Test und Verlässlichkeit Foliensatz 4: Themenspezifische Verteilungen

Prof. G. Kemnitz

Institut für Informatik, TU Clausthal

7. Mai 2025

Inhalt Foliensatz 4

Verteilung
1.1 Charakteristische Größen
1.2 Lineare Transformationen, Summen
1.3 Verteilung von Zählwerten
1.4 Messfehler
Schätzungen
2.1 Binomialverteilung
2.2 Poisson-Verteilung
2.3 Poisson-verteilte Zählwerte
2.4 Normalverteilung
2.5 Normalverteilte Zählwerte
2.6 Zukünftige für bekannte ZW
2.7 Erforderliche Zählwertgröße
Abhängigkeiten
3.1 Varianzerhöhung
3.2 Haftfehlernachweis
3.3 Mischverteilung
3.4 Varianzvergrößerung
3.5 Beispiele für Mischvert.
3.6 Tschebytscheffsche Ungl.
3.7 Exkurs Defektanteil
Pareto-Verteilung
4.1 Eigenschaften
4.2 Anwendungen
4.3 Schaden durch MF

4.2 Themenspezifische Verteilungen

Zählwerte, daraus abgeleitete Eintrittsraten und weitere interssierende Kenngrößen haben viele oder sogar nicht abzählbar viele mögliche Werte. Das Modell dafür sind Verteilungen, Dichtefunktionen und Verteilungsfunktionen, die den Werten Wahrscheinlichkeiten zuordnen.

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

- Kenngrößen von Verteilungen, lineare Transformationen, Summen, ...
- Berechnung der Verteilung 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.

Zufallsvariablen werden durch ihre Verteilungsfunktion beschrieben:

$$F_X(x) = \mathbb{P}[X \leq x] \tag{4.1}$$

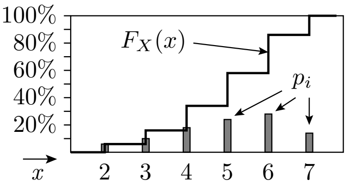
Diskrete Zufallsvariablen zusätzlich durch ihre Verteilung:

$$\mathbb{P}[X = x_i] = p_i \tag{4.2}$$

Stetige Zufallsvariablen zusätzlich durch ihre Dichtefunktion:

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} \tag{4.3}$$

4.4 Diskrete Verteilung



Zufallsvariable X kann nur abzählbare Werten x_i annehmen, z.B.:

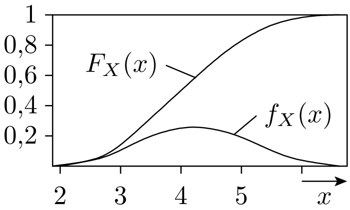
x_i	2	3	4	5	6	7
$\mathbb{P}[X = x_i] = p_i$	6%	10%	18%	24%	28%	14%
$F_X(x) = \mathbb{P}[X \leq x_i]$	6%	16%	34%	58%	86%	100%

Anwendbar auf Zählwerte:

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

$\mathbb{P}[X = x_i]$ Verteilung der diskreten Zufallsvariablen X.
$F_X(x)$ Verteilungfunktion 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. Dichtefunktion:

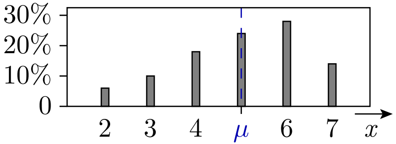
(4.3) $$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

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

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

Charakteristische Größen

4.7 Erwartungswert



Mittlerer Eintrittswert, Berechnung aus der

- Verteilung einer diskreten Zufallsvariablen:

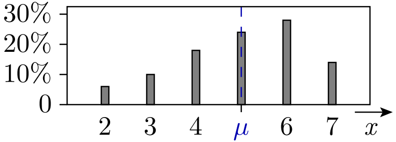
$$\mathbb{E}\left[X\right] = \mu = \sum_{i=1}^{\#x} p_i \cdot x_i \tag{4.4}$$

- Dichte einer stetigen Zufallsvariablen:

$$\mathbb{E}\left[X\right] = \mu = \int_{x_{\min}}^{x_{\max}} f_X\left(x\right) \cdot x \cdot \mathrm{d}x \tag{4.5}$$

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

4.8 Erwartungswert am Beispiel



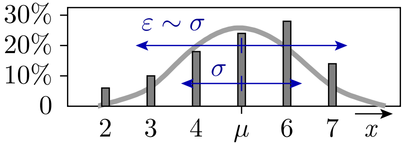
x_i	2	3	4	5	6	7
$\mathbb{P}[X = x_i] = p_i$	6%	10%	18%	24%	28%	14%
$F_X(x) = \mathbb{P}[X \le x_i]$	6%	16%	34%	58%	86%	100%

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

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

X	Zufallsvariable.
x_i	Realisierung (möglichen Wert) i der Zuvallsvariablen X.
$\#x$	Anzahl der möglichen Realisierungen der Zufallsvariablen X.
p_i	Eintrittswahrscheinlichkeit der Realisierung x_i.
$\mathbb{E}[\ldots]$, μ	Erwartungswert.

4.9 Varianz, Standardabweichung



Die Standardabweichung σ ist ein Maß für die Breite des wahrscheinlichen Bereichs $\mu \mp \varepsilon$. Die Varianz ist mittlere quadratische Abweichung vom Erwartungswert μ und das Quadrat der Standardabweichung:

$$\mathrm{Var}\left[X\right] = \sigma^2 = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^2\right] \tag{4.6}$$

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

$$\mathrm{Var}\left[X\right] = \sigma^2 = \int_{x_{\min}}^{x_{\max}} f_X\left(x\right) \cdot \left(x - \mathbb{E}\left[X\right]\right)^2 \cdot \mathrm{d}x \tag{4.8}$$

$$\mathrm{sd}\left[X\right] = \sigma = \sqrt{\mathrm{Var}\left[X\right]} \tag{4.9}$$

$\mathrm{Var}\left[\ldots\right]$, σ^2 Varianz.
$\mathrm{sd}\left[\ldots\right]$, σ Standardabweichung.

4.10 Steinerscher Verschiebungssatz

Alternative Berechnung der Varianz:

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \tag{4.10}$$

Für diskrete Zufallsvariablen:

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

Kontrolle durch Nachrechnen:

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

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

$\#x$	Anzahl der möglichen Realisierungen der Zufallsvariablen X.
x_i	Realisierung (möglichen Wert) i der Zuvallsvariablen 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%
$F_X(x) = \mathbb{P}[X \le x_i]$	6%	16%	34%	58%	86%	100%

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

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

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

- Varianz nach Verschiebesatz:

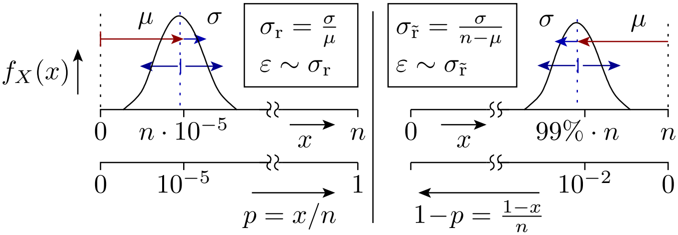
(4.11) $$\mathrm{Var}[X] = \sum_{i=1}^{\#x} p_i \cdot (x_i)^2 - \mathbb{E}[X]^2$$

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

- Standardabweichung:

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

4.12 Varianzkoeffizient



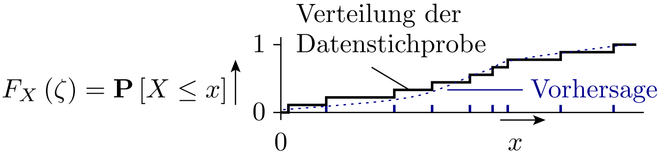
Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten

$$\begin{aligned} \sigma_{\mathrm{r}} &= \frac{\sigma}{\mu} && \text{zweckmäßiges Mass für } \mu \le n/2 \\ \sigma_{\tilde{\mathrm{r}}} &= \frac{\sigma}{n-\mu} && \text{zweckmäßiges Mass für } \mu > n/2 \end{aligned} \qquad (4.12)$$

Maße der relativen Schätzgenauigkeit. Skalierungs- und verschiebungs-invariant. Für Zählwerte und Eintrittswahrscheinlichkeiten gleich.

p	Eintrittswahrscheinlichkeit, Anzahl der Zählversuche.
σ_{r}, $\sigma_{\tilde{\mathrm{r}}}$	Varianzkoeffizient für das Eintreten bzw. Nichteintreten des Zählwerts.
μ, σ	Erwartungswert, Standardabweichung.
ε	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.

4.13 Kontrolle mit Datenstichproben



Wahrscheinlichkeitsmodelle dienen zur Vorhersage künftig zu erwartender Häufigkeiten von Datenmerkmalen. Umgekehrt dienen erhobene Daten zur Kontrolle der Wahrscheinlichkeitsmodelle.

Für eine Datenstichprobe einer Zufallsvariable X

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

ist die geschätzte Verteilungsfunktion eine Treppenfunktion. Aufsteigende Sortierung der Daten. An jedem Wert v_i Erhöhung um $\#v$.

$\#v$	Größe der Datenstichprobe.
v_i	Wert i der Datenstichprobe.

4.14 Schätzer für Kenngrößen

Im weiteren verwendeter Schätzer für den Erwartungswert:

$$\hat{\mathbb{E}}\left[X\right] = \hat{\mu} = \frac{1}{\#v} \cdot \sum_{i=1}^{\#v} v_i \tag{4.13}$$

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

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

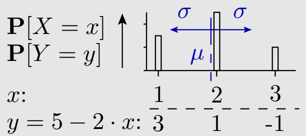
$$\hat{\mathrm{sd}}\left[X\right] = \hat{\sigma} = \sqrt{\hat{\mathrm{Var}}\left[X\right]} \tag{4.15}$$

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.
$\hat{\ldots}$	Schätzwert.
$\mathrm{Var}\left[\ldots\right]$, σ^2	Varianz.

Lineare Transformationen, Summen

4.15 Lineare Transformation



Multiplikationen und Additionen einer Zufallsvariablen mit reellen Zahlen

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

ordnet die Wahrscheinlichkeiten der Originalwerte den transformierten Werten zu (siehe auch Zählwert und Eintrittswahrscheinlichkeit):

$$\mathbb{P}\left[y = ax + b\right] = \mathbb{P}\left[x\right] \tag{4.16}$$

Der Erwartungswert wird genau wie jeder andere Wert transformiert:

$$\mathbb{E}\left[a \cdot X + b\right] = a \cdot \mathbb{E}\left[X\right] + b \tag{4.17}$$

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

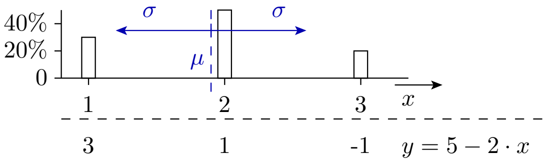
$$\mathrm{Var}\left[a \cdot X + b\right] = a^2 \cdot \mathrm{Var}\left[X\right] \tag{4.18}$$

$$\mathrm{sd}\left[a \cdot X + b\right] = \sqrt{\mathrm{Var}\left[a \cdot X + b\right]} = |a| \cdot \mathrm{sd}\left[X\right] \tag{4.19}$$

Varianzkoeffizient zusätzlich skalierungsinvariant.

$a,\ b$ Beliebige reele Zahlen.

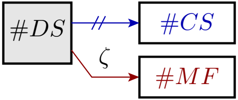
4.16 Veranschaulichung am Beispiel



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

$$\begin{aligned}
\mathbb{E}[X] &= 0{,}3 \cdot 1 + 0{,}5 \cdot 2 + 0{,}3 \cdot 3 = 1{,}9 \\
\operatorname{Var}[X] &= 0{,}3 \cdot 1^2 + 0{,}5 \cdot 2^2 + 0{,}3 \cdot 3^2 - 1{,}9^2 = 0{,}49 \\
\mathbb{E}[Y] &= 0{,}3 \cdot 3 + 0{,}5 \cdot 1 + 0{,}2 \cdot (-1) = 1{,}2 \\
\operatorname{Var}[Y] &= 0{,}3 \cdot 3^2 + 0{,}5 \cdot 1^2 + 0{,}2 \cdot (-1)^2 - 1{,}2^2 = 1{,}96 \\
\mathbb{E}[Y] &= 5 - 2 \cdot \mathbb{E}[X] = 5 - 2 \cdot 1{,}9 = 1{,}2 \checkmark \\
\operatorname{Var}[Y] &= (-2)^2 \cdot \operatorname{Var}[X] = 4 \cdot 0{,}49 = 4 \cdot 0{,}5 - 4 \cdot 0{,}1 = 1{,}96 \checkmark
\end{aligned}$$

4.17 Summe von Zählereignissen



Zufallsgröße X:	$\#CS$
Anzahl Zählversuche:	$\#DS$
Summanden X_i:	jeder DS
Eintrittswahrsch. p_i:	$1-\zeta$

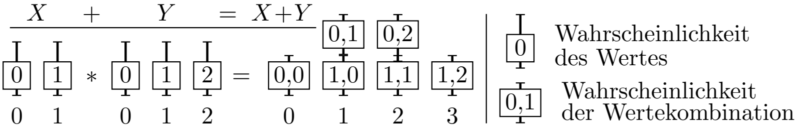
Ein einzelner Zählversuch, im Beispiel, ob eine erbrachte Service-Leistung (DS) eine Fehlfunktion (MF) ist, hat die Bernoulli-Verteilungen:

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

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

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

4.18 Verteilung der Summe



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

$$\begin{aligned}
\mathbb{P}[X+Y=0] &= \mathbb{P}[X=0]\cdot\mathbb{P}[Y=0] = 0{,}4\cdot 0{,}3 = 0{,}12\\
\mathbb{P}[X+Y=1] &= \mathbb{P}[X=0]\cdot\mathbb{P}[Y=1] + \mathbb{P}[X=1]\cdot\mathbb{P}[Y=0] = 0.42\\
\mathbb{P}[X+Y=2] &= \mathbb{P}[X=0]\cdot\mathbb{P}[Y=2] + \mathbb{P}[X=1]\cdot\mathbb{P}[Y=1] = 0{,}4\\
\mathbb{P}[X+Y=3] &= \mathbb{P}[X=1]\cdot\mathbb{P}[Y=2] = 0{,}06
\end{aligned}$$

4.19 Erwartungswert und Varianz

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

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \tag{4.20}$$

$$\operatorname{Var}[X+Y] = \operatorname{Var}[X] + \operatorname{Var}[Y] \tag{4.21}$$

Kontrolle am Beispiel:

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%	$42\% \cdot 1 + 40\% \cdot 2 + 6\% \cdot 3 = 1{,}4$

$$\begin{aligned}
\operatorname{Var}[X] &= 40\% \cdot 0 + 60\% \cdot 1^2 - 0{,}6^2 &&= 0{,}24\\
\operatorname{Var}[Y] &= 60\% \cdot 1 + 10\% \cdot 2^2 - 0{,}8^2 &&= 0{,}36\\
\operatorname{Var}[X+Y] &= 42\% \cdot 1 + 40\% \cdot 2^2 + 6\% \cdot 3 - 1{,}4^2 &&= 0{,}60
\end{aligned}$$

X, Y	Zufallsvariablen.
$\mathbb{E}[\ldots]$	Erwartungswert von ...
$\operatorname{Var}[\ldots]$	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 bedeuten (Abschn. 2.1.5):

- 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*:

$$\mathrm{Var}\left[X + Y\right] = \mathrm{Var}\left[X\right] + \mathrm{Var}\left[Y\right] + 2 \cdot \mathrm{Cov}\left[X, Y\right] \tag{4.22}$$

$$\mathrm{Cov}\left[X, Y\right] = \mathbb{E}\left[(X - \mathbb{E}\left[X\right]) \cdot (Y - \mathbb{E}\left[Y\right])\right] \tag{4.23}$$

Wir werden Abhängigkeiten zwischen Zählwerten statt dessen durch eine aus Daten abschätzbare Varianzvergrößerung beschreiben.

$\mathrm{Cov}[X,Y]$ Kovarianz der beiden Zuvallsvariablen.

* Kontrolle durch Nachrechnen folgt später als Übungsaufgabe.

Verteilung von Zählwerten

4.22 Verteilung von Zählwerten

$$\boxed{\#F} \xrightarrow{p_i} \boxed{\#DF}$$

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

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

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.

4.23 Erwartungswert und Varianz je Summand

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

Erwartungswerte der Einzelereignisse:

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

Varianzen nach dem Verschiebungssatz:

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

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

4.24 Erwartungswert und Varianz der Summe

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

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

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

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

Abhängigkeiten werden später durch einen Faktor κ gegenüber einer aus dem Erwartungswert abschätzbaren garantierbaren Obergrenze für unabhängige Zählwerte modelliert (Abschn. 4.2.1).

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

4.25 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 $S_{i+1} = X_{i+1} + S_i$:

$$\begin{aligned}(1 - p_{i+1}) \cdot (\, &\mathbb{P}[S_i = 0]\,, \mathbb{P}[S_i = 1]\,, ..., \mathbb{P}[X_i = i - 1]\,, \quad 0 \,)\\ + p_{i+1} \cdot (\, &0, \quad \mathbb{P}[S_i = 0]\,, ..., \mathbb{P}[X_i = i - 2], \mathbb{P}[X_i = i - 1]\,)\end{aligned}$$

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.

