



Test und Verlässlichkeit

Foliensatz 4:

Themenspezifische Verteilungen

Prof. G. Kemnitz

Institut für Informatik, TU Clausthal
17. Dezember 2024



Inhalt Foliensatz 4

Verteilung

- 1.1 Charakteristische Größen
- 1.2 Summen, lineare Transformationen
- 1.3 Verteilung von Zählwerten
- 1.4 Messfehler

Näherungen

- 2.1 Binomialverteilung
- 2.2 Poisson-Verteilung
- 2.3 Bereichsschätzung Pois
- 2.4 Defektanteil

- 2.5 Normalverteilung
- 2.6 Bereichsschätzung NVT
- 2.7 Schätzen von Zählwerten
- 2.8 Varianzerhöhung

Mischverteilung

- 3.1 Eigenschaften
- 3.2 Anwendungen
- 3.3 Tschebyscheffsche Ungl.

Pareto-Verteilung

- 4.1 Eigenschaften
- 4.2 Anwendungen
- 4.3 Schaden durch MF

4.2 Themenspezifische Verteilungen

Wahrscheinlichkeiten sind die besten Vorhersagen der Eintrittshäufigkeiten für Zufallsexperimente mit zwei möglichen Ergebnissen.

Zählwerte, daraus abgeleitete Eintrittsraten und weitere interessierende Kenngrößen haben viele oder sogar nicht abzählbar viele mögliche Werte. Das Modell dafür sind Verteilungen.

Verteilungen erlauben Bereichsschätzungen, Genauigkeitsvorhersagen, ... Unsere Themen:

- Verteilungen: Kenngrößen, lineare Transformationen, Summen, ...
- Berechnung der Verteilungen für Zählwerte,
- Näherungsverteilungen für Zählwerte und Bereichsschätzungen,
- Genauigkeitsbetrachtungen, Schätzfehler.

- Pareto-Verteilung für Sachverhalte, bei denen ein kleiner Ursachen die Mehrheit der Wirkungen erzielt.



Verteilung



4.3 Grundbegriffe

- 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 annimmt.
- Das einfachste Zufallsexperiment ist der Bernoulliversuch. Er hat nur die mögliche Ergebnisse 0 und 1, die Verteilung

$$\begin{aligned}\mathbb{P}[X = 0] &= 1 - p \\ \mathbb{P}[X = 1] &= p\end{aligned}\tag{4.1}$$

und dient auch zur Definition der Wahrscheinlichkeit (Gl. 3.1).

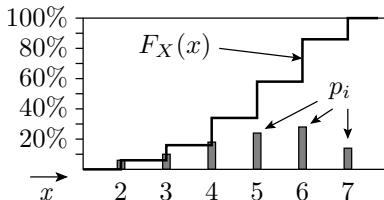
- Zufallsexperimente mit (sehr) vielen möglichen Ergebnissen werden durch ihre Verteilungsfunktion beschrieben:

$$F_X(x) = \mathbb{P}[X \leq x]\tag{4.2}$$

4.4 Diskrete Verteilung

Die Zufallsvariable X kann nur abzählbare 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%



Anwendbar auf Zählwerte:

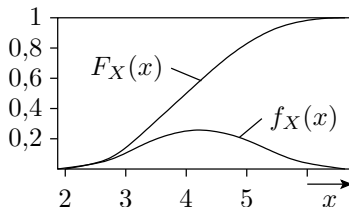
- Anzahl der Fehlfunktionen (aufgetretene, erkannte, ...),
- Anzahl der Fehler (entstandene, beseitigte, vermiedene, ...), ...

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

4.6 Stetige Verteilungen

Zufallsvariable X ist stetig und hat im Intervall $a \leq X \leq b$ unendlich viele Ausprägungen. Beschreibbar auch durch ihre Dichtefunktion:

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (4.3)$$



Auch als Näherung für diskrete Verteilungen mit sehr vielen möglichen Werten.

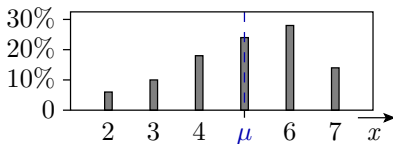
-
- $F_X(x)$ Verteilungsfunktion der Zufallsvariablen X .
 - $f_X(x)$ Dichtefunktion der Zufallsvariablen X .



Charakteristische Größen

4.7 Erwartungswert

Mittlerer Eintrittswert bei einer großen Versuchsanzahl:



- Diskrete Zufallsvariable:

$$\mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i \quad (4.4)$$

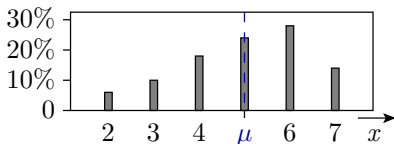
- Stetige Zufallsvariable:

$$\mathbb{E}[X] = \mu = \int_{x_{\min}}^{x_{\max}} f_X(x) \cdot x \cdot dx \quad (4.5)$$

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 .



4.8 Erwartungswert am Beispiel



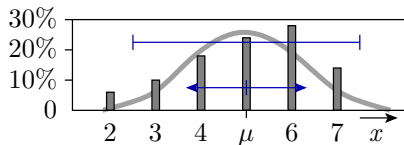
x_i	2	3	4	5	6	7
$\mathbb{P}[X = x_i] = p_i$	6%	10%	18%	24%	28%	14%
$F_X(x) = \mathbb{P}[X \leq x_i]$	6%	16%	34%	58%	86%	100%

$$(4.4) \quad \mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i$$

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

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 .
$\mathbb{E}[\dots], \mu$	Erwartungswert.

4.9 Varianz, Standardabweichung



wahrscheinlicher Bereich

Bereich $\mu \mp \sigma$

Die Standardabweichung ist ein Maß für die Breite des wahrscheinlichen Bereichs. Die Varianz ist mittlere quadratische Abweichung vom Erwartungswert und das Quadrat der Varianz:

$$\text{Var}[X] = \sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \quad (4.6)$$

$$\text{Var}[X] = \sigma^2 = \sum_{i=1}^{\#x} p_i \cdot (x_i - \mathbb{E}[X])^2 \quad (4.7)$$

$$\text{Var}[X] = \sigma^2 = \int_{x_{\min}}^{x_{\max}} f_X(x) \cdot (x - \mathbb{E}[X])^2 \cdot dx \quad (4.8)$$

$$\text{sd}[X] = \sigma = \sqrt{\text{Var}[X]} \quad (4.9)$$

$\text{Var}[\dots], \sigma^2$ Varianz.

$\text{sd}[\dots], \sigma$ Standardabweichung.



4.10 Steinerscher Verschiebungssatz

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

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \quad (4.10)$$

Für diskrete Zufallsvariablen:

$$\text{Var}[X] = \sum_{i=1}^{\#x} p_i \cdot (x_i)^2 - \mathbb{E}[X]^2 \quad (4.11)$$

Kontrolle durch Nachrechnen:

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

(Gl. 4.11) oft handlicher, aber bei kleinen Differenzen großer Werte numerische Probleme.

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

4.11 Varianz am Beispiel

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: $\mathbb{E}[X] = \mu = 5$ (siehe Gl. 4.4).

- Varianz nach

$$(4.7) \quad \text{Var}[X] = \sigma^2 = \sum_{i=1}^{\#x} p_i \cdot (x_i - \mathbb{E}[X])^2$$

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

- Varianz nach Verschiebesatz:

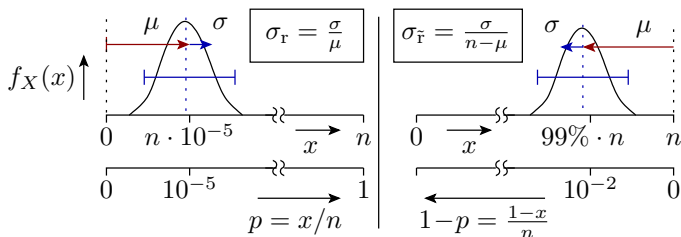
$$(4.11) \quad \text{Var}[X] = \sum_{i=1}^{\#x} p_i \cdot (x_i)^2 - \mathbb{E}[X]^2$$

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

- Standardabweichung:

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

4.12 Varianzkoeffizient



Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten

$$\begin{aligned}
 \sigma_r &= \frac{\sigma}{\mu} && \text{zweckmäßiges Mass für } \mu \leq n/2 \\
 \sigma_{\bar{r}} &= \frac{\sigma}{n-\mu} && \text{zweckmäßiges Mass für } \mu > n/2
 \end{aligned}
 \tag{4.12}$$

Maße für relative Breite wahrscheinlicher Bereich im Verhältnis zum Abstand vom Minimum bzw. Maximum eines Zählwerts x und auch dessen Eintritts- bzw. Nicht-Eintrittswahrscheinlichkeit (Schätzgenauigkeit).

p	Eintrittswahrscheinlichkeit, Anzahl der Zählversuche.
$\sigma_r, \sigma_{\bar{r}}$	Varianzkoeffizient für das Eintreten bzw. Nichteintreten des Zählwerts.
μ, σ	Erwartungswert, Standardabweichung.

4.13 Erwartungswert einer Datenstichprobe

Wahrscheinlichkeitsmodelle dienen zur Vorhersage künftig zu erwartender Häufigkeiten von Datenmerkmalen. Umgekehrt dienen erhobene Daten zur Kontrolle der Wahrscheinlichkeitsmodelle. Wir werden im weiteren zwischen Modell und Daten vergleichen:

- Erwartungswert, Varianz und
- Verteilungsparameter.

Für eine Datenstichprobe einer Zufallsvariable X

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

ist der im weiteren verwendete Schätzer für den Erwartungswert:

$$\hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \quad (4.13)$$

$\#v$	Größe der Datenstichprobe.
v_i	Wert i der Datenstichprobe.
$\hat{\cdot}$	Schätzwert.
$\mathbb{E}[\dots], \mu$	Erwartungswert.

4.14 Varianz, Standardabw. Datenstichproben

Schätzer der Varianz ist die mittlere quadratische Abweichung vom geschätzten Erwartungswert:

$$\hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v - 1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2 \quad (4.14)$$

$$\hat{\text{sd}}[X] = \hat{\sigma} = \sqrt{\hat{\text{Var}}[X]} \quad (4.15)$$

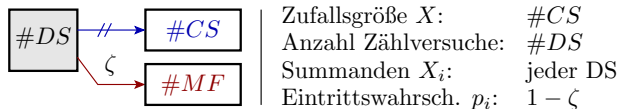
Der Divisor ist um eins kleiner als die Anzahl der Datenwerte $\#v$, d.h. Abschätzungen von Varianz und Standardabweichung erfordern mindestens eine Datenstichprobe $\#v \geq 2$.

$\#v$	Größe der Datenstichprobe.
v_i	Wert i der Datenstichprobe.
$\hat{\cdot}$	Schätzwert.
$\text{Var}[\dots], \sigma^2$	Varianz.



Summen, lineare Transformationen

4.15 Verteilung von Zählwerten



Ein einzelner Zählversuch, im Beispiel, ob eine erbrachte Serviceleistung (DS) eine Fehlfunktion (MF) ist, hat die Bernoulli-Verteilungen

$$\mathbb{P}[X_i = 0] = 1 - \zeta$$

$$\mathbb{P}[X_i = 1] = \zeta$$

Die Anzahl aller Fehlfunktionen ist die Summe der Ergebnisse aller $\#DS$ Einzelversuche:

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

Die Häufigkeit einer Fehlfunktion als Anteil der eingetretenen Fehlfunktionen $\frac{X}{\#DS}$ ist eine lineare Transformation mit $\#DS + 1$ möglichen Werten, von denen nur Werte nahe ζ wahrscheinlich sind.



4.16 Lineare Transformation

Lineare Transformationen sind Multiplikationen und Additionen einer Zufallsvariablen mit reellen Zahlen:

$$Y = a \cdot X + b$$

Die (diskrete) Verteilung ordnet die Wahrscheinlichkeiten der Originalwerte den transformierten Werten zu:

$$\mathbb{P}[y = ax + b] = \mathbb{P}[x] \quad (4.16)$$

Der Erwartungswert wird genau wie jeder andere Wert transformiert:

$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \quad (4.17)$$

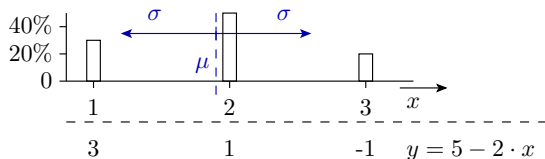
Varianz und Standardabweichung sind verschiebungs- und spiegelungsinvariant, d.h. um Quadrat bzw. Betrag der Skalierung größer:

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X] \quad (4.18)$$

$$\text{sd}[a \cdot X + b] = \sqrt{\text{Var}[a \cdot X + b]} = a \cdot \text{sd}[X] \quad (4.19)$$

Kontrollen durch Nachrechnen folgen später als Übungsaufgaben.

4.17 Veranschaulichung 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

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

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

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

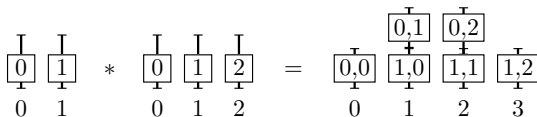
$$\text{Var}[Y] = 0,3 \cdot 3^2 + 0,5 \cdot 1^2 + 0,2 \cdot (-1)^2 - 1,2^2 = 1,96$$

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

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



4.18 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):

v	0	1	μ
$\mathbb{P}[X = v]$	0,4	0,6	0,6

v	0	1	2	μ
$\mathbb{P}[Y = v]$	0,3	0,6	0,1	0,8

$$\mathbb{P}[X + Y = 0] = \mathbb{P}[X = 0] \cdot \mathbb{P}[Y = 0] = 0,4 \cdot 0,3 = 0,12$$

$$\mathbb{P}[X + Y = 1] = \mathbb{P}[X = 0] \cdot \mathbb{P}[Y = 1] + \mathbb{P}[X = 1] \cdot \mathbb{P}[Y = 0] = 0,42$$

$$\mathbb{P}[X + Y = 2] = \mathbb{P}[X = 0] \cdot \mathbb{P}[Y = 2] + \mathbb{P}[X = 1] \cdot \mathbb{P}[Y = 1] = 0,4$$

$$\mathbb{P}[X + Y = 3] = \mathbb{P}[X = 1] \cdot \mathbb{P}[Y = 2] = 0,6 \cdot 0,1 = 0,06$$



4.19 Erwartungswert und Varianz

Für Summen unabhängiger Zufallsvariablen sind Erwartungswert und Varianz gleich der Summe der Erwartungswerte bzw. Varianzen der Summanden:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad (4.20)$$

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \quad (4.21)$$

v	0	1	2	3	μ
$\mathbb{P}[X = v]$	40%	60%			$40\% \cdot 0 + 60\% \cdot 1 = 0,6$
$\mathbb{P}[Y = v]$	30%	60%	10%		$60\% \cdot 1 + 10\% \cdot 2 = 0,8$
$\mathbb{P}[Y + Y = v]$	12%	42%	40%	6%	$42\% \cdot 1 + 40\% \cdot 2 + 6\% \cdot 3 = 1,4$

$$\text{Var}[X] = 40\% \cdot 0 + 60\% \cdot 1^2 - 0,6^2 = 0,24$$

$$\text{Var}[Y] = 60\% \cdot 1 + 10\% \cdot 2^2 - 0,8^2 = 0,36$$

$$\text{Var}[X + Y] = 42\% \cdot 1 + 40\% \cdot 2^2 + 6\% \cdot 3 - 1,4^2 = 0,60$$

X, Y Zufallsvariablen.

