# Kapitel 2

# Spezielle Verteilungen einer Variablen

In diesem Kapitel werden wir einige häufig benutzte Verteilungen, die von einer Variablen abhängen, vorstellen.

## 2.1 Binomial-Verteilung

Binomial-Verteilungen treten auf, wenn man die betrachteten Ereignisse in zwei Klassen mit den Eigenschaften $A$ und $\bar{A}$ zerlegen kann, die mit komplementären Wahrscheinlichkeiten auftreten:

| Eigenschaft | Wahrscheinlichkeit |
|---|---|
| $A$ | p |
| $\bar{A}$ | 1-p |

Wie groß ist die Wahrscheinlichkeit $W_k^n$, bei $n$ Ereignissen $k$ mit der Eigenschaft $A$ zu erhalten?

**Beispiele:**

- Aus einer Übungsaufgabe: Die Wahrscheinlich ein Ei zu finden ist $p$. Wie groß ist die Wahrscheinlichkeit bei $n$ versteckten Eiern $k$ zu finden. Die Kenntnis der entsprechenden Wahrscheinlichkeitsverteilung wird uns helfen, den Fehler in der Abschätzung der Effizienz zu bestimmen.

- Wie groß ist die Wahrscheinlichkeit, dass sich in einem System mit $n$ Spins $k$ in Richtung eines vorgegebenen Magnetfeldes einstellen? Die Wahrscheinlichkeit für jeden einzelnen Spin ist abhängig von Temperatur und Feldstärke: $p = f(T, B)$.

- Es seien $n$ Teilchen in einer Box mit Volumen $V$. Wie groß ist die Wahrscheinlichkeit, $k$ davon in einem Teilvolumen $V_1$ zu finden? Die Wahrscheinlichkeit für jedes einzelne Teilchen ist offensichtlich $p = V_1/V$.

- Das Galton-Brett ist eine Anordnung von Nägeln wie in Abb. 2.1 gezeigt. Man setzt eine Kugel auf den obersten Nagel, von dem sie zufällig nach rechts oder

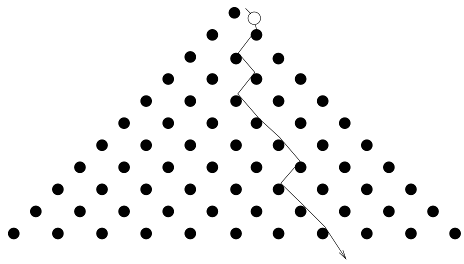


Abbildung 2.1: Galton-Brett.

links auf einen Nagel der nächsten Reihe fällt und so weiter. Wenn alles schön symmetrisch ist, fällt die Kugel jeweils mit gleicher Wahrscheinlichkeit nach links oder rechts: $p = 0.5$.

- Am Computer kann man dem Galton-Brett auch einen beliebigen Parameter $p$ zuordnen: Man würfelt $n$-mal im Intervall $[0, 1]$ und ermittelt die Anzahl $k$, für die die Zufallszahl kleiner als $p$ ist (das ist zum Beispiel, wie häufig die Kugel nach links gefallen ist).

**Herleitung der Binomial-Verteilung:** Es gibt verschiedene Kombinationen, in einer Gesamtheit von $n$ Ereignissen $k$ mit der Eigenschaft $A$ zu erhalten, die sich durch die Reihenfolge des Auftretens von $A$ unterscheiden. Zum Beispiel gibt es für $n = 3$ und $k = 2$ offensichtlich 3 mögliche Kombinationen:

$$\begin{array}{ccc} 1 & 2 & 3 \\ \hline \\ A & A & \bar{A} \\ A & \bar{A} & A \\ \bar{A} & A & A \end{array} \tag{2.1}$$

Jede einzelne Kombination zu festen Zahlen $n$ und $k$ hat die gleiche Wahrscheinlichkeit. Diese ergibt sich als Produkt der Wahrscheinlichkeiten, jeweils für ein bestimmtes Ereignis die Eigenschaft $A$ oder $\bar{A}$ zu haben. Zum Beispiel würde man in der ersten Zeile von (2.1) $p \cdot p \cdot (1 - p) = p^2(1 - p)$ erhalten. Allgemein ergibt sich:

$$p^k \cdot (1 - p)^{n-k}. \tag{2.2}$$

Um dieses Produkt der Wahrscheinlichkeiten zu bilden, muss die Wahrscheinlichkeit für das Auftreten von $A$ unabhängig davon sein, wie häufig $A$ bereits gezählt wurde. Zum Beipiel müssen bei einer Ziehung aus einer endlichen Anzahl von schwarzen und weissen Kugeln die Kugeln immer wieder zurückgelegt werden, damit die Wahrscheinlichkeiten für schwarz und weiss sich nicht ändern.

