Test und Verlässlichkeit Foliensatz 6: Verteilungen und zusicherbare Kenngrößen

Prof. G. Kemnitz

8. Juni 2018

Inhaltsverzeichnis

1 Verteilungen

Kennwerte schätzen

Zur Bewertung der Verlässlichkeit und der Maßnahmen zu ihrer Sicherung sind

- zu erwartende Bereiche für Zählwerte (Fehler, FF, ...),
- Wahrscheinlichkeiten (Fehlerauftrittswahrscheinlichkeiten, Fehlernachweiswahrscheinlichkeiten, ...) und
- Häufigkeitsverteilungen, z.B. für die zu erwartende Fehleranzahl in Abhängigkeit von der mittleren Anzahl der SL je FF

zu schätzen. Grundlange dafür sind Zufallsexperimente, Zufallsgrößen, deren **Verteilungen**, Erwartungswerte, ...

Verteilung

Eine Verteilung weist möglichen Werten einer Zufallsvariablen Werte zu:

- Häufigkeitsverteilung: Zählwerte, Umfragewerte, ...,
- Wahrscheinlichkeitsverteilung: Wahrscheinlichkeiten.

Häufigkeitsverteilungen werden empirisch erfasst. Mit Wahrscheinlichkeitsverteilungen lassen sich Häufigkeitsverteilungen annähern, vorhersagen, auf ähnliche Sachverhalte übertragen. Wichtige Kenngrößen von Verteilungen sind **Erwartungswert**, **Varianz** und **Standardabweichung**.

1.1 Erwartungswert, Varianz, ...

Erwartungswert

Bei einer Wahrscheinlichkeitsverteilung sind die möglichen Ergebnisse[1] ein Zahlenbereich (abzählbar oder stetig). Jedem Wert dieses Bereiches ist eine Wahrscheinlichkeit zugeordnet.

mögliche Ergebnisse x_i	2	3	4	5	6	Summe
Wahrscheinlichkeit p_i	4%	12%	29%	37%	18%	100%

Der **Erwartungswert** $E(X)$ (X – Zufallsgröße) ist der mit ihren Auftrittswahrscheinlichkeiten gewichtete Mittelwert:

$$E(X) = \sum_{i=1}^{N} p_i \cdot x_i;\ E(X) = \int_0^{\infty} h(x) \cdot dx \tag{1}$$

(N – Anzahl der möglichen Ergebnisse). Für das Beispiel:

$$2 \cdot 4\% + 3 \cdot 12\% + 4 \cdot 29\% + 5 \cdot 37\% + 6 \cdot 18\% = 4{,}53$$

Varianz und Standardabweichung

Die **Varianz** $D^2(X)$ ist die mittlere quadratische Abweichung vom Erwartungswert:

$$D^2(X) = \sum_{i=1}^{N} p_i \cdot (x_i - E(X))^2 \tag{2}$$

Die **Standardabweichung** $\sqrt{D^2(X)}$ ist die Wurzel aus der Varianz und ein Maß dafür, wie stark die Ergebnisse eines Zufallsexperiments um ihren Erwartungswert streuen.

mögliche Ergebnisse x_i	2	3	4	5	6	Summe
Wahrscheinlichkeit p_i	4%	12%	29%	37%	18%	100%

$$\begin{aligned} D^2(X) &= 4\% \cdot (2-4{,}53)^2 + 12\% \cdot (3-4{,}53)^2 + 29\% \cdot (4-4{,}53)^2 \\ &+ 37\% \cdot (5-4{,}53)^2 + 18\% \cdot (6-4{,}53)^2 = 1{,}09 \\ \sqrt{D^2(X)} &= 1{,}04 \end{aligned}$$

Verschiebungssatz

Die Varianz ist gleichfalls die Differenz aus dem Erwartungswert der Quadrate und dem Quadrat des Erwartungswertes[2]:

$$D^2(X) = E\left(X^2\right) - (E(X))^2 \tag{3}$$

Herleitung:

$$\sum_{i=1}^{N} p_i \cdot (x_i - E(X))^2 = \sum_{i=1}^{N} p_i \cdot \left(x_i^2 - 2 \cdot x_i \cdot E(X) + E(X)^2\right)$$

$$\underbrace{\sum_{i=1}^{N} p_i \cdot x_i^2}_{E\left(X^2\right)} + E(X) \cdot \left(E(X) \cdot \underbrace{\sum_{i=1}^{N} p_i}_{1} - 2 \cdot \underbrace{\sum_{i=1}^{N} p_i \cdot x_i}_{E(X)} \right)$$

Für das Beispiel zuvor:

$$D^2(X) = 4\% \cdot 2^2 + 12\% \cdot 3^2 + 29\% \cdot 4^2 + 37\% \cdot 5^2 + 18\% \cdot 6^2 - 4{,}53^2 = 1{,}09$$

[1] Die möglichen Ergebnisse werden als Realisierungen bezeichnet.
[2] Bei begrenzter Rechengenauigkeit u.U. numerisch problematisch.

Ein weiteres Beispiel

Gegeben ist die Verteilung in der nachfolgenden Tabelle:

Wert	5	6	8	11	22
Wahrscheinlichkeit	0,1	0,2	0,4	0,2	0,1

Wie groß sind

1. der Erwartungswert,
2. die Varianz und
3. die Standardabweichung?

Lösung

Wert	5	6	8	11	22
Wahrscheinlichkeit	0,1	0,2	0,4	0,2	0,1

1. Erwartungswert:
$$E(X) = 0{,}1 \cdot 5 + 0{,}2 \cdot 6 + 0{,}4 \cdot 8 + 0{,}2 \cdot 11 + 0{,}1 \cdot 22 = 9{,}3$$
2. Varianz nach Gl. 2:
$$\begin{aligned} D^2(X) &= 0{,}1 \cdot (5-9{,}3)^2 + 0{,}2 \cdot (6-9{,}3)^2 + 0{,}4 \cdot (8-9{,}3)^2 \\ &+ 0{,}2 \cdot (11-9{,}3)^2 + 0{,}1 \cdot (22-9{,}3)^2 = 21{,}4 \end{aligned}$$
Varianz nach Verschiebungssatz:
$$\begin{aligned} D^2(X) &= 0{,}1 \cdot 5^2 + 0{,}2 \cdot 6^2 + 0{,}4 \cdot 8^2 + 0{,}2 \cdot 11^2 \\ &+ 0{,}1 \cdot 22^2 - 9{,}3^2 = 21{,}4 \end{aligned}$$
3. Standardabweichung: $\sqrt{D^2(X)} = \sqrt{21{,}4} = 4{,}63$

Erwartungswert und Varianz einer Datenstichprobe

Für eine Datenstichprobe einer Zufallsgröße X
$$w = (w_1, w_2, \ldots, w_{N_\mathrm{S}})$$
ist der Schätzer für den Erwartungswert der Mittelwert:
$$E_\mathrm{S}(X) = \frac{1}{N_\mathrm{S}} \cdot \sum_{i=1}^{N_\mathrm{S}} w_i \tag{4}$$
Der Schätzer für die Varianz ist die mittlere quadratische Abweichung vom geschätzten Mittelwert:
$$D_\mathrm{S}^2(X) = \frac{1}{N_\mathrm{S}-1} \cdot \sum_{i=1}^{N_\mathrm{S}} (w_i - E(X))^2 \tag{5}$$
Der Quotion ist um eins kleiner als die Stichprobengröße N_S, d.h die Abschätzung der Varianz erfordert mindestens Stichprobengröße $N_\mathrm{S} = 2$.

1.2 Lineare Transformationen, ...

Lineare Transformation

Lineare Transformationen sind die Multiplikation und Addition einer Zufallsgröße mit reellen Zahlen. Der Erwartungswert vergrößert und verschiebt sich um dieselben Werte:
$$E(a \cdot X + b) = a \cdot E(X) + b$$
Bei der Varianz entfällt die Verschiebung und der Skalierungsfaktor geht im Quadrat ein[3]:
$$D^2(a \cdot X + b) = a^2 \cdot D^2(X) \tag{6}$$
Die Varianz ist insbesondere verschiebungsinvariant und bleibt bei einer Spiegelung der Verteilung gleich:
$$D^2(-X) = (-1)^2 \cdot D^2(X) = D^2(X)$$

[3] Die Kontrolle der Gleichung ist eine Übungsaufgabe.

Kontrolle am Beispiel

Realisierungen x von X	1	2	3
Realisierungen y von $Y=5-2X$	3	1	-1
$P(Y=y)=P(X=x)$	0,3	0,5	0,2

$$\begin{aligned}
E(X) &= 0{,}3+1+0{,}6=1{,}9\\
D^2(X) &= 0{,}3+2+1{,}8-1{,}9^2=0{,}49\\
E(Y) &= 0{,}9+0{,}5-0{,}2=1{,}2\\
D^2(Y) &= 2{,}7+0{,}5+0{,}2-1{,}2^2=1{,}96\\
E(Y) &= 5-2\cdot E(X)\\
D^2(Y) &= (-2)^2\cdot D^2(X)
\end{aligned}$$

Summe von Zufallsgrößen

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

x	1	3	4
$P(X=x)$	0,1	0,4	0,5

y	2	3	4
$P(Y=y)$	0,3	0,6	0,1

$$\begin{aligned}
P(X+Y=3) &= P(X=1)\cdot P(Y=2)\\
P(X+Y=4) &= P(X=1)\cdot P(Y=3)\\
P(X+Y=5) &= P(X=1)\cdot P(Y=4)+P(X=3)\cdot P(Y=2)\\
P(X+Y=6) &= P(X=3)\cdot P(Y=3)+P(X=4)\cdot P(Y=2)\\
P(X+Y=7) &= P(X=3)\cdot P(Y=4)+P(X=4)\cdot P(Y=3)\\
P(X+Y=8) &= P(X=4)\cdot P(Y=4)
\end{aligned}$$

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

$$E(X+Y)=E(X)+E(Y)$$

Die Varianz ist die Summe der Varianzen plus doppelte Kovarianz:

$$D^2(X+Y)=D^2(X)+D^2(Y)+2\cdot\operatorname{Cov}(X,Y) \tag{7}$$

mit der Kovarianz[4]:

$$\operatorname{Cov}(X,Y)=E\left((X-E(X))\cdot(Y-E(Y))\right) \tag{8}$$

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

$$D^2(X+Y)=D^2(X)+D^2(Y)$$

Gemessener Wert und Messfehler

In der Messtechnik gilt für jeden gemessenen Wert:

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

(X – Messwert; X_{F} – Messfehler). Alle drei Größen haben einen Erwartungswert und eine Varianz. Mit dem Messwert und dem Messfehler als unabhängige Zufallsgrößen, gilt für diese:

$$\begin{aligned}
E(X_{\mathrm{M}}) &= E(X_{\mathrm{F}})+E(X)\\
D^2(X_{\mathrm{M}}) &= D^2(X_{\mathrm{F}})+D^2(X)
\end{aligned}$$

- $E(X_{\mathrm{F}})$ – Maß für den systematischen Messfehler
- $\sqrt{D^2(X_{\mathrm{F}})}$ – Standardabweichung und Maß für den zufälligen Messfehler.