$\mathbb{E}[\dots]$ Erwartungswert von ...

$\text{Var}[\dots]$ Varianz von ...



4.20 Abhängigkeiten

Abhängigkeit bedeutet, dass die Eintrittswahrscheinlichkeit eines Summanden X vom zufälligen Ergebnis eines anderen Summanden Y abhängt. Für $X, Y \in \{0, 1\}$ als Zufallsvariablen für den Nachweis von zwei Haftfehlern Fehlern wurden (Abschn. 2.1.5) z.B. folgende mögliche Abhängigkeiten behandelt:

- impliziter Fehlernachweis: wenn $Y = 1$ dann $X = 1$ und
- identischer Fehlernachweis $Y = X$.

Abhängigkeiten haben keinen Einfluss auf den Erwartungswert und erhöhen die Varianz um die doppelte Kovarianz*:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y] \quad (4.22)$$

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \quad (4.23)$$

Wir werden Abhängigkeiten zwischen Zählwerten statt dessen durch eine aus Daten abschätzbare Varianzvergrößerung κ beschreiben (siehe später Abschn. 4.2.1).

$\text{Cov}[X, Y]$ Kovarianz der beiden Zufallsvariablen.

* Kontrolle durch Nachrechnen folgt später als Übungsaufgabe.

4.22 Summen von Zufallsvariablen



Komplexe Zufallsereignisse lassen sich oft als Summe linearer Transformationen einfacher zu untersuchender Ereignisse beschreiben.

Ein zufälliger Zählwert X , z.B. die Anzahl der korrekt ausgeführten Service-Leistungen, ist eine Summe elementarer Zählereignisse X_i :

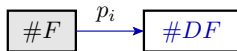
$$X = \sum_{i=1}^n X_i$$

Die Summanden haben die Bernoulli-Verteilungen (vergl. Gl. 4.1):



Verteilung von Zählwerten

4.23 Verteilung von Zählwerten



Ein einzelner Zählversuch, im Beispiel, der Nachweis eines Fehlers (DF) hat die Bernoulli-Verteilung

$$\begin{aligned}\mathbb{P}[X_i = 0] &= 1 - p_i \\ \mathbb{P}[X_i = 1] &= p_i\end{aligned}\tag{4.24}$$

Die Anzahl der Zählwerte, im Beispiel der nachweisbaren Fehler, ist die Summe der Ergebnisse der n Einzelversuche:

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

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



4.24 Erwartungswert und Varianz der Summanden

$$(4.24) \quad \begin{aligned} \mathbb{P}[X_i = 0] &= 1 - p_i \\ \mathbb{P}[X_i = 1] &= p_i \end{aligned}$$

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.

4.25 Erwartungswert und Varianz der Summe

Der Erwartungswert der Summe ist die Summe der Erwartungswerte (Gl. 4.17):

$$\mathbb{E}[X] = \sum_{i=1}^n p_i \quad (4.25)$$

Für unabhängige Zählwerte ist die Varianz der Summe gleich der Summe der Varianzen der Summanden, Kovarianz null, Gl. 4.20):

$$\text{Var}[X] = \sum_{i=1}^n p_i \cdot (1 - p_i) \quad (4.26)$$

Abhängigkeiten werden später durch die Varianzerhöhung κ modelliert
4.2.1.

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



4.26 Berechnung der Verteilung

- Beginne mit der Verteilung von $S_1 = X_1$:

$$\mathbb{P}[S_1 = 0] = 1 - p_1$$

$$\mathbb{P}[S_1 = 1] = p_1$$

- Wiederhole für $i = 1$ bis $n - 1$

Berechne Verteilung von $S_{i+1} = X_{i+1} + S_i$:

$$(1 - p_{i+1}) \cdot (\mathbb{P}[S_i = 0], \mathbb{P}[S_i = 1], \dots, \mathbb{P}[X_i = i - 1], \quad 0 \quad) \\ + p_{i+1} \cdot (\quad 0, \quad \mathbb{P}[S_i = 0], \dots, \mathbb{P}[X_i = i - 2], \mathbb{P}[X_i = i - 1])$$

Beispiel:

	p_i	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\mathbf{P}[S_1 = X_1 = k]$	30%	70%	30%			
$\mathbf{P}[S_2 = X_1 + X_2 = k]$	50%	35%	50%	15%		
$\mathbf{P}[S_3 = X_1 + X_2 + X_3 = k]$	40%	21%	44%	29%	6%	
$\mathbf{P}[S_4 = X_1 + X_2 + X_3 + X_4 = k]$	10%	18,9%	41,7%	30,5%	8,3%	0,6%

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

$\mathbb{P}_i[S_i = k]$ Verteilung der Summe der ersten i Zählerwerte.



4.27 Erwartungswert und Varianz für das Beispiel

i	0	1	2	3	4
p_i	30%	50%	40%	10%	
$\mathbf{P}[S_4 = i]$	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. 4.25 ist die Berechnung kürzer:

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

Die Varianz beträgt nach dem Verschiebungssatz Gl. 4.10:

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

Vereinfachte Berechnung nach Gl. 4.26:

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



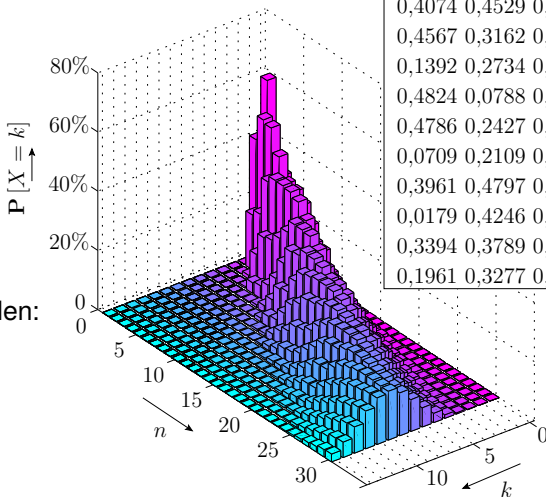
4.28 Beispiel einer Zählverteilung

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

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

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

Wahrscheinlicher Bereich ca. 5 bis 15.





Messfehler

4.30 Gemessener Wert und Messfehler

Die Summenbeziehungen für Zufallsgrößen sind auch für andere Aufgabenstellung aus dem Bereich Test und Verlässlichkeit nützlich.

In der Messtechnik gilt für jeden gemessenen Wert:

$$X_M = X + X_F \quad (4.27)$$

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

$$\mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F] \quad (4.28)$$

$$\text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F] \quad (4.29)$$

- $\mathbb{E}[X_F]$ – 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 4.1: Messfehler

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

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

Welchen Erwartungswert und welche Standardabweichung hat der zu messende Wert?

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?

$$(4.28) \quad \mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F]$$

$$(4.29) \quad \text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F]$$

$$(4.9) \quad \text{sd}[X] = \sigma = \sqrt{\text{Var}[X]}$$

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

$$\begin{aligned} \mathbb{E}[R] &= \mathbb{E}[R_M] - \mathbb{E}[R_F] = 1010 \Omega - 12 \Omega = 998 \Omega \\ \text{Var}[R] &= \text{Var}[R_M] - \text{Var}[R_F] = (11,18 \Omega)^2 - (5 \Omega)^2 = 100 \Omega^2 \\ \text{sd}[R] &= 10 \Omega \end{aligned}$$

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

R Zu messender Widerstandswert.



Zusammenfassung

4.32 Zufallsvariable und Verteilung

Ein Zufallsexperiment ordnet einer Zufallsvariablen zufällig Werte zu.

Die Verteilungsfunktion gibt die Wahrscheinlichkeit an, dass der zugeordnete Wert X maximal x ist:

$$(4.2) \quad F_X(x) = \mathbb{P}[X \leq x]$$

Eine diskrete Zufallsvariable große kann nur abzählbar viele Werte x_i annehmen. Die Zuordnungshäufigkeit beschreibt die Verteilung $\mathbb{P}[X = x_i] = p_i$.

Eine stetige Zufallsgröße hat Intervall $x_{\min} \leq X \leq x_{\max}$ unendlich viele Ausprägungen und statt der Verteilung eine Dichtefunktion:

$$(4.3) \quad f_X(x) = \frac{dF_X(x)}{dx}$$

Ein einzelner Zählversuch hat nur die möglichen Ergebnisse null und eins und eine Bernoulli-Verteilung:

$$(4.1) \quad \begin{aligned} \mathbb{P}[X = 0] &= 1 - p \\ \mathbb{P}[X = 1] &= p \end{aligned}$$

4.33 Kenngrößen von Zufallsvariablen

Erwartungswert:

$$(4.4) \quad \mathbb{E}[X] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i$$

$$(4.5) \quad \mathbb{E}[X] = \mu = \int_{x_{\min}}^{x_{\max}} f_X(x) \cdot x \cdot dx$$

Varianz:

$$(4.6) \quad \text{Var}[X] = \sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$$(4.7) \quad \text{Var}[X] = \sigma^2 = \sum_{i=1}^{\#x} p_i \cdot (x_i - \mathbb{E}[X])^2$$

$$(4.8) \quad \text{Var}[X] = \sigma^2 = \int_{x_{\min}}^{x_{\max}} f_X(x) \cdot (x - \mathbb{E}[X])^2 \cdot dx$$

Berechnung der Varianz nach dem Verschiebesatz:

$$(4.10) \quad \text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$(4.11) \quad \text{Var}[X] = \sum_{i=1}^{\#x} p_i \cdot (x_i)^2 - \mathbb{E}[X]^2$$

4.34 Maße der Schätzgenauigkeit

Standardabweichung als Maß der absoluten Breite des wahrscheinlichen Bereichs und der absoluten Schätzgenauigkeit:

$$(4.9) \quad \text{sd}[X] = \sigma = \sqrt{\text{Var}[X]}$$

Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten als Maße der relativen Bereichsbreite und der relativen Schätzgenauigkeit:

$$(4.12) \quad \begin{array}{ll} \sigma_r &= \frac{\sigma}{\mu} & \text{zweckmäßiges Maß für } \mu \leq n/2 \\ \sigma_{\bar{r}} &= \frac{\sigma}{n-\mu} & \text{zweckmäßiges Maß für } \mu > n/2 \end{array}$$

4.35 Schätzen der Kenngrößen

Wahrscheinlichkeitsmodelle dienen zur Vorhersage künftig zu erwartender Häufigkeiten von Datenmerkmalen. Umgekehrt dienen erhobene Daten zur Kontrolle der Wahrscheinlichkeitsmodelle und zum Schätzen von Kenngrößen der Zufallsvariablen.

Schätzer für den Erwartungswert:

$$(4.13) \quad \hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i$$

Schätzer für die Varianz:

$$(4.14) \quad \hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2$$

Standardabweichung:

$$(4.15) \quad \hat{\text{sđ}}[X] = \hat{\sigma} = \sqrt{\hat{\text{Var}}[X]}$$

4.36 Lineare Transformationen, Summe

- Verteilung, Erwartungswert und Varianz von $Y = a \cdot X + b$:

$$(4.16) \quad \mathbb{P}[y = ax + b] = \mathbb{P}[x]$$

$$(4.17) \quad \mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b$$

$$(4.18) \quad \text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$$

$$(4.19) \quad \text{sd}[a \cdot X + b] = \sqrt{\text{Var}[a \cdot X + b]} = a \cdot \text{sd}[X]$$

- Erwartungswert und Varianz der Summe unabhängiger Zufallsvariablen:

$$(4.20) \quad \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$(4.21) \quad \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

- Abhängigkeiten vergrößern die Varianz um die doppelte Kovarianz. (Wir berücksichtigen Abhängigkeiten später anders):

$$(4.22) \quad \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \cdot \text{Cov}[X, Y]$$

$$(4.23) \quad \text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])]$$

4.37 Unabhängige Zählwerte

Erwartungswert und Varianz:

$$(4.25) \quad \mathbb{E}[X] = \sum_{i=1}^n p_i$$

$$(4.26) \quad \text{Var}[X] = \sum_{i=1}^n p_i \cdot (1 - p_i)$$

Berechnung der Verteilung:

	p_i	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\mathbf{P}[S_1 = X_1 = k]$	30%	70%	30%			
$\mathbf{P}[S_2 = X_1 + X_2 = k]$	50%	35%	50%	15%		
$\mathbf{P}[S_3 = X_1 + X_2 + X_3 = k]$	40%	21%	44%	29%	6%	
$\mathbf{P}[S_4 = X_1 + X_2 + X_3 + X_4 = k]$	10%	18,9%	41,7%	30,5%	8,3%	0,6%

4.38 Messwerte und Messfehler

$$(4.27) \quad X_M = X + X_F$$

$$(4.28) \quad \mathbb{E}[X_M] = \mathbb{E}[X] + \mathbb{E}[X_F]$$

$$(4.29) \quad \text{Var}[X_M] = \text{Var}[X] + \text{Var}[X_F]$$



Näherungen



Binomialverteilung



4.39 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:

$$\mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (4.30)$$

Erwartungswert einer Binomialverteilung:

$$\mathbb{E}[X] = \mu = n \cdot p \quad (4.31)$$

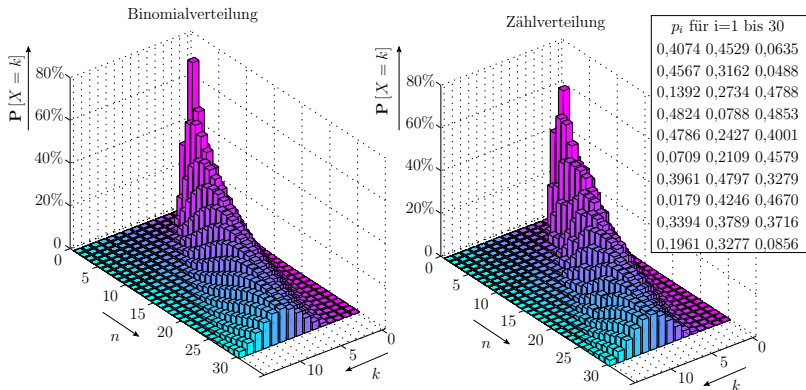
Varianz und Standardabweichung einer Binomialverteilung:

$$\text{Var}[X] = \sigma^2 = n \cdot p \cdot (1-p) \quad (4.32)$$

$$\text{sd}[X] = \sigma = \sqrt{n \cdot p \cdot (1-p)} \quad (4.33)$$

n Anzahl der Zählversuche, maximaler Zählwert.
 p Eintrittswahrscheinlichkeit.