Die Wahrscheinlichkeit für das Auftreten irgendeiner Kombination mit $k$-mal der Eigenschaft $A$ ist die Summe der Wahrscheinlichkeiten der einzelnen Kombinationen (in (2.1) also die Summe der Wahrscheinlichkeiten der 3 Zeilen, das ist $3\,p^2(1 - p)$). Da jede dieser Wahrscheinlichkeiten gleich ist, muss man also nur die Anzahl der möglichen Kombinationen bestimmen.

Um nun allgemeiner die Anzahl der Kombinationen mit $k$-mal der Eigenschaft $A$ zu bestimmen, beginnt man damit, zunächst $k$ unterscheidbare Ereignisse $A_1, \ldots, A_k$ auf $n$ Stellen zu verteilen. In (2.1) würden sich die beiden $A$ in einer Spalte durch einen Index 1 und 2 ($A_1,\ A_2$) unterscheiden, dessen Vertauschung dann zu einer Verdoppelung der Möglichkeiten führt (von 3 auf 6). Um nun die Anzahl der Anordnungen bei $k$ Ereignissen zu bestimmen, kann man die Ereignisse nacheinander auf die jeweils noch freien Plätze verteilen:

| | | |
|---|---|---|
| $A_1$ | $n$ | Möglichkeiten (alle Plätze sind noch frei), |
| $A_2$ | $n-1$ | Möglichkeiten (ein Platz ist bereits mit $A_1$ besetzt), |
| | $\ldots$ | |
| $A_k$ | $n-(k-1)$ | Möglichkeiten ($k-1$ Plätze sind von $A_1$ bis $A_{k-1}$ besetzt). |

Das sind insgesamt

$$n \cdot (n-1) \ldots \cdot n-(k-1) = \frac{n!}{(n-k)!} \tag{2.3}$$

Möglichkeiten, von der jede aber in $k!$ Anordnungen der $A_i$ auftreten (in (2.1) gibt es für die 2 $A$-Ereignisse jeweils 2 Permutationen). Da nach der Reihenfolge nicht unterschieden wird, ergibt sich schließlich für die Gesamtzahl der Kombinationen, die Eigenschaft $A$ $k$-mal auf $n$ Ereignisse zu verteilen:

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} \tag{2.4}$$

Der Ausdruck $\binom{n}{k}$ beschreibt die Binomialkoeffizienten, die sich bekanntlich mit dem Pascalschen Dreieck darstellen lassen:

$$\begin{array}{c|ccccccccc}
n & & & & & & & & & \\
0 & & & & & 1 & & & & \\
1 & & & & 1 & & 1 & & & \\
2 & & & 1 & & 2 & & 1 & & \\
3 & & 1 & & 3 & & 3 & & 1 & \\
4 & 1 & & 4 & & 6 & & 4 & & 1 \\
 & & . & & . & & . & & . & \\
k \rightarrow & & & & & & & & &
\end{array}$$

Damit ergibt sich die Binomial-Verteilung:

$$W_k^n = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \tag{2.5}$$

**Normierung:** Es ist einfach zu sehen, dass die Normierung

$$\sum_{k=0}^{n} W_k^n = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = 1 \tag{2.6}$$

richtig ist, weil die Summe gerade der Formel für $(a+b)^n$ mit $a = p$ und $b = 1-p$ entspricht:

$$\sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = (p+1-p)^n = 1^n = 1 \tag{2.7}$$

**Mittelwert:**

$$\begin{aligned}
\langle k \rangle &= \sum_{k=0}^{n} k \cdot W_k^n \\
&= \sum_{k=0}^{n} k \cdot \frac{n!}{(n-k)!k!} \cdot p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!} \cdot p^k \cdot (1-p)^{n-k} \\
&= np \cdot \sum_{k=1}^{n} \frac{(n-1)!}{[(n-1)-(k-1)]!(k-1)!} \cdot p^{k-1} \cdot (1-p)^{n-1-(k-1)} \qquad \text{mit } n-k=(n-1)-(k-1) \\
&= np \cdot \sum_{k'=0}^{n'} \frac{n'!}{(n'-k')!k'!} \cdot p^{k'} \cdot (1-p)^{n'-k'} = np \qquad \text{mit } n'=n-1;\ k'=k-1
\end{aligned} \tag{2.8}$$

Die letzte Zeile benutzt die Normierung der Summe auf 1. Damit ergibt sich für den Mittelwert von $k$:

$$\langle k \rangle = np \tag{2.9}$$

Zum Beipiel ist für $p = 0.5$ wie zu erwarten $\langle k \rangle = n/2$.