[4] Die Kontrolle der Gleichungen sind Übungsaufgaben.

Beispielaufgabe

Der gemessene Wert einer Widerstands-Charge ist im Mittel $E(R_\mathrm{M}) = 1010\,\Omega$ und hat eine Standartabweichung von $\sqrt{D^2(R_\mathrm{M})} = 11{,}18\,\Omega$. Die Messung habe einen systematischen Fehler von $E(R_\mathrm{F}) = 12\,\Omega$ und eine Standardabweichung von $\sqrt{D^2(R_\mathrm{F})} = 5\,\Omega$. Welchen Erwartungswert und welche Standardabweichung hat der (tatsächliche) Messwert?

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

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

1.3 Verteilung von Zählwerten

Verteilung von Zählwerten

Zählwerte sind eine Summe von Einzelereignissen (z.B. Anzahl der korrekt ausgeführten oder fehlerhaft ausgeführten Service-Leistungen). Die Einzelereignisse können null oder eins sein und haben eine Bernoulli-Verteilung:

k	0	1
$P(X_i = k)$	$1-p_i$	p_i

(X_i – Zufallsgröße Einzelereignis i; p_i – Eintrittswahrscheinlichkeit $X_i = 1$). Für N zu zählende Versuchsergebnisse ist die Anzahl der eingetretenen Ereignisse die Summe der Zufallsgrößen X_i:

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

k	0	1
$P(X_i = k)$	$1-p_i$	p_i

Der Erwartungswert der Einzelereignisse ist

$$E(X_i) = (1-p_i)\cdot 0 + p_i \cdot 1 = p_i$$

Varianz nach Verschiebungssatz:

$$D^2(X_i) = (1-p_i)\cdot 0^2 + p_i \cdot 1^2 - p^2 = p_i \cdot (1-p_i)$$

Der Erwartungswert der Summe ist die Summe der Erwartungswerte:

$$E(X) = \sum_{i=1}^{N} p_i \tag{9}$$

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

$$D^2(X) = \sum_{i=1}^{N} p_i \cdot (1-p_i) \tag{10}$$

Für die Verteilung gilt, dass bei Hinzunahme eines weiteren Experiments i sich mit Wahrscheinlichkeit p_i der Zählwert um eins erhöht und mit der Wahrscheinlichkeit $1-p_i$ gleich bleibt:

$$P_i(X=k) = p_i \cdot P_{i-1}(X=k-1) + (1-p_i)\cdot P_{i-1}(X=k)$$

Berechnung der Verteilung:

i	p_i	$X=0$	$X=1$	$X=2$	$X=3$	$X=4$
1	30%	70%	30%			
2	50%	35%	50%	15%		
3	40%	21%	44%	29%	6%	
4	10%	18,9%	41,7%	30,5%	8,3%	0,6%

$P_1(X=0) = 1 - p_1$
$P_1(X=1) = p_1$

`Wiederhole für` $i = 2$ `bis` N
$P_i(X=0) = P_{i-1}(X=0) \cdot (1-p_i) \qquad P_i(X=i) = P_{i-1}(X=i-1) \cdot p_i$

`Wiederhole für` $k = 1$ `bis` $i-1$
$P_i(X=k) \quad = P_{i-1}(X=k) \cdot (1-p_i)$
$\qquad\qquad + P_{i-1}(X=k-1) \cdot p_i$

(i – Anzahl der berücksichtigten Summanden; k – Zählwert).

Erwartungswert und Varianz für das Beispiel

i	p_i	$X{=}0$	$X{=}1$	$X{=}2$	$X{=}3$	$X{=}4$
1	30%	70%	30%			
2	50%	35%	50%	15%		
3	40%	21%	44%	29%	6%	
4	10%	18,9%	41,7%	30,5%	8,3%	0,6%

Nach Gl. 1 beträgt der Erwartungswert der Summe aller $N = 4$ Summanden:

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

Als Summe aller p_i nach Gl. 9 ist die Berechung kürzer:

$$E(X) = 30\% + 50\% + 40\% + 10\% = 1{,}3$$

Die Varianz beträgt nach dem Verschiebungssatz Gl. 3:

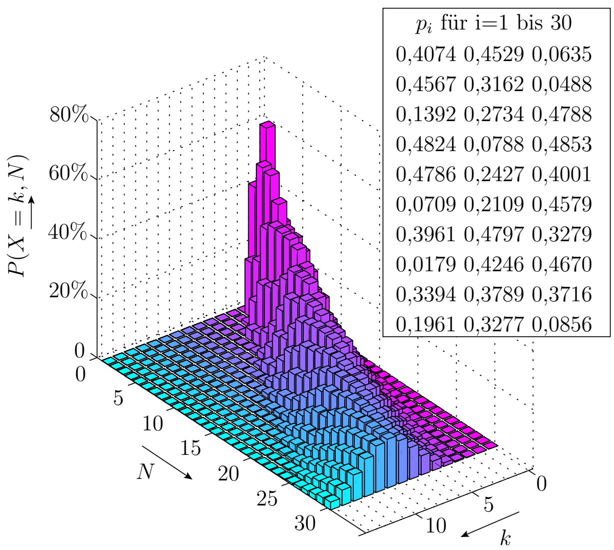
$$18{,}9\% \cdot 0^2 + 41{,}7\% \cdot 1^2 + 30{,}5\% \cdot 2^2 + 8{,}3\% \cdot 3^2 + 0{,}6\% \cdot 4^2 - 1{,}3^2 = 0{,}79$$

Die vereinfachte Berechnung nach Gl. 10 lautet:

$$D^2(X) = 0{,}3 \cdot 0{,}7 + 0{,}5 \cdot 0{,}5 + 0{,}4 \cdot 0{,}6 + 0{,}1 \cdot 0{,}9 = 0{,}79$$

Mit Matlab berechnete Zählverteilung

Das nachfolgende Säulendiagramm zeigt eine mit Matlab schrittweise berechnete Zählverteilung. Die Eintrittswahrscheinlichkeiten der Zählereignisse siehe Kasten im Bild. Erwartungswert und Varianz für alle 30 Summanden betragen $E(X)$=7,05 und $D^2(X) = 2{,}19$:



1.4 Binomialverteilung

Binomialverteilung

Für den Sonderfall, dass gleichwahrscheinliche Ereignisse gezählt werden (alle $p_i = p$), ist die Summe der gezählten Ereignisse binomialverteilt

$$P(X = k) = \binom{N}{k} \cdot p^k \cdot (1-p)^{N-k} \tag{11}$$

mit dem Erwartungswert

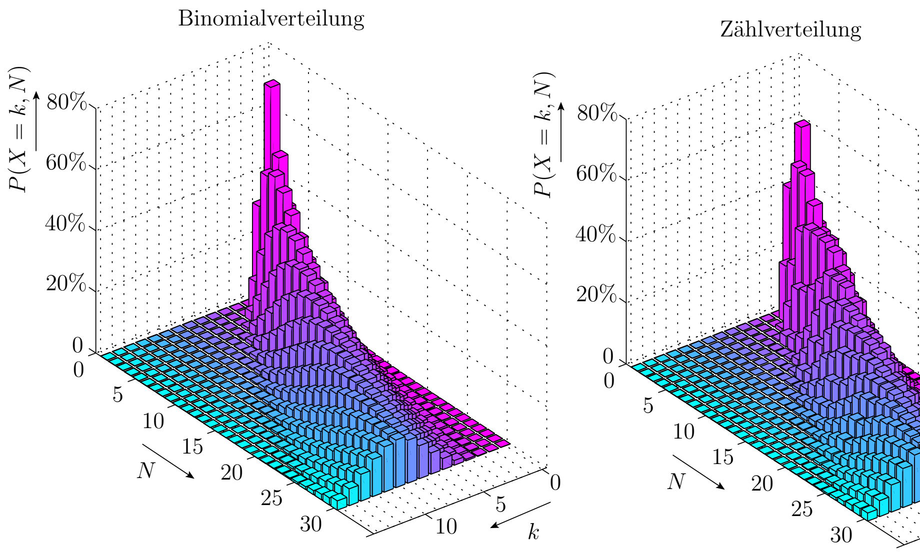
$$E(X) = N \cdot p$$

und der Varianz

$$D^2(X) = N \cdot p \cdot (1-p) \tag{12}$$

(N – Anzahl der gezählten Ereignisse, die 0 oder 1 sein können).

Binomialverteilung vs. allgemeine 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 $p = \frac{1}{N} \cdot \sum_{i=1}^{N} p_i$ und demselben N nähert eine Zählverteilung gut an, berechnet sich aus nur den zwei Parametern N und p und hat eine etwas kleinere Standardabweichung.

Beispielaufgabe

Die mittlere Nachweiswahrscheinlichkeit von 10 Fehlern sei 30%. Wie groß ist die Wahrscheinlichkeit, dass mindestens zwei Fehler nachgewiesen werden?

Zur Kontrolle:

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

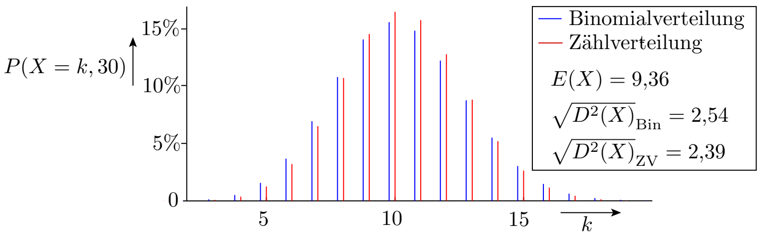
Varianzobergrenze

Satz

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

$$D^2(X)_{\text{Bin}} = N \cdot p \cdot (1-p) \geq D^2(X)_{\text{ZV}} = \sum_{i=1}^{N} p_i \cdot (1-p_i) \tag{13}$$

Für die beiden Verteilungen der Folie zuvor gilt für $N = 30$:



Beweis

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

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

Varianz der Zählverteilung:

$$\begin{aligned} D^2(X)_{\text{ZV}} &= \sum_{i=1}^{N} (p+\delta_i) \cdot (1-p-\delta_i) \\ &= \underbrace{N \cdot p \cdot (1-p)}_{D^2(X)_{\text{Bin}}} - (1-2p) \cdot \underbrace{\sum_{i=1}^{N} \delta_i}_{0} - \underbrace{\sum_{i=1}^{N} \delta_i^2}_{\geq 0} \\ D^2(X)_{\text{ZV}} &\leq D^2(X)_{\text{Bin}} \checkmark \end{aligned}$$

1.5 Poisson-Verteilung

Poisson-Verteilung

Beim Zählen vieler seltener Ereignisse, z.B. der Fehlfunktionen bei Millionen von Service-Anforderungen, von denen nur wenige eintreten, streben die Eintrittswahrscheinlichkeit der Einzelereignisse und die Abweichung zwischen Erwartungswert und Varianz gegen null:

$$\begin{aligned} p_i &\rightarrow 0 \\ E(X_i) - D^2(X_i) = p_i - p_i \cdot (1-p_i) = -p_i^2 &\rightarrow 0 \end{aligned}$$

Die Varianz der zu zählenden Ereignisse und die der Summe streben gegen den Erwartungswert

$$\begin{aligned} D^2(X_i) &= E(X_i) \\ D^2(X) = \sum_{i=1}^{N} D^2(X_i) &= E(X) = \sum_{i=1}^{N} E(X_i) \end{aligned}$$

Die Verteilung der Summe strebt gegen die Poisson-Verteilung.