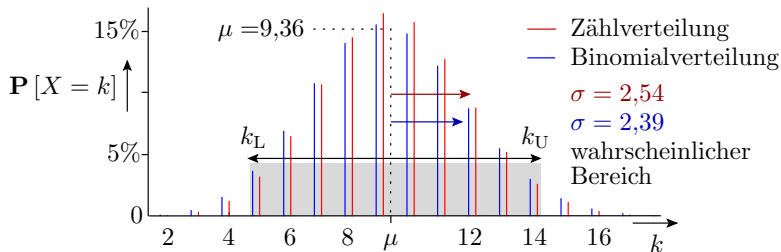
4.40 Binomialverteilung vs. allg. Zählverteilung



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.

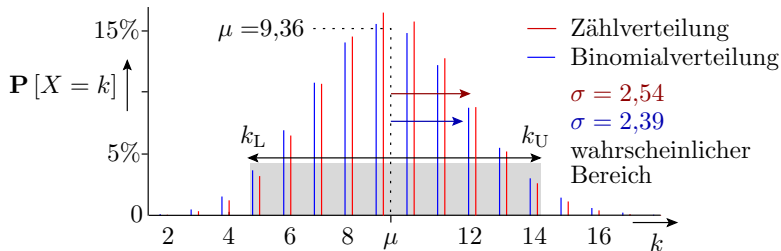
- n Anzahl der Zählversuche, maximaler Zählwert.
- p Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
- p_i Eintrittswahrscheinlichkeit Zählereignis i .

4.41 Bereichsschätzung Binomialvert.-Näherung



Bei gleicher Anzahl unabhängiger Zählversuche n liefert die Annäherung durch eine Binomialverteilung mit $p = \frac{1}{n} \cdot \sum_{i=1}^n p_i$ eine Worst-Case-Abschätzung:

- garantiert eingehaltene Irrtumswahrscheinlichkeiten bzw.
- einen mindestens garantierbaren Bereich.



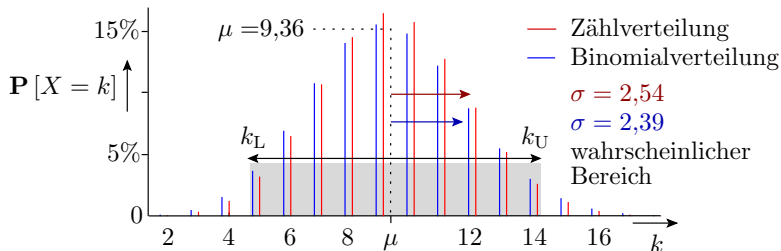
Irrtumswahrscheinlichkeiten für $k_L = 5$, $k_U = 14$:

	$\alpha_1 = \mathbb{P}[X < 5]$	$\alpha_2 = \mathbb{P}[X > 14]$
Zählverteilung	2,2%	1,9%
Binomialverteilung	2,4%	2,4%

Bereich für $\alpha_1 = 2,3\%$ und $\alpha_2 = 2\%$:

Zählverteilung	$k_L = 5$	$k_U = 14$
Binomialverteilung	$k_L = 4$	$k_U = 15$

4.43 Varianzobergrenze 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:

$$\sigma^2 = \sum_{i=1}^n p_i \cdot (1 - p_i) \leq n \cdot p \cdot (1 - p) \quad (4.34)$$

4.44 Beweis durch Nachrechnen

Ersatz der individuellen Eintrittswahrscheinlichkeiten der zu zählenden Ereignisse durch Mittelwert und Differenz zum Mittelwert:

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

Die Varianz einer Zählverteilung:

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

ist um Summe der quadratischen Abweichungen kleiner als die Varianz der Binomialverteilung mit gleichem p und n .

4.45 Schätzer für Zählwerten

Die Ungleichheit in

$$(4.34) \quad \sigma^2 = \sum_{i=1}^n p_i \cdot (1 - p_i) \leq n \cdot p \cdot (1 - p)$$

wird einen Korrekturfaktor versteckt, der später als Mass für Abhängigkeiten dienen wird:

$$\sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \quad (4.35)$$

Schätzer für Erwartungswert, Eintrittswahrscheinlichkeit und Standardabweichung für einen experimentell bestimmten Zählwert x_{AV} :

$$\hat{\mu} = x_{AV} \quad (4.36)$$

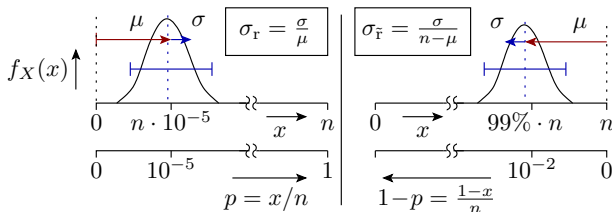
$$\hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{AV}}{n} \quad (4.37)$$

$$\hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)} \quad (4.38)$$

Zählwerte haben typ. Intervallradien von $2\sigma \dots 3\sigma$ (Abschn. 4.2.6).

$\hat{\mu}, \hat{\sigma}$	Schätzwerte für Erwartungswert und Standardabweichung.
\hat{p}	Schätzwert der Eintrittswahrscheinlichkeit.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.

4.46 Varianzkoeffizient als Genauigkeitsmaß



Der Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten ist ein Maße für die relative Breite des wahrscheinlichen Bereichs im Verhältnis zum möglichen Minimum bzw. Maximum und damit ein Maß für die relative Schätzgenauigkeit für:

$$(4.12) \quad \begin{aligned} \sigma_r &= \frac{\sigma}{\mu} && \text{zweckmäßiges Maß für } \mu \leq n/2 \\ \sigma_{\bar{r}} &= \frac{\sigma}{n-\mu} && \text{zweckmäßiges Maß für } \mu > n/2 \end{aligned}$$

$\sigma_r, \sigma_{\bar{r}}$ Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszahlwert.



Mit dem Schätzer für die Standardabweichung

$$(4.38) \quad \hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)}$$

ergeben sich als Schätzer für die Varianzkoeffizienten für die später behandelten Bereichsschätzungen unter Annäherung der Zählwertverteilung durch eine Normalverteilung:

$$\begin{aligned} \hat{\sigma}_r &= \frac{\hat{\sigma}}{x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} - \frac{1}{n}\right)} && \text{für } x_{AV} \leq \frac{n}{2} \\ \hat{\sigma}_{\bar{r}} &= \frac{\hat{\sigma}}{n-x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{n-x_{AV}} - \frac{1}{n}\right)} && \text{für } x_{AV} > \frac{n}{2} \end{aligned} \quad (4.39)$$

Für $x_{AV} \ll n$ bzw. $n - x_{AV} \ll n$ ist der Term $\frac{1}{n}$ vernachlässigbar.

$\sigma_r, \sigma_{\bar{r}}$	Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszählwert.
$\hat{\sigma}$	Geschätzte Standardabweichung des Zählwerts.
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
n	Anzahl der Zählversuche, maximaler Zählwert.



Poisson-Verteilung

4.48 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:

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

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.

λ Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).

Poisson-Verteilung allgemein

$$\mathbb{P}[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad (4.40)$$

und für Zählwerte:

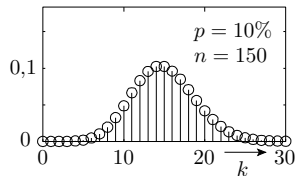
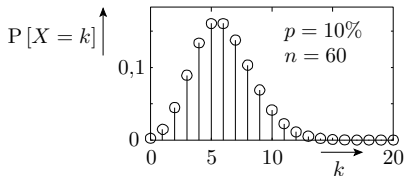
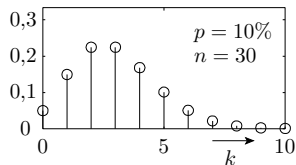
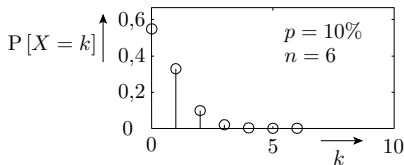
$$\mathbb{P}[X = k] = e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!} \quad (4.41)$$

Die Poisson-Verteilung hat nur einen Parameter, der gleichzeitig Erwartungswert und Varianz und das Produkt aus der Anzahl der Zählversuche und der mittleren Eintrittswahrscheinlichkeit ist.

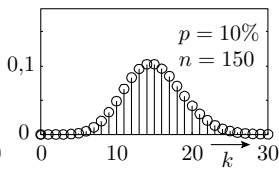
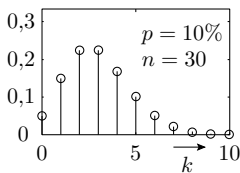
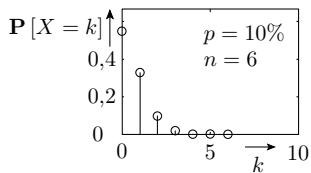
λ	Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).
n	Anzahl der Zählversuche, maximaler Zählwert.
p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

4.50 Anzahl der Zählversuche und Verteilung

$$(4.41) \quad \mathbb{P}[X = k] = e^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!}$$



- $\mathbb{P}[X = k]$ 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:

$$k_U \approx 3 \dots 5 \cdot \lambda$$

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

$$k_L \approx \frac{\lambda}{3 \dots 5}$$

Ab $\lambda = p \cdot n > 10$ zunehmend symmetrischer Bereich um den Erwartungswert (Abschn. 4.2.6 *Bereichsschätzung NVT*):

$$sr = [k_U, k_L] \approx \lambda \mp 2 \dots 3 \cdot \sqrt{\lambda}$$

k_L, k_U	Untere (lower) und obere (upper) Bereichsgrenze des Zählwerts.
sr	Symmetrischer Bereich der wahrscheinlichen Werte.
n	Anzahl der Zählversuche, maximaler Zählwert.
p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

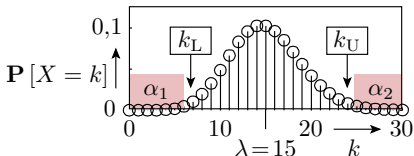


Bereichsschätzung Pois

4.52 Schätzer $k_L \xleftrightarrow{(\alpha_1)} \lambda_U$

Vorgabe k_L und α_1 . Numerische Suche λ_U , so dass

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



λ_U	$k_L=1$	$k_L=2$	$k_L=3$	$k_L=4$	$k_L=5$	$k_L=6$
$\alpha_1=0,5\%$	$\lambda = 5,298$	7,430	9,273	10,978	12,593	14,150
$\alpha_1=1\%$	$\lambda = 4,606$	6,638	8,406	10,045	11,605	13,109
$\alpha_1=2\%$	$\lambda = 3,912$	5,834	7,516	9,084	10,580	12,027
$\alpha_1=10\%$	$\lambda = 2,303$	3,890	5,323	6,681	7,993	9,275
$\alpha_1=20\%$	$\lambda = 1,609$	2,995	4,279	5,514	6,721	7,906

Beispielabschätzungen:

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

α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

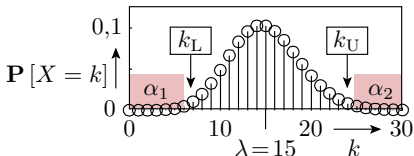
k_L, k_U Untere (lower) und obere (upper) Bereichsgrenze des Zählwerts.

λ_U Obere (upper) Bereichsgrenze des Erwartungswerts.

4.53 Schätzer $k_U \overset{(\alpha_2)}{\longleftrightarrow} \lambda_L$

Vorgabe k_U und α_2 . Numerische Suche λ_L , so dass

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



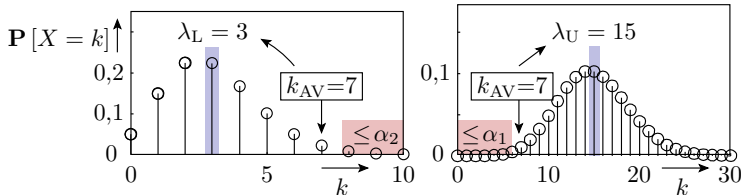
λ_L	$k_U = 0$	$k_U = 1$	$k_U = 2$	$k_U = 3$	$k_U = 4$	$k_U = 5$	$k_U = 6$
$\alpha_2 = 0,5\%$	0,005	0,103	0,338	0,672	1,078	1,537	2,037
$\alpha_2 = 1\%$	0,01	0,148	0,436	0,823	1,279	1,785	2,330
$\alpha_2 = 2\%$	0,02	0,215	0,567	1,016	1,529	2,089	2,684
$\alpha_2 = 10\%$	0,105	0,532	1,102	1,744	2,432	3,152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

Beispielabschätzungen:

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

α_2	Irrtumswahrscheinlichkeit, Werte oberhalb des geschätzten Bereichs.
k_L, k_U	Untere (lower) und obere (upper) Bereichsgrenze des Zählwerts.
λ_L	Untere (lower) Bereichsgrenze des Erwartungswerts.

4.54 Schätzer $k_{AV} \xrightarrow{(\alpha)} [\lambda_L, \lambda_U]$



$[\lambda_L, \lambda_U] = f(\alpha, k_{AV})$ für die Ist-Zählwerte $k_{AV} = 1$ bis 3:

$[\lambda_L, \lambda_U]$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$\alpha_1 = \alpha_2 = 0,5\%$	[0,10, 5,30]	[0,34, 7,43]	[0,67, 9,27]
$\alpha_1 = \alpha_2 = 1\%$	[0,15, 4,60]	[0,44, 6,64]	[0,82, 8,41]
$\alpha_1 = \alpha_2 = 2\%$	[0,22, 3,91]	[0,57, 5,83]	[1,02, 7,52]
$\alpha_1 = \alpha_2 = 10\%$	[0,53, 2,30]	[1,10, 3,89]	[1,74, 5,32]
$\alpha_1 = \alpha_2 = 20\%$	[0,82, 1,61]	[1,53, 2,99]	[2,30, 4,28]

Die Werte wurden aus den Tabellen davor übernommen.

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

$[\lambda_L, \lambda_U]$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$\alpha_1 = \alpha_2 = 0,5\%$	[1,08, 11,0]	[1,54, 12,6]	[2,04, 14,2]
$\alpha_1 = \alpha_2 = 1\%$	[1,28, 10,0]	[1,79, 11,6]	[2,33, 13,1]
$\alpha_1 = \alpha_2 = 2\%$	[1,53, 9,08]	[2,09, 10,6]	[2,68, 12,0]
$\alpha_1 = \alpha_2 = 10\%$	[2,43, 6,68]	[3,15, 7,99]	[3,89, 9,28]
$\alpha_1 = \alpha_2 = 20\%$	[3,09, 5,51]	[3,90, 6,73]	[4,73, 7,91]

k_{AV}	Ist-Zählwert (Actual count value).
α_1, α_2	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
λ_L, λ_U	Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

4.55 Aktueller Zählwert $k_{AV} = 0$

- minimaler Erwartungswert:

$$\lambda_L = 0$$

- Der maximale Erwartungswert ergibt sich über:

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

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

α_1	0,5%	1%	2%	10%	20%
λ_U	5,30	4,61	3,91	2,30	1,61%

k_{AV}	Ist-Zählwert (Actual count value).
α_1, α_2	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
λ_L, λ_U	Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

Beispiel 4.2: Poissonverteilte Anzahl der 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?*

α_1, α_2	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
k_{AV}	Ist-Zählwert (Actual count value).
λ_L, λ_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 $\alpha_1 \leq 1\%$?*

α_1	0,5%	1%	2%	10%	20%
λ_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.

-
- α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
 - k_{AV} Ist-Zählwert (Actual count value).
 - λ_L, λ_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?