**Varianz:** Die Varianz ist die mittlere quadratische Abweichung vom Mittelwert, die sich nach (1.44) zerlegen läßt:

$$\sigma^2 = \langle (k - \langle k \rangle)^2 \rangle = \langle k^2 \rangle - \langle k \rangle^2 \tag{2.10}$$

Der Erwartungswert von $k^2$ läßt sich ähnlich wie der Mittelwert bestimmen:

$$\begin{aligned}
\langle k^2 \rangle &= \sum_{k=0}^{n} k^2 \cdot W_k^n \\
&= \sum_{k=0}^{n} k^2 \cdot \frac{n!}{(n-k)!k!} \cdot p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{(n-k)!(k-1)!} \cdot p^k \cdot (1-p)^{n-k} \\
&= np \cdot \sum_{k'=0}^{n'} (k'+1) \cdot \frac{n'!}{(n'-k')!k'!} \cdot p^{k'} \cdot (1-p)^{n'-k'} \qquad (n' = n-1;\ k' = k-1) \\
&= np \cdot \left[ 1 + \sum_{k'=0}^{n'} k' \cdot \frac{n'!}{(n'-k')!k'!} \cdot p^{k'} \cdot (1-p)^{n'-k'} \right] \\
&= np \cdot [1 + (n-1)p]
\end{aligned} \tag{2.11}$$

Damit ergibt sich für die Varianz:

$$\sigma^2 = n\,p\,(1-p). \tag{2.12}$$

**Bemerkungen:** Folgende Eigenschaften der Binomial-Verteilung werden in Abb. 2.2 demonstriert:

1. Die Varianz hat für $p = 0.5$ ein Maximum:

$$\frac{d\sigma^2}{dp} = n\,[(1-p) + (-p)] = 0 \implies p = 0.5 \tag{2.13}$$

2. Die relative Breite wird mit wachsendem $n$ kleiner:

$$\frac{\sigma}{\langle k \rangle} = \frac{\sqrt{n\,p\,(1-p)}}{n\,p} = \sqrt{\frac{1-p}{n\,p}} \sim \frac{1}{\sqrt{n}} \tag{2.14}$$

