 0 \quad (3.85)$$

Gamma-Funktion:

$$\Gamma(\alpha) = \int_0^\infty e^{-z} \cdot z^{\alpha-1} \cdot dz \quad (3.86)$$

---

$\mathcal{G}(\alpha, \beta)$	Gamma-Verteilung, $\alpha$ – Formparameter, $\beta$ – Skalenparameter.
$f_X(x)$	Dichtefunktion der Zufallsvariablen $X$ .
$\Gamma(\dots)$	Gamma-Funktion.



## Gammafunktion

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

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

$\alpha$	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:

$$\Gamma(\alpha + 1) = \alpha \cdot \Gamma(\alpha) \quad (3.87)$$

$\Gamma(\dots)$

Gamma-Funktion.

$\alpha$

Formparameter der Gamma-Verteilung.



## Eigenschaften der Gamma-Verteilung

Erwartungswert:

$$\mathbb{E}[X] = \mu = \frac{\alpha}{\beta} \quad (3.88)$$

Varianz:

$$\text{Var}[X] = \sigma^2 = \frac{\alpha}{\beta^2} \quad (3.89)$$

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:

$$X_1 + X_2 \sim \mathcal{G}(\alpha_1 + \alpha_2, \beta) \quad (3.90)$$

---

$\mathcal{G}(\alpha, \beta)$	Gamma-Verteilung, $\alpha$ – Formparameter, $\beta$ – Skalenparameter.
$\mathbb{E}[X]$	Erwartungswert der Zufallsvariablen $X$ .
$\text{Var}[X]$	Varianze der Zufallsvariablen $X$ .



## Exponentialverteilung

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

$$\begin{aligned} f_X(x) &= e^{-\beta \cdot x} \cdot x^{\alpha-1} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \quad \text{für } x \geq 0 \text{ mit } \alpha = 1 \text{ und } \beta = \lambda \\ &= \lambda \cdot e^{-\lambda \cdot x} \quad \text{für } x \geq 0 \end{aligned} \quad (3.91)$$

$\lambda$  – Anzahl der zu erwartenden Ereignisse pro Zeitintervall.  
Verteilungsfunktion:

$$F_X(x) = \mathbb{P}(X \leq x) = \int_0^x f(z) \cdot dz = 1 - e^{-\lambda \cdot x} \quad \text{für } x \geq 0 \quad (3.92)$$

Erwartungswert:

$$\mathbb{E}[X] = \mu = \frac{1}{\lambda} \quad (3.93)$$

Varianz:

$$\text{Var}[X] = \sigma^2 = \frac{1}{\lambda^2} \quad \Rightarrow \quad \sigma = \mu \quad (3.94)$$

---

$\text{Exp}(\lambda)$     Exponentialverteilung,  $\lambda$  – Verteilungsparameter.



MF-Rate



## Korrektur Dichte MF-Rate etc. ab Folie 1.145

Die Nachweislänge als Wahrscheinlichkeit, dass ein Fehler mit  $N$  Tests nachgewiesen wird, ist gleich der zu erwartenden Fehlerüberdeckung und in Anlehnung an Gl. 1.57 pareto-verteilt:

$$F_X(N) = \mu_{FC}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K}$$

Auf Foliensatz 1, Folie 1.145 hatten wir unter der vereinfachten Annahme, dass ein Zufallstest alle Fehler mit MF-Rate  $\zeta \geq \frac{1}{n}$  nachweist, daraus als Dichte für die MF-Rate abgeleitet:

Dichtefunktion der MF-Rate

$$h(\zeta, N) = \begin{cases} K \cdot N^K \cdot \zeta^{K-1} & 0 \leq \zeta < \frac{1}{N} \\ 0 & \text{sonst} \end{cases} \quad (1.59)$$

---

$\mu_{FC}(N)$	Zu erwartende Fehlerüberdeckung in Abhängigkeit von der Testanzahl.
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .
$K$	Formfaktor der Verteilung der Fehlfunktionsrate ( $0 < K < 1$ ).

## Korrektur $p_N$ und Dichte MF-Rate

Ausgehend von der Beseitigungswahrscheinlichkeit für einen Fehler mit MF-Rate  $\zeta_i$

$$p_i(N) = 1 - e^{-\zeta_i \cdot N} \quad (2.9)$$

beträgt die Nichtbeseitigungswahrscheinlichkeit bei Fortsetzung der Fehlerbeseitigungsiteration mit  $N - N_0$  Tests:

$$p_F(\zeta, N) = e^{-\zeta \cdot (N - N_0)}$$

Mit

$$\mu_F(N) = \mu_F \cdot \int_0^1 p_F(\zeta, N) \cdot h(\zeta) \cdot d\zeta \quad (1.55)$$

$$1 - \mu_{FC}(N) = \left(\frac{N}{N_0}\right)^{-K} \quad (1.57)$$

muss Dichte der MF-Rate  $h(\zeta)$  ...