4.26 Erwartungswert und Varianz 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 Berechung 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:

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

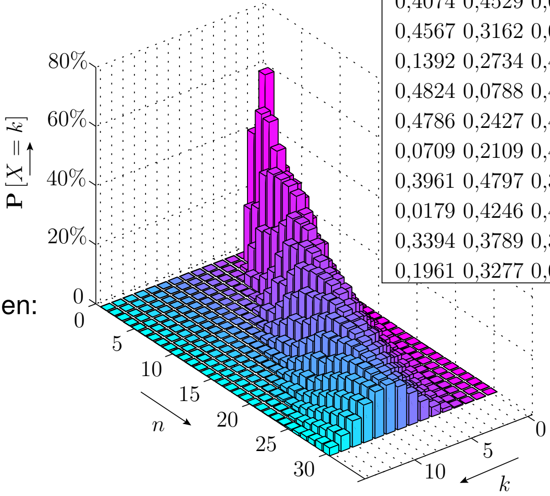
4.27 Beispiel einer Zählverteilung

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

$$\mathbb{E}[X] = 7{,}05$$
$$\sqrt{\mathrm{Var}[X]} = 2{,}19$$

Wahrscheinlicher Bereich ca. 5 bis 15.

p_i für i=1 bis 30		
0,4074	0,4529	0,0635
0,4567	0,3162	0,0488
0,1392	0,2734	0,4788
0,4824	0,0788	0,4853
0,4786	0,2427	0,4001
0,0709	0,2109	0,4579
0,3961	0,4797	0,3279
0,0179	0,4246	0,4670
0,3394	0,3789	0,3716
0,1961	0,3277	0,0856



n Anzahl der Zählversuche, maximaler Zählwert.

Messfehler

4.29 Gemessener Wert und Messfehler

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

In der Messtechnik gilt für jeden gemessenen Wert:

$$X_{\mathrm{M}} = X + X_{\mathrm{F}} \tag{4.27}$$

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

$$\mathbb{E}\left[X_{\mathrm{M}}\right] = \mathbb{E}\left[X\right] + \mathbb{E}\left[X_{\mathrm{F}}\right] \tag{4.28}$$

$$\mathrm{Var}\left[X_{\mathrm{M}}\right] = \mathrm{Var}\left[X\right] + \mathrm{Var}\left[X_{\mathrm{F}}\right] \tag{4.29}$$

- $\mathbb{E}\left[X_{\mathrm{F}}\right]$ – zu erwartender (systematischer) Messfehler
- $\mathrm{sd}\left[X_{\mathrm{F}}\right] = \sqrt{\mathrm{Var}\left[X_{\mathrm{F}}\right]} \sim \varepsilon$ – Standardabweichung des Messfehler, proportional zum Intervallradius ε wahrscheinlicher Messfehler.

X_{M}	Zufallsvariable gemessener Wert.
X	Zufallsvariable zu messender Wert.
X_{F}	Zufallsvariable Messfehler.
$\mathbb{E}\left[\ldots\right]$	Erwartungswert von ...
$\mathrm{Var}\left[\ldots\right]$	Varianz von ...

Beispiel 4.1: Messfehler

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

$\mathbb{E}\left[R_{\mathrm{M}}\right]=1010\,\Omega$, $\mathrm{sd}\left[R_{\mathrm{M}}\right]=11{,}18\,\Omega$, $\mathbb{E}\left[R_{\mathrm{F}}\right]=12\,\Omega$ und $\mathrm{sd}\left[R_{\mathrm{F}}\right]=5\,\Omega$.

Welchen Erwartungswert und welche Standardabweichung hat der zu messende Wert?

R_{M}	Gemessener Widerstandswert.
R_{F}	Messfehler des Widerstandswertes.
$\mathbb{E}\left[\ldots\right]$	Erwartungswert von ...
$\mathrm{Var}\left[\ldots\right]$	Varianz von ...
$\mathrm{sd}\left[\ldots\right]$	Standardabweichung von ...

$\mathbb{E}[R_\mathrm{M}] = 1010\,\Omega$, $\mathrm{sd}[R_\mathrm{M}] = 11{,}18\,\Omega$, $\mathbb{E}[R_\mathrm{F}] = 12\,\Omega$ und $\mathrm{sd}[R_\mathrm{F}] = 5\,\Omega$.

Welchen Erwartungswert und welche Standardabweichung hat der zu messende Wert?

(4.28) $$\mathbb{E}[X_\mathrm{M}] = \mathbb{E}[X] + \mathbb{E}[X_\mathrm{F}]$$

(4.29) $$\mathrm{Var}[X_\mathrm{M}] = \mathrm{Var}[X] + \mathrm{Var}[X_\mathrm{F}]$$

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

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

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

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

R Zu messender Widerstandswert.

Zusammenfassung

4.31 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:

$$F_X(x) = \mathbb{P}[X \le x] \tag{4.1}$$

Eine diskrete Zufallsvariable kann nur abzählbar viele Werte x_i annehmen. Die Zuordnungshäufigkeit beschreibt die Verteilung:

$$\mathbb{P}[X = x_i] = p_i \tag{4.2}$$

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

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} \tag{4.3}$$

4.32 Kenngrößen von Zufallsvariablen

Erwartungswert:

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

(4.5) $$\mathbb{E}\left[X\right]=\mu=\int_{x_{\min}}^{x_{\max}}f_X\left(x\right)\cdot x\cdot\mathrm{d}x$$

Varianz:

(4.6) $$\mathrm{Var}\left[X\right]=\sigma^2=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)^2\right]$$

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

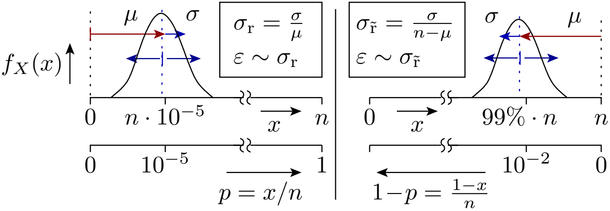
(4.8) $$\mathrm{Var}\left[X\right]=\sigma^2=\int_{x_{\min}}^{x_{\max}}f_X\left(x\right)\cdot\left(x-\mathbb{E}\left[X\right]\right)^2\cdot\mathrm{d}x$$

Berechnung der Varianz nach dem Verschiebesatz:

(4.10) $$\mathrm{Var}\left[X\right]=\mathbb{E}\left[X^2\right]-\mathbb{E}\left[X\right]^2$$

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

4.33 Maße der Schätzgenauigkeit



Standardabweichung als Mass der absoluten Schätzgenauigkeit:

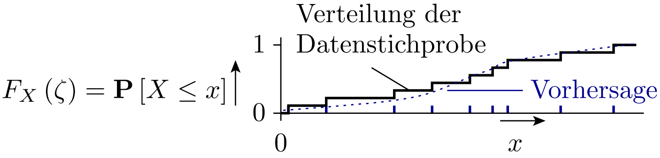
$$\text{sd}[X] = \sigma = \sqrt{\text{Var}[X]} \tag{4.9}$$

Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten als Masse der relativen Schätzgenauigkeit:

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

Varianzkoeffizienten sind verschiebungs- und skalierungsinvariant, d.h. für Zählwerte und Eintrittswahrscheinlichkeiten gleich.

4.34 Kontrolle mit Datenstichproben



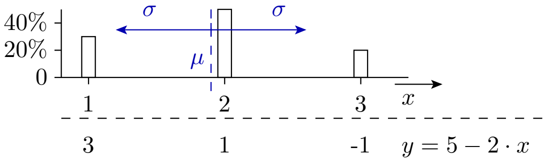
Die Verteilungsfunktion einer Stichprobe experimentell erhobenen Daten ist die relative Häufigkeit, dass das Versuchergebnis mindestens den Wert x hat. Der Vergleich, ob die Daten zum Wahrscheinlichkeitsmodell passen erfolgt in der Regel anhand von aus den Daten abschätzbaren Verteilungsparametern. Wie beschränken uns auf Schätzer für Erwartungswert, Varianz und Standardabweichung:

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

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

$$\hat{\mathrm{sd}}[X] = \hat{\sigma} = \sqrt{\hat{\mathrm{Var}}[X]} \tag{4.15}$$

4.35 Lineare Transformationen



Lineare Transformation $Y = a \cdot X + b$ sind Multiplikationen und Additionen einer Zufallsvariablen mit reellen Zahlen. Sie ändert die den Wahrscheinlichkeiten zugeordneten Werte incl. Erwartungswert und Bereichsgrenzen. Varianz und Standardabweichung sind verschiebungsinvariant und skalieren mit dem Quadrat bzw. dem Betrag:

$$\mathbb{P}[y = ax + b] = \mathbb{P}[x] \tag{4.16}$$

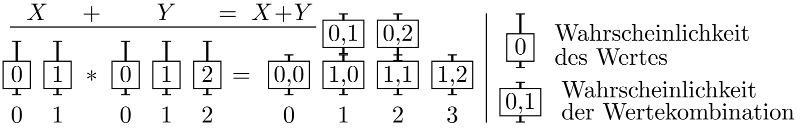
$$\mathbb{E}[a \cdot X + b] = a \cdot \mathbb{E}[X] + b \tag{4.17}$$

$$\mathrm{Var}[a \cdot X + b] = a^2 \cdot \mathrm{Var}[X] \tag{4.18}$$

$$\mathrm{sd}[a \cdot X + b] = \sqrt{\mathrm{Var}[a \cdot X + b]} = |a| \cdot \mathrm{sd}[X] \tag{4.19}$$

Varianzkoeffizient zusätzlich skalierungsinvariant.

4.36 Summen von Zufallsvariablen



Die Verteilung der Summe von Zufallsvariablen ordnet jedem der möglichen Werte der Summe die Wahrscheinlichkeit zu, dass die Summe diesen Wert hat.

Erwartungswert, Varianz der Summe unabhängiger Zufallsvariablen:

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \tag{4.20}$$

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] \tag{4.21}$$

Abhängigkeiten vergrößern die Varianz. Wir werden später Abhängigkeiten durch eine experimentell abschätzbare Varianzvergrößerung κ beschreiben. Standardabweichungen addieren sich »nach Pytagoras«.

4.37 Unabhängige Zählwerte

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%

Erwartungswert und Varianz:

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

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

4.38 Messwerte und Messfehler

Messungen und andere Formen der »experimentellen« Bestimmung von Werten (z.B. Expertenbefragungen zur Einschätzung von Risiken) haben einen systematischen und einen zufälligen Fehler, der sich oft additiv überlagert und unabhängig vom zu bestimmenden Wert ist:

$$X_{\mathrm{M}} = X + X_{\mathrm{F}} \tag{4.27}$$

$$\mathbb{E}\left[X_{\mathrm{M}}\right] = \mathbb{E}\left[X\right] + \mathbb{E}\left[X_{\mathrm{F}}\right] \tag{4.28}$$

$$\mathrm{Var}\left[X_{\mathrm{M}}\right] = \mathrm{Var}\left[X\right] + \mathrm{Var}\left[X_{\mathrm{F}}\right] \tag{4.29}$$

Später wird daraus geschlussfolgert, dass Prüftechnik viel genauer arbeiten muss als das zu testende System (Folie 5.28 *Kontrolle von analogen Merkmalen*).

Schätzungen

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

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

Binomialverteilung:

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

Erwartungswert einer Binomialverteilung:

$$\mathbb{E}\,[X] = \mu = n \cdot p \tag{4.31}$$

Varianz und Standardabweichung einer Binomialverteilung:

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

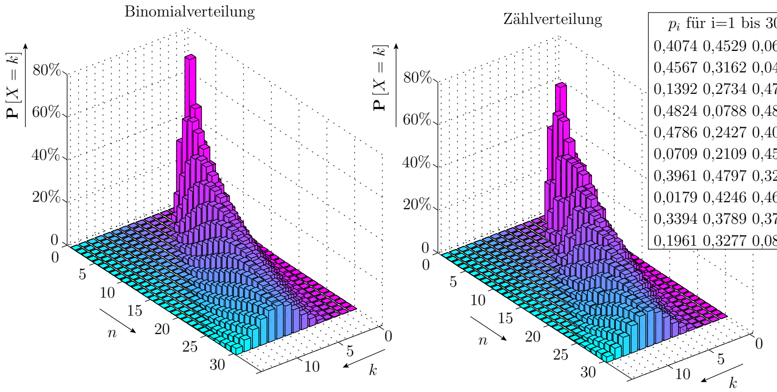
$$\mathrm{sd}\,[X] = \sigma = \sqrt{n \cdot p \cdot (1-p)} \tag{4.33}$$

---

n Anzahl der Zählversuche, maximaler Zählwert.

p Eintrittswahrscheinlichkeit.

4.40 Binomialverteilung vs. allg. Zählverteilung



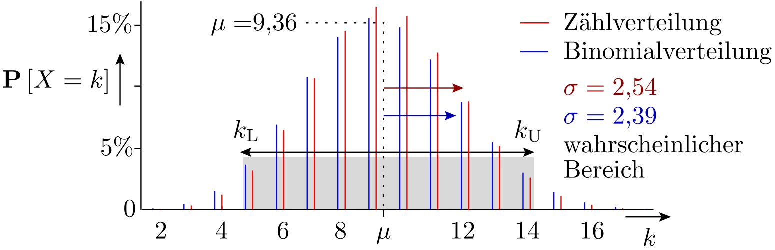
p_i für i=1 bis 30		
0,4074	0,4529	0,0635
0,4567	0,3162	0,0488
0,1392	0,2734	0,4788
0,4824	0,0788	0,4853
0,4786	0,2427	0,4001
0,0709	0,2109	0,4579
0,3961	0,4797	0,3279
0,0179	0,4246	0,4670
0,3394	0,3789	0,3716
0,1961	0,3277	0,0856

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.

---

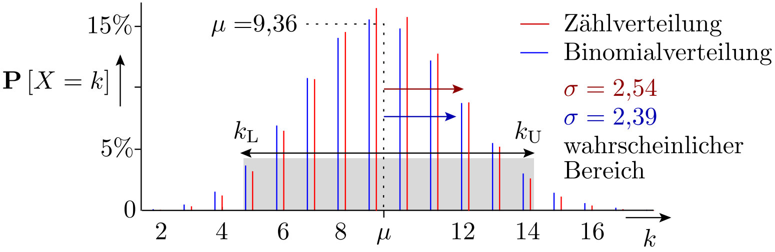
n	Anzahl der Zählversuche, maximaler Zählwert.
p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
p_i	Eintrittswahrscheinlichkeit Zählereignis i.

4.41 Bereichsschätzung Binomialvert.-Näherung



Bei gleicher Anzahl unabhängiger Zählversuche, im Beispiel $n = 30$, 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.



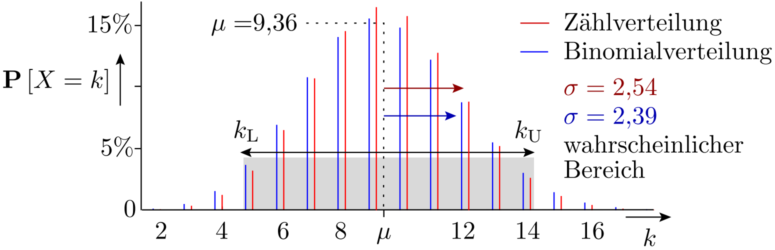
Irrtumswahrscheinlichkeiten für $k_{\mathrm{L}} = 5$, $k_{\mathrm{U}} = 14$:

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

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

Zählverteilung	$k_{\mathrm{L}} = 5$	$k_{\mathrm{U}} = 14$
Binomialverteilung	$k_{\mathrm{L}} = 4$	$k_{\mathrm{U}} = 15$

4.43 Varianzobergrenze unabhängige Zählwerte



Bei gleicher Anzahl von unabhängigen Zählwerten n und $p = \frac{1}{n} \cdot \sum_{i=1}^{n} p_i$ sind die Varianz und Standardabweichung der Binomialverteilung obere Schranken der Varianz bzw. Standardabweichung der Zählverteilung:

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

4.44 Beweis durch Nachrechnen

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

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

Bei gleicher Versuchsanzahl n ist die Varianz einer Zählverteilung:

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

um Summe der quadratischen Abweichungen der Eintrittswahrscheinlichkeiten vom Mittelwert kleiner als die Varianz der Binomialverteilung.

4.45 Schätzer für Varianz von Zählwerten

Die Ungleichheit in

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

wird in einen Korrekturfaktor $\kappa \le 1$ versteckt, der später als experimentell abschätzbares Mass für Abhängigkeiten dient:

$$\sigma^2 = \kappa \cdot n \cdot p \cdot (1-p) \tag{4.35}$$

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

$$\hat{\mu} = x_{\mathrm{AV}} \tag{4.36}$$

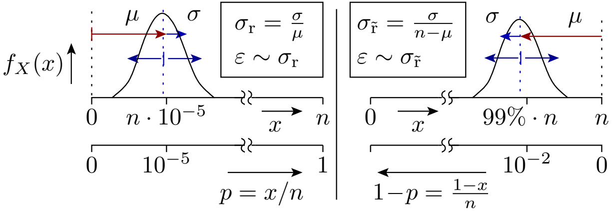
$$\hat{p} = \frac{\hat{\mu}}{n} = \frac{x_{\mathrm{AV}}}{n} \tag{4.37}$$

$$\hat{\sigma} = \sqrt{\kappa \cdot x_{\mathrm{AV}} \cdot (1-\hat{p})} \tag{4.38}$$

Intervallradius wahrsch. Bereich typ $\varepsilon \approx (2 \ldots 3) \cdot \sigma$ (Abschn. 4.2.4).

$\hat{\mu}$, $\hat{\sigma}$	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 $\kappa \le 1$.

4.46 Varianzkoeffizient als Genauigkeitsmaß



Der Varianzkoeffizient für das Eintreten bzw. Nicht-Eintreten ist das Verhältnis aus Standardabweichung zum Erwartungswert und ein skalierungsinvarianten Maß für die Schätzgenauigkeit von Zählwerten und Eintrittswahrscheinlichkeiten:

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

σ_r, $\sigma_{\tilde{r}}$ Relative Standardabweichung zum erwarteten Eintritts- bzw. Nichteintrittszählwert.

Mit dem Schätzer für die Standardabweichung

$$\hat{\sigma} = \sqrt{\kappa \cdot x_{\mathrm{AV}} \cdot (1 - \hat{p})} \tag{4.38}$$

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

$$\begin{aligned} \hat{\sigma}_{\mathrm{r}} &= \frac{\hat{\sigma}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1 - \hat{p})} \quad \text{für } \hat{p} = \frac{x_{\mathrm{AV}}}{n} \le 0{,}5 \\ \hat{\sigma}_{\tilde{\mathrm{r}}} &= \frac{\hat{\sigma}}{n - x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n - x_{\mathrm{AV}}} \cdot \hat{p}} \quad \text{für } \hat{p} = \frac{x_{\mathrm{AV}}}{n} > 0{,}5 \end{aligned} \tag{4.39}$$

Für $x_{\mathrm{AV}} \ll n$ bzw. $n - x_{\mathrm{AV}} \ll n$ sind die Terme $1 - \hat{p}$ bzw. $\hat{p}$ praktisch eins.