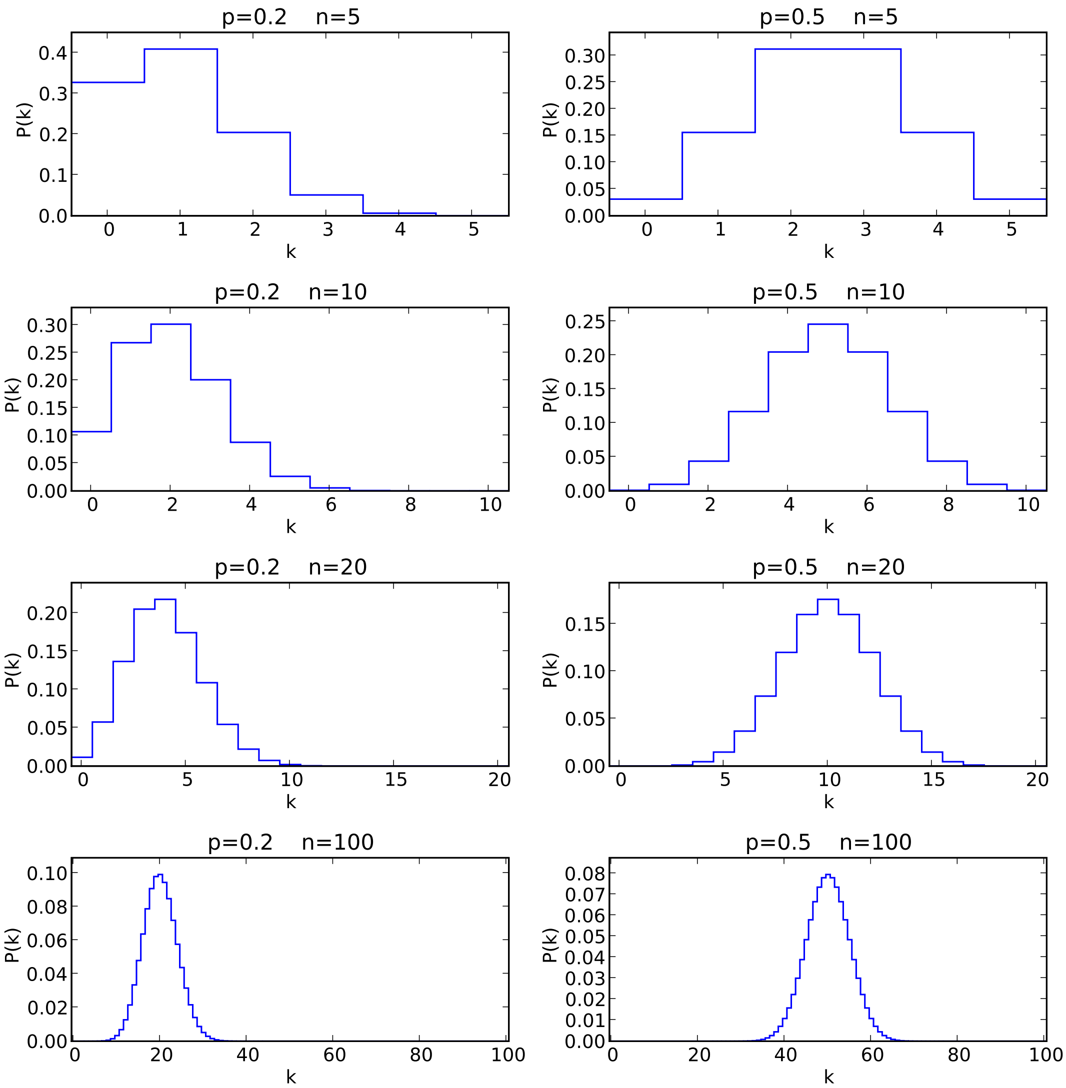



Abbildung 2.2: Beispiele von Binomial-Verteilungen mit verschiedenen Parametern $n$ und $p$ .

3. Für große $n$ und $n\,p$ ($p$ nicht zu klein) nähert sich die Binomial-Verteilung der Normalverteilung mit $\mu = np$ und $\sigma^2 = n\,p\,(1-p)$ an (das ergibt sich aus dem 'Zentralen Grenzwertsatz', siehe Abschnitt 2.6):

$$W_k^n \to W(k;\, n, p) = \frac{1}{\sqrt{2\pi\, np(1-p)}} \exp\left(-\frac{(k-np)^2}{2np(1-p)}\right) \tag{2.15}$$

## 2.2 Multinomial-Verteilung

Die Multinomial-Verteilung ist die natürliche Erweiterung der Definition der Binomial-Verteilung: Gegeben seien $l$ Klassen von Ereignissen $A_j$ ($j = 1, \ldots, l$) mit den Eigenschaften $j$ und den Wahrscheinlichkeiten $p_j$, die sich gegenseitig ausschliessen und erschöpfend sind:

$$E = \sum_{j=1}^{l} A_j; \qquad A_i \cap A_j = \varnothing \;\; \emptyset\; i \neq j. \tag{2.16}$$

Daraus folgt für die Summe der Wahrscheinlichkeiten aller Klassen:

$$\sum_{j=1}^{l} p_j = 1 \tag{2.17}$$

Die Wahrscheinlichkeit, bei $n$ Ereignissen gleichzeitig $k_1$ mit der Eigenschaft $A_1$, $k_2$ mit der Eigenschaft $A_2$ ... und $k_l$ mit der Eigenschaft $A_l$ usw. zu erhalten, ist

$$W^n_{k_1,k_2,\ldots,k_l} = n! \prod_{j=1}^{l} \frac{p_j^{k_j}}{k_j!} \tag{2.18}$$

Jedes der $n$ Ereignisse ist jeweils in einer der $l$ Klassen, so dass gilt:

$$\sum_{j=1}^{l} k_j = n. \tag{2.19}$$

Das bedeutet, dass die Faktoren in (2.18) nicht unabhängig voneinander sind. Der vollständige Beweis der Formel (2.18) kann durch Induktion von $l-1$ auf $l$ durchgeführt werden.

Für $l = 2$ erhält man die Binomial-Verteilung wieder ($k_1 = k$; $k_2 = n-k$):

$$W^n_{k_1,k_2} = n! \frac{p_1^{k_1}}{k_1!} \cdot \frac{p_2^{k_2}}{k_2!} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = W_k^n \tag{2.20}$$

Die Multinomial-Verteilung ist eine Verteilung mit mehreren Variablen (die $k_j$), die wir eigentlich erst im nächsten Kapitel besprechen. Im Vorgriff geben wir im Folgenden Parameter der Verteilung an, die zum Teil erst später (wie die Kovarianzmatrix) definiert werden.

**Normierung:** Unter Berücksichtigung der Bedingungen (2.17) und (2.19) ergibt sich für die Normierung:

$$\sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1}\dots\sum_{k_{l-1}=0}^{n-k_1-k_2-\dots k_{l-2}} W^n_{k_1,k_2,\dots,k_l} = 1 \qquad \text{mit } k_l = n - \sum_{j=1}^{l-1} k_j \text{ und } p_l = 1 - \sum_{j=1}^{l-1} p_j \tag{2.21}$$

**Mittelwert:** Der Mittelwert jeder einzelnen Variablen ist:

$$\langle k_j \rangle = np_j \qquad (j = 1, \dots, l) \tag{2.22}$$

**Varianz:** Die Varianzen der einzelnen Variablen ergeben sich entsprechend der Binomial-Verteilung:

$$\sigma_i^2 = np_i(1 - p_i) \tag{2.23}$$

Bei mehreren Variablen treten auch Kovarianzen auf, die Korrelationen beschreiben (siehe Kapitel 3):

$$\mathrm{cov}_{\mathrm{ij}} = -np_ip_j \tag{2.24}$$

Das Minuszeichen bedeutet eine negative Korrelation zwischen $k_i$, $k_j$ (eine Änderung einer Variablen bewirkt tendentiell eine Änderung der anderen Variablen in die entgegengesetzte Richtung).