$[\lambda_L, \lambda_U]$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$\alpha_1 = \alpha_2 = 0,5\%$	[1,08, 11,0]	[1,54, 12,6]	[2,04, 14,2]
$\alpha_1 = \alpha_2 = 1\%$	[1,28, 10,0]	[1,79, 11,6]	[2,33, 13,1]
$\alpha_1 = \alpha_2 = 2\%$	[1,53, 9,08]	[2,09, 10,6]	[2,68, 12,0]
$\alpha_1 = \alpha_2 = 10\%$	[2,43, 6,68]	[3,15, 7,99]	[3,89, 9,28]
$\alpha_1 = \alpha_2 = 20\%$	[3,09, 5,51]	[3,90, 6,73]	[4,73, 7,91]

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

- α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
 k_{AV} Ist-Zählwert (Actual count value).
 λ_L, λ_U Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

Beispiel 4.3: 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.
α_1, α_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?

$[\lambda_L, \lambda_U]$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$\alpha_1 = \alpha_2 = 1\%$	[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.

λ_L, λ_U

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

ζ_L, ζ_U

Untere und obere Bereichsgrenze der geschätzten Fehlfunktionsrate.

Beispiel 4.4: 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).
α_1, α_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

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

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

Abschätzbarer Bereich der Maskierungswahrscheinlichkeit:

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

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

α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

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

p_{ML}, p_{MU} Untere und obere Bereichsgrenze der geschätzten Maskierungswahrscheinlichkeit.

Beispiel 4.5: 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.
k_{AV}	Ist-Zählwert (Actual count value).
α_1, α_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?

$[\lambda_L, \lambda_U]$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$\alpha_1 = \alpha_2 = 10\%$	[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:

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

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

ζ_L, ζ_U
 R_L, R_U

Untere und obere Bereichsgrenze der geschätzten Fehlfunktionsrate.

Untere und obere Bereichsgrenze der geschätzten Zuverlässigkeit.



Defektanteil

4.60 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 einer geringen gleichbleibenden Fehlerentstehungsrate ist die Fehleranzahl in einem Produkt poissonverteilt:

$$\mathbb{P}[X = k] = e^{-\mu_F} \cdot \frac{\mu_F^k}{k!} \quad (4.42)$$

k – Fehleranzahl. Der zu erwartende Defektanteil ist gleich der Wahrscheinlichkeit, dass die Fehleranzahl $k > 0$ ist:

$$\mu_{DL} = 1 - e^{-\mu_F} \quad (4.43)$$

Für eine geringe zu erwartende Fehleranzahl $\mu_F \ll 1$:

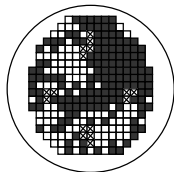
$$\mu_{DL} = 1 - \left(1 + (-\mu_F) + \frac{(-\mu_F)^2}{2!} + \dots \right)^{(\mu_F \ll 1)^*} \mu_F \quad (4.44)$$

μ_F Zu erwartende Fehleranzahl.

μ_{DL} Zu erwartender Defektanteil.

*

4.61 Prozessschwankungen, Fehlercluster



Fehlercluster auf einem Schaltkreiswafer [FSB87]

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

Örtliche und zeitliche Schwankungen der Fehlerentstehungsrate verursachen Fehlercluster (Fehlerhäufungen). Beispiele:

- Cluster von Schreibfehler in Texten,
- Fehlerhäufungen in Programmteilen,
- qualitativ niederwertige »Montagsprodukte«, ...

Fehlercluster

- bleiben nach Test und Fehlerbeseitigung erhalten,
- mindern den Fehleranteil bei gleicher Fehleranzahl,
- liefern Hinweise zur Verbesserung der Entstehungsprozesse.

4.62 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. 4.43 mit $\mu_F = FC \cdot \mu_{CF}$:

$$\mu_Y = e^{-FC \cdot \mu_{CF}} \quad (4.45)$$

Für jeden als fehlerfrei befundenen Schaltkreis müssen im Mittel

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

Schaltkreise gefertigt werden.

μ_Y	Zu erwartende Ausbeute.
FC	Fehlerabdeckung (fault coverage), Anteil der nachweisbaren Fehler.
μ_{CF}	Zu erwartende Anzahl der Fehler aus den Entstehungsprozessen.

4.63 Fertigungskosten und Fehleranzahl

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

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

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

(vergl. Gl. 2.48). Kosten je nicht als defekt aussortierter Schaltkreis:

$$C_{IC} = \frac{C_{MIC}}{\mu_Y} = C_{Tr} \cdot \#Tr \cdot e^{FC \cdot \xi_{Tr} \cdot \#Tr} \quad (4.46)$$

μ_{CF}	Zu erwartende Anzahl der Fehler aus den Entstehungsprozessen.
ξ_{Tr}	Fehlerentstehungsrate 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.

Aussprache: ξ xi.

Beispiel 4.6: 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], Fehlerabdeckung $FC = 1$.

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

ξ_{Tr}	Fehlerentstehungsrate in Fehlern je Transistor.
C_{Tr}	Transistorkosten in Euro pro Transistor.
$\#Tr$	Anzahl der Transistoren.
FC	Fehlerabdeckung (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], Fehlerabdeckung $FC = 1$.

Welche Kosten entfallen auf jeden als gut befundenen (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 je Schaltkreis wird eine Fehlerbeseitigung durch Aussortieren teuer. Alternative:

- redundante Funktionsblöcke zum Ersatz fehlerhafter Blöcke,
- Nutzung ohne fehlerhafte Blöcken, z.B. mit weniger Cache, ...

μ_{CF} Zu erwartende Anzahl der Fehler je gefertigter Schaltkreis.

C_{MIC} Zu erwartende Kosten je gefertigter Schaltkreis.

μ_Y Zu erwartende Ausbeute.

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



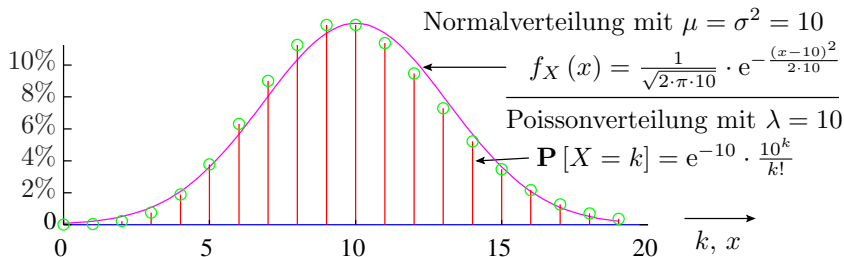
Normalverteilung

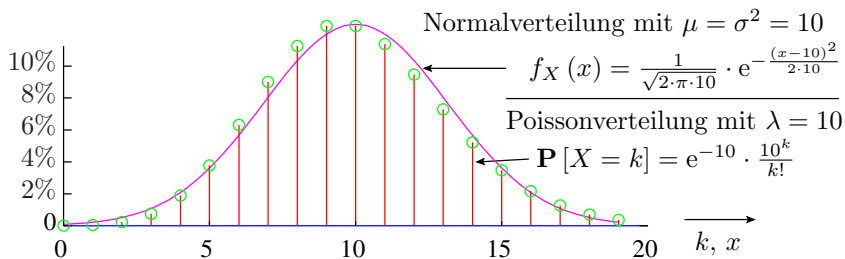
4.65 Normalverteilung

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

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \quad \text{mit } \sigma = \text{sd}[X], \mu = \mathbb{E}[X] \quad (4.47)$$

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





Bereichsschätzungen für Zählwerte über Normalverteilung sind etwa ab $\sigma \leq \mu/3$ brauchbar. Für experimentell bestimmte Zählwerte x_{AV} mit

$$(4.38) \quad \hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)}$$

müssen die Zählwerte dafür etwa in folgendem Bereich liegen:

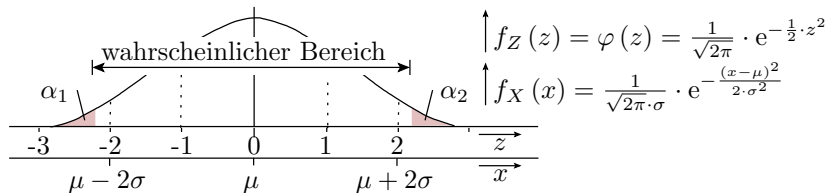
$$10 \cdot \sqrt{\kappa} \leq x_{AV} \leq (n - 10) \cdot \sqrt{\kappa} \quad (4.48)$$

$\hat{\mu}, \hat{\sigma}$	Schätzwerte für Erwartungswert und Standardabweichung.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
n	Anzahl der Zählversuche, maximaler Zählwert.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.



Bereichsschätzung NVT

4.67 Bereichsschätzung mit Normalverteilung



- Transformation einer Zufallsvariablen X mit Erwartungswert μ und Standardabweichung σ in eine Zufallsvariable Z mit Erwartungswert null und Standardabweichung eins:

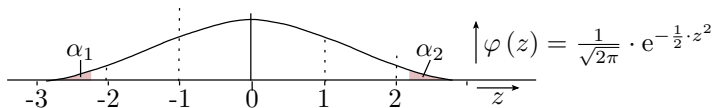
$$Z = \frac{X - \mu}{\sigma} \quad (4.49)$$

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

$$z = \frac{x - \mu}{\sigma} \quad (4.50)$$

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

4.68 Standardisierte Normalverteilung



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

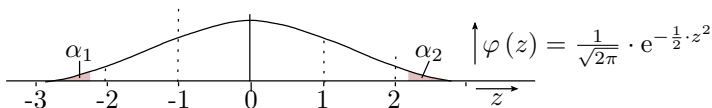
$$\Phi(z) = \int_{-\infty}^z \varphi(u) \cdot du \quad (4.51)$$

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 der Symmetrie gilt für $z < 0$:

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

4.69 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	3,9	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 $\sigma = 1$.
- x, z Werte der Zufallsvariablen X und Z .

4.70 Wahrscheinlichkeit 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

Für eine normalverteilte Zufallsvariablen X mit Erwartungswert μ , Standardabweichung σ und den Bereichsgrenzen $[x_L, x_U]$:

1 Transformation der Bereichsgrenze nach (Gl. 4.50)

$$z_L = \frac{x_L - \mu}{\sigma} \quad (4.52)$$

$$z_U = \frac{x_U - \mu}{\sigma} \quad (4.53)$$

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

3 Bestimmung der Irrtumswahrscheinlichkeiten:

$$\alpha_1 = \Phi(z_L) = 1 - \Phi(-z_L) = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right) \quad (4.54)$$

$$\alpha_2 = 1 - \Phi(z_U) = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right) \quad (4.55)$$



4.71 Wahrscheinlicher Bereich

$\alpha_{1/2}$	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

Für eine normalverteilte Zufallsvariablen X mit Erwartungswert μ , Standardabweichung σ und zugelassenen Irrtumswahrscheinlichkeiten α_1, α_2 :

1 Ablesen aus der Tabelle

$$z_L = \Phi^{-1}(\alpha_1) = -\Phi^{-1}(1 - \alpha_1) \quad (4.56)$$

$$z_U = \Phi^{-1}(1 - \alpha_2) \quad (4.57)$$

2 Transformation:

$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1) \quad (4.58)$$

$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2) \quad (4.59)$$

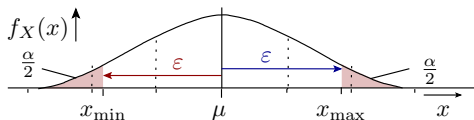
x_L, x_U Untere und obere Schranke des wahrscheinlichen Bereichs von X .

z_L, z_U Transformierte untere und obere Schranke.

μ, σ Erwartungswert, Standardabweichung.

α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

4.72 Symmetrischer Bereich



α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\varepsilon = \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (4.60)$$

$$\text{sr} = [x_L, x_U] = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (4.61)$$

$$\text{sr} = [x_L, x_U] = \mu \cdot \left(1 \mp \sigma_r \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \quad (4.62)$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen X mit den möglichen Zählwerten x .
μ, x_L, x_U	Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
ε, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
sr	Symmetrischer Bereich der wahrscheinlichen Werte.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
σ_r	Varianzkoeffizient, relative Standardabweichung zum erwarteten Eintrittszahlwert.



Beispiel 4.7: 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?*

μ, σ	Erwartungswert, Standardabweichung.
α_1, α_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 $X \geq 30$?

$$(4.55) \quad \alpha_2 = 1 - \Phi(z_U) = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right)$$

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 $X \leq 15$?*

$$(4.54) \quad \alpha_1 = \Phi(z_L) = 1 - \Phi(-z_L) = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right)$$

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?

$$(4.59) \quad x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2)$$

$\alpha_{1/2}$	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\mp \Phi^{-1}(1 - \alpha_{1/2})$	2	3	3,9	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?

$$(4.58) \quad x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1)$$

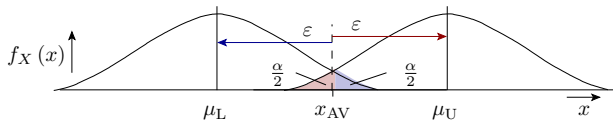
$\alpha_{1/2}$	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\mp \Phi^{-1}(1 - \alpha_{1/2})$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\begin{aligned}x_L &= 20 - 5 \cdot \Phi^{-1}(1 - 2\%) \\ &= 20 - 5 \cdot 2,05 \\ x_L &= 9,75\end{aligned}$$

4.74 Bereichsschätzung für den Erwartungswert

Der Erwartungswert μ zu einem beobachteten Zählwert x_{AV} ist

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

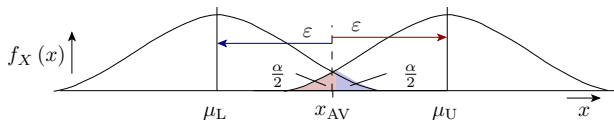


Intervallradius zwischen Erwartungswert und wahrscheinlichen Werten:

$$(4.60) \quad \varepsilon = \sigma \cdot \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen X mit den möglichen Zählwerten x .
μ_L, μ_U	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
ε	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
σ	Standardabweichung.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.

4.75 Symmetrischer Bereich



Symmetrischer Bereich mit x_{AV} als Schätzwert:

$$s_{\Gamma\mu} = [\mu_L, \mu_U] = x_{AV} \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (4.63)$$

$$s_{\Gamma\mu} = x_{AV} \cdot \left(1 \mp \sigma_r \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \quad (4.64)$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen X mit den möglichen Zählwerten x .
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
μ_L, μ_U	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
$s_{\Gamma\mu}$	Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.
σ	Standardabweichung.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
σ_r	Varianzkoeffizient, relative Standardabweichung zum erwarteten Eintrittszählwert.

Beispiel 4.8: 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.
σ	Standardabweichung.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.



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

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

$$(4.63) \quad sr_{\mu} = [\mu_L, \mu_U] = x_{AV} \mp \sigma \cdot \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	2	3	3,9	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_{μ} Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.