Die Poisson-Verteilung ist eine einparametrige Verteilung, die sich allein aus dem Erwartungswert berechnet:

$$P(X = k) = \text{Poi}(k, E(X)) = e^{-E(X)} \cdot \frac{E(X)^k}{k!}$$

Der Erwartungswert einer Zählverteilung ist das Produkt aus der Anzahl der Zählwerte N und der mittleren Eintrittswahrscheinlichkeit p:

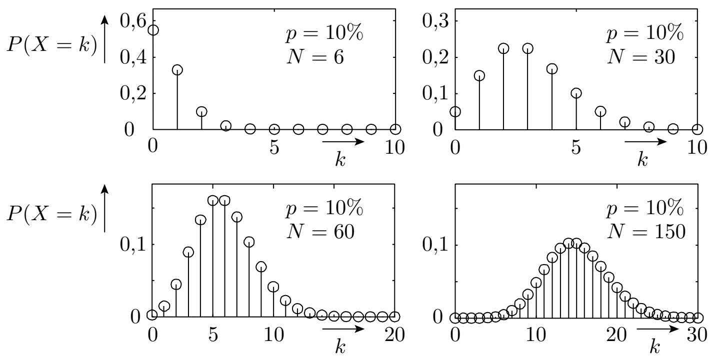
$$E(X) = \sum_{i=1}^{N} p_i = N \cdot p \;\text{ mit }\; p = \frac{1}{N} \cdot \sum_{i=1}^{N} p_i$$

Für die Zählung von Ereignissen mit geringer Eintrittswahrscheinlichkeit $p \ll 1$ lautet die Poisson-Verteilung:

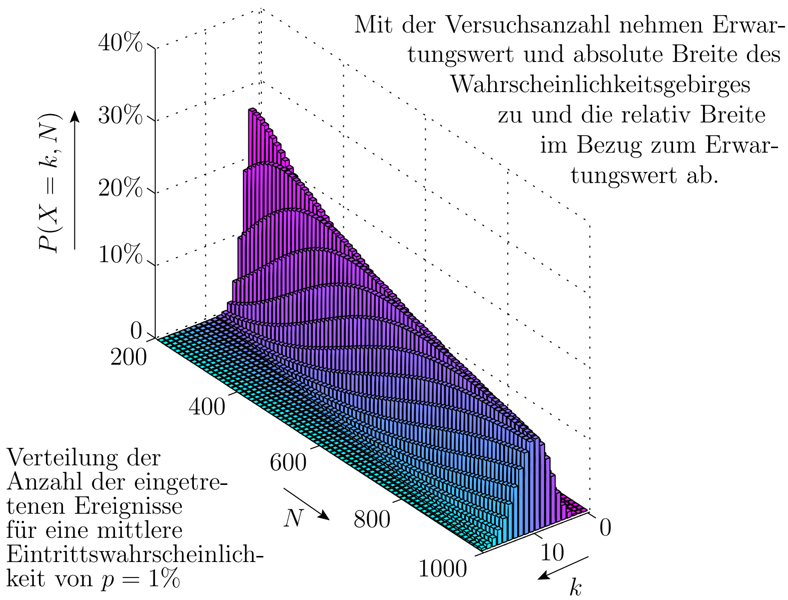
$$P(X = k) = e^{-p \cdot N} \cdot \frac{(p \cdot N)^k}{k!} \tag{14}$$

(N – Anzahl der Zählversuche; p – mittlere Eintrittswahrscheinlichkeit).

$$P(X = k) = e^{-p \cdot N} \cdot \frac{(p \cdot N)^k}{k!}$$



Der Erwartungswert $E(X) = p \cdot N$ nimmt proportional mit N zu. Die relative Breite des wahrscheinlichen Bereichs $\frac{\sqrt{D^2(X)}}{E(X)} = \frac{1}{\sqrt{p \cdot N}}$ nimmt mit N ab. Siehe auch nächste Folie.



1.6 Normalverteilung

Normalverteilung

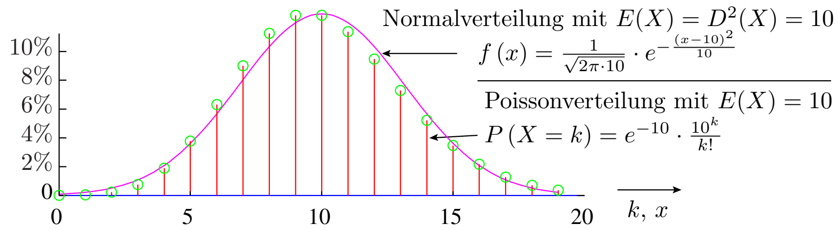
Die Summe sehr vieler unabhängiger Zufallsgrößen strebt unter sehr allgemeinen Bedingungen gegen eine Normalverteilung:

$$f(x) = \frac{1}{\sqrt{2\pi \cdot D^2(X)}} \cdot e^{-\frac{(x-E(X))^2}{D^2(X)}}$$

- kein Summand hat dominanten Einfluss und

- Erwartungswert deutlich größer als Standardabweichung.

Schließt die behandelte Zählverteilung, Binomialverteilung und Poisson-Verteilung ein.



In der Praxis gilt die Annäherung einer Zähl- durch eine Normalverteilung in der Regel bereits unter der Bedingung

$$10 \cdot \kappa \leq E(X) \leq N - 10 \cdot \kappa$$

als ausreichend genau (κ – Varianzerhöhung; $E(X) = N \cdot p$ – Erwartungswert der Zählwerte; N – Anzahl der Zählversuche; p – mittlere Eintrittswahrscheinlichkeit).

Die Annäherung einer Zählverteilung durch eine Normalverteilung eignet sich gut für Bereichsschätzungen.

Standardisierte Normalverteilung

Die standardisierte Normalverteilung hat den Erwartungswert $E(X) = 0$ und die Varianz $D^2(X) = 1$. Kumultative Verteilungsfunktion:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-z^2} \cdot dz$$

Tabelliert für $z = 0$ bis 3,9 in Schritten von 0,1:

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

Wegen der Symmetrie gilt für $z < 0$:

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

Skalierung

Normalverteilte Zufallsgrößen mit abweichendem Erwartungswert und/oder abweichender Standardabweichung werden in Zufallsgrößen mit Erwartungswert null und Standardabweichung eins transformiert:

$$Z = \frac{X - E(X)}{\sqrt{D^2(X)}}$$

Dasselbe erfolgt für alle Werte x (inkl. Unter- und Obergrenzen):

$$z = \frac{x - E(X)}{\sqrt{D^2(X)}}$$

Umrechnung zwischen Eintrittswahrscheinlichkeiten und Wertebereichen mit der Tabelle der standardisierten Normalverteilung.

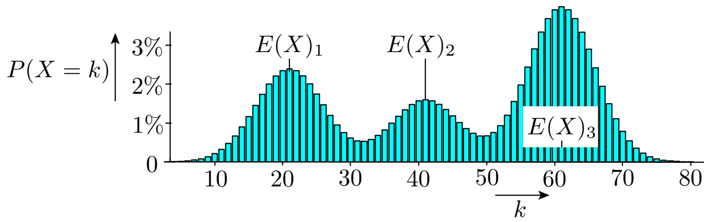
1.7 Multimodale Verteilungen

Multimodale (mehrgipflige) Verteilung

Eine multimodale Verteilung ist eine Häufigkeitsverteilung mit mehreren Gipfeln. Sie entsteht z.B. durch Mischung unterschiedlich verteilter Grundgesamtheiten. Mischung von drei Normalverteilungen mit unterschiedlichem Erwartungswert

$$P(X=k)=f(k)=0,3\cdot f_1(k)+0,2\cdot f_2(k)+0,5\cdot f_3(k)$$

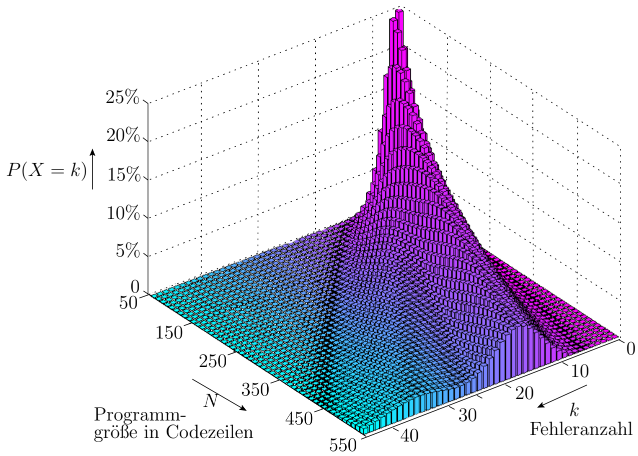
$f_i(k)$– Normalverteilungen mit Erwartungswerten $E(X)_i$ und Standardabweichung $\sqrt{D^2(X)}=5$:



Die Multimodalität deutet auf Polarisierungen der Beobachtungswerte (Zugehörigkeit zu unterschiedlichen Verteilungen). Polarisierungen können wichtige Informationen über die Natur der untersuchten Variablen liefern:

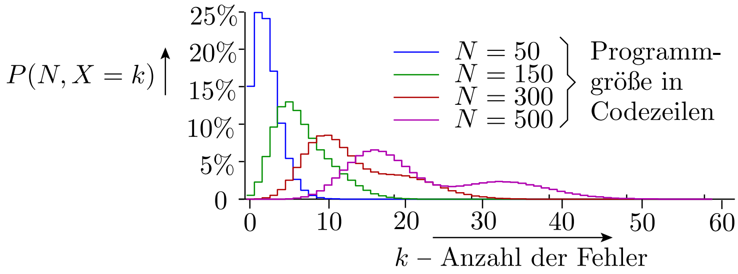
- Abhängigkeiten bei der Fehlerentstehung, bei Ausfällen beim Fehlernachweis und beim Versagen von Service-Leistungen,
- Vorliebe oder Neigung befragter Experten, z.B. bei der Einschätzung von Gefährdungen und Risiken und
- Probleme des Messverfahrens.