**Beispiele:**

- Die Häufigkeit der Buchstaben in Texten, im allgemeinen $p_i \neq p_j$, wird zur Analyse von Texten und Sprachen bestimmt.
- In Experimenten der Teilchenphysik treten in der Regel 5 Arten geladener, stabiler Teilchen mit unterschiedlichen Häufigkeiten auf (Protonen, Pionen, Kaonen, Elektronen, Myonen). Die Analyse der Häufigkeitsverteilung benötigt man zur Identifikation der Teilchen (siehe späteres Kapitel zur Entscheidung über Hypothesen).

## 2.3 Poisson-Verteilung

Der Grenzfall einer Binomialverteilung mit einer sehr großen Zahl von möglichen Ereignissen, die aber jeweils eine sehr kleine Wahrscheinlichkeit haben, führt zu der Poisson-Verteilung:

$$\lim_{\substack{n\to\infty \\ p\to 0}} W_k^n = P_k^\lambda \qquad (n \cdot p = \lambda \text{ endlich}) \tag{2.25}$$

Bei dem Grenzübergang zu sehr großen Zahlen $n$ und sehr kleinen Wahrscheinlichkeiten $p$ soll der Erwartungswert von $k$,

$$\langle k \rangle = \lambda = n \cdot p, \tag{2.26}$$

endlich bleiben.

**Beispiele:**

- Radioaktiver Zerfall: Die Zahl $n$ der radioaktiven Kerne ist bei einer Probe meistens von der Größenordnung der Loschmidt-Zahl, also sehr groß. Die Wahrscheinlichkeit, daß einer dieser Kerne in einem festen Zeitintervall $\Delta t$ zerfällt, ist dagegen sehr klein, aber die mittlere Zerfallsrate $\lambda$ ist endlich.

- Die Anzahl der Sterne, die man in einem gegebenen Ausschnitt eines Teleskops bei einer bestimmten Auflösung beobachtet, hat einen bestimmten Mittelwert $\lambda$, der klein ist gegen die Gesamtzahl der Sterne. Bei einer Himmelsdurchmusterung erwartet man Fluktuationen entsprechend einer Poisson-Verteilung. Abweichungen, eventuell als Funktion der Ausschnittgröße, können auf kosmische Strukturen hinweisen.

- Die Anzahl der Gasatome in einem Volumen von der Größenordnung einiger Atomvolumina ist Poisson-verteilt.

- Die Zahl der jährlichen tödlichen Unfälle durch Pferdetritte in der Preussischen Armee ist Poisson-verteilt.

- Die Anzahl der Druckfehler auf einer Seite eines Buches ist Poisson-verteilt.

Die Poisson-Verteilung kann durch Ausführung des Grenzüberganges (2.25) aus der Binomialverteilung abgeleitet werden. Mit $\lambda = n \cdot p$ beziehungsweise $p = \lambda/n$ gilt:

$$\begin{aligned} W_k^n &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\ &= \frac{n!}{(n-k)!\,k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\to e^{-\lambda} \text{ für } n\to\infty} \underbrace{\frac{n(n-1)\ldots(n-k-1)}{n^k\,\left(1-\frac{\lambda}{n}\right)^k}}_{\to 1 \text{ für } n\to\infty} \end{aligned} \tag{2.27}$$

Damit ergibt sich für den Limes $n \to \infty$ die Poisson-Verteilung:

$$P_k^\lambda = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \tag{2.28}$$

Ausgehend von

$$P_0^\lambda = e^{-\lambda} \tag{2.29}$$

ist vor allem zum Programmieren folgende Rekursionsformel nützlich:

$$P_{k+1}^\lambda = P_k^\lambda \cdot \frac{\lambda}{k+1} \tag{2.30}$$

**Normierung:** Die Poisson-Verteilung (2.28) ist richtig normiert:

$$\sum_{k=0}^{\infty} P_k^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^\lambda} = e^{-\lambda} \cdot e^{\lambda} = 1 \tag{2.31}$$

**Mittelwert:** Nach Konstruktion ist der Erwartungswert von $k$ gleich $\lambda$:

$$\langle k \rangle = \lambda, \tag{2.32}$$

was sich durch explizite Berechnung bestätigen läßt:

$$\langle k \rangle = \sum_{k=0}^{\infty} k \, \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \cdot e^{-\lambda} = \lambda. \tag{2.33}$$

**Varianz:** Ausgehend von der Varianz für eine Binomial-Verteilung $\sigma^2 = n\,p\,(1-p)$ erhält man mit dem Grenzübergang $p \to 0$, wobei $\lambda = np$ endlich bleibt:

$$\sigma^2 = n\,p = \lambda. \tag{2.34}$$

Die Standardabweichung ist dann

$$\sigma = \sqrt{\lambda}. \tag{2.35}$$

Breite und Mittelwert der Verteilung sind also eng miteinander verknüpft.