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



Schätzen von Zählwerten

4.77 Zu erwartender Zählwertbereich ($p \cdot n \rightarrow sr$)

Mit dem Erwartungswert und der Standardabweichung für Zählwerte:

$$(4.31) \quad \mathbb{E}[X] = \mu = n \cdot p$$

$$(4.35) \quad \sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)}$$

(siehe Binomialverteilungsnaherung) und

$$(4.58) \quad x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1)$$

$$(4.59) \quad x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2)$$

$$(4.61) \quad sr = [x_L, x_U] = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

ergeben sich folgende Bereichsgrenzen:

$$x_L = p \cdot n - \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}(1 - \alpha_1) \quad (4.65)$$

$$x_U = p \cdot n + \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}(1 - \alpha_2) \quad (4.66)$$

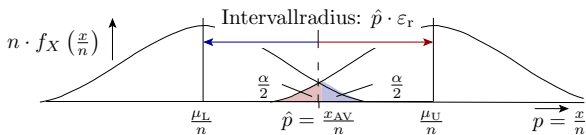
$$sr = [x_L, x_U] = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1 - p)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (4.67)$$

p Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

n Anzahl der Zählversuche, maximaler Zählwert.

μ, x_L, x_U Erwartungswert, untere und obere Schranke des wahrscheinlichen Bereichs.

4.78 Bereich der Eintrittswahrscheinlichkeit



Schätzer der Eintrittswahrscheinlichkeit für n Zählversuche:

$$(4.37) \quad \hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{AV}}{n}$$

Varianzkoeffizient des Eintritts- bzw. Nichteintrittszählwerts:

$$(4.39) \quad \begin{aligned} \hat{\sigma}_r &= \frac{\hat{\sigma}}{x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} - \frac{1}{n} \right)} && \text{für } x_{AV} \leq \frac{n}{2} \\ \hat{\sigma}_{\bar{r}} &= \frac{\hat{\sigma}}{n - x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} - \frac{1}{n} \right)} && \text{für } x_{AV} > \frac{n}{2} \end{aligned}$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen X .
μ_L, μ_U	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
$\sigma_r, \sigma_{\bar{r}}$	Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszählwert.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.

Relative Intervallradien:

$$\varepsilon_r = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} \leq \frac{n}{2} \quad (4.68)$$

$$\varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} > \frac{n}{2} \quad (4.69)$$

Symmetrischer Bereich:

$$sr_p = [p_L, p_U] = \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) \quad \text{für } x_{AV} \leq \frac{n}{2} \quad (4.70)$$

$$sr_p = [p_L, p_U] = 1 - \left(1 - \frac{x_{AV}}{n} \right) \cdot (1 \mp \varepsilon_{\bar{r}}) \quad \text{für } x_{AV} > \frac{n}{2} \quad (4.71)$$

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

4.80 Benötigte Größenordnung der Zählwerte

$$(4.68) \quad \varepsilon_r = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} \leq \frac{n}{2}$$

$$(4.69) \quad \varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} > \frac{n}{2}$$

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

$$x_{AV} \geq \frac{\kappa \cdot \left(\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right)^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) \quad \text{für } \hat{p} \leq 50\% \quad (4.72)$$

$$n - x_{AV} \geq \frac{\kappa \cdot \left(\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right)^2}{\varepsilon_{\bar{r}}^2} \cdot \hat{p} \quad \text{für } \hat{p} > 50\% \quad (4.73)$$

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.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
\hat{p}	Schätzwert der Eintrittswahrscheinlichkeit.

Beispiel 4.9: Geeignete Zählwertgröße

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

ε_r	Intervallradius relativ zum erwarteten Eintritts-Zählwert.
$\varepsilon_{\bar{r}}$	Intervallradius relativ zum erwarteten Nichteintritts-Zählwerts.
\hat{p}	Schätzwert der Eintrittswahrscheinlichkeit.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
κ	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.



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

$$(4.72) \quad x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) \text{ für } \hat{p} \leq 50\%$$

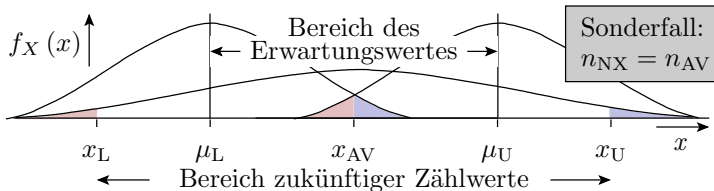
$$(4.73) \quad n - x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot \hat{p} \text{ für } \hat{p} > 50\%$$

ε_r	$p = 10\%$		$p = 50\%$		$\varepsilon_{\bar{r}}$	$p = 90\%$	
	$x_{AV.min}$	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

Brauchbare Schätzungen verlangen eine große (Nicht-) Eintrittszahl $\gtrsim 100$ und insbesondere für sehr kleine (Nicht-) Eintrittswahrscheinlichkeiten eine sehr große Versuchszahl.

- n_{min} Mindestanzahl der Zählversuche.
- $x_{AV.min}$ Minimal erforderliches Zählergebnis.
- $x_{AV.max}$ Maximal zulässiges Zählergebnis.

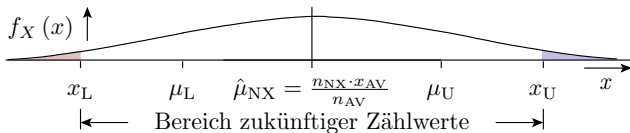
4.83 Bereich künftiger Zählwerte ($x_{AV} \rightarrow \text{sr}_{NX}$)



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.
μ_L, μ_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}	$p \cdot n_{NX}$	$\kappa \cdot n_{NX} \cdot p \cdot (1 - p)$

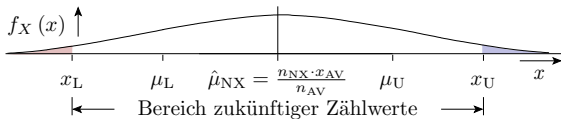
Die Abweichung der künftigen Werte von μ_{NX} :

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

Varianz der Differenz nach (Gl. 4.18):

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

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

$$(4.67) \quad \text{sr} = [x_L, x_U] = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Symmetrischer Bereich der künftigen Zählwerte:

$$\text{sr}_{\text{NX}} = [x_L, x_U] = \hat{\mu}_{\text{NX}} \mp \sqrt{\kappa \cdot n_{\text{NX}} \cdot \left(\frac{n_{\text{NX}}}{n_{\text{AV}}} + 1\right) \cdot \hat{p} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

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

(4.74)

$f_X(x)$	Dichtefunktion der Zufallsvariablen X .
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
μ_L, μ_U	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
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.
sr_{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 4.10: Bereich MF-Rate und künftige MF-Anzahl

Bei der Abarbeitung von 20.000 Service-Anforderungen wurden 100 Fehlfunktionen 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 Fehlfunktionen 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 Fehlfunktionen.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
ε_r	Relativer Intervallradius der Fehlfunktionsrate.
n_{NX}	Anzahl der Zählversuche zur Bestimmung eines künftigen Zählwerts.



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

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

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

$$(4.68) \quad \varepsilon_r = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} \leq \frac{n}{2}$$

$$(4.70) \quad sr_p = [p_L, p_U] = \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) \quad \text{für } x_{AV} \leq \frac{n}{2}$$

α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

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

$$\varepsilon_r = \Phi^{-1} (1 - 1\%) \cdot \sqrt{\frac{1}{100} - \frac{1}{20.000}} = 0,232$$

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

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

\hat{p} , sr_p Geschätzte Wahrscheinlichkeit, symmetrischer Bereich der Wahrscheinlichkeit.

$\hat{\zeta}$, sr_{ζ} Geschätzte Fehlfunktionsrate, symmetrischer Bereich der Fehlfunktionsrate.

ε_r Intervallradius relativ zum erwarteten Eintritts-Zählwert.



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

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

$$(4.72) \quad x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) \text{ für } \hat{p} \leq 50\%$$

Mindestanzahl der gezählten Fehlfunktionen:

$$x_{AV} \geq \frac{2,33^2}{0,1^2} \cdot (1 - 0,5\%) = 540 \text{ [MF]}$$

Diese Zählergebnis wird erreicht nach etwa

$$n_{AV} = \frac{x_{AV}}{\zeta} = \frac{540 \text{ [MF]}}{0,5\% \left[\frac{\text{MF}}{\text{DS}} \right]} = 108.000 \text{ [DS]}$$

[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?*

$$(4.74) \quad sr_{NX} = [x_L, x_U] = \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1 - \hat{p})} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

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

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

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

$$= 50 \mp 20,1 \text{ [MF]}$$

Relativer Intervallradius:

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

$\hat{\mu}_{NX}$ Erwartungswert des zu schätzenden Zählwerts.
 sr_{NX} Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.



Varianzerhöhung

4.87 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:

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

* Varianz von $\#X$ unabhängigen Zählwerten mit paarweise gleichem p_i . 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.

4.89 Varianzerhöhung

Die Varianzerhöhung sei definiert als Verhältnis aus tatsächlicher Varianz und der Varianz einer Binomialverteilung mit derselben Versuchszahl n und derselben mittleren Eintrittswahrscheinlichkeit p entsprechend Gl. 4.35:

$$\kappa = \frac{\text{Var}[X]}{n \cdot p \cdot (1-p)} = \frac{\sigma^2}{\mathbb{E}[X] \cdot (1-p)} \quad (4.75)$$

Die Varianz der Binomialverteilung ist eine einfach abschätzbare Obergrenze für die Varianz unabhängiger Zählwerte. Im Beispiel »paarweise identisch nachweisbare Fehler« ist die Varianzerhöhung bei übereinstimmenden Eintrittswahrscheinlichkeiten p_i $\kappa = 2$ und bei stark abweichende p_i etwas kleiner.

Analog ergibt sich für »immer #IDC identische Zählereignisse« und sonst unabhängige Varianerhöhung:

$$\kappa \leq \#IDC$$

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

4.90 Schätzen der Varianzerhöhung

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

$$(4.13) \quad \hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i$$

- Schätzen der Varianz der Zählwertstichprobe:

$$(4.14) \quad \hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2$$

- Schätzen der Eintrittswahrscheinlichkeit:

$$(4.37) \quad \hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{AV}}{n}$$

- Geschätzte Varianzerhöhung nach Gl. 4.75:

$$\hat{\kappa} = \frac{\hat{\text{Var}}[X]}{\hat{\mathbb{E}}[X] \cdot (1-\hat{p})} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot (1-\hat{p})} \quad (4.76)$$

$\hat{\mathbb{E}}[X]$	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.



Einführung

Beispiel 4.11: Varianzerhöhung

$v = 10$ Wiederholungen »Zählwertbestimmung für $n = 1.000$

Zählversuche:

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.



Einführung

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 $\hat{\kappa}$?

$$(4.13) \quad \hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i$$

$$(4.14) \quad \hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2$$

$$(4.76) \quad \hat{\kappa} = \frac{\hat{\text{Var}}[X]}{\hat{\mathbb{E}}[X] \cdot (1 - \hat{p})} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot (1 - \hat{p})}$$

$$\hat{\mu} = \frac{1}{10} \cdot \sum_{i=1}^{10} v_i = 63,1; \quad \hat{p} = \frac{\hat{\mu}}{n} = 6,3\%$$

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

$$\hat{\kappa} = \frac{135}{63,1 \cdot (1 - 6,3\%)} = 2,28$$

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



Beispiel 4.12: 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.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
[D]	Zählwert in Schadensfällen.



$x_{AV} = 100$ [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?

$$(4.74) \quad sr_{NX} = [x_L, x_U] = \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

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

α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

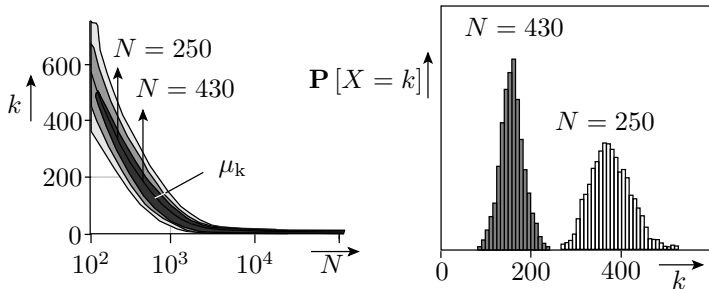
Für $n_{NX} = n_{AV}$ und $n_{AV} \gg x_{AV}$ ist $\mu_{NX} = x_{AV}$:

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

sr_{NX} Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.

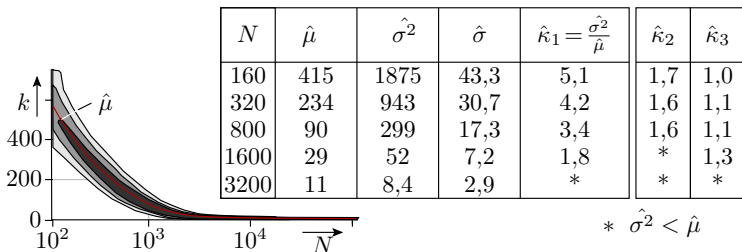
4.93 Experiment mit Haftfehlern

Kombinatorische Beispielschaltung (Benchmark c3540), simuliert mit $n = 3606$ unterschiedlich nachweisbaren Haftfehlern. 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 Anzahl der Modellfehler.
- N Anzahl der Tests.

4.94 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.

4.95 Fehlermodellierung und Vorhersagbarkeit

Für den Fehlernachweis werden Wahrscheinlichkeit $p > 50\%$ angestrebt. Relativer Intervallradius:

$$(4.69) \quad \varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} > \frac{n}{2}$$

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

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



4.96 Schlussfolgerungen

- Bei zufälliger Testauswahl hilft eine zu große Modellfehleranzahl im Verhältnis zur Testobjektgröße nicht, die Schätzgenauigkeit für die Fehlerabdeckung zu verbessern.
- Fehlermodelle, bei denen die Anzahl der Modellfehler überproportional mit der Testobjektgröße zunimmt, z.B. Kurzschlüsse und Pfadverzögerungsfehler, sind nicht zielführend (siehe später Abschn. 6.1.2 ff.).

4.97 Abschätzung mit mehreren Testsätzen

Alternative zur Verringerung des relativen Intervallradius

$$(4.69) \quad \varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n} \right]^* \right)} \quad \text{für } x_{AV} > \frac{n}{2}$$

ist eine Bestimmung der Anzahl der nicht nachweisbaren Modellfehler x_{AV} für $\#v > 1$ unterschiedliche Zufallstestsätze wie im Experiment (Folie 4.94). Vergrößert die Anzahl der Zählversuche n und Zählwerte x_{AV} um $\#v$

$$\varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{\#v \cdot (n - x_{AV})} - \frac{1}{\#v \cdot n} \right)} \quad (4.77)$$

ohne, dass die Abhängigkeiten zwischen den Zählwerten und mit ihnen die Varianzerhöhung κ in ähnlichem Maße zunehmen.

$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.
$\varepsilon_{\bar{r}}$	Intervallradius reaktiv zum erwarteten Nichteintritts-Zählwerts.
$\Phi^{-1}(\cdot)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
α, κ	Irrtumswahrscheinlichkeit, Varianzerhöhung durch Abhängigkeiten.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.