Beispiel sei ein Software-Team, in dem ein Anfänger und ein Profi gemeinsam Software-Bausteine aus N Code-Zeilen entwickeln, der Profi 66% der Bausteine mit ca. einem Fehler je 30 Codezeilen und der Anfänger 33% der Bausteine mit einem Fehler je 15 Codezeilen:



Die Wahrscheinlichkeit, dass ein Modul genau k Fehler enthält, ist 2/3 mal der Wahrscheinlichkeit, das es k Fehler enthält und vom Profi stammt plus 1/3 mal der Wahrscheinlichkeit, dass es vom Anfänger stammt:

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



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

Beispielaufgabe

In einer Modellfehlermenge aus $N = 25$ Fehlern mit einer Nachweiswahrscheinlichkeit $p = 40\%$ seien zehn Fehler identisch und die übrigen Fehler unabhängig voneinander nachweisbar. Bestimmen Sie

1. den Erwartungswert und
2. die Varianz der Anzahl der nachweisbaren Fehler.
3. Die Varianzerhöhung κ gegenüber $N = 25$ unabhängig voneinander nachweisbaren Modellfehlern ($p = 40\%$ für alle Modellfehler).
4. Eine Approximation der Verteilung als Mischverteilung aus zwei Binomialverteilungen.
5. Berechnung und Darstellung der Verteilung.

Lösung Aufgabenteil 1

Die Anzahl der nachweisbaren Fehler ist modellierbar als Summe von

- $i = 1$ bis 15 unabhängigen Zufallsgrößen mit der Verteilung

k	0	1	$E(X)_i$	$D^2(X)_i$
$P(X_i = k)$	$1-p$	p	p	$p \cdot (1-p)$

- und einer dazu unabhängigen Zufallsgröße X_{16} mit der Verteilung (Varianz siehe Seite 14):

k	0	10	$E(X)_{16}$	$D^2(X)_{16}$
$P(X_{16} = k)$	$1-p$	p	$10 \cdot p$	$10^2 \cdot p \cdot (1-p)$

1. Erwartungswert als Summe der Erwartungswerte der Summanden:

$$E(X) = 15 \cdot p + 10 \cdot p = 25 \cdot 40\% = 10$$

Lösung Aufgabenteil 2 und 3

k	0	1	$E(X)_i$	$D^2(X)_i$
$P(X_i = k)$	$1-p$	p	p	$p \cdot (1-p)$
k	0	5	$E(X)_{16}$	$D^2(X)_{16}$
$P(X_{16} = k)$	$1-p$	p	$10 \cdot p$	$10^2 \cdot p \cdot (1-p)$

2. Varianz als Summe der Varianzen der Summanden:

$$D^2(X) = 15 \cdot p \cdot (1-p) + 100 \cdot p \cdot (1-p) = 115$$

---

3. Gegenüber der Varianz der Summe von 25 unabhängigen Ereignissen mit Eintrittswahrscheinlichkeit p

$$D^2(X)_{\text{unabh}} = 25 \cdot p \cdot (1-p)$$

Varianzerhöhung: $\kappa = 115/25 = 4{,}6$

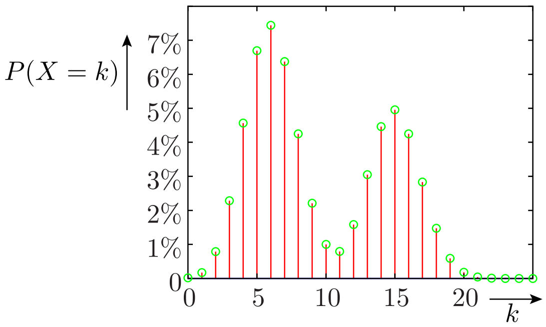
Lösung Aufgabenteil 4

2. Die Gesamtverteilung lässt sich als Mischverteilung zweier binomialverteilter Grundgesamtheiten beschreiben:

$$P\left(X=k\right)=\left(1-p\right)\cdot P_{1\ldots 15}\left(X=k\right)+p\cdot P_{1\ldots 15}\left(X+10=k\right)$$

mit

$$P_{1\ldots 15}\left(X=k\right)=\left(\begin{array}{c}15\\k\end{array}\right)\cdot 0{,}4^{k}\cdot 0{,}6^{15-k}$$



Programm zur Berechnung der Verteilung

```
clear; clf;
N=15; p=0.4;
n_k(1)=1;            % Berechnen von N über k
 n_k(k+1)=n_k(k)*(N-k+1)/k;
end
for k=0:15
 P(k+1) = n_k(k+1) *p^k *(1-p)^(N-k);
end
PP(1:16)=(1-p)*P;
PP(17:26)=[0 0 0 0 0 0 0 0 0 0];
PP(10:25) =PP(10:25) +  p*P;
stem(0:25, PP, 'r');
```

1.8 Effektive Anzahl der Zählversuche

Abhängigkeiten zwischen zu zählenden Ereignissen

Nach Gl. 13 hat das Verhältnis aus Varianz und Erwartungswert unabhängiger Zählwerte die Obergrenze die Varianz einer Binomialverteilung mit demselben Erwartungswert

$$\frac{D^{2}\left(X\right)_{\mathrm{ZV}}}{E\left(X\right)}\leq\frac{N\cdot p\cdot\left(1-p\right)}{N\cdot p}=1-p$$

(p – mittlere Eintrittswahrscheinlichkeit). Nicht nur Multimodalität, sonderen auch weniger offensichtliche Abhängigkeit können die Varianz erheblich im Vergleich zum Erwartungswert erhöhen.

Gedankenexperiment zur Varianzerhöhung

Angenommen, von den N zu addierenen Zufallsgrößen mit dem Wertebereich 0 und 1 haben immer κ denselben Wert. Dann können je κ Zufallsgrößen zu einer zusammenfasst werden mit der Verteilung:

$$\begin{array}{rcl}P\left(X_{i}=0\right)&=&1-p_{i}\\P\left(X_{i}=\kappa\right)&=&p_{i}\end{array}$$

Die Zählwerte sind dann eine Summe von N/κ unabhängigen Zufallsgrößen:

$$X=\sum_{i=1}^{N/\kappa}X_{i}$$

Der Erwartungswert der Summanden:

$$E(X_i) = 0 \cdot (1 - p_i) + \kappa \cdot p_i = \kappa \cdot p_i$$

Varianz der Summanden (nach Verschiebungssatz):

$$\begin{aligned} D^2(X_i) &= (1 - p_i) \cdot 0^2 + p_i \cdot \kappa^2 - (\kappa \cdot p_i)^2 \\ &= \kappa^2 \cdot p_i \cdot (1 - p_i) \end{aligned}$$

Erwartungswert der Summe:

$$E(X) = \sum_{i=1}^{\mathrm{N}/\kappa} \kappa \cdot p_i = N \cdot p$$

Varianz der Summe:

$$D^2(X) = \sum_{i=1}^{\mathrm{N}/\kappa} p_i \cdot \kappa^2 \cdot (1 - p_i) \leq \kappa \cdot N \cdot p \cdot (1 - p)$$

(p – mittlere Eintrittswahrscheinlichkeit). Das Verhältnis aus Varianz und Erwartungswert:

$$\frac{D^2(X)}{E(X)} = \frac{\kappa \cdot N \cdot p \cdot (1-p)}{N \cdot p} = \kappa \cdot (1 - p) \quad \text{mit } p = \frac{E(X)}{N}$$

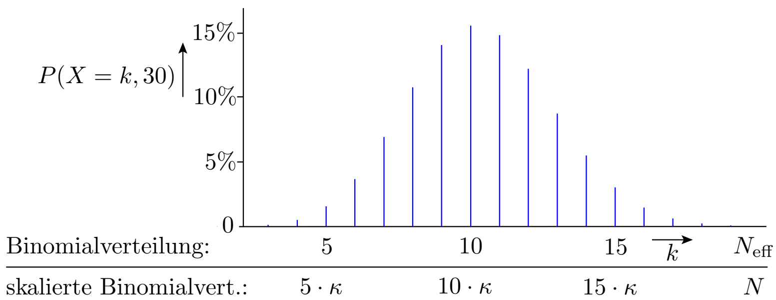
Die Rechengröße κ wird im Weiteren als Varianzerhöhung bezeichnet:

$$\kappa = \frac{D^2(X)}{E(X) \cdot \left(1 - \frac{E(X)}{N}\right)} \approx \frac{D^2_{\mathrm{S}}(X)}{E_{\mathrm{S}}(X) \cdot \left(1 - \frac{E_{\mathrm{S}}(X)}{N}\right)} \tag{16}$$

und ist aus Erwartungswert $E_{\mathrm{S}}(X)$ und Varianz $D^2_{\mathrm{S}}(X)$ einer Datenstichprobe abschätzbar.

Effektive Anzahl der Zählversuche

Eine Zählgröße aus $\frac{N}{\kappa}$ Summanden mit den Werten 0 und κ hat dieselbe Verteilung wie eine mit genauso vielen Summanden mit den Werten 0 und 1, wenn alle Zählwerte durch κ geteilt werden.



Erwartungswert und Varianz einer skalierten Binomialverteilung

Mit der Varianzerhöhung κ und der effektiven Fehleranzahl

$$N_{\text{eff}} = \frac{N}{\kappa} \tag{17}$$

ergibt sich für Varianz und Erwartungswert:

	Erwartungswert	Varianz
Binomialverteilung	$N_{\text{eff}} \cdot p$	$N_{\text{eff}} \cdot p \cdot (1-p)$
skalierte Binomialv.	$\kappa \cdot N_{\text{eff}} \cdot p = N \cdot p$	$\kappa^2 \cdot N_{\text{eff}} \cdot p \cdot (1-p)$ $= \kappa \cdot N \cdot p \cdot (1-p)$

Beispielabschätzung der effektiven Fehleranzahl

- $N = 2.000$ gezählte Werte. $N_\mathrm{S} = 10$ Wiederholungen des Zählversuchs. Ergebnisse (Zählwerte):

Versuch i	1	2	3	4	5	6	7	8	9	10
Ergebnis w_i	34	67	58	40	49	47	66	54	35	57

- Erwartungswert der Datenstichprobe nach Gl. 4:

$$E_\mathrm{S}(X) = \frac{1}{N_\mathrm{S}} \cdot \sum_{i=1}^{N_\mathrm{S}} w_i = \frac{1}{10} \cdot \sum_{i=1}^{10} w_i = 50{,}7$$

- Varianz der Datenstichprobe nach Gl. 5:

$$D_\mathrm{S}^2(X) = \frac{1}{N_\mathrm{S} - 1} \cdot \sum_{i=1}^{N_\mathrm{S}} (w_i - E(X))^2 = \frac{1}{9} \cdot \sum_{i=1}^{10} (w_i - 50{,}7)^2 = 140{,}01$$

- Geschätzter Erwartungswert: $E_\mathrm{S}(X) = 50{,}7$
- mittlere Eintrittswahrscheinlichkeit: $p = \frac{E_\mathrm{S}(X)}{N} = \frac{50{,}7}{2000} \approx 2{,}5\%$
- Geschätzte Varianz: $D_\mathrm{S}^2(X) = 140{,}01$
- κ nach Gl. 16:

$$\kappa \geq \frac{D_\mathrm{S}^2(X)}{E_\mathrm{S}(X) \cdot (1-p)} = \frac{140{,}01}{50{,}7 \cdot (1 - 2{,}5\%)} = 2{,}83$$

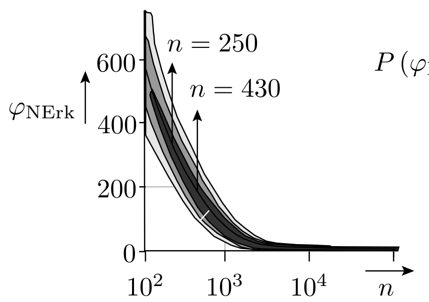
- Effektive Fehleranzahl:

$$N_\mathrm{eff} = \frac{N}{\kappa} \approx 707$$

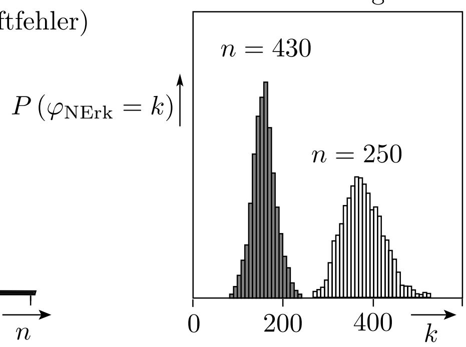
Experiment mit Haftfehlern

Kombinatorische Beispielschaltung (Benchmark c3540). 3606 simulierte, unterschiedlich nachweisbare Haftfehler. Bestimmung der Verteilung mit 1000 verschiedenen Zufallstestsätzen.