Häufig entnimmt man als Stichprobe einer Poisson-Verteilung nur einen einzigen Wert, zum Beispiel die Zählrate $N$ von Kernzerfällen in einem Zeitintervall. Dann ist $N$ der beste Schätzwert für die mittlere Zerfallsrate $\lambda$ und als Fehler wird der Schätzwert für die Standardabweichung benutzt:

$$\hat{\sigma} = \sqrt{N}. \tag{2.36}$$

Allerdings muss man bei der Weiterverarbeitung von Daten vorsichtig sein, weil bei Fluktuationen von $N$ nach unten ein kleinerer Fehler folgt als bei Fluktuationen nach oben (siehe Diskussion bei 'Likelihood-Methode').

**Bemerkungen:** Folgende Eigenschaften sind charakteristisch für die Poisson-Verteilung (siehe Abb. 2.3):

1. Die Varianz ist gleich dem Mittelwert.
2. Für kleine Mittelwerte $\lambda$ (nahe 1) ergibt sich eine asymmetrische Verteilung.
3. Für wachsende $\lambda$ wird die Verteilung immer symmetrischer und nähert sich einer Gauss-Verteilung mit Mittelwert und Varianz $\lambda$ (das ergibt sich wieder aus dem 'Zentralen Grenzwertsatz', siehe Abschnitt 2.6):

$$P_k^{\lambda} \to P(k;\, \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{(k-\lambda)^2}{2\lambda}\right) \tag{2.37}$$

## 2.4 Gleichverteilung

Der einfachste, aber durchaus wichtige, Fall einer Wahrscheinlichkeitsverteilung einer kontinuierlichen Variablen ist die Gleichverteilung:

$$f(x) = c = const \tag{2.38}$$

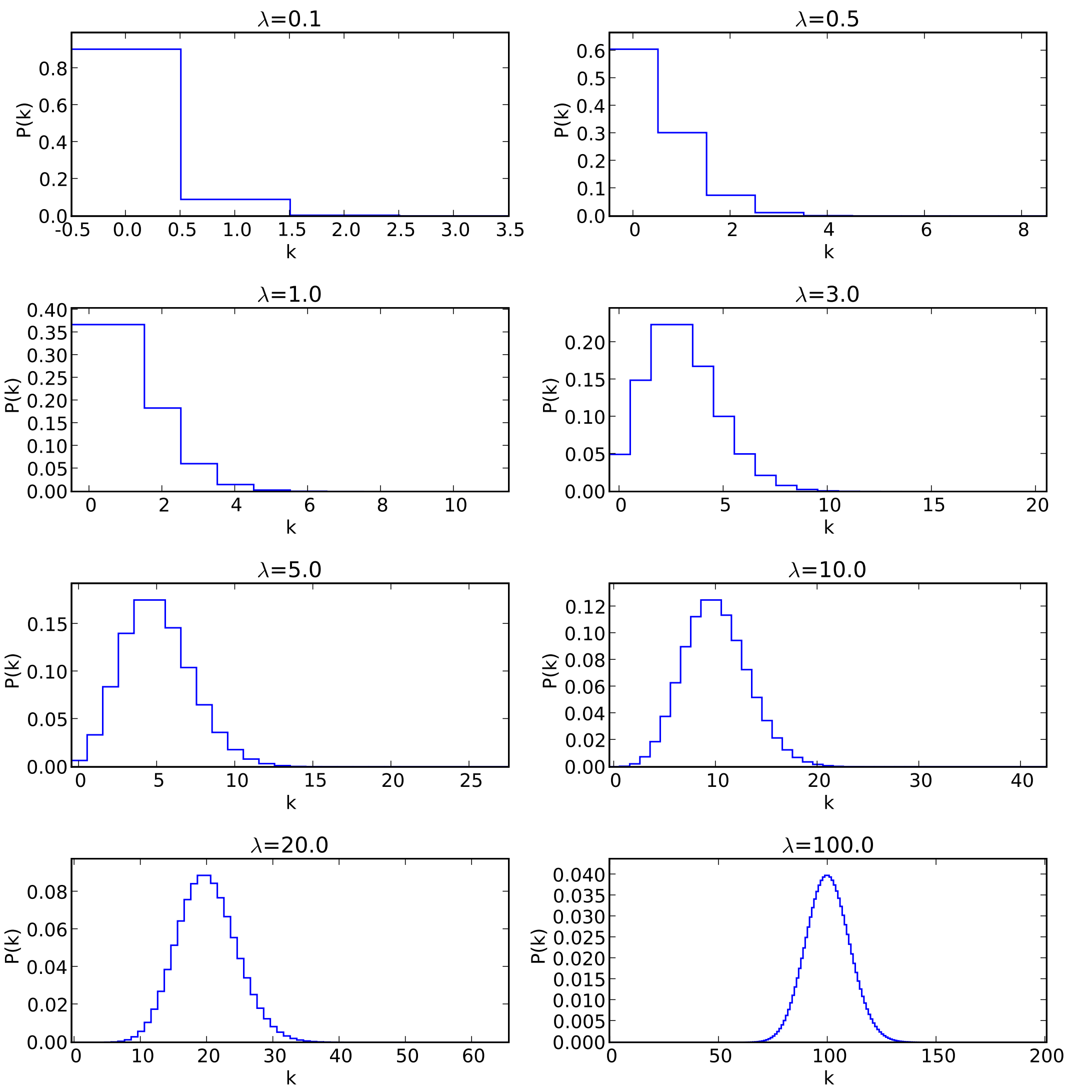
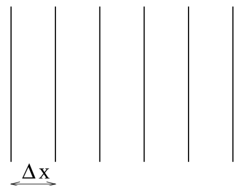