---

$p_F(\zeta, N)$	Wahrscheinlichkeit, dass Fehler mit MF-Rate $\zeta$ nach $N$ Tests noch vorhanden sind.
$\zeta$	Fehlerfunktionsrate eines einzelnen Fehlers.
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .
$\mu_F(N)$	Zu erwartende Anzahl der Fehler, die nach $N$ Tests nicht erkannt und beseitigt sind.



... muss Dichte der MF-Rate folgendes Integral erfüllen:

$$\mu_{FC}(N) = 1 - \int_0^1 e^{-\zeta \cdot (N - N_0)} \cdot h(\zeta) \cdot d\zeta$$

Mit einer pareto-verteilten Nachweislänge, d.h.  $FC$  nach

$$1 - \mu_{FC}(N) = \left(\frac{N}{N_0}\right)^{-K} \quad (1.57)$$

$$\left(\frac{N}{N_0}\right)^{-K} = \int_0^1 e^{-\zeta \cdot (N - N_0)} \cdot h(\zeta) \cdot d\zeta$$

ist die Lösung für  $N \gg 1$  eine Gamma-verteilte Dichtefunktion:

$$h(\zeta) = \frac{N_0^K}{\Gamma(K)} \cdot \zeta^{K-1} \cdot e^{-N_0 \cdot \zeta} \quad (3.95)$$

mit Formfaktor  $K$  und Skalenparameter  $N_0$ . Der Skalenparameter ist dabei die Anzahl der Tests, für die alle erkannten Fehler beseitigt sind, vor dem Test  $N_0$  und nach dem Test  $N$ .

$\mu_F(N)$	Zu erwartende Anzahl der Fehler, die nach $N$ Tests nicht erkannt und beseitigt sind.
$\mu_{FC}(N)$	Zu erwartende Fehlerüberdeckung in Abhängigkeit von der Testanzahl.
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .
$K$	Formfaktor der Verteilung der Fehlfunktionsrate ( $0 < K < 1$ ).

### Beweis durch Nachrechnen

$$\left(\frac{N}{N_0}\right)^{-K} \stackrel{!}{=} \frac{N_0^K}{\Gamma(K)} \cdot \int_0^1 e^{-\zeta \cdot (N-N_0)} \cdot \zeta^{K-1} \cdot e^{-N_0 \cdot \zeta} \cdot d\zeta$$

$$N^{-K} \stackrel{!}{=} \frac{1}{\Gamma(K)} \cdot \int_0^1 e^{-\zeta \cdot N} \cdot \zeta^{K-1} \cdot d\zeta$$

Substitution  $\zeta = \frac{z}{N}$ ,  $d\zeta = \frac{dz}{N}$ :

$$N^{-K} \stackrel{!}{=} \frac{1}{\Gamma(K)} \cdot \int_0^N \left(\frac{z}{N}\right)^{K-1} \cdot e^{-z} \cdot \frac{dz}{N}$$

$$= \frac{1}{\Gamma(K)} \cdot N^{-K} \cdot \underbrace{\int_0^N z^{K-1} \cdot e^{-z} \cdot dz}_{\Gamma(K) \text{ für } N \gg 1} = N^{-K}$$

$\Gamma(K)$	Gamma-Funktion: $\Gamma(K) = \int_0^\infty e^{-z} \cdot z^{K-1} \cdot dz$ .
$K$	Formfaktor der Verteilung der Fehlfunktionsrate ( $0 < K < 1$ ).
$N_0$	Testanzahl, für die vorher alle Fehler beseitigt wurden, also für $FC = 0$ .
$N$	Anzahl der Tests, für die erkannten Fehler beseitigt werden, incl. $N_0$ .

## Zu erwartende Fehleranzahl und MF-Rate

Mittlere MF-Rate je Fehler ist der Erwartungswert der gamma-verteilten MF-Rate, also Formfaktor  $K$  durch Skalenparameter  $N$ :

$$\bar{\zeta}(N) = \frac{K}{N}$$

Eingesetzt in

$$\zeta_F(N) = \mu_F(N) \cdot \bar{\zeta}(N) \quad (1.54)$$

ergibt sich für Gl. 1.60 die korrigierte Version ohne  $(K + 1)$ :

$$\zeta_F(N) = \frac{\mu_F(N) \cdot K}{(K+1) \cdot N} \quad (1.60)$$

In Gl. 1.61 hatte sich der Term  $(K + 1)$  ohnehin herausgekürzt:

$$\mu_F(N_2) = \mu_F(N_1) \cdot \left(\frac{N_2}{N_1}\right)^{-K} \quad (1.58)$$

$$\frac{\zeta_F(N_2) \cdot (K+1) \cdot N_2}{K} = \frac{\zeta_F(N_1) \cdot (K+1) \cdot N_1}{K} \cdot \left(\frac{N_2}{N_1}\right)^{-K}$$

$$\zeta_F(N_2) = \zeta_F(N_1) \cdot \left(\frac{N_2}{N_1}\right)^{-(K+1)} \quad (1.61)$$