Verteilung der Anzahl der nicht erkannten Modellfehler als Funktion von n (Benchmark c3540, 3606 Haftfehler)



Verteilung für zwei Testsatzlängen



Varianzerhöhung, effektive Modellfehleranzahl

Abhängigkeiten im Fehlernachweis führen zu einer Varianzerhöhung

$$\kappa \geq \frac{D^2\left(\varphi_{\text{NErk}}\right)}{E\left(\varphi_{\text{NErk}}\right)\cdot\left(1-\frac{E\left(\varphi_{\text{NErk}}\right)}{N_{\text{PF}}}\right)}$$

Für den Versuch auf der Folie sind die Nachweisabhängigkeiten offenbar bei vielen nicht nachweisbaren Fehlern erheblich und bei wenigen nicht. (Identisch nachweisbare Fehler waren aus der Haftfehlermenge entfernt.)

n	$E(\varphi_{\text{NErk}})$	$\sqrt{D^2(\varphi_{\text{NErk}})}$	$D^2(\varphi_{\text{NErk}})$	κ	N_{eff}
160	415	43,3	1875	5,1	706
320	234	30,7	943	4,3	837
800	90	17,3	299	3,4	1057
1600	29	7,2	52	1,8	2001
3200	11	2,9	8,4	1	3606

Simulation mit Fehlerstichproben

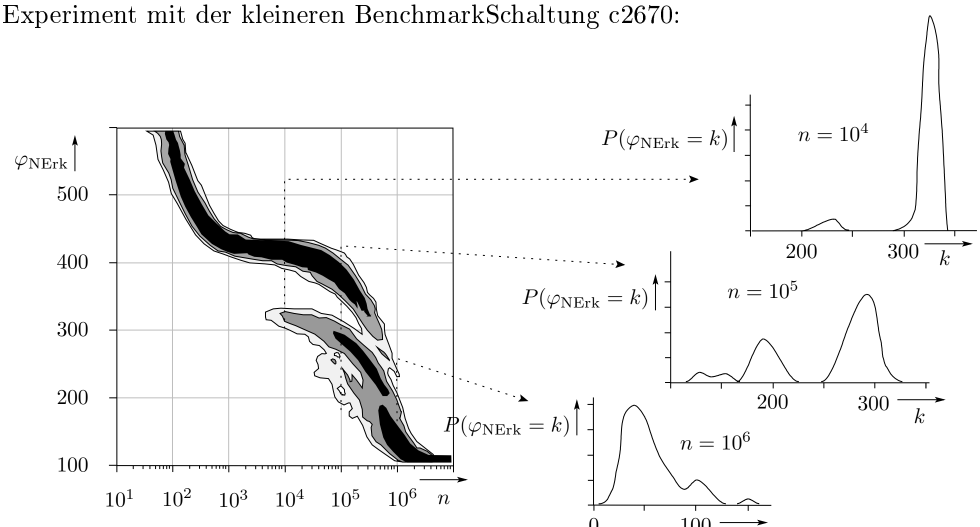
Im nachfolgenden Versuch wird eine zufällige Fehlerstichprobe von $N = 1000$ bzw. 300 der 3606 Haftfehler simuliert. Das naheliegende Ergebnis ist eine Verringerung der Nachweisabhängigkeiten, erkennbar an einer effektiven Fehleranzahl, die näher an der tatsächlich simulierten Fehleranzahl liegt.

n	160	320	800	1600	3200
N_{eff} für $N = 1000$	594	629	630	1000	1000
N_{eff} für $N = 300$	297	268	277	231	300

Bei der Stichprobe von 1000 Fehlern ist die effektive Fehleranzahl im ungünstigste Fall fast halb so groß und bei 300 Modellfehler mindestens 77% der Anzahl der simulierten Fehler.

Dasselbe Experiment mit einer anderen Schaltung

Dasselbe Experiment mit der kleineren BenchmarkSchaltung c2670:



Im Bereich von $n = 10^4$ bis 10^6 multimodale Verteilung. Offenbar ca. 80 sehr ähnlich nachweisbare Fehler mit $p_{\text{Erk}} \approx 10^5$.

2 Kennwerte schätzen

Kennwerte schätzen

Zur Bewertung der Verlässlichkeit und der Maßnahmen zu ihrer Sicherung sind

- zu erwartende Bereiche für Zählwerte (Fehler, FF, ...),

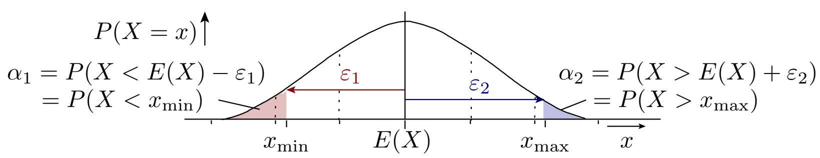
- Wahrscheinlichkeiten (Fehlerauftrittswahrscheinlichkeiten, Fehlernachweiswahrscheinlichkeiten, ...),
- Verteilungen,
- erforderliche Testsatzlängen,
- erzielbare Zuverlässigkeiten, Sicherheiten, ...

zu schätzen. Schätzergebnisse sind mit Unsicherheiten und Irrtumswahrscheinlichkeiten behaftet. Vertrauenswürdige und genaue Schätzungen verlangen eine große Anzahl von Versuchen, Zählwerten, ...

2.1 Bereichschätzung

Bereichschätzung

Bestimmung wahrscheinlicher Realisierungsbereiche unter Vorgabe einer kleinen zulässigen Irrtumswahrscheinlichkeit:



α_1	Irrtumswahrscheinlichkeit, dass Werte unterhalb des geschätzten Bereichs liegen.
α_2	Irrtumswahrscheinlichkeit, dass Werte oberhalb des geschätzten Bereichs liegen.
$\varepsilon_{1/2}$	Intervallradius, Abstand der unteren / oberen Bereichsgrenze vom Erwartungswert.

Bei $\alpha_1 = 0$ / $\alpha_2 = 0$ wird nur eine Ober- / Untergrenze geschätzt.

Bei den im weiteren behandelten Abschätzungen wird in der Regel von einer Zählverteilung ausgegangen mit der Varianz

$$D^2(X) = \kappa \cdot E(X) \cdot \left(1 - \frac{E(X)}{N}\right)$$

(κ – Abhängigkeitskoeffizient; N – Anzahl der Zählversuche) bzw. für $E(X) \ll N$ und $\kappa = 1$:

$$D^2(X) = E(X) \quad E(X) \ll N;\ \kappa = 1$$

Annäherung der Zählverteilung für die Schätzaufgaben:

- Normalverteilung: $10 \le \frac{E(X)}{\kappa} \le N - 10$
- Poissonverteilung: $E(X) < 10 \ll N$

Bei geringer Streuung ist auch dann noch eine Bereichschätzung möglich, wenn die Verteilung nicht bekannt, multimodal, ... ist, und zwar unter Nutzung der Tschebyscheffschen Ungleichung (siehe später).

Bereichsschätzung für normalverteilte Zufallsgrößen

Bereichsschätzungen für normalverteilte Zufallsgrößen werden auf die für die standardisierte Normalverteilung

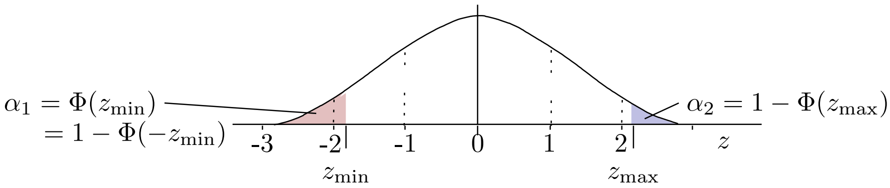
$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2} \cdot dx$$

zurückgeführt. Dazu werden die Zufallsgrößen mit abweichendem Erwartungswert und/oder abweichender Standardabweichung in eine Zufallsgröße mit Erwartungswert null und Standardabweichung eins transformiert:

$$Z = \frac{X - E(X)}{\sqrt{D^2(X)}}$$

Dasselbe erfolgt für alle Werte x (inkl. Unter- und Obergrenzen):

$$z = \frac{x - E(X)}{\sqrt{D^2(X)}}$$



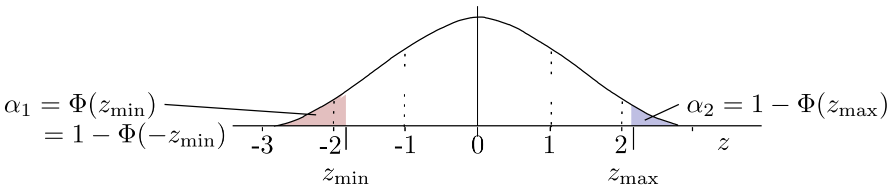
Die nachfolgende Tabelle zeigt die Wahrscheinlichkeitszuordnung der standardisierten Normalverteilung

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2} \cdot dx$$

für $z = 0$ bis 3,9 in Schritten von 0,1:

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

Ablesen von Irrtums- und Bereichswahrscheinl.