Abbildung 2.3: Beispiele von Poisson-Verteilungen mit verschiedenen Parametern $\lambda$
.

**Beispiele:**

- Der Winkel eines Uhrzeigers nimmt mit gleicher Wahrscheinlichkeit einen Wert zwischen 0° und 360° an.
- Viele Detektoren für Strahlung haben eine Streifenstruktur, die eine Koordinate innerhalb einer Streifenbreite festlegt:




  Bei homogener Einstrahlung ist die Koordinate des Auftreffens des Teilchens innerhalb eines Streifens gleichverteilt.

- Rundungsfehler sind gleichverteilt in dem Rundungsintervall.

**Normierung:**

$$1 = \int_{x_1}^{x_2} f(x)\,dx = \int_{x_1}^{x_2} c\,dx = c\,(x_2 - x_1) = c\,\Delta x \implies c = \frac{1}{\Delta x} \tag{2.39}$$

Zum Beispiel ergibt sich für eine Uhr:

$$f(\varphi) = \frac{1}{360^\circ} \tag{2.40}$$

**Mittelwert:**

$$\bar{x} = \langle x \rangle = \frac{1}{\Delta x}\int_{x_1}^{x_2} x\,dx = \frac{1}{2}\frac{x_2^2 - x_1^2}{x_2 - x_1} = \frac{x_1 + x_2}{2} \tag{2.41}$$

**Varianz:**

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{3}\frac{x_2^3 - x_1^3}{x_2 - x_1} - \left(\frac{1}{2}\frac{x_2^2 - x_1^2}{x_2 - x_1}\right)^2 = \frac{(\Delta x)^2}{12} \tag{2.42}$$

Die Standardabweichung ist dann

$$\sigma = \frac{\Delta x}{\sqrt{12}}. \tag{2.43}$$

Das heisst, die Standardabweichung ist um eine Faktor $\sqrt{12} \approx 3.5$ besser als das Raster einer Messung.

**Verteilungsfunktion:** Die Verteilungsfunktion steigt linear mit $x$ an:

$$F(x) = \frac{1}{\Delta x}\int_{x_1}^{x} d\xi = \frac{x - x_1}{\Delta x} \tag{2.44}$$

## 2.5 Normalverteilung

Die in der Statistik am häufigsten benutzte Verteilung ist die Gauss- oder Normalverteilung. Wir haben bereits gesehen, dass diese Verteilung aus den Binomial- und Poisson-Verteilungen im Grenzfall großer Zahlen ($n$ bzw. $\lambda$) folgt. Wir werden weiter unten den 'zentralen Grenzwertsatz' besprechen, der solche Grenzübergänge noch allgemeiner behandelt.

Eine Normalverteilung ergibt sich, wenn viele kleine Änderungen $\epsilon_i$ aufsummiert werden. Anschaulich kann man sich das zum Beispiel anhand des Galton-Brettes (Abb. 2.1) klar machen: Die Kugel entscheidet $n$-mal, ob Sie links oder rechts von einem Nagel fällt, entsprechend einem Versatz um $\epsilon_i = \pm\Delta\epsilon$. Die Verteilung der Auftrefforte unter dem Brett $x = \sum_{i=1}^{n} \epsilon_i$ nähert sich einer Normalverteilung im Grenzfall großer $n$.

Die Normalverteilung $N(\mu, \sigma)$ ist durch die beiden Parameter Mittelwert $\mu$ und Varianz $\sigma^2$ gegeben:

$$f(x) = f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.45}$$

**Normierung:** Die Normierung wird durch den Faktor $(\sqrt{2\pi}\sigma)^{-1}$ sichergestellt, was sich mit folgendem bestimmten Integral ergibt:

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}} \tag{2.46}$$

**Mittelwert:** Der Mittelwert ergibt sich aus:

$$\langle x \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \tag{2.47}$$

Zur Berechnung des Integrals setzt man $x = (x-\mu)+\mu$ und erhält damit die beiden Integrale:

$$\langle x \rangle = \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}_{=0} + \mu \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}_{=1} = \mu \tag{2.48}$$

Das linke Integral verschwindet, weil sich die Beiträge für $x-\mu < 0$ und die für $x-\mu > 0$ gerade aufheben.