σ_{r}, $\sigma_{\tilde{\mathrm{r}}}$	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.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.

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

$$n \to \infty$$
$$p_i \to 0$$
$$\operatorname{Var}[X_i] = p_i \cdot (1 - p_i) \to p_i$$

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

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

Die Verteilung der Summe strebt gegen die Poisson-Verteilung:

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

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

$$\mathbb{P}[X = k] = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!} \tag{4.40}$$

und für Zählwerte:

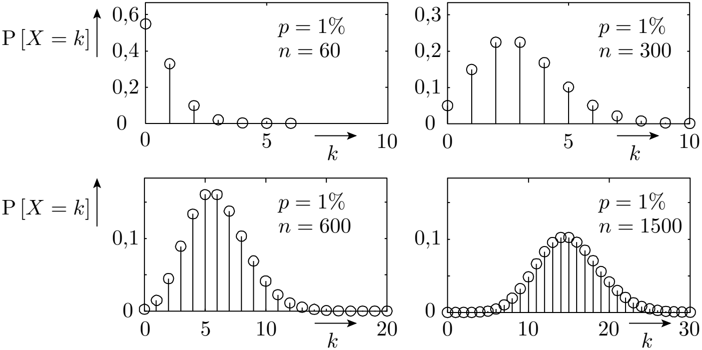
$$\mathbb{P}[X = k] = \mathrm{e}^{-p \cdot n} \cdot \frac{(p \cdot n)^k}{k!} \tag{4.41}$$

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.

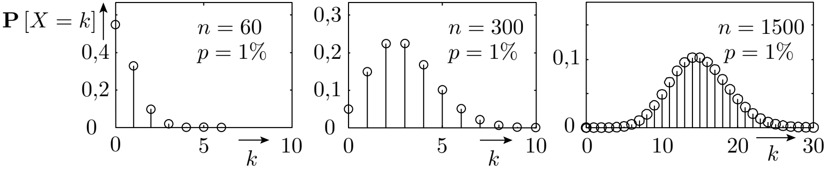
λ	Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).
n	Anzahl der Zählversuche, maximaler Zählwert.
p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

4.50 Anzahl der Zählversuche und Verteilung

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



$\mathbb{P}\left[X=k\right]$	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 größer null:

$$[k_{\mathrm{L}},\ k_{\mathrm{U}}] \approx [0,\ 3...5 \cdot \lambda]$$

- Für $\lambda = p \cdot n \approx 3 \ldots 10$:

$$[k_{\mathrm{L}},\ k_{\mathrm{U}}] \approx \left[\tfrac{\lambda}{3...5},\ 3 \ldots 5 \cdot \lambda\right]$$

Ab $\lambda = p \cdot n > 10$ zunehmend symmetrischer Bereich um den Erwartungswert (Abschn. 4.2.4 *Normalverteilung*):

$$\mathrm{sr} = [k_{\mathrm{U}}, k_{\mathrm{L}}] \approx \lambda \mp 2 \ldots 3 \cdot \sqrt{\lambda}$$

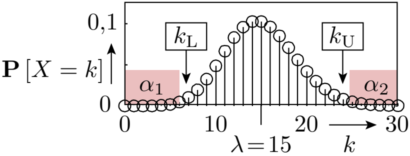
$k_{\mathrm{L}},\ k_{\mathrm{U}}$	Untere (lower) und obere (upper) Bereichsgrenze des Zählwerts.
sr	Symmetrischer Bereich der wahrscheinlichen Werte.
n	Anzahl der Zählversuche, maximaler Zählwert.
p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.

Poisson-verteilte Zählwerte

4.52 Schätzer $k_{\mathrm{L}} \overset{(\alpha_1)}{\longleftrightarrow} \lambda$

Vorgabe k_{L} und α_1. Numerische Suche λ, so dass

$$\sum_{k=0}^{k_{\mathrm{L}}-1} \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!} \leq \alpha_1 \qquad (4.42)$$



λ	$k_{\mathrm{L}}=1$	$k_{\mathrm{L}}=2$	$k_{\mathrm{L}}=3$	$k_{\mathrm{L}}=4$	$k_{\mathrm{L}}=5$	$k_{\mathrm{L}}=6$
$\alpha_1=0{,}5\%$	$\lambda=5{,}298$	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

Beispielschätzungen:

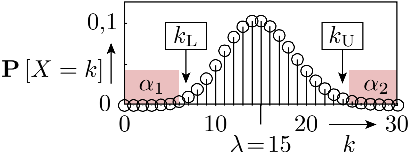
- $\lambda=7$ und $\alpha_1 \leq 1\% \Rightarrow k_{\mathrm{L}}=2$
- $k_{\mathrm{L}}=1$ und $\alpha_1=2\% \Rightarrow \lambda \geq 3{,}912$

α_1, α_2 Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
k_{L}, k_{U} Untere (lower) und obere (upper) Bereichsgrenze des Zählwerts.

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

Vorgabe k_U und α_2. Numerische Suche λ_L, so dass

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



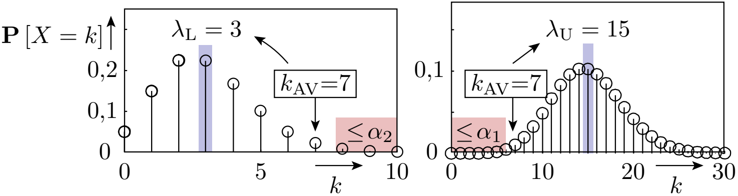
λ	$k_U=0$	$k_U=1$	$k_U=2$	$k_U=3$	$k_U=4$	$k_U=5$	$k_U=6$
$\alpha_2=0{,}5\%$	0,005	0,103	0,338	0,672	1,078	1,537	2,037
$\alpha_2=1\%$	0,01	0,148	0,436	0,823	1,279	1,785	2,330
$\alpha_2=2\%$	0,02	0,215	0,567	1,016	1,529	2,089	2,684
$\alpha_2=10\%$	0,105	0,532	1,102	1,744	2,432	3,152	3,894
$\alpha_2=20\%$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

Beispielabschätzungen:

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

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.

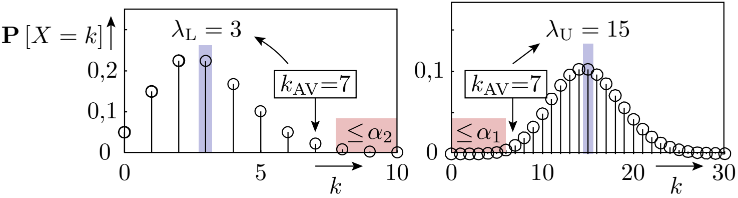
4.54 Schätzer $k_{AV} \xrightarrow{(\alpha)} [\lambda_L,\ \lambda_U]$



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

$[\lambda_L,\ \lambda_U]$	$k_{AV} = 0$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$\alpha_1 = \alpha_2 = 0{,}5\%$	[0,005, *]	[0,10, 5,30]	[0,34, 7,43]	[0,67, 9,27]
$\alpha_1 = \alpha_2 = 1\%$	[0,01, *]	[0,15, 4,60]	[0,44, 6,64]	[0,82, 8,41]
$\alpha_1 = \alpha_2 = 2\%$	[0,02, *]	[0,22, 3,91]	[0,57, 5,83]	[1,02, 7,52]
$\alpha_1 = \alpha_2 = 10\%$	[0,105, *]	[0,53, 2,30]	[1,10, 3,89]	[1,74, 5,32]
$\alpha_1 = \alpha_2 = 20\%$	[0,223, *]	[0,82, 1,61]	[1,53, 2,99]	[2,30, 4,28]

Werte aus Tabellen davor übernommen, * mit (Gl. 4.42) keine Obergrenze bestimmbar, aber λ_U sicher nicht größer, als für $k_{AV} = 1$.



$[\lambda_L, \lambda_U] = f(\alpha, k_{AV})$ 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.

Beispiel 4.2: Wenige beobachtete Schadenfälle

In einem Nutzungjahr sind 5 Schadensfälle eingetreten.

Auf welchen Bereich der zu erwartenden Anzahl der Schadensfälle lässt sich 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.

In einem Nutzungjahr sind 5 Schadensfälle eingetreten.

Auf welchen Bereich der zu erwartenden Anzahl der Schadensfälle lässt sich für die nächsten 10 Nutzungsjahre mit den Irrtumswahrscheinlichkeiten $\alpha_1 = \alpha_2 = 1\%$ *schließen?*

Annahme Poissonverteilung. Erwartungswertbereich aus der Tabelle ablesen. Für 10 Jahre zehnfache zu erwartende Schadensanzahl.

$[\lambda_L, \lambda_U]$	$k_{AV} = 4$	$k_{AV} = 5$	$k_{AV} = 6$
$\alpha_1 = \alpha_2 = 0{,}5\%$	[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: Kein Schadensfall in der Vergangenheit

In den vergangen 10 Jahren ist kein Schaden eingetreten.

Wie groß ist die zu erwartende Anzahl der Schadensfälle in den nächsten 2 Jahren. Irrtumswahrschieinlichkeit $\alpha = 2\%$ *(*$\alpha_1 = \alpha_2$*).*

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.

In den vergangen 10 Jahren ist kein Schaden eingetreten.

Wie groß ist die zu erwartende Anzahl der Schadensfälle in den nächsten 2 Jahren. Irrtumswahrschieinlichkeit $\alpha = 2\%$ ($\alpha_1 = \alpha_2$).

$[\lambda_L,\ \lambda_U]$	$k_{AV} = 0$	$k_{AV} = 1$	$k_{AV} = 2$	$k_{AV} = 3$
$\alpha_1 = \alpha_2 = 0{,}5\%$	[0,005, *]	[0,10, 5,30]	[0,34, 7,43]	[0,67, 9,27]
$\alpha_1 = \alpha_2 = 1\%$	[0,01, *]	[0,15, 4,60]	[0,44, 6,64]	[0,82, 8,41]

Bereich des Erwartungswerts für 10 Jahre mit der Obergrenze für $k_{AV} = 1$:

$$[\lambda_L,\ \lambda_U] = [0{,}01,\ 4{,}6]$$

Für die nächsten zwei Jahre davon 1/5:

$$[\lambda_L,\ \lambda_U] = [0{,}002,\ 0{,}91]$$

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 4.4: Fehlfunktionsrate

Bei $\#DS = 10^5$ Service-Leistungen wurden $k_{\mathrm{AV}} = 3$ 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).
$\left[\frac{\mathrm{MF}}{\mathrm{DS}}\right]$	Zählwertverhältnis in Fehlfunktionen je erbrachte Service-Leistung.

Bei $\#DS = 10^5$ Service-Leistungen wurden $k_{\mathrm{AV}} = 3$ 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?

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

Abschätzbarer Bereich der Fehlfunktionsrate:

$$\zeta_{\mathrm{L}} = \frac{\lambda_{\mathrm{L}}}{\#DS} = 0{,}82 \cdot 10^{-5}\ [\mathrm{MF}/\mathrm{DS}]$$

$$\zeta_{\mathrm{U}} = \frac{\lambda_{\mathrm{U}}}{\#DS} = 8{,}41 \cdot 10^{-5}\ [\mathrm{MF}/\mathrm{DS}]$$

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

$\lambda_{\mathrm{L}}, \lambda_{\mathrm{U}}$ Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.

$\zeta_{\mathrm{L}}, \zeta_{\mathrm{U}}$ Untere und obere Bereichgrenze der geschätzten Fehlfunktionsrate.

Beispiel 4.5: Maskierungswahrscheinlichkeit

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

Welchen Bereich der Maskierungswahrscheinlichkeit ergibt sich aus diesem Versuchsergebnis für eine Irrtumswahrscheinlichkeiten $\alpha_1 = \alpha_2 = 10\%$?

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

Abschätzbarer Bereich der Maskierungswahrscheinlichkeit:

$$p_{\mathrm{ML}} = \frac{\lambda_{\mathrm{L}}}{\#MF} = 3{,}15 \cdot 10^{-4}$$
$$p_{\mathrm{MU}} = \frac{\lambda_{\mathrm{U}}}{\#MF} = 7{,}99 \cdot 10^{-4}$$

$\alpha_1,\ \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
$\lambda_{\mathrm{L}},\ \lambda_{\mathrm{U}}$	Minimaler bzw. maximaler Erwartungswert für einen poisson-verteilten Ist-Wert.
$p_{\mathrm{ML}},\ p_{\mathrm{MU}}$	Untere und obere Bereichsgrenze der geschätzten Maskierungswahrscheinlichkeit.
$\#MF$	Anzahl der Fehlfunktionen (Number of malfunctions).

Beispiel 4.6: Zuverlässigkeitsbereich

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

Auf welchen Bereich der Zuverlässigkeit kann nach diesem Versuchsergebnis mit den Irrtumswahrscheinlichkeiten $\alpha_1 = \alpha_2 = 10\%$ geschlussfolgert 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 $k_{\mathrm{AV}} = 6$ Fehlfunktionen beobachtet.

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

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

- Abschätzbarer Bereich der MF-Rate:

$$\zeta_{\mathrm{L}} = 3{,}89 \cdot 10^{-3}\ [\mathrm{MF}/\mathrm{DS}]$$
$$\zeta_{\mathrm{U}} = 9{,}28 \cdot 10^{-3}\ [\mathrm{MF}/\mathrm{DS}]$$

- Daraus folgender Bereich der Zuverlässigkeit:

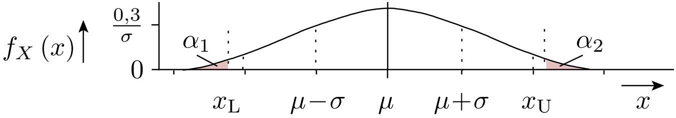
$$R_{\mathrm{L}} = \frac{1}{\zeta_{\mathrm{U}}} = 108\ [\mathrm{DS}/\mathrm{MF}]$$
$$R_{\mathrm{U}} = \frac{1}{\zeta_{\mathrm{L}}} = 257\ [\mathrm{DS}/\mathrm{MF}]$$

ζ_{L}, ζ_{U} Untere und obere Bereichgrenze der geschätzten Fehlfunktionsrate.

R_{L}, R_{U} Untere und obere Bereichgrenze der geschätzten Zuverlässigkeit.

Normalverteilung

4.61 Normalverteilung



Die Summe sehr vieler unabhängiger Zufallsvariablen strebt unter sehr allgemeinen Bedingungen gegen eine Normalverteilung:

- keine dominanten Summanden, ...
- für Zählwerte, poisson- und binomialverteilte Zufallsvariablen:

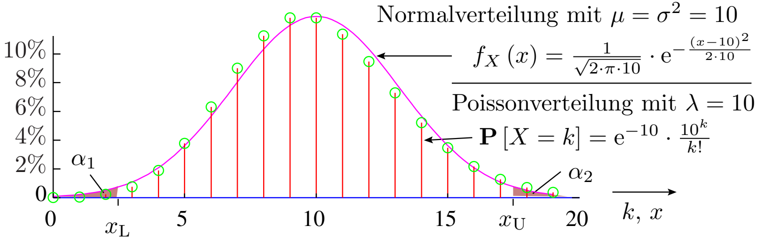
$$3\sigma \leq \mu \leq n - 3\sigma \tag{4.43}$$

Dichtefunktion:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \mathrm{e}^{-\frac{(x-\mu)^2}{2\cdot\sigma^2}} \tag{4.44}$$

μ, σ	Erwartungswert, Standardabweichung.
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.
n	Anzahl der Zählversuche, maximaler Zählwert.

4.62 NVT-Annäherung für Poissonverteilung

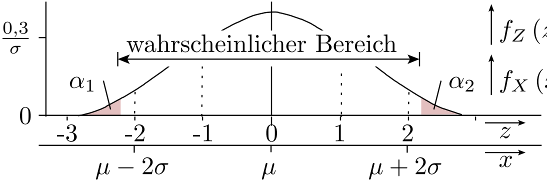


Poissonverteilung leichte Unsymmetrie. Größere $\mu = \lambda$ bessere Übereinstimmung. Hinreichende Ähnlichkeit der Beziehungen:

$$\alpha_1 = f(x_L) \quad \text{und} \quad \alpha_2 = f(x_U)$$

$\mu,\ \sigma$	Erwartungswert, Standardabweichung der Normalverteilung.
λ	Parameter der Poisson-Verteilung (Erwartungswert und gleichzeitig Varianz).
$x_L,\ x_U$	Untere und obere Schranke des wahrscheinlichen Bereichs von X.
$\alpha_1,\ \alpha_2$	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.
NDT	Normalverteilung (Normal distribution).

4.63 Standardisierte Normalverteilung



$$f_Z(z) = \varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2}\cdot z^2}$$

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

Die lineare Transformation

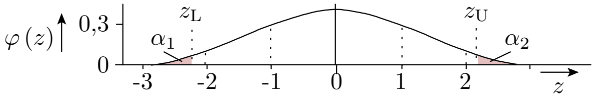
$$Z = \frac{X-\mu}{\sigma} \tag{4.45}$$

wandet normalverteilte in standardisiert normalverteilte Zufallsvariablen mit Erwartungswert $\mu = 0$ und Standardabweichung $\sigma = 1$ um. Dichtefunktion:

$$\varphi(z) = \frac{2}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{z^2}{2}} \tag{4.46}$$

μ, σ	Erwartungswert, Standardabweichung.
X, Z	normal verteilte, standardisiert normalverteilte Zufallsvariable.
$\varphi(z)$	Dichtefunktion der standardisierten Normalverteilung.