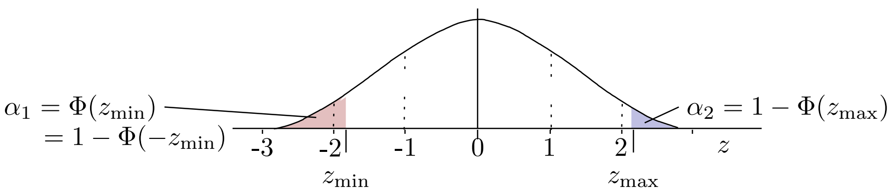


- Untergrenze: $x_{\min} = E(X) - 1{,}8 \cdot \sqrt{D^2(X)} \Rightarrow z_{\min} = -1{,}8$
- Obergrenze: $x_{\max} = E(X) + 2{,}2 \cdot \sqrt{D^2(X)} \Rightarrow z_{\max} = 2{,}2$

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

$$\begin{aligned} \alpha_1 &= 1 - \Phi^{-1}(1{,}8) = 3{,}59\% \\ \alpha_2 &= 1 - \Phi^{-1}(2{,}2) = 1{,}39\% \\ 1 - \alpha_1 - \alpha_2 &= 91{,}02\% \end{aligned}$$

Einseitige Bereichsschätzung



Für ein α_1 wird das $z_{\min}$ oder für ein α_2 das $z_{\max}$ gesucht. Umstellung der tabellierten Funktion nach:

$$\begin{aligned} z_{\min} &= -\Phi^{-1}(1-\alpha_1) \\ z_{\max} &= \Phi^{-1}(1-\alpha_2) \end{aligned}$$

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

Berechnung unterer und oberer Bereichsgrenzen:

$$\begin{aligned} z_{\min} &= -\Phi^{-1}(1-\alpha_1) \\ x_{\min} &= E(X) - \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_1) \\ z_{\max} &= -\Phi^{-1}(1-\alpha_2) \\ x_{\max} &= E(X) + \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_2) \end{aligned}$$

Symmetrische Bereichsschätzung ($\alpha_1 = \alpha_2 = \frac{\alpha}{2}$)

$$\begin{aligned} P(-z_{\min} \le Z \le z_{\max}) &= \Phi(z) - \Phi(-z) = 2 \cdot \Phi(z) - 1 \ge 1 - \alpha \\ -z_{\min} = z_{\max} &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \end{aligned}$$

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

Bereich für $E(X) \neq 0$ und/oder $D^2(X) \neq 1$:

$$x_{\min\,/\,\max} = E(X) \mp \sqrt{D^2(X)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Beispielaufgaben

Eine Zufallsgröße X hat den Erwartungswert $E(X) = 20$ und die Standardabweichung $\sqrt{D^2(X)} = 5$.

1. Mit welcher Wahrscheinlichkeit ist $X \ge 30$?
2. Mit welcher Wahrscheinlichkeit ist $X \le 15$?
3. Mit welcher Wahrscheinlichkeit liegt X im Bereich $17{,}5 \pm 7{,}5$?
4. Welche Schranke $x_{\max}$ überschreitet X nur mit einer Irrtumswahrscheinlichkeit von $\alpha_2 \le 1\%$?
5. Welche Schranke $x_{\min}$ unterschreitet X nur mit einer Irrtumswahrscheinlichkeit von $\alpha_1 \le 2\%$?
6. In welchem (symmetrischen) Bereich liegt der Wert von x mit 99%-iger Sicherheit?

Lösung Aufgabenteil 1

1. ... Mit welcher Wahrscheinlichkeit ist $X \ge 30$?

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

Gegeben: $E(X) = 20$, $\sqrt{D^2(X)} = 5$, $x_{\min} = 30$. Gesucht $P(X \ge x_{\min})$.

$$\begin{aligned} P(X \ge x_{\min}) &= 1 - \Phi\left(\frac{x_{\min} - E(X)}{\sqrt{D^2(X)}}\right) = \\ &= 1 - \Phi\left(\frac{30-20}{5}\right) = 1 - \Phi(2) = 0{,}0228 \end{aligned}$$

Lösung Aufgabenteil 2 und 3

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

2. Mit welcher Wahrscheinlichkeit ist $X \leq 15$? ($E(X) = 20$, $\sqrt{D^2(X)} = 5$)

$$P(X < 15) = \Phi\left(\frac{15-20}{5}\right) = 1 - \Phi(1) = 0{,}1587$$

---

3. Mit welcher Wahrscheinlichkeit liegt X im Bereich $17{,}5 \pm 7{,}5$? ($E(X) = 20$, $\sqrt{D^2(X)} = 5$)

$$\begin{aligned} P(10 < X < 25) &= \Phi\left(\frac{25-20}{5}\right) - \Phi\left(\frac{10-20}{5}\right) \\ &= \Phi(1) - (1 - \Phi(2)) = 0{,}8185 \end{aligned}$$

Lösung Aufgabenteil 4 und 5

α	2,28%	0,13%	0	2%	1%	0,5%	0,2%	0,1%
$\Phi^{-1}(1-\alpha)$	2	3	4	2,05	2,33	2,57	2,88	3,10

4. Obergrenze $x_{\max}$ für Irrtumswahrscheinlichkeit $\alpha_2 = 1\%$, $E(X) = 20$ und $\sqrt{D^2(X)} = 5$:

$$\begin{aligned} x_{\max} &= E(X) + \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_2) \\ &= 20 + 5 \cdot \Phi^{-1}(1 - 1\%) = 31{,}65 \end{aligned}$$

---

5. Untergrenze $x_{\min}$ für Irrtumswahrscheinlichkeit $\alpha_1 = 2\%$, $E(X) = 20$ und $\sqrt{D^2(X)} = 5$:

$$\begin{aligned} x_{\min} &= E(X) - \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_1) \\ &= 20 - 5 \cdot \Phi^{-1}(1 - 2\%) = 9{,}75 \end{aligned}$$

Lösung Aufgabenteil 6

4. Symmetrischer Bereich mit 99%-iger Sicherheit:

$$x_{\min}, x_{\max} = E(X) \mp \sqrt{D^2(X)} \cdot \Phi^{-1}\left(1 - \frac{1\%}{2}\right)$$

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

$$\begin{aligned} x_{\min}, x_{\max} &= 20 \mp 5 \cdot 2{,}57 \\ x_{\min} &= 7{,}15 \\ x_{\min} &= 32{,}85 \end{aligned}$$

Bereichsschätzungen für Zählergebnisse

Die Varianz einer Zählverteilung ist überschlagsweise die einer Binomialverteilung mit demselben Erwartungswert erhöht um den Abhängigkeitskoeffizienten κ:

$$D^2(X) \approx \kappa \cdot E(X) \cdot \left(1 - \frac{E(X)}{N}\right)$$

$$\begin{aligned} x_{\min} &= E(X) - \sqrt{\kappa \cdot E(X) \cdot \left(1 - \frac{E(X)}{N}\right)} \cdot \Phi^{-1}(1-\alpha_1) \\ x_{\max} &= E(X) + \sqrt{\kappa \cdot E(X) \cdot \left(1 - \frac{E(X)}{N}\right)} \cdot \Phi^{-1}(1-\alpha_2) \end{aligned}$$

Für eine symmetrische Bereichsschätzung ($\alpha_1 = \alpha_2 = \frac{\alpha}{2}$):

$$x_{\min\,/\,\max} = E(X) \mp \sqrt{\kappa \cdot E(X) \cdot \left(1 - \frac{E(X)}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Beispielaufgabe

Die zu erwartende Anzahl der Fehlfunktionen bei der Abarbeitung von $N_{\mathrm{SL}} = 20.000\,\mathrm{SL}$ sei $E(N_{\mathrm{FF}}) = 100\,\mathrm{FF}$. In welchem symmetrischen Bereich um den Erwartungswert liegt N_{FF} mit einer Irrtumswahrscheinlichlichkeit von $\alpha = 1\%$, wenn die Fehlfunktionen voneinander unabhängig auftreten ($\kappa = 1$)?

Lösung

$N_{\mathrm{SL}} = 20.000\,\mathrm{SK}$, $E(N_{\mathrm{FF}}) = 100\,\mathrm{FF}$, Irrtumswahrscheinlichlichkeit $\alpha = 1\%$, $\kappa = 1$:

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

$$\begin{aligned} N_{\mathrm{FF.\,min\,/\,max}} &= E(N_{\mathrm{FF}}) \mp \sqrt{\kappa \cdot E(N_{\mathrm{FF}}) \cdot \left(1 - \frac{E(N_{\mathrm{FF}})}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \\ &= 100 \mp \sqrt{1 \cdot 100 \cdot \left(1 - \frac{100}{20000}\right)} \cdot 2{,}57 \\ &= 100 \mp 25{,}6 \end{aligned}$$

2.2 Kleine Zählwerte

Kleine Zählwerte

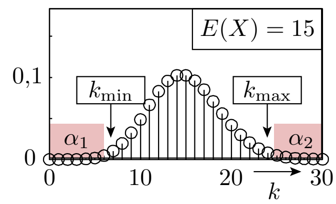
Für weniger als $10 \ll N$ eingetretene / nicht eingetretene Zählereignisse ist deren Anzahl näherungsweise poissonverteilt. Die Abschätzung einer Unter-, einer Obergrenze oder eines Bereichs basiert auf der kumulativen Poisson-Verteilung:

$$\mathrm{poiscdf}(E(X), k) = P(X \le k) = \sum_{i=0}^{k} e^{-E(X)} \cdot \frac{E(X)^i}{i!}$$

Berechnung mit Matlab/Octave:

```
function [P] = poiscdf(E, k)
 w=1; s=1;
 for i=1:k
  w = w*E/i; s = s + w;
 end
 P =e^(-E)*s;
```

Obere und untere Bereichsgrenze



- Untere Bereichsgrenze:
$$\mathrm{poiscdf}(E(X), k_{\min} - 1) \le \alpha_1$$
- Obere Bereichsgrenze:
$$1 - \mathrm{poiscdf}(E(X), k_{\max}) < \alpha_2$$
- symmetrischer Bereich:
$$\left(\mathrm{poiscdf}(E(X), k_{\min} - 1) \le \frac{\alpha}{2}\right) \wedge \left(1 - \mathrm{poiscdf}(E(X), k_{\max}) < \frac{\alpha}{2}\right)$$

Eine garantierbare Untergrenze $k_{\min} \ge 1$ verlangt $E(X) \ge -\ln(\alpha)$, z.B. für $\alpha = 1\%$ $E(X) > 4{,}6$.

Numerische Suche des größten $k_{\min}$

$$\mathrm{poiscdf}\left(E\left(X\right), k_{\min} - 1\right) \leq \alpha_1$$

```
function  [k_min] = pois_k_min(alpha1, E)
 assert (poiscdf(E, 0)<=alpha)
 k_min = 1
 while (poiscdf(E, k_min-1)<=alpha1)
   k_min = k_min + 1;
 end
 k_min = k_min - 1;
```

Suchablauf für $\alpha_1 = 1\%$ und $E\left(X\right) = 9$ (pois_k_min(0.01, 9):

```
k_min= 0 alpha1=0.0123%
k_min= 1 alpha1=0.1234%
k_min= 2 alpha1=0.6232%
k_min= 3 alpha1=2.1226% (poiscdf (9, 2) ≤ 1%)
```

Garantierbare Untergrenze: $k_{\min} = 3$. Nur ein Drittel von $E\left(X\right) = 9$.