Beispiel 4.13: Bereich der Modellfehlerabdeckung

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

$$n = 1.000 \text{ [F]}, x_{AV} = 1.000 - 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

- a) *In welchem Bereich liegt die Modellfehlerabdeckung 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	Anzahl der Modellfehler.
[F]	Zählwert in Modellfehler.
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.



$$n = 1.000 \text{ [F]}, x_{AV} = 1.000 - 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

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

$$(4.77) \quad \varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{\#v \cdot (n - x_{AV})} - \frac{1}{\#v \cdot n} \right)}$$

$$(4.71) \quad sr_p = [p_L, p_U] = 1 - \left(1 - \frac{x_{AV}}{n} \right) \cdot (1 \mp \varepsilon_{\bar{r}}) \text{ für } x_{AV} > \frac{n}{2}$$

α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\varepsilon_{\bar{r}} = 2,33 \cdot \sqrt{2 \cdot \left(\frac{1}{32} - \frac{1}{1000} \right)} = 58\%$$

$$sr_p = 1 - \left(1 - \frac{1}{32} \right) \cdot (1 \mp 58\%) = [94,9\%, 98,7\%]$$

$\varepsilon_{\bar{r}}$ Relativer Intervallradius der Anzahl der nichtnachweisbaren Modellfehler.
 sr_p Geschätzter symmetrischer Bereich der Fehlerüberdeckung.



$$n = 1.000 \text{ [F]}, x_{AV} = 1.000 - 32 \text{ [F]}, \kappa \leq 2, \alpha = 2\%$$

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

$$(4.77) \quad \varepsilon_{\bar{r}} = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{\#v \cdot (n - x_{AV})} - \frac{1}{\#v \cdot n} \right)}$$

Umstellung nach $\#v$:

$$\begin{aligned} \#v &\geq \kappa \cdot \left(\frac{1}{(n - x_{AV})} - \frac{1}{n} \right) \cdot \left(\frac{(\Phi^{-1} (1 - \frac{\alpha}{2}))}{\varepsilon_{\bar{r}}} \right)^2 \\ &= 2 \cdot \left(\frac{1}{32} - \frac{1}{1000} \right) \cdot \left(\frac{2,33}{0,1} \right)^2 = 35 \end{aligned}$$

Die Modellfehlerabdeckung muss für mindestens 35 unterschiedliche Zufallstestsätze bestimmt und gemittelt werden.



Zusammenfassung

4.99 Näherungen für Zählwertverteilungen

Für unabhängige Zählwerte lassen sich aus den Eintrittswahrscheinlichkeiten p_i aller Zählversuche, Erwartungswert, Varianz und Verteilung berechnen. Es sind aber in der Regel weder die einzelnen p_i bekannt noch kann für Unabhängigkeit garantiert werden. Dann hilft die Annäherung durch bekannte Verteilungen:

- Binomialverteilung: Sonderfall, wenn alle p_i gleich sind.
- Poissonverteilung: Sonderfall seltene Ereignisse, d.h. Zählwerte 0 bis etwa 10 und sehr viel Zählversuche.
- Normalverteilung: Näherung für mindestens 10 eingetretene und 10 nicht eingetretene Ereignisse. Dann einfacher zu handhaben als Binomialverteilung.

Alle drei Näherungen benötigen die individuellen p_i nicht, tun sich aber schwer mit Abhängigkeiten zwischen Zählereignissen. Bei der Annäherung durch Poissonverteilungen wurden mögliche Abhängigkeiten ignoriert, bei der Normalverteilungnäherung in Form der Varianzerhöhung κ berücksichtigt.

4.100 Binomialverteilung

Die Binomialverteilungsnäherung

$$(4.30) \quad \mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$$(4.31) \quad \mathbb{E}[X] = \mu = n \cdot p$$

$$(4.32) \quad \text{Var}[X] = \sigma^2 = n \cdot p \cdot (1-p)$$

liefert eine Abschätzung für die Standardabweichung aus Eintrittswahrscheinlichkeit p oder einem experimentell bestimmten Zählwerten x_{AV} :

$$(4.35) \quad \sigma = \sqrt{\kappa \cdot n \cdot p \cdot (1-p)}$$

$$(4.38) \quad \hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot \left(1 - \frac{x_{AV}}{n}\right)}$$

Daraus leiteten sich Varianzkoeffizienten für das Eintreten bzw. Nichteintreten als Maße der Schätzgenauigkeit ab:

$$(4.39) \quad \begin{aligned} \hat{\sigma}_r &= \frac{\hat{\sigma}}{x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} - \frac{1}{n}\right)} && \text{für } x_{AV} \leq \frac{n}{2} \\ \hat{\sigma}_{\bar{r}} &= \frac{\hat{\sigma}}{n-x_{AV}} = \sqrt{\kappa \cdot \left(\frac{1}{n-x_{AV}} - \frac{1}{n}\right)} && \text{für } x_{AV} > \frac{n}{2} \end{aligned}$$

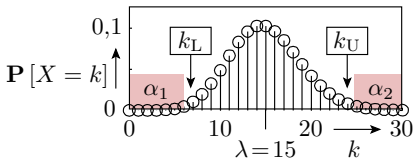
4.101 Poissonverteilung

Varianz gleich Erwartungswert
gleich λ . Verteilungsfunktion:

$$(4.40) \quad \mathbb{P}[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Tabellierbar für Schätzaufgaben in der Form:

- Minimaler Erwartungswert λ_U , damit Zählwerte k mit Wahrscheinlichkeit $\geq 1 - \alpha_1$ mindestens k_L ist und
- maximaler Erwartungswert λ_L , damit Zählwert k mit Wahrscheinlichkeit $\geq 1 - \alpha_2$ nicht größer k_U ist.

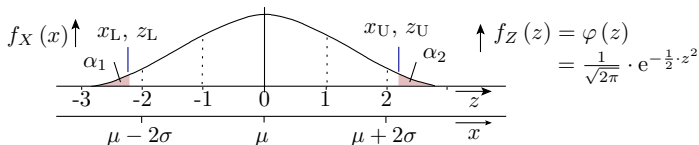


λ_U	$k_L=1$	$k_L=2$	$k_L=3$	$k_L=4$	$k_L=5$	$k_L=6$
$\alpha_1 = 1\%$	4,606	6,638	8,406	10,045	11,605	13,109

λ_L	$k_U=0$	$k_U=1$	$k_U=2$	$k_U=3$	$k_U=4$	$k_U=5$	$k_U=6$
$\alpha_2 = 1\%$	0,01	0,148	0,436	0,823	1,279	1,785	2,330

Für einen experimentellen Ist-Wert $x_{AV} = 2$ und $\alpha_1 = \alpha_2 = 1\%$ liegt z.B. der Erwartungswert im Bereich $0,436 \leq \lambda \leq 6,638$. Genauere Vorhersagen verlangen größere Zählwerte.

4.102 Normalverteilung, Bereichszugehörigkeit



z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
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

■ Transformation Bereichsgrenzen:

$$(4.52) \quad z_L = \frac{x_L - \mu}{\sigma}$$

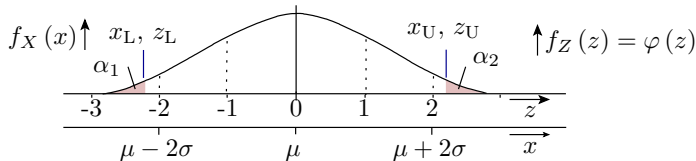
$$(4.53) \quad z_U = \frac{x_U - \mu}{\sigma}$$

■ Bestimmung der Irtumswahrscheinlichkeiten mit der Tabelle:

$$(4.54) \quad \alpha_1 = \Phi(z_L) = 1 - \Phi(-z_L) = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right)$$

$$(4.55) \quad \alpha_2 = 1 - \Phi(z_U) = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right)$$

4.103 Normalverteilung, Bereichsgröße



Inverse standardisierte Normalverteilung:

$\alpha_{1/2}$	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1 - \alpha_{1/2})$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

- Ablesen der Bereichsgrenze z_L und z_U aus der Tabelle.

- (4.56)
$$z_L = \Phi^{-1}(\alpha_1) = -\Phi^{-1}(1 - \alpha_1)$$

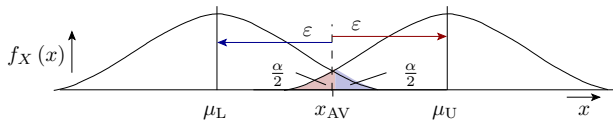
- (4.57)
$$z_U = \Phi^{-1}(1 - \alpha_2)$$

- Transformation in die Bereichsgrenzen der Zufallsvariablen X :

- (4.58)
$$x_L = \mu - \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1)$$

- (4.59)
$$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha_2)$$

4.104 Erwartungswertbereich um einen Ist-Wert



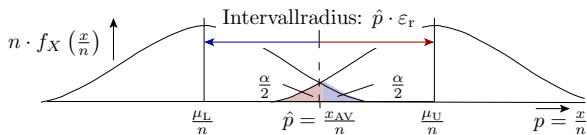
α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

Symmetrischer Bereich mit x_{AV} als Schätzwert:

$$(4.63) \quad sr_{\mu} = [\mu_L, \mu_U] = x_{AV} \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

$$(4.64) \quad sr_{\mu} = x_{AV} \cdot \left(1 \mp \sigma_r \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)$$

4.105 Mittleren Eintrittswahrscheinlichkeit



Schätzer der Eintrittswahrscheinlichkeit für n Zählversuche:

$$(4.37) \quad \hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{AV}}{n}$$

Relative Intervallradien:

$$(4.68) \quad \varepsilon_r = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{x_{AV}} \left[-\frac{1}{n}\right]^*\right)} \quad \text{für } x_{AV} \leq \frac{n}{2}$$

$$(4.69) \quad \varepsilon_{\bar{r}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\kappa \cdot \left(\frac{1}{n - x_{AV}} \left[-\frac{1}{n}\right]^*\right)} \quad \text{für } x_{AV} > \frac{n}{2}$$

Symmetrischer Bereich:

$$(4.70) \quad sr_p = [p_L, p_U] = \frac{x_{AV}}{n} \cdot (1 \mp \varepsilon_r) \quad \text{für } x_{AV} \leq \frac{n}{2}$$

$$(4.71) \quad sr_p = [p_L, p_U] = 1 - \left(1 - \frac{x_{AV}}{n}\right) \cdot (1 \mp \varepsilon_{\bar{r}}) \quad \text{für } x_{AV} > \frac{n}{2}$$

4.106 Benötigte Größenordnung der Zählwerte

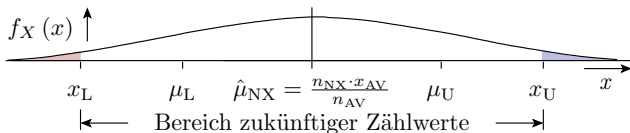
$$(4.72) \quad x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot (1 - \hat{p}) \text{ für } \hat{p} \leq 50\%$$

$$(4.73) \quad n - x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_r^2} \cdot \hat{p} \text{ für } \hat{p} > 50\%$$

Beispiel für $\Phi^{-1}(1 - \frac{\alpha}{2}) = 2$, d.h. $\alpha = 4,52\%$:

ε_r	$p = 10\%$		$p = 50\%$		$\varepsilon_{\tilde{r}}$	$p = 90\%$	
	$x_{AV.min}$	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

4.107 Bereich künftiger Zählwerte



Abschätzung des wahrscheinlichen Bereichs zukünftiger Zählwerte aus einem Istzählwert:

$$(4.74) \quad \text{sr}_{NX} = [x_L, x_U] = \hat{\mu}_{NX} \mp \sqrt{\kappa \cdot n_{NX} \cdot \left(\frac{n_{NX}}{n_{AV}} + 1\right) \cdot \hat{p} \cdot (1 - \hat{p}) \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}$$

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

Vergrößerung des Bereichs gegenüber einer Schätzung mit bekannter Eintrittswahrscheinlichkeit p und n_{NX} Zählversuchen um den Faktor:

$$\sqrt{1 + \frac{n_{NX}}{n_{AV}}}$$

4.108 Varianzerhöhung

Abhängigkeiten von Zählwerten erhöhen bei gleichem Erwartungswert die Varianz. Experimentelle Abschätzung der Varianzerhöhungen:

$$(4.76) \quad \hat{\kappa} = \frac{\hat{\text{Var}}[X]}{\mathbb{E}[X] \cdot (1 - \hat{p})} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot (1 - \hat{p})}$$

Ein Varianzerhöhung $\kappa > 1$ bedeutet, dass für dieselbe Schätzgenauigkeit ε_R und α in

$$(4.72) \quad x_{AV} \geq \frac{\kappa \cdot (\Phi^{-1}(1 - \frac{\alpha}{2}))^2}{\varepsilon_R^2} \cdot (1 - \hat{p}) \quad \text{für } \hat{p} \leq 50\%$$

die erforderliche Zählwertgröße x_{AV} und damit auch die Anzahl der Zählversuche n κ -mal so groß sein muss.

4.109 Modellfehlerüberdeckung

Zwischen Modellfehlern gibt es auch, wenn identisch nachweisbare Fehler zusammengefasst sind, erhebliche Nachweisabhängigkeiten durch gemeinsame Anregungs- und Beobachtungsbedingungen, erkennbar an einer großen Varianzerhöhung.

Ab einer gewissen Größe verbessern mehr Modellfehler die Schätzgenauigkeit nicht weiter. Das ist ein wichtiger Aspekt für die Fehlermodellierung.

Bessere Schätzgenauigkeit lässt sich hier durch Mittelung der Fehlerabdeckungen mehrerer Zufallstestsätze erreichen.

4.110 Fehleranzahl und Defektanteil

In einem idealen Prozess mit gleichbleibender Fehlerentstehungsrate ist die Fehleranzahl je Objekt poissonverteilt:

$$(4.42) \quad \mathbb{P}[X = k] = e^{-\mu_F} \cdot \frac{\mu_F^k}{k!}$$

Der zu erwartende Fehleranteil ist Wahrscheinlichkeit Fehlerzahl null:

$$(4.43) \quad \mu_{DL} = 1 - e^{-\mu_F}$$

Für geringe Werte ist der zu erwartende Fehleranteil gleich der zu erwartenden Fehleranzahl.

4.111 Fehlercluster

Örtliche und zeitliche Schwankungen der Fehlerentstehungsrate verursachen Fehlercluster (Fehlerhäufungen). Beispiele:

- Cluster von Schreibfehler in Texten,
- Fehlerhäufung in Software-Teilen, ...
- Fehlercluster auf Schaltkreis-Wavern, ...

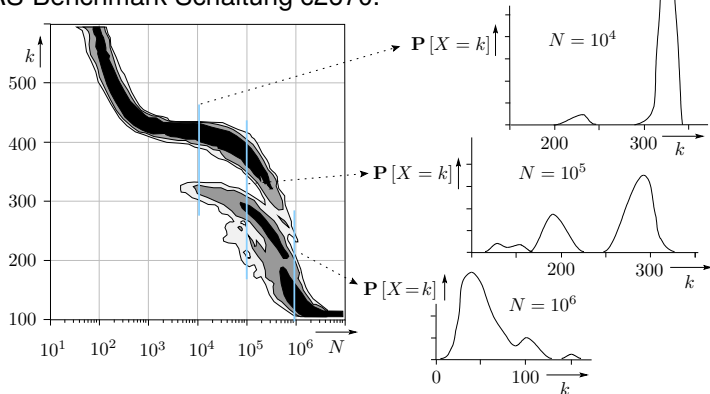
Fehlercluster mindern den zu erwartenden Fehleranteil bei gleicher zu erwartender Fehleranzahl und liefern Hinweise auf Möglichkeiten zur Verbesserung der Entstehungsprozesse.