$\bar{\zeta}(N)$  Mittlere Fehlfunktionsrate je Fehler als Funktion der Testanzahl  $N$ .



### Beispiel 3.16: Fehleranzahl und MF-Rate

Die MF-Rate eines Systems beträgt nach der effektiven Testlänge  $N_1 = 10^5$  Tests  $\zeta_F(N_1) = 10^{-4} \left[ \frac{MF}{DS} \right]$ . Gamma-verteilte MF-Raten mit den Formfaktoren  $K \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$ .

- Wie groß ist die zu erwartende Anzahl der nicht beseitigten Fehler  $\mu_F$  für Testlänge  $N_1$ ?
- Wie groß ist die zu erwartende Anzahl der nicht beseitigten Fehler für die zehnfache effektive Testlänge  $N_2 = 10^6$ ?
- Wie groß ist die MF-Rate für die zehnfache effektive Testlänge  $N_2 = 10^6$ ?

---

$N_1, N_2$	Testanzahl mit bekannter / gesuchter Fehlfunktionsrate oder Fehleranzahl.
$\zeta_F(N)$	Fehlfunktionsrate durch Fehler in Abhängigkeit von der Testanzahl.
$\left[ \frac{MF}{DS} \right]$	Zählwertverhältnis in Fehlfunktionen je erbrachte Service-Leistung.
$K$	Formfaktor der Verteilung der Fehlfunktionsrate ( $0 < K < 1$ ).
$\mu_F(N)$	Zu erwartende Anzahl der Fehler, die nach $N$ Tests nicht erkannt und beseitigt sind.



Die MF-Rate eines Systems beträgt nach der effektiven Testlänge  $N_1 = 10^5$  Tests  $\zeta_F(N_1) = 10^{-4} \left[ \frac{MF}{DS} \right]$ . Gamma-verteilte MF-Raten mit den Formfaktoren  $K \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$ .

a) *Wie groß ist die zu erwartende Anzahl der nicht beseitigten Fehler  $\mu_F$  für Testlänge  $N_1$ ?*

$$\zeta_F(N) = \frac{\mu_F(N) \cdot K}{N} \quad (1.60)$$

$$\mu_F(N_1) = \zeta_F(N_1) \cdot \frac{N_1}{K} = 10^4 \cdot \frac{10^5}{K} = \frac{10}{K}$$

$K$	0,3	0,4	0,5	0,6	0,7
$\mu_F(10^5)$	33,3	25	20	16,7	14,3



Die MF-Rate eines Systems beträgt nach der effektiven Testlänge  $N_1 = 10^5$  Tests  $\zeta_F(N_1) = 10^{-4} \left[ \frac{MF}{DS} \right]$ . Gamma-verteilte MF-Raten mit den Formfaktoren  $K \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$ .

b) *Wie groß ist die zu erwartende Anzahl der nicht beseitigten Fehler für die zehnfache effektive Testlänge  $N_2 = 10^6$ ?*

$$\mu_F(N_2) = \mu_F(N_1) \cdot \left( \frac{N_2}{N_1} \right)^{-K} \quad (1.58)$$

$$\mu_F(10^6) = \mu_F(10^5) \cdot \left( \frac{10^6}{10^5} \right)^{-K} = \frac{10}{K} \cdot 10^{-K}$$

$K$	0,3	0,4	0,5	0,6	0,7
$\mu_F(10^5)$	33,3	25	20	16,7	14,3
$\mu_F(10^6)$	16,7	9,9	6,3	4,2	2,9



Die MF-Rate eines Systems beträgt nach der effektiven Testlänge  $N_1 = 10^5$  Tests  $\zeta_F(N_1) = 10^{-4} \left[ \frac{MF}{DS} \right]$ . Gamma-verteilte MF-Raten mit den Formfaktoren  $K \in \{0,3, 0,4, 0,5, 0,6, 0,7\}$ .

c) *Wie groß sind ist die MF-Rate für die zehnfache effektive Testlänge  $N_2 = 10^6$ ?*

$$\zeta_F(N_2) = \zeta_F(N_1) \cdot \left( \frac{N_2}{N_1} \right)^{-(K+1)} \quad (1.61)$$

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

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

Die Vorhersagewerte liegen deutlich näher beieinander als bei der zu erwartenden Fehleranzahl.