4.64 Tabelle der Verteilungsfunktion



Bereichsschätzung benötigen die Verteilungsfunktion

$$\Phi(z) = \int_{-\infty}^{z} \varphi(u) \cdot \mathrm{d}u \qquad (4.47)$$

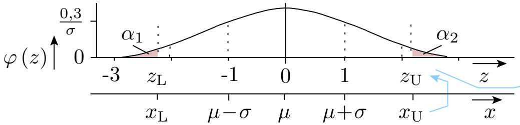
Tabellenfunktion $\Phi(z)$ für $z \geq 0$ mit einer Nachkommastelle:

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

Wegen Symmetrie gilt für $z < 0$:

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

4.65 Irtumwahrscheinl. zur Bereichsgrenze



$\alpha_2 = 1 - 0{,}9821 = 1{,}79\%$

$z_U = 2{,}1$

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

$$z_L = \frac{x_L - \mu}{\sigma} \tag{4.48}$$

$$z_U = \frac{x_U - \mu}{\sigma} \tag{4.49}$$

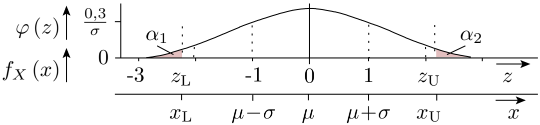
$$\alpha_1 = 1 - \Phi(-z_L) = 1 - \Phi\left(\frac{\mu - x_L}{\sigma}\right) \tag{4.50}$$

$$\alpha_2 = 1 - \Phi(z_U) = 1 - \Phi\left(\frac{x_U - \mu}{\sigma}\right) \tag{4.51}$$

---

z_L, z_U Transformierte untere und obere Bereichsgrenze.

4.66 Bereichsgrenze zur Irrtumswahrscheinl.



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

$\alpha_2 = 1\%$ $\quad z_U = 2{,}3\ldots 2{,}4 \longrightarrow x_U = \mu + \sigma \cdot z_U$

- Suche ausreichend kleines z_L für Ungleichung $1 - \Phi(-z_L) \leq \alpha_1$
- und ausreichend großes z_U für Ungleichung $1 - \Phi(z_U) \leq \alpha_2$
- Transformation $[z_L, z_U]$ nach $[x_L, x_U]$.

z_L, z_U	Transformierte untere und obere Bereichsgrenze.
μ, σ	Erwartungswert, Standardabweichung.
α_1, α_2	Irrtumswahrscheinlichkeit, Wert unterhalb bzw. oberhalb des geschätzten Bereichs.

4.67 Tabelle der inversen standardisierte NVT

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:

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

$$z_{\mathrm{L}} = \Phi^{-1}\left(\alpha_1\right) = -\Phi^{-1}\left(1-\alpha_1\right) \tag{4.52}$$

$$z_{\mathrm{U}} = \Phi^{-1}\left(1-\alpha_2\right) \tag{4.53}$$

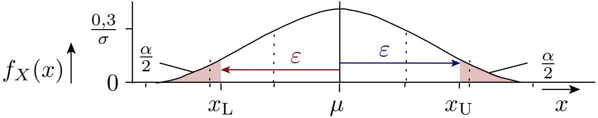
$$x_{\mathrm{L}} = \mu + \sigma \cdot z_{\mathrm{L}} = \mu - \sigma \cdot \Phi^{-1}\left(1-\alpha_1\right) \tag{4.54}$$

$$x_{\mathrm{U}} = \mu + \sigma \cdot z_{\mathrm{U}} = \mu + \sigma \cdot \Phi^{-1}\left(1-\alpha_2\right) \tag{4.55}$$

$x_{\mathrm{L}}, x_{\mathrm{U}}$ Untere und obere Schranke des wahrscheinlichen Bereichs von X.

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

4.68 Symmetrischer Bereich



Sonderfall $\alpha_1 = \alpha_2$, siehe auch Prozesszentrierung (Folie 2.107):

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

$$\varepsilon = \sigma \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{4.56}$$

$$\mathrm{sr} = [x_{\mathrm{L}}, x_{\mathrm{U}}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{4.57}$$

$f_X(x)$	Dichtefunktion der Zufallsvariablen X mit den möglichen Zählwerten x.
μ, x_{L}, x_{U}	Erwartungswert, untere und oberen Bereichsgrenze.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.
ε, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
sr	Symmetrischer Bereich der wahrscheinlichen Werte.
$\Phi^{-1}(..)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.

Beispiel 4.7: Bereichschätzung Normalverteilung

Zufallsvariable X, $\mu = 20$, $\sigma = 5$. Ein Beispiele für jede der sechs behandelten Typen von Schätzaufgaben.

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

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

c) *Welche obere Schranke x_{U} wird nur mit einer Irrtumswahrscheinlichkeit $\alpha_2 \leq 1\%$ überschritten?*

d) *Welche untere Schranke x_{L} wird nur mit einer Irrtumswahrscheinlichkeit $\alpha_1 \leq 2\%$ unterschritten?*

e) *Symmetrischer $\varepsilon = 3\sigma$ Bereich und Irrtumswahrscheinlichkeit?*

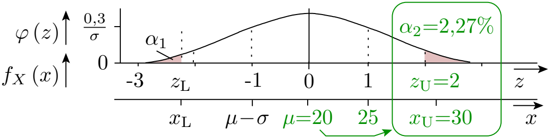
f) *Symmetrischer für Irrtumswahrscheinlichkeit $\alpha = 2\%$?*

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.
$\Phi(z)$	Verteilungsfunktion der standardisierten Normalverteilung.

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

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

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



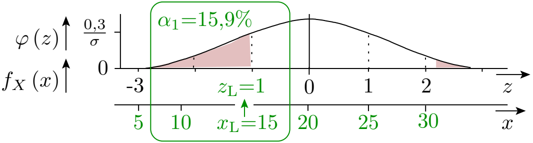
z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0,5000	0,5398	0,5793	0,6179	0,6554	0,6915	0,7257	0,7580	0,7881	0,8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0,9772	0,9821	0,9861	0,9893	0,9918	0,9938	0,9953	0,9965	0,9974	0,9981
3,...	0,9987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

$$\alpha_2 = 1 - \Phi\left(\frac{30-20}{5}\right) = 1 - \Phi(2) = 2{,}27\%$$

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

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

(4.50) $$\alpha_1 = 1 - \Phi\left(-z_{\mathrm{L}}\right) = 1 - \Phi\left(\frac{\mu - x_{\mathrm{L}}}{\sigma}\right)$$



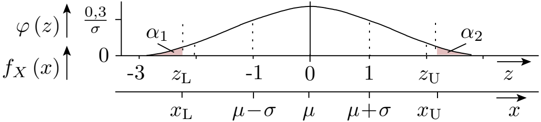
z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
0,...	0,5000	0,5398	0,5793	0,6179	0,6554	0,6915	0,7257	0,7580	0,7881	0,8159
1,...	0,8413	0,8643	0,8849	0,9032	0,9192	0,9332	0,9452	0,9554	0,9641	0,9713
2,...	0,9772	0,9821	0,9861	0,9893	0,9918	0,9938	0,9953	0,9965	0,9974	0,9981
3,...	0,9987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

$$\alpha_1 = 1 - \Phi\left(-\frac{15-20}{5}\right) = 1 - \Phi(1) = 15{,}9\%$$

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

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

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



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

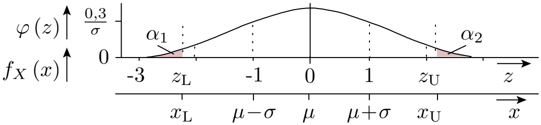
$$x_U = 20 + 5 \cdot 2{,}33 = 31{,}65$$

$\Phi^{-1}(..)$ 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?*

(4.54) $$x_{\mathrm{L}} = \mu + \sigma \cdot z_{\mathrm{L}} = \mu - \sigma \cdot \Phi^{-1}(1 - \alpha_1)$$



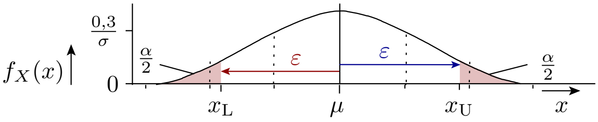
$\alpha_{1/2}$	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1-\alpha_{1/2})$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$x_{\mathrm{L}} = 20 - 5 \cdot 2{,}05 = 9{,}75$$

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

e) *Symmetrischer* $\varepsilon = 3\sigma$ *Bereich und Irrtumswahrscheinlichkeit?*

(4.57) $$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$



Sonderfall $\alpha_1 = \alpha_2$, siehe auch Prozesszentrierung (Folie 2.107):

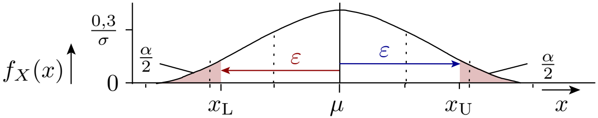
α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\mathrm{sr}_\mathrm{Z} = [z_\mathrm{L}, z_\mathrm{U}] = \mp 3\sigma$$
$$\alpha = 0{,}26\%$$

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

f) *Symmetrischer für Irrtumswahrscheinlichkeit* $\alpha = 2\%$*?*

(4.57)
$$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$



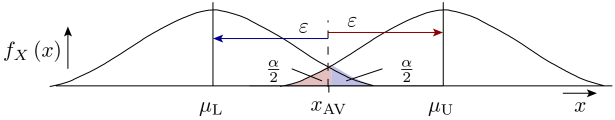
α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\begin{aligned}\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] &= \mu \mp 2{,}33 \cdot \sigma\\ &= 20 \mp 11{,}65\\ &= [8{,}35,\ 31{,}65]\end{aligned}$$

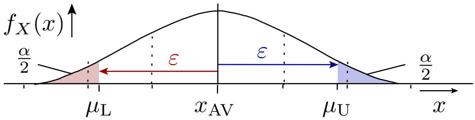
4.70 Symmetrischer Bereich des Erwartungswerts

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

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

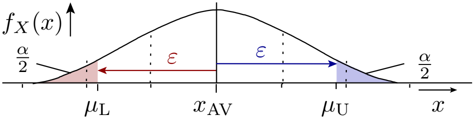


Verschiebung der Erwartungswerte beider Dichtfunktionen nach x_{AV}:



Gleicher Intervallradius wie bei symmetrischen Bereich um μ.

$f_X(x)$	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.



Der experimentell bestimmte Werte x_{AV} ist jetzt der Erwartungswert und der Intervallradius ε, in dem mit einer Irrtumswahrscheinlichkeit α der zu schätzende Erwartungswert liegt, berechnet sich wie der Intervallradius wahrscheinlicher Werte um einen Erwartungswert:

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

Symmetrischer Bereich mit x_{AV} als Schätzwert (vergl. :

$$\hat{\mathrm{sr}}_\mu = [\mu_L, \mu_U] = x_{AV} \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \qquad (4.58)$$

$f_X(x)$	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.
sr_μ	Symmetrischer Bereich des wahrscheinlichen Erwartungswerts.
σ	Standardabweichung.
$\Phi^{-1}(..)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.

Beispiel 4.8: Bereichschätzung Erwartungswert

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

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

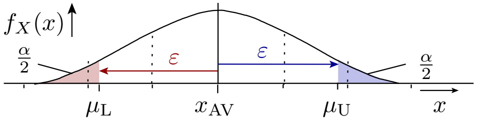
---

x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
σ	Standardabweichung.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.

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

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

(4.58) $$\hat{\mathrm{sr}}_\mu = [\mu_\mathrm{L}, \mu_\mathrm{U}] = x_{\mathrm{AV}} \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$



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

$$\mathrm{sr}_\mu = 100 \mp 10 \cdot 2{,}33 = 100 \mp 23{,}3$$

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

Normalverteilte Zählwerte

4.73 Zu erwartender Zählwertbereich $(p \cdot n \to \mathrm{sr})$

Mit dem Erwartungswert und der Standardabweichung für Zählwerte:

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

(4.35) $$\sigma^2 = \kappa \cdot n \cdot p \cdot (1-p)$$

(siehe Binomialverteilungsnäherung) und

(4.54) $$x_\mathrm{L} = \mu + \sigma \cdot z_\mathrm{L} = \mu - \sigma \cdot \Phi^{-1}(1-\alpha_1)$$

(4.55) $$x_\mathrm{U} = \mu + \sigma \cdot z_\mathrm{U} = \mu + \sigma \cdot \Phi^{-1}(1-\alpha_2)$$

(4.57) $$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1-\tfrac{\alpha}{2}\right)$$

ergeben sich folgende Bereichsgrenzen:

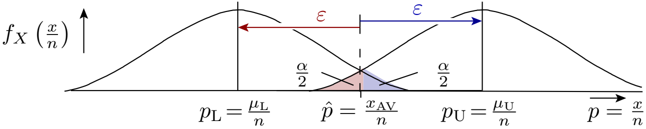
$$x_\mathrm{L} = p \cdot n - \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1-\alpha_1) \quad (4.59)$$

$$x_\mathrm{U} = p \cdot n + \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1-\alpha_2) \quad (4.60)$$

$$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}\left(1-\tfrac{\alpha}{2}\right) \quad (4.61)$$

p	Mittlere Eintrittswahrscheinlichkeit der zu zählenden Ereignisse.
n	Anzahl der Zählversuche, maximaler Zählwert.
μ, x_L, x_U	Erwartungswert, untere und oberen Bereichsgrenze.

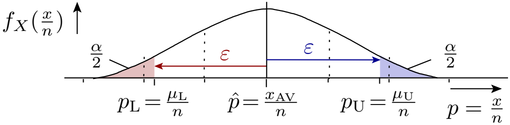
4.74 Eintrittswahrscheinlichkeit

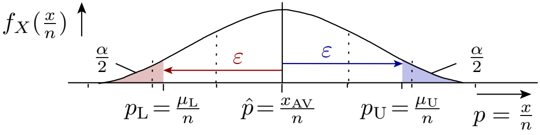


Die zu erwartende (mittlere) Eintrittswahrscheinlichkeit ist zu erwartender Zählwert durch Anzahl der Zählversuche n. Schätzwert für einen einzelnen Zählwert x_{AV} als Erwartungswert:

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

Zur Abschätzung des Intervallradius, um den der Schätzwert vom Erwartungswert abweichen kann, Verschiebung der Erwartungswerte beiden Dichtefunktionen nach x_{AV}/n:





Intervallradius:

(4.38) $$\hat{\sigma} = \sqrt{\kappa \cdot x_{AV} \cdot (1 - \hat{p})}$$

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

$$\hat{\varepsilon} = \frac{\sqrt{\kappa \cdot x_{AV} \cdot (1-\hat{p})}}{n} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \qquad (4.62)$$

Wahrscheinlicher Bereich der Eintrittswahrscheinlichkeit:

$$\hat{\mathrm{sr}}_{\mathrm{p}} = [p_L, p_U] = \hat{p} \mp \hat{\varepsilon}$$

$$\hat{\mathrm{sr}}_{\mathrm{p}} = \frac{x_{AV} \mp \sqrt{\kappa \cdot x_{AV} \cdot (1-\hat{p})}}{n} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \qquad (4.63)$$

$\hat{\mathrm{sr}}_{\mathrm{p}}$	Geschätzter symmetrischer Bereich der Eintrittswahrscheinlichkeit.
p_L, p_U	Unter und ober Schranke der geschätzter Eintrittswahrscheinlichkeit.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
$\Phi^{-1}(..)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.

Zukünftige für bekannte ZW

4.76 Zukünftige für bekannte Zählwerte

Verkettung von zwei Zählexperimenten mit gleicher Eintrittswahrsch. p:

1. Bestimmung eines aktuellen Zählwert x_{AV} für n_{AV} Zählversuche.
2. Bestimmung zukünftiger Zählwerte x_{UC} für n_{UC} Zählversuche.

	Zufallsvariable	Erwartungswert	Varianz nach (Gl. 4.35)
1	X_{AV}	$n_{\mathrm{AV}} \cdot p$	$n_{\mathrm{AV}} \cdot \kappa \cdot p \cdot (1-p)$
2	X_{UC}	$n_{\mathrm{UC}} \cdot p$	$n_{\mathrm{UC}} \cdot \kappa \cdot p \cdot (1-p)$

Die zufällige Eintrittshäufigkeit X/n_{NX} des Gesamtversuchs ist die zufällige geschätzte Eintrittshäufigkeit aus Versuch 1 plus der zufällige Abweichung vom Schätzwert aus Versuch 1:

$$\begin{aligned} \frac{X}{n_{\mathrm{UC}}} &= \frac{X_{\mathrm{AV}}}{n_{\mathrm{AV}}} + \left(\frac{X_{\mathrm{UC}}}{n_{\mathrm{UC}}} - \frac{x_{\mathrm{AV}}}{n_{\mathrm{AV}}}\right) \\ X &= \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot X_{\mathrm{AV}} + X_{\mathrm{UC}} - \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \end{aligned} \tag{4.64}$$

Summen und Linearkombinationen normalverteilter Zufallsvariablen sind gleichfalls normalverteilt.

4.77 Erwartungswert und Varianz

	Zufallsvariable	Erwartungswert	Varianz nach (Gl. 4.35)
1	X_{AV}	$n_{\text{AV}} \cdot p$	$n_{\text{AV}} \cdot \kappa \cdot p \cdot (1-p)$
2	X_{UC}	$n_{\text{UC}} \cdot p$	$n_{\text{UC}} \cdot \kappa \cdot p \cdot (1-p)$

(4.64) $$X = \frac{n_{\text{UC}}}{n_{\text{AV}}} \cdot X_{\text{AV}} + X_{\text{UC}} - \frac{n_{\text{UC}} \cdot x_{\text{AV}}}{n_{\text{AV}}}$$

$$\mathbb{E}[X] = \mu = n_{\text{UC}} \cdot p + n_{\text{NX}} \cdot p - n_{\text{UC}} \cdot p = n_{\text{NX}} \cdot p$$

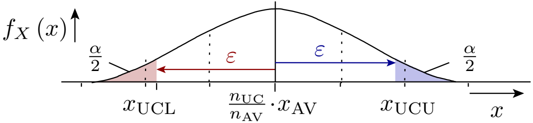
$$\operatorname{Var}[X] = \sigma^2 = \left(\frac{n_{\text{UC}}}{n_{\text{AV}}}\right)^2 \cdot n_{\text{AV}} \cdot \kappa \cdot p \cdot (1-p) + n_{\text{UC}} \cdot \kappa \cdot p \cdot (1-p) + 0$$

$$\sigma^2 = n_{\text{UC}} \cdot \kappa \cdot \left(1 + \frac{n_{\text{UC}}}{n_{\text{AV}}}\right) \cdot p \cdot (1-p) \tag{4.65}$$

Zur Abschätzung zukünftiger Zählwerte für einen bekannten Zählwert ist die Varianz gegenüber einer Abschätzung mit bekannter Eintrittswahrscheinlichkeit um einen Faktor $1 + \frac{n_{\text{UC}}}{n_{\text{AV}}}$ größer.