Mischverteilung

4.112 Anzahl der nachweisbare Fehler c2670

Bestimmung der Verteilung der Anzahl der nachweisbaren Fehler wie auf (Folie 4.93) für die kleinere ISCAS-Benchmark-Schaltung c2670:



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

Wie ist das möglich?



4.113 Mischverteilung

Aus einer Grundgesamtheit gemischter Objekte mit unterschiedlichen Verteilungen wählt eine diskrete Zufallsvariable Y mit Verteilung

$$\mathbb{P}[Y = j] = h_j$$

zufällig ein Objekt X_i mit Verteilung F_{X_j} aus. Die resultierende Verteilungsfunktion, Verteilung bzw. Dichte ergeben sich durch gewichtete Mittelwertbildung für alle Werte, die die Zufallsgröße annehmen kann:

$$F_X(x) = \mathbb{P}[X \leq x] = \sum_{j=1}^{\#Y} h_j \cdot F_{X_j}(x) \quad (4.78)$$

$$\mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i) \quad (4.79)$$

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{i=1}^{\#Y} h_j \cdot f_{X_j}(x) \quad (4.80)$$



4.114 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 von Schäden aus unterschiedlichen Schadensklassen mit unterschiedlichen Kostenverteilungen.
- Anzahl der nicht nachweisbaren Fehler, wenn große Fehlergruppen sehr ähnliche Nachweisbedingungen haben.
- ...



Eigenschaften

4.115 Varianzvergrößerung durch Mischung

Der Erwartungswert ist der gewichtete Mittelwert:

$$\mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \quad (4.81)$$

Varianz bei abweichenden Erwartungswerten $\mu = \mu_j + \delta_j$:

$$\begin{aligned} \text{Var}[X] &= \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{E}[(X_i - (\mu_j + \delta_j))^2] \\ &= \sum_{j=1}^{\#Y} h_j \left(\underbrace{\mathbb{E}[(X_j - \mu_j)^2]}_{\sigma_j^2} - 2 \cdot \delta_j \cdot \underbrace{\mathbb{E}[(X_j - \mu_j)]}_0 + \underbrace{\mathbb{E}[\delta_j^2]}_{\delta_j^2} \right) \\ \text{Var}[X] &= \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \quad (4.82) \end{aligned}$$

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

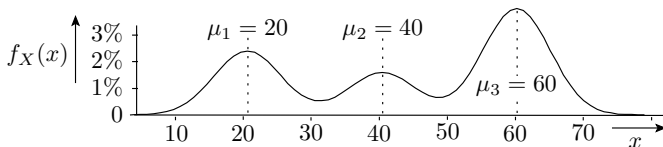
4.116 Multimodale Verteilung

Beim Mischen von Grundgesamtheiten mit deutlich abweichenden Erwartungswerten entstehen multimodale Verteilungen. Beispiel: Mischen von drei Normalverteilungen nach Gl. 4.80 mit:

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

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

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





$$(4.81) \quad \mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j$$

$$(4.82) \quad \text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2$$

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

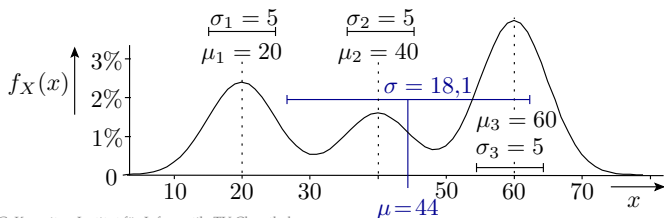
Erwartungswert:

$$\mu = 0,3 \cdot 20 + 0,2 \cdot 40 + 0,5 \cdot 60 = 44$$

Varianz, Standardabweichung:

$$\sigma^2 = 25 + 0,3 \cdot (20 - 44)^2 + 0,2 \cdot (40 - 44)^2 + 0,5 \cdot (60 - 44)^2 = 329$$

$$\sigma = 18,1$$





Anwendungen



Beispiel 4.14: 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, $p = 60\%$

- a) *Mischverteilung als Zusammensetzung aus zueinander verschobenen Binomialverteilungen?*
- b) *Erwartungswert und Varianz?*
- c) *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?*

$$(4.30) \quad \mathbb{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$$(4.79) \quad \mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i)$$

Binomialverteilung ohne die 10 nur gemeinsam nachweisbaren Fehler:

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

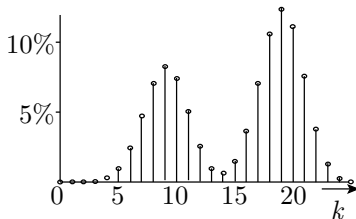
Mit den 10 nur gemeinsam nachweisbaren Fehlern verschieben sich alle Wahrscheinlichkeiten um 10 Realisierung:

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

$n = 25$, 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?

$$(4.81) \quad \mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j$$

$$(4.82) \quad \text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2$$

Erwartungswert:

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

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

Varianz als Summe der Varianzen der Summanden:

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

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



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

c) *Standardabweichung und Varianzerhöhung?*

$$(4.9) \quad \text{sd}[X] = \sigma = \sqrt{\text{Var}[X]}$$

$$(4.75) \quad \kappa = \frac{\text{Var}[X]}{n \cdot p \cdot (1-p)} = \frac{\sigma^2}{\mathbb{E}[X] \cdot (1-p)}$$

Standardabweichung:

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

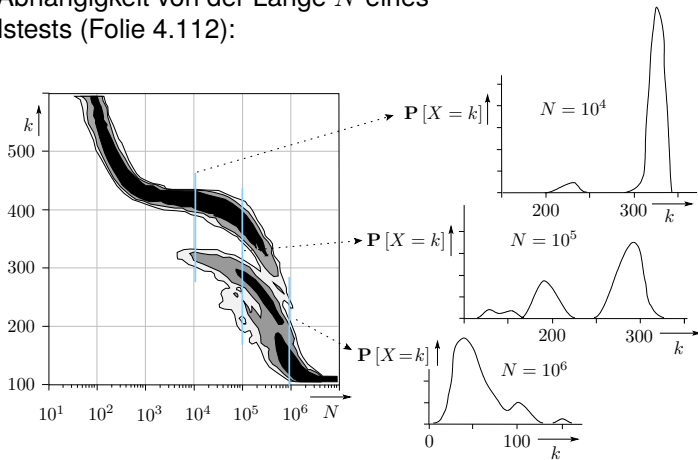
Varianzerhöhung:

$$\kappa = \frac{27,6}{25 \cdot (1 - 60\%)} = 2,76$$

Varianzerhöhung wie wenn immer 2 bis 3 Modellfehler identisch nachweisbar wären.

4.120 Verteilung nicht nachweisbare Fehler c2670

... in Abhängigkeit von der Länge N eines
Zufallstests (Folie 4.112):

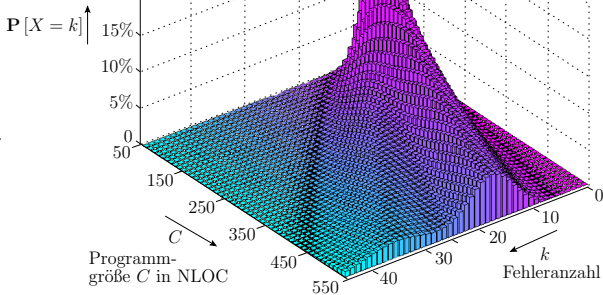


Im Bereich von $N = 10^4$ bis 10^6 multimodale Verteilung. Offenbar ca. 80 sehr ähnlich nachweisbare Fehler mit MF-Rate $\zeta_i \approx 10^{-5} \left[\frac{\text{MF}}{\text{DS}} \right]$.

4.121 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:

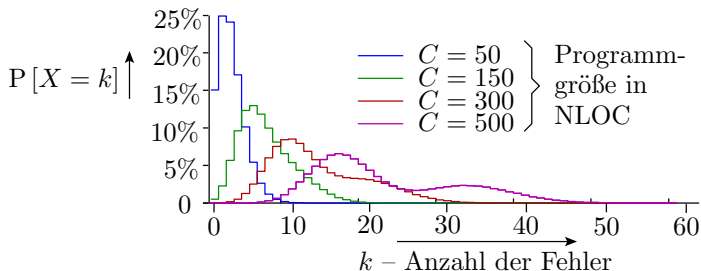
$$\begin{aligned}
 \mathbb{P}[X = k] &= \underbrace{\frac{2}{3} \cdot e^{-\frac{C}{30}} \cdot \frac{\left(\frac{C}{30}\right)^k}{k!}}_{\text{Pois}\left(\lambda = \frac{C}{30}\right)} \\
 &+ \underbrace{\frac{1}{3} \cdot e^{-\frac{C}{15}} \cdot \frac{\left(\frac{C}{15}\right)^k}{k!}}_{\text{Pois}\left(\lambda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





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. 4.79):

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



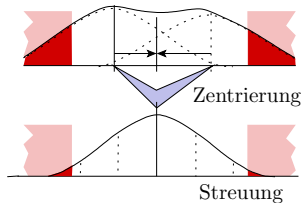
Die Polarisierung nimmt mit der Größe der Software-Bausteine, die vom Profi und vom Anfänger getrennt entwickelt werden, zu.

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.

4.123 Objekte 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, ..., Verlust der Zentrierung und eventuell neue Polarisierungen.



(Folie 2.100 *Prozesszentrierung*).



4.125 Multimodalität und Fehlervermeidung

Polarisierungen (mehrere Gipfel) liefern auch allgemein wichtige Informationen über Schwachstellen und Ansatzmöglichkeiten für Verbesserungen:

- 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, ...

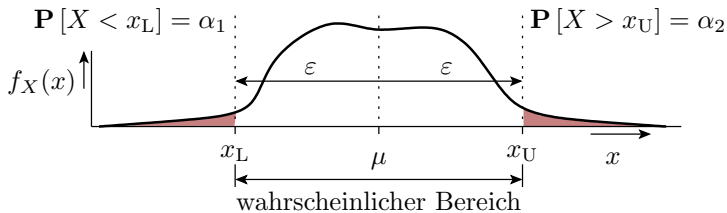
Wenn die Zufallsvariable ein Gütemaß ist, hat man es offenbar mit einer zufälligen Mischung von besser und schlechter funktionierenden Prozessabläufen zu tun. Dann ist es natürlich interessant, warum der Entstehungsprozess mal besser und mal schlechter funktioniert, um das schlechter Funktionierende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

Ein gereifter Entstehungsprozess ist oft daran zu erkennen, dass die überwachten Parameter näherungsweise normalverteilt sind.



Tschebyscheffsche Ungl.

4.126 Bereichsschätzung Verteilung unbekannt

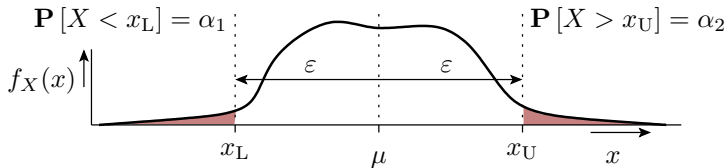


Die Bestimmung eines wahrscheinlichen Bereichs $[x_L, x_U]$

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

x_L, x_U	Untere und obere Schranke des wahrscheinlichen Bereichs von X .
α_1, α_2	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$\mathbb{P}[\dots]$	Eintrittswahrscheinlichkeit des Ereignisses ...
$f_X(x)$	Dichtefunktion der Zufallsvariablen X .
μ	Erwartungswert.
ε	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.

4.127 Das schwache Gesetz der großen Zahlen



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

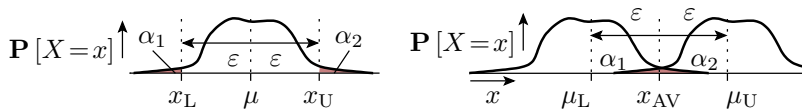
$$\mathbb{P}[|X - \mu| \geq \epsilon] = \alpha \leq \frac{\sigma^2}{\epsilon^2} \quad (4.83)$$

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

$$\epsilon \leq \sqrt{\frac{\sigma^2}{\alpha}} = \frac{\sigma}{\sqrt{\alpha}} \quad (4.84)$$

X	Zufallsvariable.
ϵ, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.

4.128 Bereichsschätzung



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

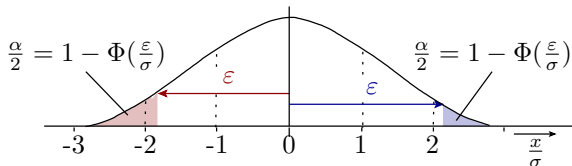
$$sr = [x_L, x_U] = \mu \mp \frac{\sigma}{\sqrt{\alpha}} \quad (4.85)$$

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

$$sr_\mu = [\mu_L, \mu_U] = x_{AV} \mp \frac{\sigma}{\sqrt{\alpha}} \quad (4.86)$$

ε, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
α	Irrtumswahrscheinlichkeit Werte außerhalb des geschätzten Bereichs.
μ_L, μ_U	Untere und obere Schranke des wahrscheinlichen Bereichs des Erwartungswerts.
x_L, x_U	Untere und obere Schranke des wahrscheinlichen Bereichs von X .
sr	Symmetrischer Bereich der wahrscheinlichen Werte.
sr_μ	Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.

4.129 Vergleich Intervallradius Normalverteilung



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

$$(4.60) \quad \varepsilon = \sigma \cdot \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

Intervallradius Tschebyscheffsche Ungleichung:

$$(4.84) \quad \varepsilon \leq \sqrt{\frac{\sigma^2}{\alpha}} = \frac{\sigma}{\sqrt{\alpha}}$$

α	4,55%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10
$1/\sqrt{\alpha}$	4,68	19,60		5	7,07	10	15,8	22,4



Beispiel 4.15: Tschebyscheffsche Ungleichung

Aus eine Stichprobe gemessener Widerstandswerte in $k\Omega$

$$R_i : 10,3, 10,5, 9,7, 8,9, 10,1, 11,0, 10,2, 9,5$$

soll auf den möglichen Bereich des Erwartungswertes geschlussfolgert werden. Zugelassene Irrtumswahrscheinlichkeit 2%.