# 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.

Kenngrößen:

- 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) und Fehler, die bei Beseitigungsversuchen anderer Fehler entstehen.

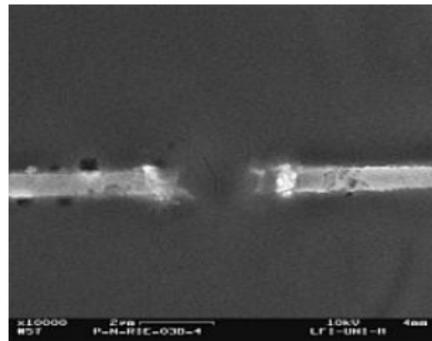
---

MF	Fehlfunktion.
NS	Keine Service-Leistung.

### 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  $L$ : Zeit vom Beanspruchungsbeginn bis zum Ausfall.  
Verteilungsfunktion:

$$F_L(t) = \mathbb{P}[L \leq t]$$

- Überlebenswahrscheinlichkeit:

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

- Ausfallrate  $\lambda$ : Relative Abnahme der Überlebenswahrscheinlichkeit mit der Zeit:

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

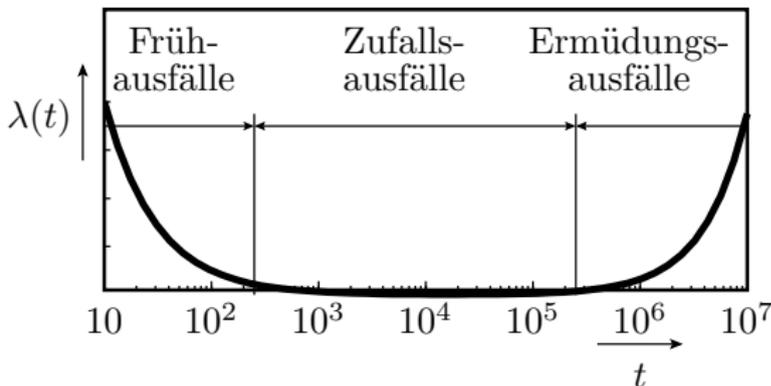
- Mittlere Lebensdauer:

$$\mu_L = \int_0^{\infty} V(t) \cdot dt$$

---

$L$	Lebensdauer, Zufallsvariable.
$V(t)$	Überlebenswahrscheinlichkeit als Funktion der Lebensdauer.
$\lambda(t)$	Ausfallrate als Funktion der Zeit.
$\mu_L$	Zu erwartende Lebensdauer.

### Ausfallphasen



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

$\lambda(t)$  Ausfallrate als Funktion der Zeit.

## 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\*:

$$V(t) = e^{-\lambda \cdot t} \quad (3.96)$$

$$F_L(t) = 1 - e^{-\lambda \cdot t} \quad (3.97)$$

(Exponentialverteilung vergl. Gl. 3.92). Mittlere Lebensdauer (Gl. 3.93):

$$\mu_L = \int_0^{\infty} V(t) \cdot dt = \frac{1}{\lambda} \quad (3.98)$$

Maßeinheit der Ausfallrate: fit (failure in time)

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

---

$V(t)$  Überlebenswahrscheinlichkeit als Funktion der Lebensdauer.

fit Failure In Time, Anzahl der Ausfälle in  $10^9$  Stunden.

\* 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:

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

Mit einer konstanten Ausfallrate  $\lambda_i$  für alle Komponenten (Gl. 3.97):

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

ist die Gesamtausfallrate die Summe der Ausfallraten aller Komponenten:

$$\lambda_{\text{Sys}} = \sum_{i=1}^{\#C} \lambda_i \quad (3.99)$$

---

$V(t)$	Überlebenswahrscheinlichkeit als Funktion der Lebensdauer.
$\#C$	Anzahl der für das Funktionieren erforderlichen Bausteine.
$\lambda_{\text{Sys}}$	Ausfallrate des Systems.
$\lambda_i$	Ausfallrate Komponente $i$ .



## Ausfallraten in der Hauptnutzungsphase nach [KRG96]

Bauteil	Ausfallrate in fit	Bauteil	Ausfallrate in fit
diskrete SC	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 Halbleiterbauteile (Semiconductor components).

IC Integrierte Schaltkreise (Integrated circuits).