- x_{AV}, n_{AV} Ist-Zählwert, Anzahl der zugehörigen Zählversuche.
- x_{UC}, n_{UC} Zukünftige (Upcoming) Zählwerte, Anzahl der zugehörigen Zählversuche.
- μ, σ Erwartungswert, Standardabweichung.
- p, κ Eintrittswahrscheinlichkeit, Varianzerhöhung.
- X_{AV}, X_{UC} Zufallsvariable für den aktuellen Wert zur Schätzung von p, für zukünftige Zählwerte.

4.78 Bereich zukünftiger Zählwerte



Schätzer Eintrittswahrscheinlichkeit und Erwartungswert:

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

$$\hat{\mu} = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot x_{\mathrm{AV}} \qquad (4.66)$$

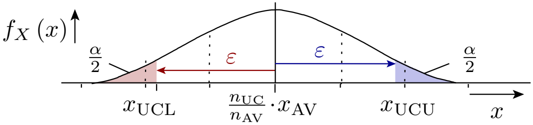
Schätzwert Standardabweichung, Intervallradius:

(4.65) $$\sigma^2 = n_{\mathrm{UC}} \cdot \kappa \cdot \left(1 + \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}}\right) \cdot p \cdot (1-p)$$

$$\hat{\sigma} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}}\right) \cdot \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \cdot \left(1 - \frac{x_{\mathrm{AV}}}{n_{\mathrm{AV}}}\right)} \qquad (4.67)$$

---

x_{AV}, n_{AV} Ist-Zählwert, Anzahl der zugehörigen Zählversuche.
x_{UC}, n_{UC} Zukünftige (Upcoming) Zählwerte, Anzahl der zugehörigen Zählversuche.
μ, σ, ε Erwartungswert, Standardabweichung, Intervallradius.



Intervallradius:

(4.67) $$\hat{\sigma} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\text{UC}}}{n_{\text{AV}}}\right) \cdot \frac{n_{\text{UC}} \cdot x_{\text{AV}}}{n_{\text{AV}}} \cdot \left(1 - \frac{x_{\text{AV}}}{n_{\text{AV}}}\right)}$$

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

$$\hat{\varepsilon} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\text{UC}}}{n_{\text{AV}}}\right) \cdot \frac{n_{\text{UC}} \cdot x_{\text{AV}}}{n_{\text{AV}}} \cdot \left(1 - \frac{x_{\text{AV}}}{n_{\text{AV}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \tag{4.68}$$

Wahrscheinlicher Bereich zukünftiger Zählwerte:

$$\hat{\text{sr}}_{\text{UC}} = \frac{n_{\text{UC}}}{n_{\text{AV}}} \cdot x_{\text{AV}} \mp \hat{\varepsilon} \tag{4.69}$$

Im Vergleich zu einer Abschätzung mit bekanntem Erwartungswert sind Intervallradius und Bereich um den Faktor $1 + \frac{n_{\text{UC}}}{n_{\text{AV}}}$ größer.

---

x_{AV}, n_{AV}	Ist-Zählwert, Anzahl der zugehörigen Zählversuche.
x_{UC}, n_{UC}	Zukünftige (Upcoming) Zählwerte, Anzahl der zugehörigen Zählversuche.
μ, σ	Erwartungswert, Standardabweichung.
$\hat{.}$, sr	Schätzwert, symmetrischer Bereich.

Beispiel 4.9: Bereich MF-Rate und künftige MF-Anzahl

Bei der Abarbeitung von 20.000 Service-Anforderungen wurden 100 Fehlfunktionen beobachtet. Keine Abhängigkeiten. Zugelassene Irrtumswahrscheinlichkeit 2%.

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

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

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

n_{AV}	Anzahl der Zählversuche, mit denen x_{AV} bestimmt wurde.
$[\mathrm{DS}]$	Zählwert in erbrachten Service-Leistungen.
x_{AV}	Experimentell bestimmter Ist-Zählwert, Schätzwert für den Erwartungswert.
$[\mathrm{MF}]$	Zählwert in Fehlfunktionen.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.
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.

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

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

(4.70) $$\hat{\varepsilon}_{\mathrm{r}} = \frac{\hat{\varepsilon}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \text{ für } \hat{p} \le 0{,}5$$

(4.72) $$\hat{\mathrm{sr}}_{\mathrm{p}} = [p_{\mathrm{L}}, p_{\mathrm{U}}] = \frac{x_{\mathrm{AV}}}{n} \cdot (1 \mp \hat{\varepsilon}_{\mathrm{r}}) \text{ für } \hat{p} \le 0{,}5$$

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

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

$$\hat{\varepsilon}_{\mathrm{r}} = \Phi^{-1}(1-1\%) \cdot \sqrt{\frac{1}{100} - \frac{1}{20.000}} = 0{,}232$$

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

$\Phi^{-1}(..)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
$\hat{p}$, $\hat{\mathrm{sr}}_{\mathrm{p}}$	Schätzer für Erwartungswert und symmetrischen Bereich der Wahrscheinlichkeit.
$\hat{\zeta}$, $\hat{\mathrm{sr}}_{\zeta}$	Schätzer für Erwartungswert und symmetrischen Bereich der Fehlfunktionsrate.
ε_{r}	Geschätzter Intervallradius realtiv zum erwarteten Eintritts-Zählwert.

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

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

$$\hat{\varepsilon} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}}\right) \cdot \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \cdot \left(1 - \frac{x_{\mathrm{AV}}}{n_{\mathrm{AV}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \tag{4.68}$$

$$\hat{\mathrm{sr}}_{\mathrm{UC}} = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot x_{\mathrm{AV}} \mp \hat{\varepsilon} \tag{4.69}$$

$$\frac{n_{\mathrm{AV}}}{n_{\mathrm{AV}}} = \frac{10.000\ [\mathrm{DS}]}{20.000\ [\mathrm{DS}]} = 0{,}5$$

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

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

$$\hat{\varepsilon} = \sqrt{(1 + 0{,}5) \cdot 0{,}5 \cdot 100 \cdot (1 - 0{,}5\%)} \cdot 2{,}33 = 27{,}8$$

$$\hat{\mathrm{sr}}_{\mathrm{UC}} = 50 \mp 27{,}8$$

$\hat{\mu}_{\mathrm{NX}}$ Erwartungswert des zu schätzenden Zählwerts.

$\mathrm{sr}_{\mathrm{NX}}$ Symmetrischer Bereich zukünftiger Zählergebnisse zu einem bekannten Ist-Zählwert.

Erforderliche Zählwertgröße

4.81 Relativer Intervallradius

Die Schätzgenauigkeit wird durch Irrtumswahrscheinlichkeit und den relativen Intervallradius bezogen auf die Eintritts- bzw. die Nichteintritts-wahrscheinlichkeit beschrieben:

$$\hat{\sigma}_{\mathrm{r}} = \frac{\hat{\sigma}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1-\hat{p})} \quad \text{für } \hat{p} = \frac{x_{\mathrm{AV}}}{n} \leq 0{,}5$$

(4.39)

$$\hat{\sigma}_{\tilde{\mathrm{r}}} = \frac{\hat{\sigma}}{n-x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}} \cdot \hat{p}} \quad \text{für } \hat{p} = \frac{x_{\mathrm{AV}}}{n} > 0{,}5$$

Relativer Intervallradius:

$$\hat{\varepsilon}_{\mathrm{r}} = \frac{\hat{\varepsilon}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \text{ für } \hat{p} \leq 0{,}5 \tag{4.70}$$

$$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n-x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \text{ für } \hat{p} > 0{,}5 \tag{4.71}$$

Symmetrischer Bereich:

$$\hat{\mathrm{sr}}_{\mathrm{p}} = [p_{\mathrm{L}}, p_{\mathrm{U}}] = \frac{x_{\mathrm{AV}}}{n} \cdot (1 \mp \hat{\varepsilon}_{\mathrm{r}}) \text{ für } \hat{p} \leq 0{,}5 \tag{4.72}$$

$$\hat{\mathrm{sr}}_{\mathrm{p}} = [p_{\mathrm{L}}, p_{\mathrm{U}}] = 1-\left(1-\frac{x_{\mathrm{AV}}}{n}\right) \cdot (1 \mp \hat{\varepsilon}_{\tilde{\mathrm{r}}}) \text{ für } \hat{p} > 0{,}5 \tag{4.73}$$

$\hat{\varepsilon}_{\mathrm{r}}$, $\hat{\varepsilon}_{\tilde{\mathrm{r}}}$ Geschätzter Intervallradius realtiv zum erwarteten Eintritts- bzw. Nichteintritts-Zählwert.

$\hat{\mathrm{sr}}_{\mathrm{p}}$ Geschätzter symmetrischer Bereich der Eintrittswahrscheinlichkeit.

n, x_{AV} Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.

4.82 Erforderliche Zählwertgröße

(4.70) $\hat{\varepsilon}_{\mathrm{r}} = \frac{\hat{\varepsilon}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ für $\hat{p} \le 0{,}5$

(4.71) $\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n-x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ für $\hat{p} > 0{,}5$

Mindestzählwert eingetretene und nicht eingetretene Ereignisse:

$$x_{\mathrm{AV}} \ge \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\mathrm{r}}^{-2} \cdot (1-\hat{p}) \quad \text{für } \hat{p} \le 0{,}5 \tag{4.74}$$

$$n - x_{\mathrm{AV}} \ge \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\tilde{\mathrm{r}}}^{-2} \cdot \hat{p} \quad \text{für } \hat{p} > 0{,}5 \tag{4.75}$$

Grobabschätzung: $\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \approx 5\ldots 10$, $0{,}5 \le (1-\hat{p}) \le 1$ und $0{,}5 \le \hat{p} \le 1$. Geeigneter Zählwertbereich (ACR, Folie 1.18) bedeutet:

$$5\ldots 10 \cdot \varepsilon_{\mathrm{r}}^{-2} \le x_{\mathrm{AV}} \le n - 5\ldots 10 \cdot \varepsilon_{\mathrm{r}}^{-2}$$

Erforderliche Anzahl der Zählversuche (Gl. 4.37):

$$n \ge x_{\mathrm{AV}}/\hat{p}$$

ACR	Brauchbare Schätzwerte nur bei geeigneten Zählwertgrößen.
ε_{r}, $\varepsilon_{\tilde{\mathrm{r}}}$	Intervallradius realtiv zum erwarteten Eintritts- bzw. Nichteintritts-Zählwert.
n, x_{AV}	Anzahl der Zählversuche, Experimentell bestimmter Ist-Zählwert.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.

Beispiel 4.10: Erforderlicher Zählwertgröße

Relative Intervallradien ε_{r} bzw. $\varepsilon_{\tilde{\mathrm{r}}}$ von 20% und 2% für geschätzte Eintrittswahrscheinlichkeiten $\hat{p} \in \{10\%, 50\%, 90\%\}$. Irrtumswahrscheinlichkeit $\alpha = 4{,}52\%$, d.h. $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2$, $\kappa = 1$.

Erforderliche Zählwertgröße?

ε_{r}	Intervallradius realtiv zum erwarteten Eintritts-Zählwert.
$\hat{\varepsilon}_{\tilde{\mathrm{r}}}$	Geschätzter Intervallradius realtiv zum erwarteten Nichteintritts-Zählwerts.
$\hat{p}$	Schätzwert der Eintrittswahrscheinlichkeit.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.
$\Phi^{-1}(..)$	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.

Relative Intervallradien ε_{r} bzw. $\varepsilon_{\tilde{\mathrm{r}}}$ von 20% und 2% für geschätzte Eintrittswahrscheinlichkeiten $\hat{p} \in \{10\%, 50\%, 90\%\}$. Irrtumswahrscheinlichkeit $\alpha = 4{,}52\%$, d.h. $\Phi^{-1}\left(1-\frac{\alpha}{2}\right) = 2$, $\kappa = 1$.

Erforderliche Zählwertgröße?

(4.74) $$x_{\mathrm{AV}} \geq \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\mathrm{r}}^{-2} \cdot (1-\hat{p}) \quad \text{für } \hat{p} \leq 0{,}5$$

(4.75) $$n - x_{\mathrm{AV}} \geq \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\tilde{\mathrm{r}}}^{-2} \cdot \hat{p} \quad \text{für } \hat{p} > 0{,}5$$

ε_{r}	$\hat{p} = 10\%$		$\hat{p} = 50\%$		$\varepsilon_{\tilde{\mathrm{r}}}$	$\hat{p} = 90\%$	
	$x_{\mathrm{AV.min}}$	$n_{\min}$	$x_{\mathrm{AV.min}}$	$n_{\min}$		$x_{\mathrm{AV.max}}$	$n_{\min}$
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-) Eintrittsanzahl $\gtrsim 100$ und insbesondere für sehr kleine (Nicht-) Eintrittswahrscheinlichkeiten eine sehr große Versuchsanzahl.

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

Zusammenfassung

4.84 Schätzen einer Zählwertverteilung

Für unabhängige Zählwerte lassen sich Verteilung, Erwartungswert und Varianz aus den Eintrittswahrscheinlichkeiten p_i aller Zählversuche 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. Auch als Näherung für binomial- und poisson-verteilte Zufallsvariablen.

Alle drei Näherungen benötigen die individuellen p_i nicht.

Mögliche Abhängigkeiten zwischen den Zählwerten und deren Berücksichtigung in Bereichsschätzungen behandelt erst der Folgeabschnitt.

4.85 Binomialverteilung

Die Binomialverteilungsnäherung

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

$$\mathbb{E}\left[X\right]=\mu=n\cdot p \tag{4.31}$$

$$\operatorname{Var}\left[X\right]=\sigma^{2}=n\cdot p\cdot(1-p) \tag{4.32}$$

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

$$\sigma^{2}=\kappa\cdot n\cdot p\cdot(1-p) \tag{4.35}$$

$$\hat{\mu}=x_{\mathrm{AV}} \tag{4.36}$$

$$\hat{p}=\frac{\hat{\mu}}{n}=\frac{x_{\mathrm{AV}}}{n} \tag{4.37}$$

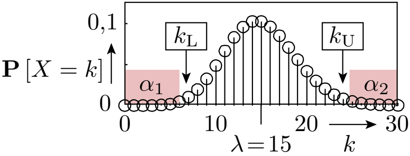
$$\hat{\sigma}=\sqrt{\kappa\cdot x_{\mathrm{AV}}\cdot(1-\hat{p})} \tag{4.38}$$

Varianzkoeffizienten für das Eintreten bzw. Nichteintreten von Zählwerten als Maße der Schätzgenauigkeit ab:

$$\begin{aligned}\hat{\sigma}_{\mathrm{r}}&=\frac{\hat{\sigma}}{x_{\mathrm{AV}}}=\sqrt{\frac{\kappa}{x_{\mathrm{AV}}}\cdot(1-\hat{p})}&&\text{für } \hat{p}=\frac{x_{\mathrm{AV}}}{n}\le 0{,}5\\ \hat{\sigma}_{\tilde{\mathrm{r}}}&=\frac{\hat{\sigma}}{n-x_{\mathrm{AV}}}=\sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}}\cdot\hat{p}}&&\text{für } \hat{p}=\frac{x_{\mathrm{AV}}}{n}> 0{,}5\end{aligned} \tag{4.39}$$

4.86 Poissonverteilung

Varianz gleich Erwartungswert gleich λ. Verteilungsfunktion:



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

Tabellierbar für Schätzaufgaben in der Form:

- Minimaler Erwartungswert λ, damit Zählwerte k mit Wahrscheinlichkeit $\geq 1-\alpha_1$ mindesten k_L ist:

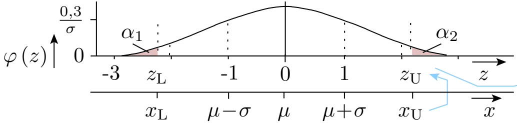
λ	$k_L=1$	$k_L=2$	$k_L=3$	$k_L=4$	$k_L=5$	$k_L=6$
$\alpha_1=1\%$	4,606	6,638	8,406	10,045	11,605	13,109

- maximaler Erwartungswert λ_L, damit Zählwert k mit Wahrscheinlichkeit $\geq 1-\alpha_2$ nicht größer k_U ist.

λ	$k_U=0$	$k_U=1$	$k_U=2$	$k_U=3$	$k_U=4$	$k_U=5$	...
$\alpha_2=1\%$	0,01	0,148	0,436	0,823	1,279	1,785	...

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.

4.87 NDT: Irrtumswahrsch. für Bereichsgrenzen



$\alpha_2 = 1 - 0{,}9821 = 1{,}79\%$

$z_U = 2{,}1$

z	...,0	...,1	...,2	...,3	...,4	...,5	...,6	...,7	...,8	...,9
2,...	0,9772	0,9821	0,9861	0,9893	0,9918	0,9938	0,9953	0,9965	0,9974	0,9981

- Transformation Bereichsgrenzen:

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

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

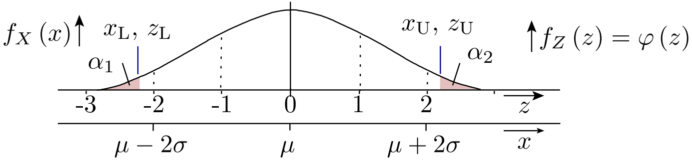
- Bestimmung der Irtumswahrscheinlichkeiten mit der NDT-Tabelle:

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

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

NDT Normalverteilung (Normal distribution).