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

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

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



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

a) *Ohne Kenntnisse der Verteilung über die Tschebyscheffsche Ungleichung?*

$$(4.13) \quad \hat{\mathbb{E}}[X] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i$$

$$(4.14) \quad \hat{\text{Var}}[X] = \hat{\sigma}^2 = \frac{1}{\#v-1} \cdot \sum_{i=1}^{\#v} (v_i - \hat{\mu})^2$$

$$(4.85) \quad \text{sr} = [x_L, x_U] = \mu \mp \frac{\sigma}{\sqrt{\alpha}}$$

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

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

Bereich des Erwartungswerts:

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



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

b) *Unter der Annahme, dass die Widerstandswerte normalverteilt sind?*

$$(4.61) \quad sr = [x_L, x_U] = \mu \mp \sigma \cdot \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

α	0,26%	0	4%	2%	1%	0,4%
$\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$	3	3,9	2,05	2,33	2,57	2,88

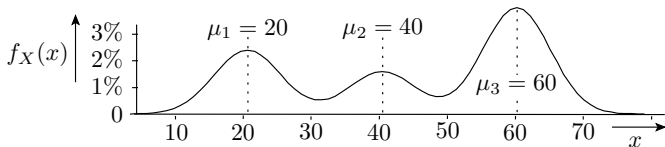
$$\begin{aligned} sr(R) &= \hat{\mu} \mp \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \hat{\sigma} \\ &= 10,025 \text{ k}\Omega \mp 2,33 \cdot 647 \Omega \\ &= [8,5 \text{ k}\Omega, 11,5 \text{ k}\Omega] \end{aligned}$$

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

4.131 Misch- und multimodale Verteilungen



Mischung diskreter Verteilungen:

$$(4.79) \quad \mathbb{P}[X = x_i] = \sum_{j=1}^{\#Y} h_j \cdot \mathbb{P}(X_j = x_i)$$

Mischung stetiger Verteilungen:

$$(4.80) \quad f_X(x) = \frac{dF_X(x)}{dx} = \sum_{i=1}^{\#Y} h_j \cdot f_{X_j}(x)$$

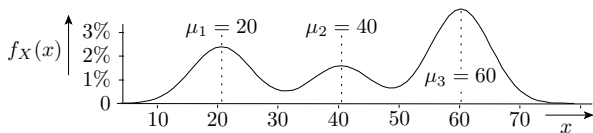
Erwartungswert:

$$(4.81) \quad \mathbb{E}[X] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j$$

Varianz:

$$(4.82) \quad \text{Var}[X] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2$$

4.133 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, ...

4.134 Bereichsschätzung

Bereichsschätzung für beliebige incl. multimodale Verteilungen:

- tschebyscheffsche Ungleichung:

$$(4.83) \quad \mathbb{P}[|X - \mu| \geq \varepsilon] = \alpha \leq \frac{\sigma^2}{\varepsilon^2}$$

- garantierbarer Intervallradius:

$$(4.84) \quad \varepsilon \leq \sqrt{\frac{\sigma^2}{\alpha}} = \frac{\sigma}{\sqrt{\alpha}}$$

- garantierbarer Bereiche für Realisierungen und Erwartungswerte:

$$(4.85) \quad \text{sr} = [x_L, x_U] = \mu \mp \frac{\sigma}{\sqrt{\alpha}}$$

$$(4.86) \quad \text{sr}_\mu = [\mu_L, \mu_U] = x_{AV} \mp \frac{\sigma}{\sqrt{\alpha}}$$



Pareto-Verteilung



4.135 Pareto-Prinzip*

Ein kleiner Teil der Ursachen verursacht Mehrheit der Wirkungen:

- ein kleiner Teil der Fehler die Mehrheit der Fehlfunktionen,
- ein kleiner Teil der Fehlfunktionen den meisten Schaden,
- ein kleiner Teil der Tests weist die Mehrheit der Fehler nach.

Auch nach der Beseitigung der dominanten Ursachen gibt es in der Regel neue dominante Ursachen.

Ein Beispiel ist die Verteilung der erforderlichen Testanzahl X für den Nachweis eines zufälligen Fehlers mit zufälligen Eingaben. Die Wahrscheinlichkeit, dass die erforderliche Nachweislänge X nicht größer N ist, ist die zu erwartende Fehlerabdeckung mit N Tests nach (Gl. 3.14)

$$F_X(N) = \mathbb{P}[X \leq N] = \mu_{\text{FC}}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K} \quad \text{für } N \geq N_0 \quad (4.87)$$

N_0 Skalenparameter, hier Testanzahl, für die erkennbare Fehler bereits beseitigt sind.

$K > 0$ Formfaktor 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.



4.136 Pareto-Verteilung

Die Pareto-Verteilung

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

ist die Verallgemeinerung auf stetige Zufallsgrößen z.B. die Zeit bis zum erstmaligem Auftreten eines Problems oder die Schadenshöhe.
Verteilungsfunktion:

$$F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^K & \text{sonst} \end{cases} \quad (4.88)$$

Dichtefunktion für $x \geq x_{\min}$:

$$f_X(x) = \frac{K \cdot x_{\min}^K}{x^{K+1}} \quad (4.89)$$

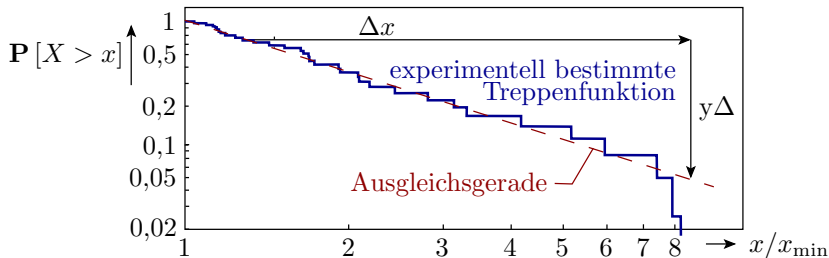
-
- $K > 0$ Formfaktor der Pareto-Verteilung.
 $x_{\min} > 0$ Skalenparameter der Pareto-Verteilung.



Eigenschaften

4.137 Eigenschaften der Pareto-Verteilung

Typisch für Zufallsprozesse, bei denen ein kleiner Teil der Ursachen für den überwiegenden Teil der Wirkungen verantwortlich ist.



Der Formfaktor K ergibt sich aus dem Abfall der Ausgleichsgeraden der doppelt-logarithmisch dargestellten Gegenwahrscheinlichkeit der Verteilungsfunktion:

$$1 - F_X(x) = \mathbb{P}[X > x] = \left(\frac{x}{x_{\min}}\right)^{-K} \quad (4.90)$$

$$K = -\frac{\log(\Delta y)}{\log(\Delta x)} \quad (4.91)$$

4.138 Erwartungswert und Varianz

- 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 $K > 1$:

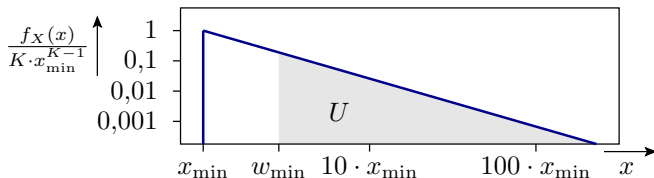
$$\mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K-1} \quad (4.92)$$

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

$$\text{Var}[X] = \sigma^2 = x_{\min}^2 \cdot \frac{K}{(K-2)(1-K)^2} \quad (4.93)$$

Die pareto-verteilte Fehlernachweislänge für Zufallstests (Gl. 4.87) mit $0 < K < 1$ hat keinen Erwartungswert. Es gibt keine mittlere Anzahl von Tests für den Nachweis eines beliebigen Fehlers. Das untermauert die These, dass ein System auch nach langer Nutzung noch Fehler enthalten kann, die noch nie eine Fehlfunktion verursacht haben.

4.139 Formfaktor für die Pareto-20%-80%-Regel



Der Anteil der Ursachen U mit der größten Wirkung:

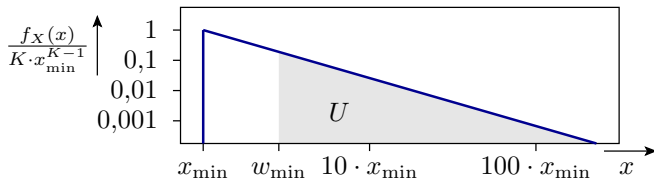
$$U = \int_{w_{\min}}^{\infty} f(x) \cdot dx = \int_{w_{\min}}^{\infty} \frac{K \cdot x_{\min}^K}{x^{K+1}} \cdot dx = \left(\frac{x_{\min}}{w_{\min}} \right)^K \quad (4.94)$$

hat mindestens die Wirkung

$$w_{\min} = x_{\min} \cdot U^{-\frac{1}{K}}$$

und für $K > 1$ insgesamt die Wirkung:

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



Die Gesamtwirkung aller Ursachen ist der Erwartungswert:

$$(4.92) \quad \mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K-1}$$

Anteilige Wirkung der Ursachen mit der größten Wirkung:

$$W = \frac{\mathbb{E}[X|X \geq w_{\min}]}{\mathbb{E}[X]} = \left(\frac{x_{\min}}{w_{\min}}\right)^{K-1} = \left(U^{\frac{1}{K}}\right)^{K-1} = U^{\frac{K-1}{K}} \quad (4.95)$$

$$K = \frac{1}{1 - \frac{\log(W)}{\log(U)}} \quad (4.96)$$

Für $U = 20\%$ der Ursachen $W = 80\%$ der Wirkungen: $K = 1,161$

U Der Anteil der Ursachen mit der größten Wirkung.

W Anteil an der Gesamtwirkung.

w_{\min} Mindestwirkung des Anteils der Ursachen mit der größten Wirkung.



Anwendungen

4.141 Verteilung der Fehlernachweislänge

Die Hypothese, dass die Nachweislänge pareto-verteilt ist

$$(4.87) \quad F_X(N) = \mathbb{P}[X \leq N] = \mu_{FC}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K} \quad \text{für } N \geq N_0$$

basiert auf dem Erfahrungswert, dass bei einem Zufallstest eine Verringerung des Anteils der nicht nachweisbaren Fehler um eine Dekade in der Regel eine Erhöhung der Testsatzlänge um mehr als eine Dekade erfordert (Abschn. 2.2.2):

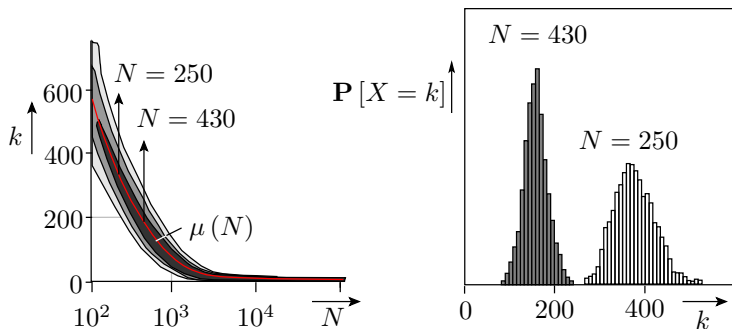
$$(3.14) \quad \mu_{FC}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K}$$

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

Wie brauchbar ist die Näherung?

$\mu_{FC}(N)$	Zu erwartende Fehlerabdeckung 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 Dichte der Fehlfunktionsrate ($0 < K < 1$).

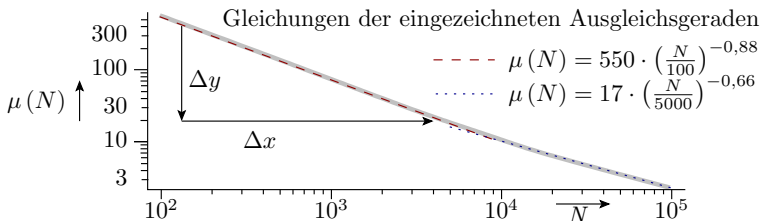
4.142 Benchmarkschaltung c3540 (Folie 4.93)



Experimentell für Haftfehler 3606 simulierte Haftfehler bestimmte Verteilungen der Anzahl der nicht nachweisbaren Fehler mit 1.000 verschiedenen Zufallstests (Folie 4.93). Der Erwartungswert μ in Abhängigkeit von der Testanzahl N ist rot eingezeichnet.

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

4.143 Pareto-Näherung



Doppelt logarithmische Darstellung der Anzahl der nicht nachweisbaren Hauffehler, Ausgleichsgerade. Formfaktor:

$$(4.91) \quad K = -\frac{\log(\Delta y)}{\log(\Delta x)}$$

Keine perfekte Pareto-Verteilung. Der Formfaktor ändert sich mit der Größenordnung der Testanzahl, ... Weitere Forschung erforderlich.

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



Schaden durch MF

4.144 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: Flugzeugabsturz, Atomkrieg, ... (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.



4.145 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

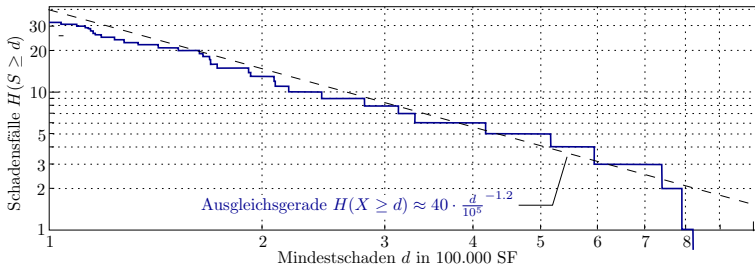
- Anzahl der Schadensfälle: 33
- Gesamtschadenssumme: 9.458.208 SF

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..

4.146 Pareto-Näherung

Anzahl der Versicherungsfälle mit einem Schaden größer s :



Pareto-Verteilung der Schadenshöhe ab $d_{\min} = 10.000$ 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. 4.92):

$$\mu_d = d_{\min} \cdot \frac{K}{K-1} = \frac{12}{12-1} \cdot d_{\min} = 600.000 \text{ SF}$$

Eine Pareto-Verteilung hat erst für $K > 2$ eine Varianz. Alle im Kapitel behandelten Bereichsschätzungen incl. über die Tschebyscheffsche Ungleichung nicht anwendbar. 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.

μ_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

4.148 Zusammenfassung

Verteilungsfunktion und Dichte der Pareto-Verteilung:

$$(4.88) \quad F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^K & \text{sonst} \end{cases}$$

$$(4.89) \quad f_X(x) = \frac{K \cdot x_{\min}^K}{x^{K+1}}$$

Erwartungswert erst ab $k > 1$:

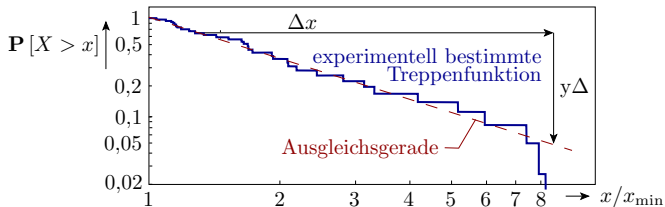
$$(4.92) \quad \mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K-1}$$

Varianz erst ab $k > 2$:

$$(4.93) \quad \text{Var}[X] = \sigma^2 = x_{\min}^2 \cdot \frac{K}{(K-2)(1-K)^2}$$

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

4.149 Experimentell Untermauerung



Für die doppelt log. dargestellte Gegenwahrscheinlichkeit der Verteilungsfunktion:

$$(4.90) \quad 1 - F_X(x) = \mathbb{P}[X > x] = \left(\frac{x}{x_{\min}}\right)^{-K}$$

muss es eine brauchbare Geradenannäherung geben:

$$(4.91) \quad K = -\frac{\log(\Delta y)}{\log(\Delta x)}$$

Mit experimentell erhobenen Zählwerten untersucht für

- die Nachweislänge für Haftfehler ($K = 0,66 \dots 0,88$) und
- Haftpflichtschäden für Fahrzeuge ($K = 1,2$).