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

## Ausfallrate einer Baugruppe

Bauteiltyp $i$	Anzahl $\#C_i$	Ausfallrate $\lambda_i$	$\#C_i \cdot \lambda_i$
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

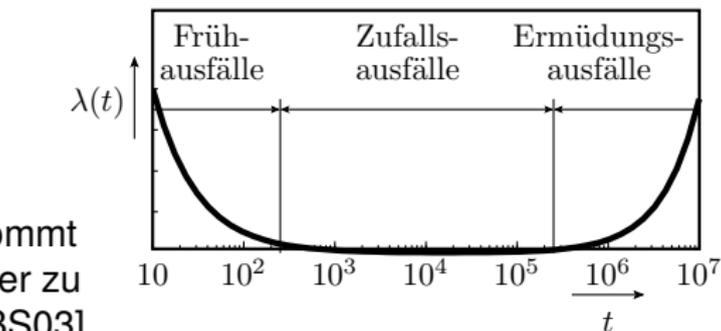
- Im Mittel 1 Ausfall in  $2 \cdot 10^5$  Stunden ( $\approx 23$  Jahre) Betriebsdauer.
- Von den heutigen PCs, Handys, ... fallen pro Jahr und hundert Stück nur wenige aus.
- Nach 2 ... 5 Jahren erste Ermüdungsausfälle, z.B. durch Austrocknung von Elektrolytkondensatoren.



# 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 einen 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?



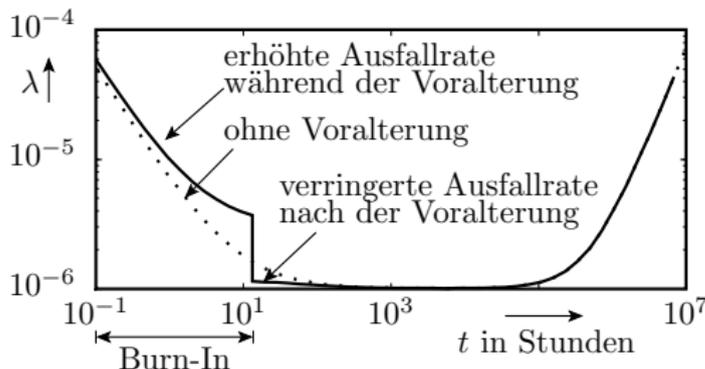
⇒ Voralterung

(Problem existiert auch bei Mechanik.)

$\lambda(t)$  Ausfallrate als Funktion der Zeit.

[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.

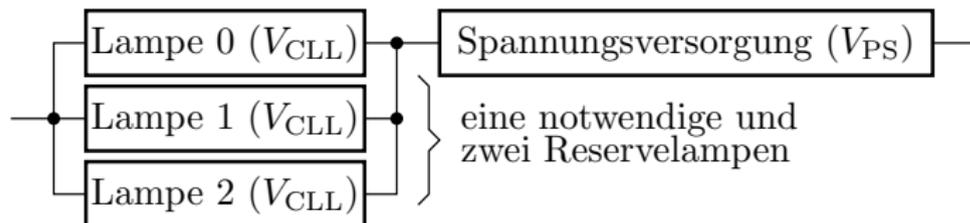
Reserveeinheiten sind erforderlich für

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

## 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:



$$V_{CL}(t) = (1 - (1 - V_L(t))^3) \cdot V_{PS}(t) \wedge$$

$V_{CL}(t)$  Überlebenswahrscheinlichkeit der Flurbeleuchtung.

$V_{CLL}(t)$  Überlebenswahrscheinlichkeit einer einzelnen Flurbeleuchtungslampe.

$V_{PS}(t)$  Überlebenswahrscheinlichkeit der Spannungsversorgung.

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



### Kalte, warme und heiße Reserve

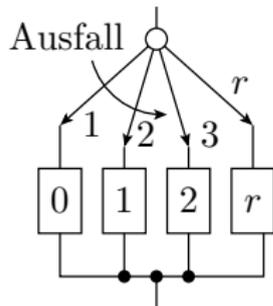
- Heiße Reserve: Reservekomponenten arbeiten parallel (z.B. Mehrversionssystem) und fallen mit derselben Wahrscheinlichkeit wie das aktive System aus.
- Kalte Reserve: Reservekomponenten werden geschont und funktionieren idealerweise noch alle zum Ausfallzeitpunkt der aktiven Komponente.
- Warme Reserve: 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« eingeschraubt 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	$\mu_{L.0}$
2	$\mu_{L.1}$
3	$\mu_{L.2}$
...	...



Die zu erwartenden Lebensdauern aller Komponenten addieren sich\*:

$$\mu_{L.Sys} = \sum_{i=0}^r \mu_{L.i} \quad (3.100)$$

$\mu_{L.Sys}$  Zu erwartende Lebensdauer des Systems.