4.88 NDT: Bereichsgrenzen für Irrtumswahrsch.

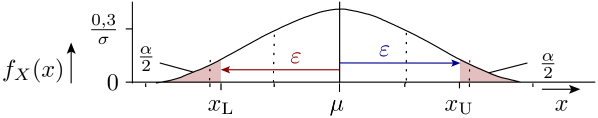


Inverse standardisierte Normalverteilung:

$\alpha_{1/2}$	2,27%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}\left(1-\alpha_{1/2}\right)$	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.
- (4.52) $$z_L = \Phi^{-1}(\alpha_1) = -\Phi^{-1}(1-\alpha_1)$$
  (4.53) $$z_U = \Phi^{-1}(1-\alpha_2)$$
- Transformation in die Bereichsgrenzen der Zufallsvariablen X:
  (4.54) $$x_L = \mu + \sigma \cdot z_L = \mu - \sigma \cdot \Phi^{-1}(1-\alpha_1)$$
  (4.55) $$x_U = \mu + \sigma \cdot z_U = \mu + \sigma \cdot \Phi^{-1}(1-\alpha_2)$$

4.89 NDT: Symmetrischer Bereich

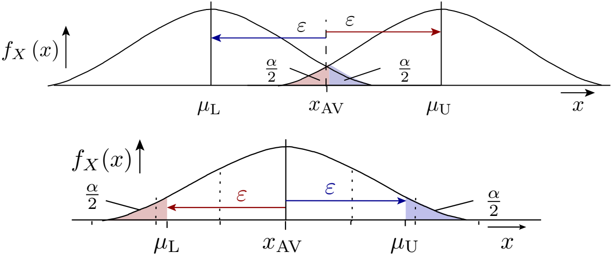


α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

$$\varepsilon = \sigma \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{4.56}$$

$$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{4.57}$$

4.90 NDT: Bereichsschätzung Erwartungswert

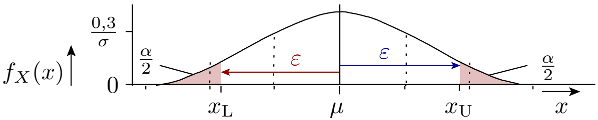


α	4,54%	0,26%	0	4%	2%	1%	0,4%	0,2%
$\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$	2	3	3,9	2,05	2,33	2,57	2,88	3,10

Symmetrischer Bereich um ein experimentell bestimmten Wert x_{AV}:

(4.58) $$\hat{\mathrm{sr}}_\mu = [\mu_L, \mu_U] = x_{AV} \mp \sigma \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$$

4.91 Normalverteilte Zählwerte



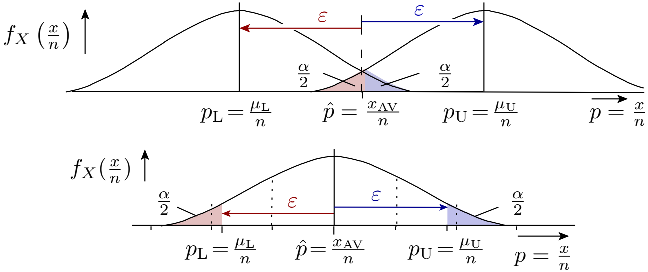
Bereichsgrenzen und symmetrischer Bereich für eine bekannte Eintrittswahrscheinlichkeit p:

(4.59) $$x_\mathrm{L} = p \cdot n - \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1-\alpha_1)$$

(4.60) $$x_\mathrm{U} = p \cdot n + \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}(1-\alpha_2)$$

(4.61) $$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = p \cdot n \mp \sqrt{\kappa \cdot n \cdot p \cdot (1-p)} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$$

4.92 Schätzen der Eintrittswahrscheinlichkeit



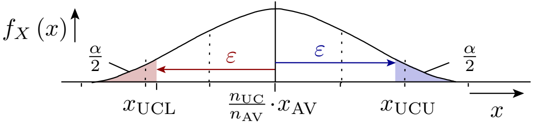
Erwartungswert, Intervallradius und symmtrischer Bereich zu einem experimentell bestimmten Zählwert x_{AV}:

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

(4.62) $$\hat{\varepsilon} = \frac{\sqrt{\kappa \cdot x_{\mathrm{AV}} \cdot (1-\hat{p})}}{n} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

(4.63) $$\hat{\mathrm{sr}}_{\mathrm{p}} = \frac{x_{\mathrm{AV}} \mp \sqrt{\kappa \cdot x_{\mathrm{AV}} \cdot (1-\hat{p})}}{n} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

4.93 Bereich zukünftiger Zählwerte



Linearkombination von zwei Zufallsvariablen:

1. skaliertes Ergebnis der Erwartungswertschätzung: $\frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot X_{\mathrm{AV}}$
2. Differenz zwischen zukünftigem Werte und Erwartungswert.

$$X = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot X_{\mathrm{AV}} + X_{\mathrm{UC}} - \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \tag{4.64}$$

Schätzer Erwartungswert, Intervallradius und Bereich:

$$\hat{\mu} = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot x_{\mathrm{AV}} \tag{4.66}$$

$$\hat{\varepsilon} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}}\right) \cdot \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \cdot \left(1 - \frac{x_{\mathrm{AV}}}{n_{\mathrm{AV}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \tag{4.68}$$

$$\hat{\mathrm{sr}}_{\mathrm{UC}} = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot x_{\mathrm{AV}} \mp \hat{\varepsilon} \tag{4.69}$$

4.94 Erforderliche Zählwertgröße

Relativer Intervallradius als Maß der Schätzgenauigkeit:

(4.70) $$\hat{\varepsilon}_{\mathrm{r}} = \frac{\hat{\varepsilon}}{x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{x_{\mathrm{AV}}} \cdot (1-\hat{p})} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \text{ für } \hat{p} \le 0{,}5$$

(4.71) $$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n-x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \text{ für } \hat{p} > 0{,}5$$

Erforderlicher Bereich der Zählwerte:

(4.74) $$x_{\mathrm{AV}} \ge \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\mathrm{r}}^{-2} \cdot (1-\hat{p}) \quad \text{für } \hat{p} \le 0{,}5$$

(4.75) $$n - x_{\mathrm{AV}} \ge \kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\tilde{\mathrm{r}}}^{-2} \cdot \hat{p} \quad \text{für } \hat{p} > 0{,}5$$

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

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

Abhängigkeiten

Varianzerhöhung

4.95 Varianzerhöhung durch Abhängigkeiten

Abhängigkeiten erhöhen Varianz und Standardabweichung und den Intervallradius der wahrscheinlichen Bereiche.

Wenn z.B. immer $a = 2, 3, \ldots$ Zählereignisse identisch nachweisbar sind, ist das beschreibbar durch eine Summe von $\#X/a$ unabhängige Zufallsvariablen mit den möglichen Werten 0 und a:

$$X = \sum_{i=1}^{\#X/a} X_i \quad \text{mit} \quad \mathbb{P}\left[X_i = k\right] = \begin{cases} 1 - p_i & k = 0 \\ p_i & k = a \end{cases}$$

Erwartungswert der Summanden:

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

Varianz der Summanden (nach Verschiebungssatz):

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

$\#X$ Anzahl der Zufallsvariablen.

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

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

Die Varianz der Summe ist a mal so groß wie * (* – Summe von $\#X$ unabhängige Zufallsvariablen, von denen je a gleichen Eintrittswahrscheinlichkeiten p_i haben):

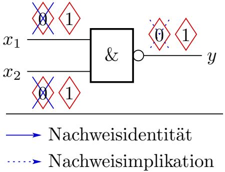
$$\mathrm{Var}[X] = \sigma^2 = \sum_{i=1}^{\#X/a} a^2 \cdot p_i \cdot (1-p_i) = a \cdot \left(\underbrace{\sum_{i=1}^{\#X/a} a \cdot p_i \cdot (1-p_i)}_{*} \right)$$

$$= \kappa \cdot \#X \cdot p \cdot (1-p) \quad \text{mit } \kappa \le a$$

Varianzvergrößerung auf max. die a–fache Varianz einer Binomialverteilung mit gleicher Versuchsanzahl und dem Mittelwert aller p_i als Eintrittswahrscheinlichkeit.

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

4.97 Varianzerhöhung



x_2	x_1	$\overline{x_2 \wedge x_1}$	sa0(x_1)	sa1(x_1)	sa0(x_2)	sa1(x_2)	sa0(y)	sa1(y)
0	0	1	1	1	1	1	0	1
0	1	1	1	1	1	0	0	1
1	0	1	1	0	1	1	0	1
1	1	0	1	0	1	0	0	1

Die Varianzerhöhung sei definiert als Verhältnis aus tatsächlicher Varianz und der Varianz einer Binomialverteilung mit derselben Versuchanzahl n und der mittleren Eintrittswahrscheinlichkeit p:

$$\kappa = \frac{\sigma^2}{n \cdot p \cdot (1-p)} = \frac{\sigma^2}{\mu \cdot \left(1 - \frac{\mu}{n}\right)} \tag{4.76}$$

Es gibt die vielfältigsten Nachweisabhängigkeiten:

- Identität, Implikation (vergl. Haftfehlernachweis Abschn. 2.1.5),
- gemeinsame Nachweisbedingungen, ...

μ, σ Erwartungswert, Standardabweichung.
n, κ Anzahl der Zählwerte, Varianzerhöhung.

4.98 Schätzen der Varianzerhöhung

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

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

- Schätzen der Varianz der Zählwertstichprobe:

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

- Geschätze Varianzerhöhung nach Gl. 4.76:

$$\hat{\kappa} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot \left(1 - \frac{\hat{\mu}}{n}\right)} \tag{4.77}$$

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

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.

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

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

(4.77) $$\hat{\kappa} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot \left(1 - \frac{\hat{\mu}}{n}\right)}$$

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

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

$$\hat{\kappa} = \frac{135}{63{,}1 \cdot \left(1 - \frac{63{,}1}{1000}\right)} = 2{,}28$$

Die Abhängigkeiten erhöhen die Varianz so, als ob mehr als 2 Zählereignisse 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.

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

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

---

x_{AV}, n_{AV}	Ist-Zählwert, Anzahl der zugehörigen Zählversuche.
x_{UC}, n_{UC}	Zukünftige (Upcoming) Zählwerte, Anzahl der zugehörigen Zählversuche.
μ, σ	Erwartungswert, Standardabweichung.
κ	Varianzerhöhung durch Abhängigkeiten, für unabhängige Zählwerte $\kappa \leq 1$.
[D]	Zählwert in Schadensfällen.

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

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

(4.68) $$\hat{\varepsilon} = \sqrt{\kappa \cdot \left(1 + \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}}\right) \cdot \frac{n_{\mathrm{UC}} \cdot x_{\mathrm{AV}}}{n_{\mathrm{AV}}} \cdot \left(1 - \frac{x_{\mathrm{AV}}}{n_{\mathrm{AV}}}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

(4.69) $$\hat{\mathrm{sr}}_{\mathrm{UC}} = \frac{n_{\mathrm{UC}}}{n_{\mathrm{AV}}} \cdot x_{\mathrm{AV}} \mp \hat{\varepsilon}$$

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

$n_{\mathrm{UC}}/n_{\mathrm{AV}} = 1$, $x_{\mathrm{AV}}/n_{\mathrm{AV}} \to 0$, $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2{,}33$:

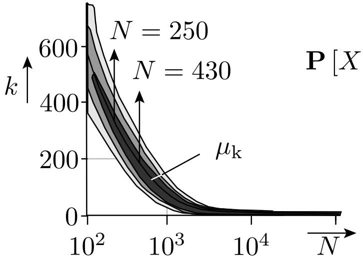
$$\begin{aligned} \hat{\mathrm{sr}}_{\mathrm{UC}} &= x_{\mathrm{AV}} \mp \sqrt{2 \cdot 2 \cdot x_{\mathrm{AV}}} \cdot 3{,}33 \\ &= 100 \mp 20 \cdot 2{,}33 = [53{,}4,\ 146{,}6] \end{aligned}$$

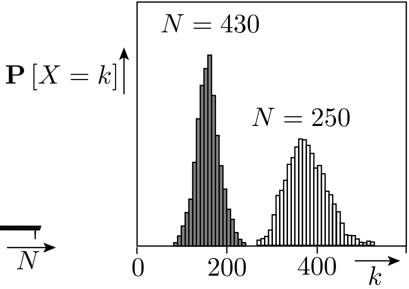
$\hat{.}$, sr   Schätzwert, symmetrischer Bereich.

Haftfehlernachweis

4.101 Experiment mit Haftfehlern

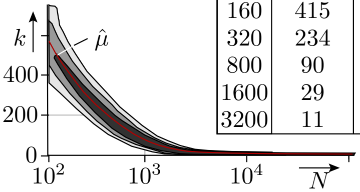
Kombinatorische Beispielschaltung (Benchmark c3540), simuliert mit $n = 3606$ unterschiedlich nachweisbaren Haftfehlern. Zählwert k ist die Anzahl der nicht nachweisbaren Haftfehler. Abschätzung der Verteilung $\mathbb{P}[X = k]$ mit einer Stichprobe von $\#v = 1000$ Zählwerten für verschiedene Zufallstestsätze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





k	Anzahl der nicht nachweisbaren Modellfehler.
n	Anzahl der Modellfehler.
N	Anzahl der Tests.

4.102 Varianzerhöhung im Experiment



N	$\hat{\mu}$	$\hat{\sigma^2}$	$\hat{\sigma}$	$\hat{\kappa}_1 = \frac{\hat{\sigma^2}}{\hat{\mu}}$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
160	415	1875	43,3	5,1	1,7	1,0
320	234	943	30,7	4,2	1,6	1,1
800	90	299	17,3	3,4	1,6	1,1
1600	29	52	7,2	1,8	*	1,3
3200	11	8,4	2,9	*	*	*

$*\ \hat{\sigma^2} < \hat{\mu}$

$\hat{\kappa}_1$: 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.

4.103 Fehlermodellierung und Vorhersagbarkeit

Testanzahl N	160	320	800
$\hat{\kappa}_1/3606 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$1{,}4 \cdot 10^{-3}$	$1{,}2 \cdot 10^{-3}$	$0{,}9 \cdot 10^{-4}$
$\hat{\kappa}_2/1000 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$1{,}7 \cdot 10^{-3}$	$1{,}6 \cdot 10^{-3}$	$1{,}6 \cdot 10^{-3}$
$\hat{\kappa}_3/300 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$3{,}3 \cdot 10^{-3}$	$3{,}6 \cdot 10^{-3}$	$3{,}6 \cdot 10^{-3}$

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

$$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n - x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n - x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ für } \hat{p} > 0{,}5 \tag{4.71}$$

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 $\varepsilon_{\tilde{\mathrm{r}}}$ als Maß der Schätzgenauigkeit nicht unbegrenzt durch mehr Modellfehler verringern.

$\hat{\varepsilon}_{\tilde{\mathrm{r}}}$	Geschätzter Intervallradius realtiv zum erwarteten Nichteintritts-Zählwerts.
n, p, κ	Anzahl der Zählwerte, Eintrittswahrscheinlichkeit, Varianzerhöhung.
$\Phi^{-1}(..)$	Inverse Funktion zur Verteilungsfunktion der standardisierten Normalverteilung.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.

4.104 Schlussfolgerungen

Testanzahl N	160	320	800
$\hat{\kappa}_1/3606 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$1{,}4 \cdot 10^{-3}$	$1{,}2 \cdot 10^{-3}$	$0{,}9 \cdot 10^{-4}$
$\hat{\kappa}_2/1000 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$1{,}7 \cdot 10^{-3}$	$1{,}6 \cdot 10^{-3}$	$1{,}6 \cdot 10^{-3}$
$\hat{\kappa}_3/300 \sim \hat{\varepsilon}_{\tilde{\mathrm{r}}}^2$	$3{,}3 \cdot 10^{-3}$	$3{,}6 \cdot 10^{-3}$	$3{,}6 \cdot 10^{-3}$

- 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 (siehe später Abschn. 6.1.2 ff.) sind vermutlich für Zufallstests nicht zielführend, weil die potentielle Zunahme der Schätzgenauigkeit mit der Modellfehleranzahl durch die wachsenen Nachweisabhängigkeiten zwischen den größere Modellfehlern zunichte gemacht wird.

4.105 Abschätzung mit mehreren Testsätzen

(4.71) $\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n-x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n-x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ für $\hat{p} > 0{,}5$

Alternative zur Verringerung des relativen Intervallradius für das Nichteintreten mit mehr Modellfehlern ist wie im Experiment (Folie 4.102) eine Bestimmung der Anzahl der nicht nachweisbaren Modellfehler x_{AV} für $\#v > 1$ unterschiedliche Zufallstestsätze. Vergrößert die Anzahl der Zählversuche n und die Zählwerte x_{AV} um $\#v$

$$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \sqrt{\frac{\kappa}{\#v} \cdot \left(\frac{1}{(n-x_{\mathrm{AV}})} - \frac{1}{n}\right)} \cdot \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{4.78}$$

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

$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.
$\hat{\varepsilon}_{\tilde{\mathrm{r}}}$	Geschätzter relativer Intervallradius der Anzahl der nichtnachweisbaren Fehler.
$\Phi^{-1}(..)$	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%.

$n = 1.000$ [F], $x_{\mathrm{AV}} = 1.000 - 32$ [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 Zufallstestsätz 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, dass der Wert nicht im geschätzten Bereich liegt.
$\#v$	Anzahl der unterschiedlichen Zufallstestsätze.

$n = 1.000$ [F], $x_{\mathrm{AV}} = 1.000 - 32$ [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* $(\#v = 1)$*?*

(4.71) $$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \frac{\hat{\varepsilon}}{n - x_{\mathrm{AV}}} = \sqrt{\frac{\kappa}{n - x_{\mathrm{AV}}} \cdot \hat{p}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ für } \hat{p} > 0{,}5$$

(4.73) $$\hat{\mathrm{sr}}_{\mathrm{p}} = [p_{\mathrm{L}}, p_{\mathrm{U}}] = 1 - \left(1 - \frac{x_{\mathrm{AV}}}{n}\right) \cdot (1 \mp \hat{\varepsilon}_{\tilde{\mathrm{r}}}) \text{ für } \hat{p} > 0{,}5$$

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

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

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

$\hat{\varepsilon}_{\tilde{\mathrm{r}}}$ Geschätzter relativer Intervallradius der Anzahl der nichtnachweisbaren Fehler.