Numerische Suche kleinsten $k_{\max}$

$$1 - \mathrm{poiscdf}\left(E\left(X\right), k_{\max}\right) \leq \alpha_2$$

```
function  [k_max] = pois_k_max(alpha, E)
 k_max = 0;
 while (poiscdf(E, k_max) <= 1-alpha)
   k_max = k_max + 1;
 end
```

Suchablauf für $\alpha_1 = 1\%$ und $E\left(X\right) = 9$ (pois_k_max(0.01, 9):

```
k_max= 0 alpha2=99.9877%
k_max= 1 alpha2=99.8766%
...
k_max=16 alpha2=1.1106%
k_max=17 alpha2=0.5320% (1 − poiscdf (9, 17) ≤ 1%)
```

Die garantierbare Obergrenze ist $k_{\max} = 17$ und fast doppelt so groß wie der Erwartungswert $E\left(X\right) = 9$.

Beispielaufgabe

1. Die mittlere Eintrittswahrscheinlichkeit einer Fehlfunktion sei $p = 10^{-5}$ je Service-Anforderung. Bis zu wie vielen Service-Anforderungen N ist die Wahrscheinlichkeit, dass keine Fehlfunktion eintritt, mindestens 99%?
2. Wie wahrscheinlich ist es, dass eine poisson-verteilte Zufallsgröße mit dem Erwartungswert $E\left(X\right) =$ 0,6 keinen Wert größer 3 hat?

Lösung Aufgabenteil 1

Für $p = 10^{-5}$ soll $P\left(X = 0\right) \geq 99\%$ sein:

$$\begin{aligned} P\left(X=0\right) &= e^{-p\cdot N} \geq 99\% \\ N &\leq \frac{1}{p}\cdot\left(-\ln\left(99\%\right)\right) \approx 10^3 \end{aligned}$$

Lösung Aufgabenteil 2

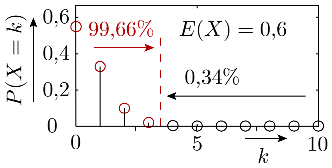
Gegeben: $E(\varphi) = 0{,}6$, $k_{\mathrm{max}} = 3$. Gesucht:

$$P(X \leq 3) = \sum_{k=0}^{3} e^{-0,6} \cdot \frac{0,6^k}{k!}$$

Programmtechnische Lösung:

```
printf('P(k<=3)=%4.2f%%\n', 100*poiscdf(0.6, 3));
>> P(k<=3)=99.66%
```

Ergebnis graphisch:

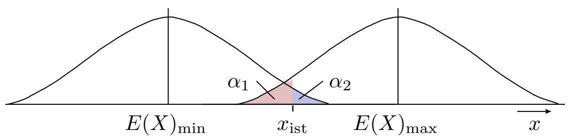


2.3 $E(X)$ und Eintrittsw.

Bereichsschätzung für den Erwartungswert

Der Erwartungswert zu einem beobachteten Ereignis ist

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



Für normalverteilte Zufallsgrößen:

$$\begin{aligned} E(X)_{\mathrm{min}} &= x_{\mathrm{ist}} - \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_2) \\ E(X)_{\mathrm{max}} &= x_{\mathrm{ist}} + \sqrt{D^2(X)} \cdot \Phi^{-1}(1-\alpha_1) \end{aligned}$$

Normalverteilte Zählwerte

Für Zählwerte liefert ein beobachteter Wert eine Abschätzung für die Varianz:

$$D^2(X) \approx \kappa \cdot x_{\mathrm{ist}} \cdot \left(1 - \frac{x_{\mathrm{ist}}}{N}\right)$$

(N – Anzahl der Zählversuche; κ – Abhängigkeitskoeffizient).

$$\begin{aligned} E(X)_{\mathrm{min}} &\approx x_{\mathrm{ist}} - \sqrt{\kappa \cdot x_{\mathrm{ist}} \cdot \left(1 - \frac{x_{\mathrm{ist}}}{N}\right)} \cdot \Phi^{-1}(1-\alpha_2) \\ E(X)_{\mathrm{max}} &= x_{\mathrm{ist}} + \sqrt{\kappa \cdot x_{\mathrm{ist}} \cdot \left(1 - \frac{x_{\mathrm{ist}}}{N}\right)} \cdot \Phi^{-1}(1-\alpha_1) \end{aligned}$$

Für eine symmetrische Bereichsschätzung ($\alpha_1 = \alpha_2 = \frac{\alpha}{2}$):

$$E(X)_{\mathrm{min\,/\,max}} = x_{\mathrm{ist}} \mp \sqrt{\kappa \cdot x_{\mathrm{ist}} \cdot \left(1 - \frac{x_{\mathrm{ist}}}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Beispielaufgabe

1. In welchen Bereich liegt der Erwartungswert für ein Zählergebnis $x_{\text{ist}} = 20$ bei $N = 100$ Zählversuchen?
2. Bei der Abarbeitung von $N_{\text{SL}} = 20.000\,\text{SL}$ wurden $N_{\text{FF}} = 100\,\text{FF}$ beobachtet. In welchem symmetrischen Bereich liegt der Erwartungswert $E(N_{\text{FF}})$ mit Irrtumswahrscheinlichkeit $\alpha = 1\%$, wenn die Fehlfunktionen voneinander unabhängig auftreten ($\kappa = 1$)?

Lösung Aufgabenteil 1

$$\begin{aligned} E(X)_{\text{min / max}} &= x_{\text{ist}} \mp \sqrt{\kappa \cdot x_{\text{ist}} \cdot \left(1 - \frac{x_{\text{ist}}}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \\ &= 20 \mp 2{,}57 \cdot \sqrt{20 \cdot \left(1 - \frac{20}{100}\right)} = [9{,}72,\ 30{,}28] \end{aligned}$$

Lösung Aufgabenteil 2

Erwartungswert für $x_{\text{ist}} = 100\,\text{FF}$, $N = 20.000\,\text{SL}$, $\alpha = 1\%$, $\kappa = 1$:

$$\begin{aligned} E(N_{\text{FF}})_{\text{min / max}} &= x_{\text{ist}} \mp \sqrt{\kappa \cdot x_{\text{ist}} \cdot \left(1 - \frac{x_{\text{ist}}}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \\ &= 100 \mp \sqrt{1 \cdot 100 \cdot \left(1 - \frac{100}{20000}\right)} \cdot 2{,}57 \\ &= 100 \mp 25{,}6 \end{aligned}$$

Schätzen von Eintrittswahrscheinlichkeiten

Die Eintrittswahrscheinlichkeit eines Zählwertes ist der Quotient aus der Anzahl der Zählwerte und der Versuchsanzahl N. Für ein experimentell bestimmten Zählwert x_{ist} betragen

- der Schätzwert

$$p_{\text{S}} = \frac{x_{\text{ist}}}{N}$$

- die Untergrenze

$$p_{\text{min}} = \frac{E(X)_{\text{min}}}{N}$$

- und die Obergrenze

$$p_{\text{max}} = \frac{E(X)_{\text{max}}}{N}$$

Je größer die Irrtumswahrscheinlichkeit α und die Anzahl N der Zählversuche desto kleiner ist der angebbare Bereich $[p_{\text{min}}, p_{\text{max}}]$ für die Wahrscheinlichkeit.

Mindestwert für das Zählergebnis

Für eine normalverteilte Zählgröße gilt bei $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$:

$$E(X)_{\text{min / max}} = x_{\text{ist}} \mp \sqrt{\kappa \cdot x_{\text{ist}} \cdot \left(1 - \frac{x_{\text{ist}}}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Relativer Intervallradius:

$$\begin{aligned}\varepsilon_{\text{rel}} &= \frac{E(X)_{\max} - x_{\text{ist}}}{x_{\text{ist}}} = \frac{x_{\text{ist}} - E(X)_{\min}}{x_{\text{ist}}} \\ &= \sqrt{\frac{\kappa}{x_{\text{ist}}} \cdot \left(1 - \frac{x_{\text{ist}}}{N}\right)} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\end{aligned}$$

Das Mindestzählergebnis, damit der relative Intervallradius unter Vorgabe einer Irrtumswahrscheinlichkeit α nicht größer als ε_{rel} ist beträgt:

$$x_{\text{ist.min}} = \frac{1}{\frac{1}{N} + \frac{\varepsilon_{\text{rel}}^2}{\kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2}}$$

Für $N \gg x_{\text{ist}}$ vereinfacht sich die Abschätzung der Mindestanzahl der Zählversuche zu:

$$x_{\text{ist.min}} = \frac{1}{\frac{1}{N} + \frac{\varepsilon_{\text{rel}}^2}{\kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2}} \approx \frac{\kappa \cdot \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{\varepsilon_{\text{rel}}^2} \tag{18}$$

Mindestanzahl der Zählversuche

Zur Schätzung der Eintrittswahrscheinlichkeit $p_{\text{S}} = \frac{x_{\text{ist}}}{N}$ mit einem relativen Intervallradius

$$p \in p_{\text{S}} \cdot (1 \mp \varepsilon_{\text{rel}})$$

und einer Irrtumswahrscheinlichkeit α muss die Versuchsanzahl N solange erhöht werden, bis der Zählwert

$$x_{\text{ist.min}} = \frac{1}{\frac{1}{N} + \frac{\varepsilon_{\text{rel}}^2}{\kappa \cdot \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2}} \approx \frac{\kappa \cdot \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{\varepsilon_{\text{rel}}^2}$$

erreicht ist.

Beispielaufgaben

Wie groß ist die zu erwartende Anzahl N der Zählversuche, um zu zeigen, dass die Eintrittswahrscheinlichkeit $p = 1\% \cdot (1 \mp 10\%)$ beträgt? Abhängigkeiten zwischen den Zählergebnissen seien zu vernachlässigen ($\kappa = 1$). Die zulässige Irrtumswahrscheinlichkeit sei $\alpha = 1\%$.