$r$  Anzahl der Reserveeinheiten.

$\mu_L$  Zu erwartende Lebensdauer Komponente  $i$ .

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

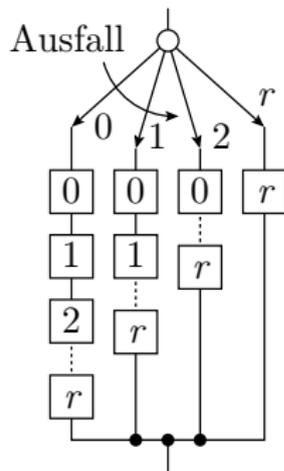
## Zu erwartende Lebensdauer bei heiße Reserve

Alle noch lebenden Komponenten können gleichermaßen ausfallen:

$$\mu_{L,i} = \frac{1}{\sum_{j=1}^{i+1} \lambda_j}$$

Komponenten mittlerer Lebensdauer  $\mu_{L,i} = \mu_{L,K}$ :

Phase	mittlere Lebensdauer
1	$\mu_{L,K} / (r + 1)$
2	$\mu_{L,K} / r$
...	...



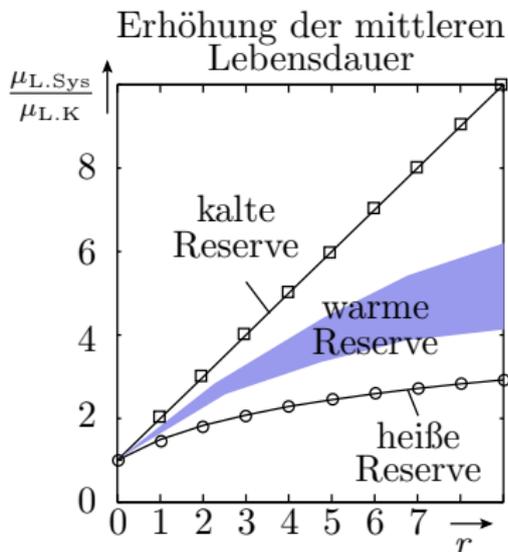
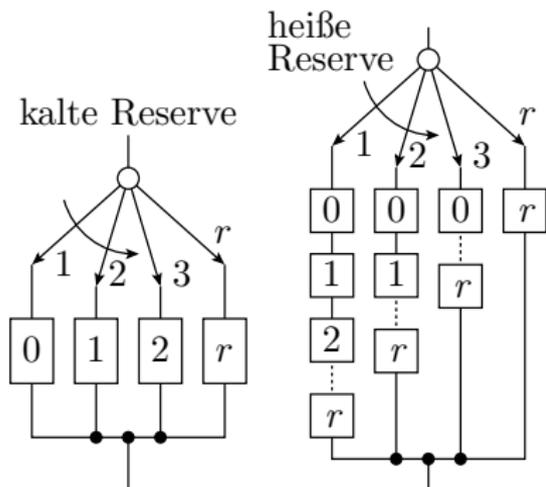
Zu erwartende Gesamtlebensdauer als Summe der zu erwartenden Lebensdauern aller Phasen:

$$\mu_{L, Sys} = \mu_{L,K} \sum_{i=0}^r \frac{1}{i+1} \quad (3.101)$$

$\mu_{L, Sys}$  Zu erwartende Lebensdauer des Systems.

$\mu_{L,K}$  Zu erwartende Lebensdauer der einzelnen Komponenten.

### 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 (Getriebe Schmierstoffe, Drucker 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  $\tau$ : Zeit zwischen den Wartungen, z.B. TÜV für KFZ

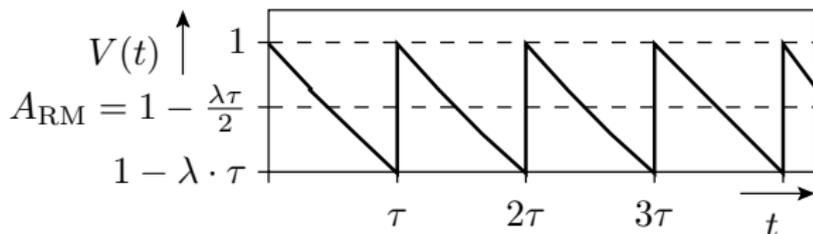
$\tau = 2$  Jahre.

---

$\tau$

Wartungsintervall.