$\hat{\mathrm{sr}}_{\mathrm{p}}$ Geschätzter Bereich der Eintrittswahrscheinlichkeit, hier der Fehlerabdeckung.

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

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

(4.78) $$\hat{\varepsilon}_{\tilde{\mathrm{r}}} = \sqrt{\tfrac{\kappa}{\#v} \cdot \left(\tfrac{1}{(n-x_{\mathrm{AV}})} - \tfrac{1}{n}\right)} \cdot \Phi^{-1}\left(1 - \tfrac{\alpha}{2}\right)$$

Umstellung nach $\#v$:

$$\#v \ge \kappa \cdot \left(\tfrac{1}{(n-x_{\mathrm{AV}})} - \tfrac{1}{n}\right) \cdot \left(\frac{\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)}{\varepsilon_{\tilde{\mathrm{r}}}}\right)^2$$

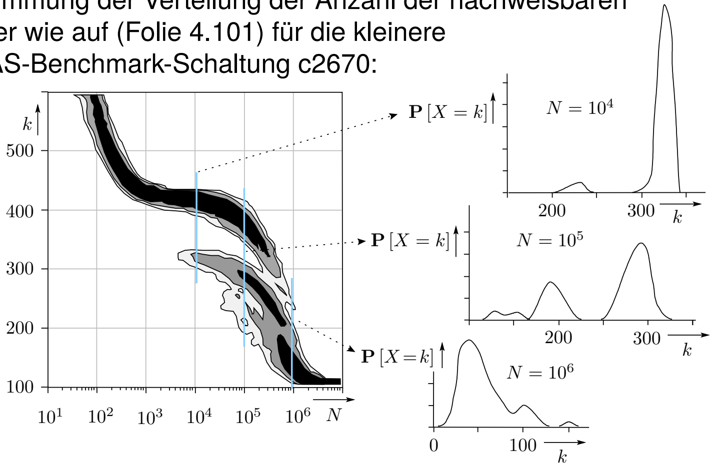
$$= 2 \cdot \left(\tfrac{1}{32} - \tfrac{1}{1000}\right) \cdot \left(\frac{2{,}33}{0{,}1}\right)^2 = 35$$

Die Modellfehlerabdeckung muss mit mindestens 35 unterschiedlichen Zufallstestsätzen bestimmt und gemittelt werden.

Mischverteilung

4.107 Anzahl der nachweisbare Fehler c2670

Bestimmung der Verteilung der Anzahl der nachweisbaren Fehler wie auf (Folie 4.101) für die kleinere ISCAS-Benchmark-Schaltung c2670:



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

Wie ist das möglich?

4.108 Mischverteilung

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

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

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

$$F_X\left(x\right)=\mathbb{P}\left[X\leq x\right]=\sum_{j=1}^{\#Y}h_j\cdot F_{X_j}\left(x\right) \tag{4.79}$$

$$\mathbb{P}\left[X=x_i\right]=\sum_{j=1}^{\#Y}h_j\cdot\mathbb{P}\left(X_j=x_i\right) \tag{4.80}$$

$$f_X\left(x\right)=\frac{\mathrm{d}F_X\left(x\right)}{\mathrm{d}x}=\sum_{i=1}^{\#Y}h_j\cdot f_{X_j}\left(x\right) \tag{4.81}$$

---

h_j Wahrscheinlichtkeit, dass ein Objekt mit Verteilungsfunktion j ausgewählt wird.

4.109 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.
- ...

Varianzvergrößerung

4.110 Varianzvergrößerung durch Mischen

Der Erwartungswert ist der gewichtete Mittelwert:

$$\mathbb{E}\left[X\right] = \mu = \sum_{j=1}^{\#Y} h_j \cdot \mu_j \tag{4.82}$$

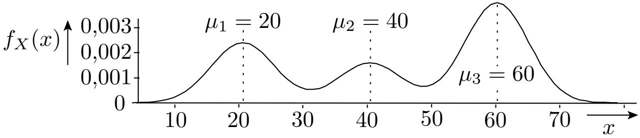
Varianz bei abweichenden Erwartungswerten $\mu = \mu_j + \delta_j$:

$$\begin{aligned}\operatorname{Var}\left[X\right] = \sigma^2 &= \sum_{j=1}^{\#Y} h_j \cdot \mathbb{E}\left[\left(X_j - \left(\mu_j + \delta_j\right)\right)^2\right] \\ &= \sum_{j=1}^{\#Y} h_j \left(\underbrace{\mathbb{E}\left[\left(X_j - \mu_j\right)^2\right]}_{\sigma_j^2} - 2 \cdot \delta_j \cdot \underbrace{\mathbb{E}\left[\left(X_j - \mu_j\right).\right]}_{0} + \underbrace{\mathbb{E}\left[\delta_j^2\right]}_{\delta_j^2}\right)\end{aligned}$$

$$\operatorname{Var}\left[X\right] = \sigma^2 = \sum_{j=1}^{\#Y} h_j \cdot \sigma_j^2 + \sum_{j=1}^{\#Y} h_j \cdot \delta_j^2 \tag{4.83}$$

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

4.111 Multimodale Verteilung

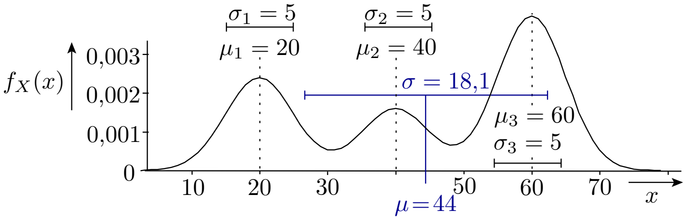


Beim Mischen von Grundgesamtheiten mit deutlich abweichenden Erwartungswerten entstehen multimodale Verteilungen. Beispiel Mischen von drei normalverteilten Grundgesamtheiten:

h_j	0,3	0,2	0,5
μ_j	20	40	60
σ_j	5	5	5

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

$\varphi(z)$ Dichtefunktion der standardisierten Normalverteilung.
h_j Auswahlhäufigkeit von Objekten mit Dichtefunktion j.
μ_j, σ_j Erwartungswert und Standardabweichung der Dichtefunktion j.



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

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

h_i	0,3	0,2	0,5

.

Erwartungswert:

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

Varianz, Standardabweichung:

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

$$\sigma = 18{,}1$$

Beispiele für Mischvert.

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

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

Binomialverteilung ohne die 10 nur gemeinsam nachweisbaren Fehler:

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

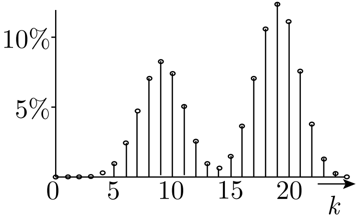
Mit den 10 nur gemeinsam nachweisbaren Fehlern verschieben sich alle Wahrscheinlichkeiten um 10 Realisierung:

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

$n = 25$, 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?*

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

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

Erwartungswert:

$$\begin{aligned}\mathbb{E}[X] &= 25 \cdot p = 25 \cdot 60\% = 15 \\ &\overset{!}{=} (1-p) \cdot \mathbb{E}[X_0] + p \cdot \mathbb{E}[X_1] = (1-0{,}6) \cdot 9 + 0{,}6 \cdot 19 = 15\surd\end{aligned}$$

Varianz als Summe der Varianzen der Summanden:

$$\begin{aligned}\mathrm{Var}[X] &= 15 \cdot p \cdot (1-p) + 10^2 \cdot p \cdot (1-p) = 115 \cdot p \cdot (1-p) = 27{,}6 \\ &\overset{!}{=} \underbrace{n \cdot p \cdot (1-p)}_{15 \cdot 0{,}4 \cdot 0{,}6} + \underbrace{(1-p) \cdot (\mathbb{E}[X_0] - \mathbb{E}[X])^2}_{+\ 0{,}4 \cdot (9-15)^2} + \underbrace{p \cdot (\mathbb{E}[X_1] - \mathbb{E}[X])^2}_{+\ 0{,}6 \cdot (19-15)^2} = 27{,}6\surd\end{aligned}$$

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

c) *Standardabweichung und Varianzerhöhung?*

(4.9)
$$\mathrm{sd}\left[X\right] = \sigma = \sqrt{\mathrm{Var}\left[X\right]}$$

(4.76)
$$\kappa = \frac{\sigma^2}{n\cdot p\cdot(1-p)} = \frac{\sigma^2}{\mu\cdot\left(1-\frac{\mu}{n}\right)}$$

Standardabweichung:

$$\mathrm{sd}\left[X\right] = \sqrt{27{,}6} = 5{,}25$$

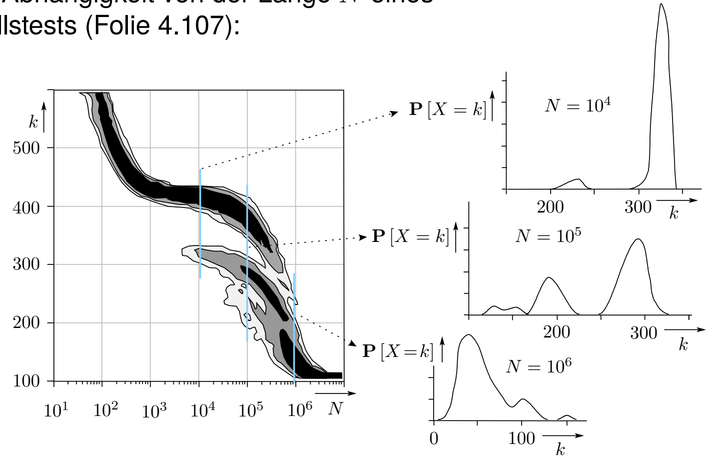
Varianzerhöhung:

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

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

4.115 Verteilung nicht nachweisbare Fehler c2670

... in Abhängigkeit von der Länge N eines Zufallstests (Folie 4.107):

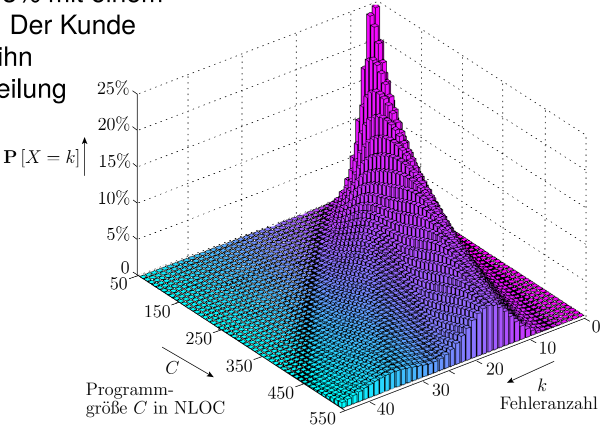


Im Bereich von $N = 10^4$ bis 10^6 multimodale Verteilung. Offenbar ca. 80 sehr ähnlich nachweisbare Fehler mit MF-Rate $\zeta_i \approx 10^{-5}\ \left[\frac{\text{MF}}{\text{DS}}\right]$.

4.116 Unterschiedlich gute Programmierer

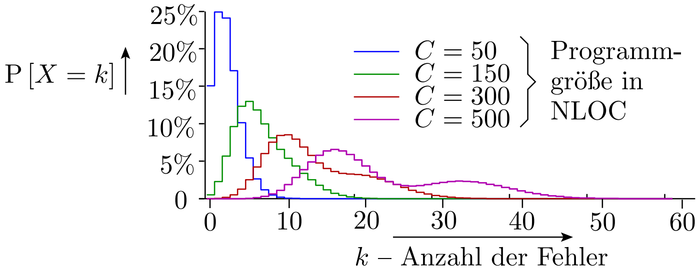
Ein Anfänger und ein Profi entwickeln Software-Bausteine aus C Netto Lines of Code (NLOC), der Profi 66% mit ca. einem Fehler je 30 NLOC und der Anfänger 33% mit einem Fehler je 15 NLOC. Der Kunde weiß nicht, wer für ihn programmiert. Verteilung der Fehleranzahl:

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



Die Wahrscheinlichkeit, dass ein Modul genau k Fehler enthält, ist 2/3 mal die Wahrscheinlichkeit, das 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. 4.80):

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



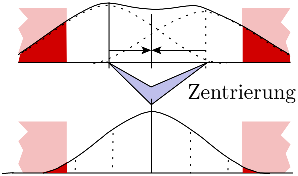
Die Polarisierung nimmt mit der Größe der Software-Bausteine, die vom Profi und vom Anfänger getrennt entwickelt werden, zu.

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.

4.118 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 2.107 *Prozesszentrierung*).

4.120 Multimodalität und Fehlervermeidung

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

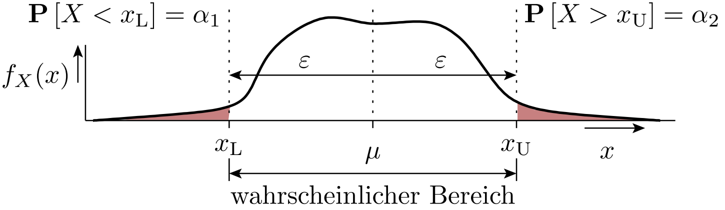
- Abhängigkeiten bei der Fehlerentstehung, bei Ausfällen beim Fehlernachweis und beim Versagen von Service-Leistungen,
- Vorliebe oder Neigung befragter Experten, z.B. bei der Einschätzung von Gefährdungen und Risiken,
- Probleme eines Messverfahrens, ...

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.

Tschebytscheffsche Ungl.

4.121 Bereichsschätzung Verteilung unbekannt

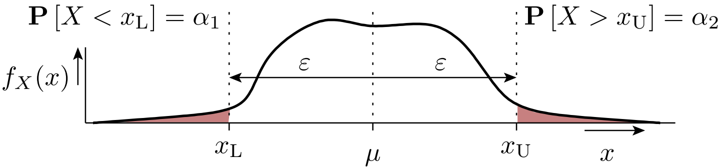


Die Bestimmung eines wahrscheinlichen Bereichs $[x_L, x_U]$

- 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.
$\mathbb{P}[\ldots]$	Eintrittswahrscheinlichkeit des Ereignisses ...
$f_X(x)$	Dichtefunktion der Zufallsvariablen X.
μ	Erwartungswert.
ε	Intervallradius, Abstand zwischen Bereichsgrenzen und Erwartungswert.

4.122 Das schwache Gesetz der großen Zahlen



Nach der tschebytscheffschen Ungleichung ist die Irrtumswahrscheinlichkeit α, das der Wert einer Zufallsvariablen mehr als ein Intervallradius ε von seinem Erwartungswert abweicht, nicht größer als das Verhältnis aus Varianz σ^2 und dem Quadrat des Intervallradius ε^2:

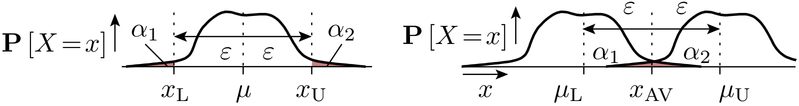
$$\mathbb{P}[|X-\mu| \geq \varepsilon] = \alpha \leq \frac{\sigma^2}{\varepsilon^2} \tag{4.84}$$

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

$$\varepsilon \leq \sqrt{\frac{\sigma^2}{\alpha}} = \frac{\sigma}{\sqrt{\alpha}} \tag{4.85}$$

X	Zufallsvariable.
ε, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.

4.123 Bereichsschätzung



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

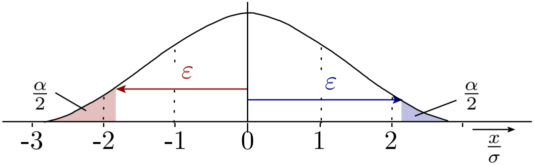
$$\mathrm{sr} = [x_L, x_U] = \mu \mp \frac{\sigma}{\sqrt{\alpha}} \tag{4.86}$$

Wahrscheinlicher Bereich des Erwartungswerts um einen experimentell bestimmten Wert x_{AV}:

$$\mathrm{sr}_\mu = [\mu_L, \mu_U] = x_{AV} \mp \frac{\sigma}{\sqrt{\alpha}} \tag{4.87}$$

ε, μ, σ	Intervallradius, Erwartungswert, Standardabweichung.
α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.
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 Wert, Schätzwert für den Erwartungswert.

4.124 Vergleich Intervallradius Normalverteilung



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

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

Intervallradius Tschebytscheffsche Ungleichung:

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

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

Beispiel 4.15: Tschebytscheffsche Ungleichung

Aus eine Stichprobe gemessener Widerstandswerte in kΩ

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

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

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

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

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

α	Irrtumswahrscheinlichkeit, dass der Wert nicht im geschätzten Bereich liegt.
R_i	Widerstandswerte in kΩ.

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

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

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

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

(4.86) $$\mathrm{sr} = \left[x_{\mathrm{L}}, x_{\mathrm{U}}\right] = \mu \mp \frac{\sigma}{\sqrt{\alpha}}$$

$$\hat{\mu} = \frac{1}{8}\left(10{,}3 + \ldots\right)\,\mathrm{k\Omega} = 10{,}025\,\mathrm{k\Omega}$$

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

Bereich des Erwartungswerts:

$$\mathrm{sr}_{\mathrm{R.Tscheb}} = 10{,}025\,\mathrm{k\Omega} \mp \frac{647\,\Omega}{\sqrt{2\%}} = \left[5{,}3\,\mathrm{k\Omega},\ 14{,}8\,\mathrm{k\Omega}\right]$$

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?*

(4.57) $$\mathrm{sr} = [x_\mathrm{L}, x_\mathrm{U}] = \mu \mp \sigma \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

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

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

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

Exkurs Defektanteil

4.126 Defektanteil und Fehleranzahl

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

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

Bei einer geringen gleichbleibenden Fehlerentstehungsrate ist die Fehleranzahl je Produkt poissonverteilt:

$$\mathbb{P}\left[X = k\right] = \mathrm{e}^{-\mu_{\mathrm{F}}} \cdot \frac{\mu_{\mathrm{F}}^k}{k!} \tag{4.88}$$

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

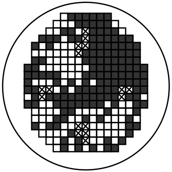
$$\mu_{\mathrm{DL}} = 1 - \mathrm{e}^{-\mu_{\mathrm{F}}} \tag{4.89}$$

Für eine geringe zu erwartende Fehleranzahl $\mu_{\mathrm{F}} \ll 1$:

$$\mu_{\mathsf{DL}} = 1 - \left(1 + (-\mu_{\mathrm{F}}) + \frac{(-\mu_{\mathrm{F}})^2}{2!} + \dots\right) \overset{(\mu_{\mathrm{F}} \ll 1)^*}{=} \mu_{\mathrm{F}} \tag{4.90}$$

μ_{F} Zu erwartende Fehleranzahl.
μ_{DL} Zu erwartender Defektanteil.
* Taylor-Reihe $\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$. Approximation bis zum linearen Glied.