Zur Kontrolle

$\kappa = 1$, $\alpha = 1\%$, $p = 1\% \cdot (1 \mp 10\%)$

Zu erreichendes Zählergebnis:

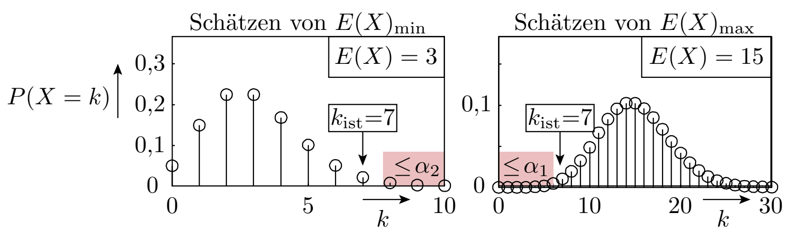
$$x_{\text{ist.min}} = \frac{\kappa \cdot \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{\varepsilon_{\text{rel}}^2} = 661$$

Dafür zu erwartende Anzahl der Zählversuche:

$$N \approx \frac{x_{\text{ist.min}}}{p} = 66100$$

2.4 Wahrsch. seltener Ereig.

Bereich des Erwartungswerts für seltene Ereiegnisse



Untere Bereichsgrenze:

$$\text{poiscdf}\left(E(X)_{\min}, k_{\text{ist}}\right) \geq 1 - \alpha_2$$

Obere Bereichsgrenze (Voraussetzung $k_{\text{ist}} > 0$):

$$\text{poiscdf}\left(E(X)_{\max}, k_{\text{ist}} - 1\right) \leq \alpha_1$$

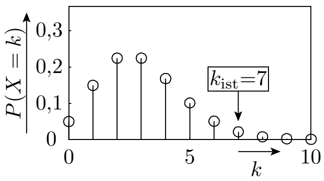
Numerische Berechnung der Untergrenze

```
function [E_min] = pois_EUG(alpha, k_ist)
 EE_min = k_ist+1; dE=E_min;
 while (1)
  dE=dE/2
  a = 1-poiscdf(E_min, k_ist);
  if (a > alpha)            E_min = E_min - dE;
  elseif (a< 0.99* alpha)   E_min = E_min + dE;
  else                      return
  end
 end
```

Für $k_{\mathrm{ist}} = 7$, $\alpha_2 = 1\%$

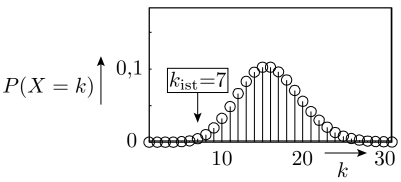
```
1-poiscdf(8.000, 7)=54.704%,
1-poiscdf(4.000, 7)=5.113%,
1-poiscdf(2.000, 7)=0.110%, ...
```

$E(X)_{\min} = 2{,}91$



Berechnung der Obergrenze

```
function [E_max] = pois_EOG(alpha, k_ist)
 E_max = 10*k_ist+(1/alpha); dE=E_max;
 while (1)
  dE=dE/2; a = poiscdf(E_max, k_ist-1);
  if (a < alpha)     E_max = E_max - dE;
  else               E_max = E_max + dE;
  end if
  if dE<0.001 return
  end % if
 end % while
end % function
```



Für $k_{\mathrm{ist}} = 7$, $\alpha_2 = 1\%$

poiscdf(85.000, 6)=0.000%
poiscdf(42.500, 6)=0.000%
poiscdf(21.250, 6)=0.010%
poiscdf(10.625, 6)=9.540%
poiscdf(15.938, 6)=0.417%
...

$E(X)_{\max} \approx 14{,}57$

Tabellierung der Erwartungswertbereiche

	$\alpha_1=\alpha_2=0{,}5\%$		$\alpha_1=\alpha_2=1\%$		$\alpha_1=\alpha_2=2\%$		$\alpha_1=\alpha_2=10\%$		$\alpha_1=\alpha_2=20\%$	
x_{ist}	$E(X)_{\min}$	$E(X)_{\max}$	$E(X)_{\min}$	$E(X)_{\max}$	$E(X)_{\min}$	$E(X)_{\max}$	$E(X)_{\min}$	$E(X)_{\max}$	$E(X)_{\min}$	$E(X)_{\max}$
0	0,005	–	0,01	–	0,02	–	0,105	–	0,223	–
1	0,103	5,298	0,148	4,606	0,215	3,912	0,532	2,303	0,824	1,609
2	0,338	7,430	0,436	6,638	0,567	5,834	1,102	3,890	1,534	2,995
3	0,672	9,273	0,823	8,406	1,016	7,516	1,744	5,323	2,296	4,279
4	1,078	10,978	1,279	10,045	1,529	9,084	2,432	6,681	3,089	5,514
5	1,537	12,593	1,785	11,605	2,089	10,580	3,152	7,993	3,903	6,721
6	2,037	14,150	2,330	13,109	2,684	12,027	3,894	9,275	4,733	7,906

Beispielaufgabe

1. Mit $N = 10^5$ Service-Anforderungen wurden drei Fehlfunktionen beobachtet. Auf welche Unter- und Obergrenze für die Wahrscheinlichkeit einer Fehlfunktion je Service-Aufruf p lässt sich mit $\alpha = 1\%$ schließen?
2. Ein Kontrolle hat von 1.000 FF 4 FF nicht erkannt. In welchem Bereich liegt mit einer Irrtumswahrscheinlichkeit von $\alpha = 1\%$ die Maskierungswahrscheinlichkeit?

Lösung Aufgabenteil 1

- $E(X)_{\text{min / max}}$ für $k_{\text{ist}} = 3$ und $\alpha = 1\%$ aus der Tabelle ablesen:

	$\alpha_1 = \alpha_2 = 0{,}5\%$		$\alpha_1 = \alpha_2 = 1\%$		...
X_{ist}	$E(X)_{\text{min}}$	$E(X)_{\text{max}}$	$E(X)_{\text{min}}$	$E(X)_{\text{max}}$	...
2	0,338	7,430	0,436	6,639	...
3	0,672	9,274	0,823	8,405	...

- Der gesuchte Wahrscheinlichkeitsbereich ist der Erwartungswertbereich geteilt durch $N = 10^5$ Service-Anforderungen:

$$p_{\text{min}} = 8{,}23 \cdot 10^{-6}; \quad p_{\text{max}} = 8{,}405 \cdot 10^{-5}$$

Lösung Aufgabenteil 2

- $E(X)_{\text{min / max}}$ für $k_{\text{ist}} = 4$ und $\alpha = 1\%$ aus der Tabelle ablesen:

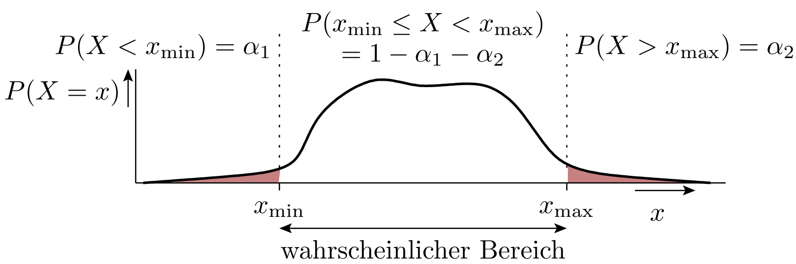
	$\alpha_1 = \alpha_2 = 0{,}5\%$		$\alpha_1 = \alpha_2 = 1\%$		...
X_{ist}	$E(X)_{\text{min}}$	$E(X)_{\text{max}}$	$E(X)_{\text{min}}$	$E(X)_{\text{max}}$	...
2	0,338	7,430	0,436	6,639	...
3	0,672	9,274	0,823	8,405	...
4	1,078	10,978	1,279	10,045	...

- Die Maskierungswahrscheinlichkeit ist die Maskierungsanzahl durch 1.000:

$$p_{\text{min}} = 1{,}28 \cdot 10^{-3}; \quad p_{\text{max}} = 1{,}00 \cdot 10^{-2}$$

2.5 Verteilungen unbekannt

Verteilungen unbekannt



Die Bestimmung eines wahrscheinlichen Intervalls $[x_{\text{min}}, x_{\text{max}}]$, in dem der Wert einer Zufallsgröße X mit einer Wahrscheinlichkeit $1 - \alpha$ liegt, ist auch möglich, wenn die Verteilung nicht bekannt, multimodal, ... (beliebig) ist. Voraussetzung ist eine hinreichend kleine Varianz im Vergleich zum Abstand zwischen dem Erwartungswert und den Wertebereichsgrenzen.

Das schwache Gesetz der großen Zahlen

Nach der tschebytscheffschen Ungleichung:

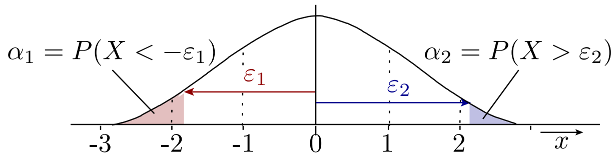
$$P(|x - E(X)| \ge \varepsilon) \le \frac{D^2(X)}{\varepsilon^2}$$

ist die Wahrscheinlichkeit, das der Wert einer Zufallsgröße mehr als ein Intervallradius ε von seinem Erwartungswert abweicht, nicht größer als das Verhältnis der Varianz zum Quadrat des Intervallradius ε. Bei Zulassen einer Irrtumswahrscheinlichkeit α beträgt der Intervallradius mindestens:

$$\varepsilon \ge \sqrt{\frac{D^2(X)}{\alpha}}$$

Ausgehend von einer bekannten Realisierung x_{ist} beschränkt das den Bereich des Erwartungswerts auf $E(X) \in x_{\text{ist}} \pm \varepsilon$. Bei bekanntem oder vermutetem Erwartungswert $E(X)$ ist der wahrscheinliche Bereich, in dem $1 - \alpha$ der experimentellen Ergebnisse liegen werden $x_{\text{min / max}} \in E(X) \pm \varepsilon$.

Zum Vergleich: Intervallradius bei Normalverteilung



Bei einer standardisierten Normalverteilung beträgt der Intervallradius für $\alpha_1 = \alpha_2 = \alpha$:

$$\varepsilon = \varepsilon_1 = \varepsilon_2 = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Für Normalverteilungen mit einer Standardabweichung $\neq 1$ ist er um die Standardabweichung größer:

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

Irrtumswahrscheinlichkeit α		10%	5%	2%	1%	0,5%	0,2%
$\varepsilon_{\mathrm{DX1}} = \frac{\varepsilon}{\sqrt{D^2(X)}}$	beliebige Verteilung	3,16	4,47	7,07	10	14,1	22,4
	Normalverteilung	1,65	1,96	2,33	2,75	2,81	3,10

Beispielaufgabe

Gegeben sei eine Stichprobe gemessener Widerstandswerte in kΩ:

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

Aus dieser Stichprobe soll

1. ohne weitere Vorkenntnisse über die Verteilung und
2. unter der Annahme, dass die Widerstandswerte normalverteilt sind,

auf den möglichen Bereich des Erwartungswertes geschlussfolgert werden. Zugelassene Irrtumswahrscheinlichkeit $\alpha = 2\%$.

Lösung

Erwartungswert und Standardabweichung der Datenstichprobe:

$$E_{\mathrm{S}}(R) = \frac{1}{8}(10{,}3 + \ldots)\ \mathrm{k}\Omega = 10{,}025\,\mathrm{k}\Omega$$

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

Auf Standardabweichung 1 normierter Intervallradius:

Irrtumswahrscheinlichkeit α		10%	5%	2%	1%	0,5%	0,2%
$\varepsilon_{\mathrm{DX1}} = \frac{\varepsilon}{\sqrt{D^2(X)}}$	beliebige Verteilung	3,16	4,47	7,07	10	14,1	22,4
	Normalverteilung	1,65	1,96	2,33	2,75	2,81	3,10

1. Ohne Kenntnis der Verteilung:

$$E(R) \in 10{,}025\,\mathrm{k}\Omega \pm 7{,}07 \cdot 647\,\Omega = [5{,}3\,\mathrm{k}\Omega,\ 14{,}8\,\mathrm{k}\Omega]$$

2. Für normalverteilte Widerstandswerte:

$$E(R) \in 10{,}025\,\mathrm{k}\Omega \pm 2{,}33 \cdot 647\,\Omega = [8{,}5\,\mathrm{k}\Omega,\ 11{,}5\,\mathrm{k}\Omega]$$