## Reparatur bei der nächsten Wartung



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

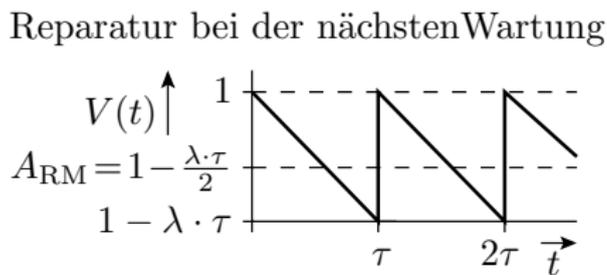
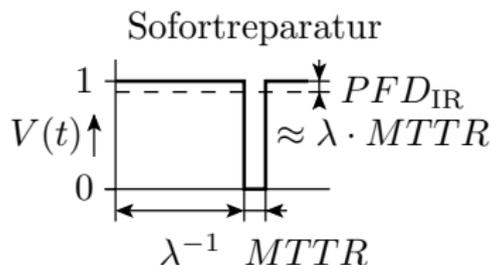
$$A_{RM} = \frac{1}{\tau} \cdot \int_0^{\tau} V(t) \cdot d\tau = \frac{1}{\tau} \cdot \int_0^{\tau} e^{-\lambda \cdot t} \cdot d\tau = \frac{1 - e^{-\lambda \cdot \tau}}{\lambda \cdot \tau}$$

mit

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

$$A_{RM} = 1 - \frac{\lambda \cdot \tau}{2}; \quad PFD_{RM} = 1 - A_{RM} = \frac{\lambda \cdot \tau}{2} \quad (3.102)$$

$V(t)$	Überlebenswahrscheinlichkeit als Funktion der Lebensdauer.
$A_{RM}$	Verfügbarkeit als mittlere Überlebenswahrsch. wenn Reparatur bei nächster Wartung.
$\tau$	Wartungsintervall.



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

$$PFD_{IR} = \frac{MTTR}{MTTF + MTTR} = \lambda \cdot MTTR \quad \text{für } \lambda \cdot MTTR \ll 1 \quad (3.103)$$

Wahrscheinlichkeit der Nichtverfügbarkeit insgesamt:

$$\begin{aligned} PFD &= \eta_{IR} \cdot PFD_{IR} + (1 - \eta_{IR}) \cdot PFD_{RM} \\ &= \eta_{IR} \cdot \lambda \cdot MTTR + (1 - \eta_{IR}) \cdot \frac{\lambda \cdot \tau}{2} \end{aligned} \quad (3.104)$$

$\eta_{IR}$	Anteil der Ausfälle, die durch Sofortreparatur beseitigt werden.
$PFD_{IR}$	Wahrscheinlichkeit der Nichtverfügbarkeit bei Sofortreparatur.
$PFD_{RM}$	Wahrscheinlichkeit Nichtverfügbarkeit, wenn Reparatur bei nächster Wartung.
$MTTR$	Mittlere Reparaturzeit (Mean time to repair).
$PFD$	Wahrscheinlichkeit der Nichtverfügbarkeit bei Anforderung.

### Beispiel 3.17: PFD

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

---

$$\lambda = 10^{-6} \text{ h}^{-1}, \eta_{\text{IR}} = 75\%, \tau = 2 \cdot 10^3 \text{ h}, MTTR = 4 \text{ h}.$$

*Wie groß ist die Wahrscheinlichkeit PFD der Nichtverfügbarkeit durch Ausfall?*

---

$\lambda$	Ausfallrate (Failure rate).
$\eta_{\text{IR}}$	Anteil der Ausfälle, die durch Sofortreparatur beseitigt werden.
$\tau$	Wartungsintervall.
$MTTR$	Mittlere Reparaturzeit (Mean time to repair).
$PFD$	Wahrscheinlichkeit der Nichtverfügbarkeit bei Anforderung.



$\lambda = 10^{-6} \text{ h}^{-1}$ ,  $\eta_{\text{IR}} = 75\%$ ,  $\tau = 2 \cdot 10^3 \text{ h}$ ,  $MTTR = 4 \text{ h}$ .

Wie groß ist die Wahrscheinlichkeit  $PFD$  der Nichtverfügbarkeit durch Ausfall?

$$PFD = \eta_{\text{IR}} \cdot \lambda \cdot MTTR + (1 - \eta_{\text{IR}}) \cdot \frac{\lambda \cdot \tau}{2} \quad (3.105)$$