4.127 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.

4.128 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. 4.89 mit $\mu_{\mathrm{F}} = FC \cdot \mu_{\mathrm{CF}}$:

$$\mu_{\mathrm{Y}} = \mathrm{e}^{-FC \cdot \mu_{\mathrm{CF}}} \tag{4.91}$$

Für jeden als fehlerfrei befundenen Schaltkreis müssen im Mittel

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

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.

4.129 Fertigungskosten und Fehleranzahl

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

$$\mu_{\mathrm{CF}} = \xi_{\mathrm{Tr}} \cdot \#Tr$$
$$C_{\mathrm{MIC}} = C_{\mathrm{Tr}} \cdot \#Tr$$

(vergl. Gl. 2.49). Kosten je nicht als defekt aussortierter Schaltkreis:

$$C_{\mathrm{IC}} = \frac{C_{\mathrm{MIC}}}{\mu_{\mathrm{Y}}} = C_{\mathrm{Tr}} \cdot \#Tr \cdot \mathrm{e}^{FC \cdot \xi_{\mathrm{Tr}} \cdot \#Tr} \tag{4.92}$$

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 4.16: Schaltkreiskosten

Fehlerentstehungsrate $\xi_{\mathrm{Tr}} = 10^{-6}$ [Fehler je Transistor], Fertigungskosten $C_{\mathrm{Tr}} = 10^{-6}$ [Euro je Transistor], Schaltkreisgröße $\#Tr \in \left\{10^5, 10^6, 10^7\right\}$ [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 $\xi_{\mathrm{Tr}} = 10^{-6}$ [Fehler je Transistor], Fertigungskosten $C_{\mathrm{Tr}} = 10^{-6}$ [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
$\mu_{\mathrm{CF}} = \xi_{\mathrm{Tr}} \cdot \#Tr$ in Fehlern	0,1	1	10
$C_{\mathrm{MIC}} = C_{\mathrm{Tr}} \cdot \#Tr$ in Euro	0,1	1	10
$\mu_{\mathrm{Y}} = \mathrm{e}^{-FC \cdot \mu_{\mathrm{CF}}}$	$90{,}5$	$36{,}8\%$	$4{,}54 \cdot 10^{-5}$
$C_{\mathrm{IC}} = \frac{C_{\mathrm{MIC}}}{Y}$ in Euro	$0{,}11$	$2{,}72$	$2{,}2 \cdot 10^5$

Ab $\mu_{\mathrm{CF}} \geq 2 \ldots 3$ 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.

Zusammenfassung

4.131 Varianzerhöhung

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

(4.77) $$\hat{\kappa} = \frac{\hat{\sigma}^2}{\hat{\mu} \cdot \left(1 - \frac{\hat{\mu}}{n}\right)}$$

Ein Varianzerhöhung $\kappa > 1$ bedeutet für normalverteilte Zählwerte, dass für dieselbe Schätzgenauigkeit ε_{R} und α in

(4.74) $$x_{\mathrm{AV}} \geq \kappa \cdot \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2 \cdot \varepsilon_{\mathrm{r}}^{-2} \cdot (1 - \hat{p}) \quad \text{für } \hat{p} \leq 0{,}5$$

die erforderlichen Zählwerte x_{AV} für das Eintreten bzw. das Nichteintreten und damit auch die Anzahl der Zählversuche n κ-mal so groß wie für unabhängige Zählwerte sein müssen.

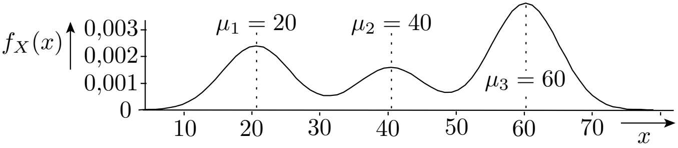
4.132 Varianzerhöhung Fehlernachweis

Zwischen Modellfehlern gibt es auch nach der Zusammenfassung identisch nachweisbare zu je einem Modellfehler erhebliche Nachweisabhängigkeiten durch gemeinsame Anregungs- und Beobachtungsbedingungen. Die Beschränkung auf zufällig ausgewählte Fehlerstichproben verringert diese Nachweisabhängigkeiten.

Daraus lässt sich die These ableiten, dass für den Zufallstests Fehlermodelle, die sehr viele Modellfehler bezogen auf die Testobjektgröße generieren, unzweckmäßig sind. Denn die Abhängigkeiten im Fehlernachweis vernichten die rechnerische Varianzminderung durch die größeren Zählwerte.

Wirkungsvoller als viele ähnlich nachweisbare Modellfehler für die Verbesserung der Schätzgenauigkeit ist eine Mittelung der Fehlerabdeckungen mehrerer Zufallstestsätze.

4.133 Misch- und multimodale Verteilungen



Mischung diskreter Verteilungen:

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

Mischung stetiger Verteilungen:

(4.81) $$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} = \sum_{i=1}^{\#Y} h_j \cdot f_{X_j}(x)$$

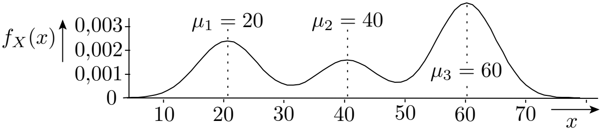
Erwartungswert:

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

Varianz:

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

4.135 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.

Multimodalität liefert Hinweise auf Verbesserungsmöglichkeiten

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

4.136 Tschebyscheffsche Ungleichung

Bereichsschätzung für beliebige incl. multimodale Verteilungen:

- tschebytscheffsche Ungleichung:

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

- garantierbarer Intervallradius:

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

- garantierbarer Bereiche für Realisierungen und Erwartungswerte:

(4.86) $$\mathrm{sr} = [x_{\mathrm{L}}, x_{\mathrm{U}}] = \mu \mp \frac{\sigma}{\sqrt{\alpha}}$$

(4.87) $$\mathrm{sr}_{\mu} = [\mu_{\mathrm{L}}, \mu_{\mathrm{U}}] = x_{\mathrm{AV}} \mp \frac{\sigma}{\sqrt{\alpha}}$$

4.137 Defektanteil, Fehlercluster

In einem idealen Prozess mit gleichbleibender Fehlerentstehungsrate ist die Fehleranzahl je Objekt poissonverteilt. Der zu erwartende Fehleranteil ist die Wahrscheinlichkeit »Fehlerzahl null«:

$$\mu_{\mathrm{DL}} = 1 - \mathrm{e}^{-\mu_{\mathrm{F}}} \tag{4.89}$$

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

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.

Pareto-Verteilung

4.138 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. 3.15)

$$F_X(N) = \mathbb{P}[X \le N] = \mu_{\mathrm{FC}}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K} \quad \text{für } N \geq N_0 \qquad (4.93)$$

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.

4.139 Pareto-Verteilung

Die Pareto-Verteilung

$$X \sim \mathrm{Par}\left(K, x_{\min}\right)$$

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

$$F_X\left(x\right) = \mathbb{P}\left[X \leq x\right] = \begin{cases} 0 & x \leq x_{\min} \\ 1 - \left(\frac{x_{\min}}{x}\right)^K & \text{sonst} \end{cases} \tag{4.94}$$

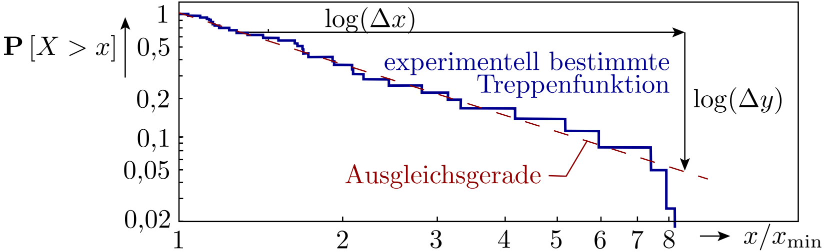
Dichtefunktion für $x \geq x_{\min}$:

$$f_X\left(x\right) = \frac{K \cdot x_{\min}^K}{x^{K+1}} \tag{4.95}$$

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

Eigenschaften

4.140 Eigenschaften der Pareto-Verteilung



Die doppelt logarithmisch dargestellte der Anzahl der noch nicht eingetretenen Ereignisse ist durch eine Ausgleichgerade annäherbar. Der Formfaktor K ergibt sich aus dem Abfall dieser Ausgleichsgeraden:

$$1 - F_X(x) = \mathbb{P}[X > x] = \left(\frac{x}{x_{\min}}\right)^{-K} \tag{4.96}$$

$$K = -\frac{\log(\Delta y)}{\log(\Delta x)} \tag{4.97}$$

4.141 Erwartungswert und Varianz

- Einen Erwartungswert

$$\mathbb{E}[X] = \int_{x_{\min}}^{\infty} \frac{K \cdot x_{\min}^{K}}{x^{K+1}} \cdot x \cdot \mathrm{d}x = \frac{K \cdot x_{\min}^{K}}{1-K} \cdot \left(\lim_{x\to\infty} x^{1-K} - x_{\min}^{1-K} \right)$$

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

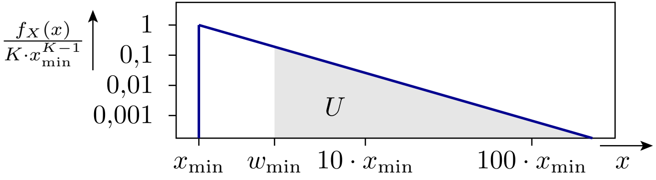
$$\mathbb{E}[X] = \mu = x_{\min} \cdot \tfrac{K}{K-1} \tag{4.98}$$

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

$$\mathrm{Var}[X] = \sigma^2 = x_{\min}^2 \cdot \tfrac{K}{(K-2)(1-K)^2} \tag{4.99}$$

Die pareto-verteilte Fehlernachweislänge für Zufallstests (Gl. 4.93) 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.

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



Der Anteil der Ursachen U mit der größten Wirkung:

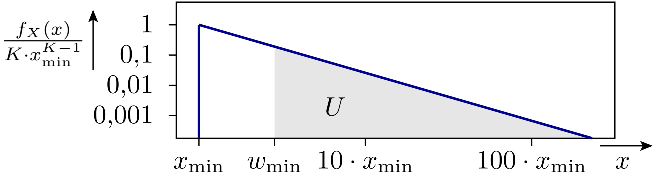
$$U = \int_{w_{\min}}^{\infty} f(x) \cdot \mathrm{d}x = \int_{w_{\min}}^{\infty} \frac{K \cdot x_{\min}^{K}}{x^{K+1}} \cdot \mathrm{d}x = \left(\frac{x_{\min}}{w_{\min}}\right)^{K} \quad (4.100)$$

hat mindestens die Wirkung

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

und für $K > 1$ insgesamt die Wirkung:

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



Zu erwartende Wirkung aller Ursachen ist der Erwartungswert:

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

Anteilige Wirkung der Ursachen mit der größten Wirkung:

$$W = \frac{\mathbb{E}[X | X \geq w_{\min}]}{\mathbb{E}[X]} = \left(\frac{x_{\min}}{w_{\min}}\right)^{K-1} = \left(U^{\frac{1}{K}}\right)^{K-1} = U^{\frac{K-1}{K}} \qquad (4.101)$$

$$K = \frac{1}{1 - \frac{\log(W)}{\log(U)}} \qquad (4.102)$$

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

4.144 Verteilung der Fehlernachweislänge

Die Hypothese, dass die Nachweislänge pareto-verteilt ist

(4.93) $$F_X(N) = \mathbb{P}[X \le N] = \mu_{\mathrm{FC}}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K} \quad \text{für } N \ge N_0$$

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):

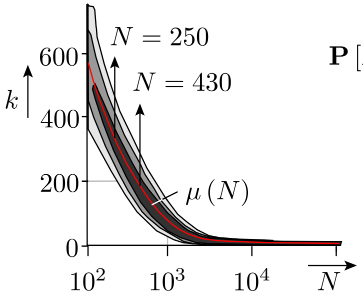
(3.15) $$\mu_{\mathrm{FC}}(N) = 1 - \left(\frac{N}{N_0}\right)^{-K}$$

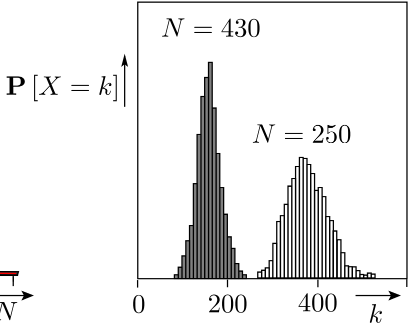
K	1	0,5	0,33	0,25
$\frac{N}{N_0}$ für $1 - \mu_{\mathrm{FC}}(N) = 0{,}1$	10	100	10^3	10^4

Wie brauchbar ist die Näherung?

$\mu_{\mathrm{FC}}(N)$ 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 $(0 < K < 1)$.

4.145 Benchmarkschaltung c3540 (Folie 4.101)

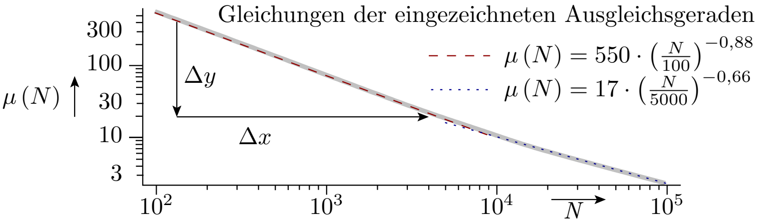




Experimentell für Haftfehler 3606 simulierte Haftfehler bestimmte Verteilungen der Anzahl der nicht nachweisbaren Fehler mit 1.000 verschiedenen Zufallstests (Folie 4.101). Der Erwartungswert μ in Abhängigkeit von der Testanzahl N ist rot eingezeichnet.

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

4.146 Pareto-Näherung



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

$$K = -\frac{\log(\Delta y)}{\log(\Delta x)} \quad (4.97)$$

Keine perfekte Pareto-Verteilung. Der Formfaktor ändert sich mit der Größenordnung der Testanzahl, ... Interessant für weitere Untersuchungen.

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

Schaden durch MF

4.147 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, ... (vergl. Foliensatz 1.12).

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

Mangels verfügbaren Schadensstatistik für IT-Fehlfunktionen betrachten wir die Verteilung von Haftpflichtschäden einer Autoversicherung aus der Schweiz*.

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

4.148 Verteilung von Haftpflichtschäden

Haftpflichtschäden über 100.000 SF einer Schweizer Autoversicherung:

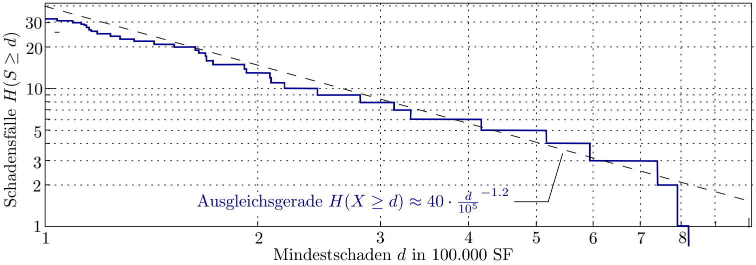
```
103.765, 109.168, 112.341, 113.800, 114.791,
115.731, 118.264, 123.464, 127.611, 133.504,
142.821, 152.270, 163.491, 164.968, 168.915,
169.346, 172.668, 191.954, 193.102, 208.522,
209.070, 219.111, 243.910, 280.302, 313.898,
330.461, 418.074, 516.218, 595.310, 742.198,
791.874, 822.787, 1.074.499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

- Anzahl der Schadensfälle: 33
- Gesamtschadenssumme: 9.458.208 SF

SF Schadenskosten in Schweizer Franken.

4.149 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. 4.98):

$$\mu_{\mathrm{d}} = d_{\min} \cdot \frac{K}{K-1} = \frac{1{,}2}{1{,}2-1} \cdot d_{\min} = 600.000\,\mathrm{SF}$$

Eine Varianz hat eine Pareto-Verteilung erst ab $K > 2$. Alle im Kapitel behandelten Bereichsschätzungen incl. über die Tschebytscheffsche-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 Haftpflichschäden 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

4.151 Zusammenfassung

Verteilungsfunktion und Dichte der Pareto-Verteilung:

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

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

Erwartungswert erst ab $k > 1$:

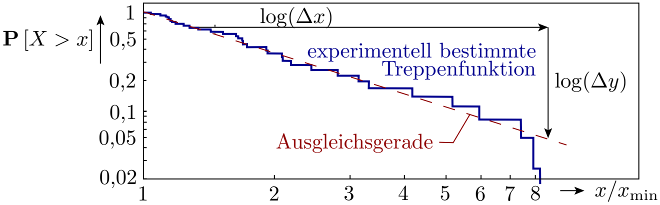
(4.98) $$\mathbb{E}[X] = \mu = x_{\min} \cdot \frac{K}{K-1}$$

Varianz erst ab $k > 2$:

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

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

4.152 Experimentell Untermauerung



Für die doppelt log. dargestellte Gegenwahrscheinlichkeit der Verteilungsfunktion:

(4.96) $$1 - F_X(x) = \mathbb{P}[X > x] = \left(\frac{x}{x_{\min}}\right)^{-K}$$

muss es eine brauchbare Geradenannäherung geben:

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

Mit experimentell erhobenen Zählwerten untersucht für

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