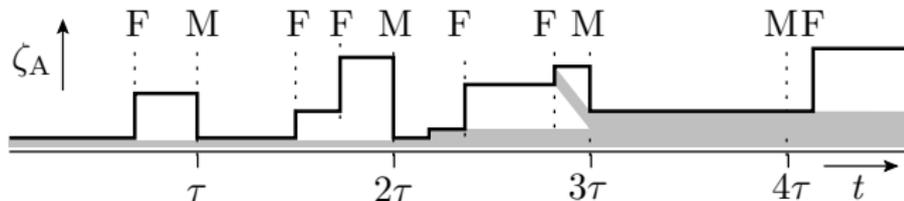
$$PFD = 10^{-6} \text{ h}^{-1} \cdot \left( \underbrace{0,75 \cdot 4 \text{ h}}_{3 \text{ h}} + \underbrace{\frac{0,25 \cdot 2 \cdot 10^3 \text{ h}}{2}}_{500 \text{ h}} \right) = 2,53 \cdot 10^{-4}$$

Zur Verringerung der  $PFD$  bieten sich an:

- Verkürzung des Wartungsintervalls  $\tau$  oder
- die Erhöhung des Anteils  $\eta_{\text{IR}}$  der sofort erkannten Ausfälle durch bessere Überwachung und Einschalttests.

Die mittlere Reparaturzeit hat wenig Einfluss und dürfte sogar größer sein, ohne dass sich die  $PFD$  dadurch wesentlich erhöht.

## Zuverlässigkeitsverlust durch Ausfälle



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

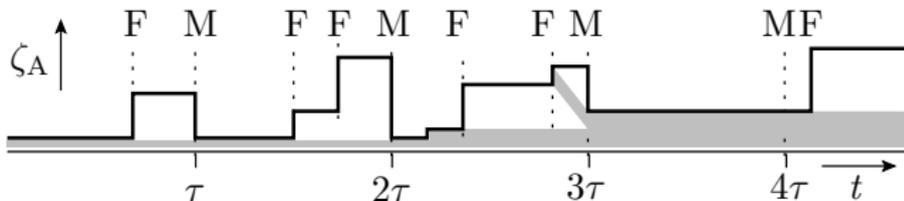
Fehler durch Ausfälle mit kleinem  $\zeta_A$

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

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

$\zeta_A$  Fehlfunktionsrate der Fehler durch Ausfälle.  
 $\tau$  Wartungsintervall.

## 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 Gesamtsystem, bevor sich durch Ausfälle verursachte Fehler häufen.

$\zeta_A$       Fehlfunktionsrate der Fehler durch Ausfälle.  
 $\tau$         Wartungsintervall.



# Zusammenfassung



## Ausfallverhalten von Hardware und Mechanik

Die Ausfallrate hat typisch eine Badewannenkurve:

- Nutzungsbeginn: Erhöhte Ausfallrate durch Frühausfälle.
- Nach den Frühausfällen näherungsweise konstante Ausfallrate.
- Nach langer Nutzungsdauer Anstieg der Ausfallrate durch Materialermüdung.

Zufallsausfälle haben konstante Ausfallrate, exponentiell abnehmende Überlebenswahrscheinlichkeit und exponentialverteilte Lebensdauer:

$$V(t) = e^{-\lambda \cdot t} \quad (3.97)$$

$$F_L(t) = 1 - e^{-\lambda \cdot t} \quad (3.98)$$

Erwartungswert der Lebensdauer:

$$\mu_L = \frac{1}{\lambda} \quad (3.99)$$

Für Systeme aus mehreren Komponenten Summation der Ausfallraten:

$$\lambda_{\text{Sys}} = \sum_{i=1}^{\#C} \lambda_i \quad (3.100)$$

Erhöhten Ausfallraten durch Frühausfälle werden durch Voralterung und Materialermüdung durch rechtzeitigen Ersatz umgangen.



## Reserveeinheiten und Wartung

Bei kalter Reserve addieren sich die Lebensdauern und ihre Erwartungswerte:

$$\mu_{L.Sys} = \sum_{i=0}^r \mu_{L.i} \quad (3.101)$$

Bei heißer Reserve vergrößert die erste Reserveeinheit die zu erwartende Lebensdauer um 50%, die zweite um ein Drittel etc.:

$$\mu_{L.Sys} = \mu_{L.K} \sum_{i=0}^r \frac{1}{i+1} \quad (3.102)$$

Regelmäßige Wartung und Fehlerbeseitigung nur zu den Wartungszeitpunkten:

$$PFD_{RM} = \frac{\lambda \cdot \tau}{2} \quad (3.103a)$$

Wenn zusätzlich zur regelmäßigen Wartung erkannte Fehler durch Ausfälle sofort beseitigt werden:

$$PFD = \eta_{IR} \cdot \lambda \cdot MTTR + (1 - \eta_{IR}) \cdot \frac{\lambda \cdot \tau}{2} \quad (3.105)$$

Fehler durch Ausfälle, die von den Wartungstests nicht erkannt werden, verringern die Zuverlässigkeit mit der Nutzungsdauer.