**Varianz:** Die Varianz ergibt sich mit Hilfe eines weiteren bestimmten Integrals:

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}\, dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}} \tag{2.49}$$

Damit erhält man:

$$\langle (x-\mu)^2 \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \sigma^2. \tag{2.50}$$

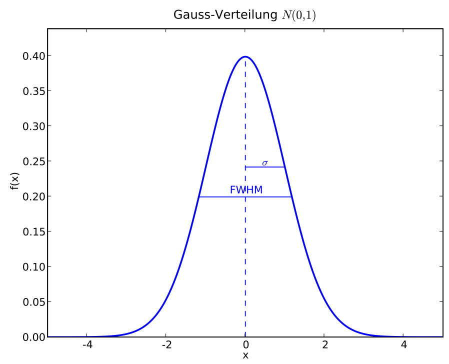



Abbildung 2.4: Standardisierte Normalverteilung $N(0,1)$.

**Standardisierte Normalverteilung:** Durch die Transformation

$$x \rightarrow \frac{x-\mu}{\sigma} \tag{2.51}$$

erhält man eine Normalverteilung $N(0,1)$ mit Mittelwert 0 und Varianz 1:

$$f(x) = f(x;0,1) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} \tag{2.52}$$

Eine standardisierte Normalverteilung ist in Abb. 2.4 gezeigt. Neben dem Mittelwert und der Standardabweichung $\sigma$ ist auch die **volle Breite auf halber Höhe** des Maximums (FWHM = full width at half maximum) gezeigt. Diese Größe ist relativ einfach (mit Lineal und Bleistift) aus einer gemessenen Verteilung zu bestimmen. Für eine Gauss-Verteilung gibt es eine feste Beziehung zwischen FWHM und $\sigma$:

$$\frac{f(0)}{2} = \frac{1}{\sqrt{2\pi}\sigma}\,\exp\left(-\frac{(FWHM/2)^2}{2\sigma^2}\right) \implies FWHM = 2\,\sigma\,\sqrt{2\,\ln 2} \approx 2.355\cdot\sigma \tag{2.53}$$

**Verteilungsfunktion:** Die Verteilungsfunktion der Normalverteilung ist nicht analytisch zu berechnen. Zahlenwerte findet man in Tabellen, in der Regel für die standardisierte Normalverteilung $N(0,1)$ als Funktion von $x$. Den Übergang zu Verteilungen $N(\mu,\sigma)$ findet man durch Skalieren von $x$ mit $\sigma$ und Verschieben um $\mu$:

$$x = \frac{x'-\mu}{\sigma} \tag{2.54}$$

Statt der Verteilungsfunktion findet man auch die sogenannte Fehlerfunktion ('error function' oder "Gauss'sches Fehlerintegral") erf$(x)$ tabelliert:

$$\begin{array}{rcl} \mathrm{erf}(x) & = & \frac{2}{\sqrt{\pi}}\int_0^x e^{-\xi^2}d\xi \\ \implies F(x) & = & \frac{1}{2}\left[1+\mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right] \end{array} \tag{2.55}$$

Tabelle 2.1: Wahrscheinlichkeiten innerhalb von $\pm n\sigma$-Bereichen einer Normalverteilung.

a)

| $n$ | $p(\pm n\sigma)$ |
|---|---|
| 1 | 0.6827 |
| 2 | 0.9545 |
| 3 | 0.9973 |
| 4 | $1 - 6.3 \cdot 10^{-5}$ |

b)

| $p(\pm n\sigma)$ | $n$ |
|---|---|
| 0.900 | 1.645 |
| 0.950 | 1.960 |
| 0.990 | 2.576 |
| 0.999 | 3.290 |

### 2.5.1 Vertrauensintervalle:

Die Verteilungsfunktion benötigt man häufig zur Bestimmung der Wahrscheinlichkeit, dass ein Ereignis innerhalb bestimmter Grenzen für $x$ liegt. Für die Beurteilung von Messergebnissen mit normalverteilten Fehlern benutzt man zum Beispiel die Wahrscheinlichkeit, in einem zentralen 'Vertrauensintervall' von $\pm n\,\sigma$ um den Mittelwert zu liegen (Abb. 2.5a, Tab. 2.1a):

$$p(\pm n\sigma) = F(\mu + n\sigma) - F(\mu - n\sigma) = \operatorname{erf}\left(\frac{n\sigma}{\sqrt{2}\sigma}\right), \tag{2.56}$$

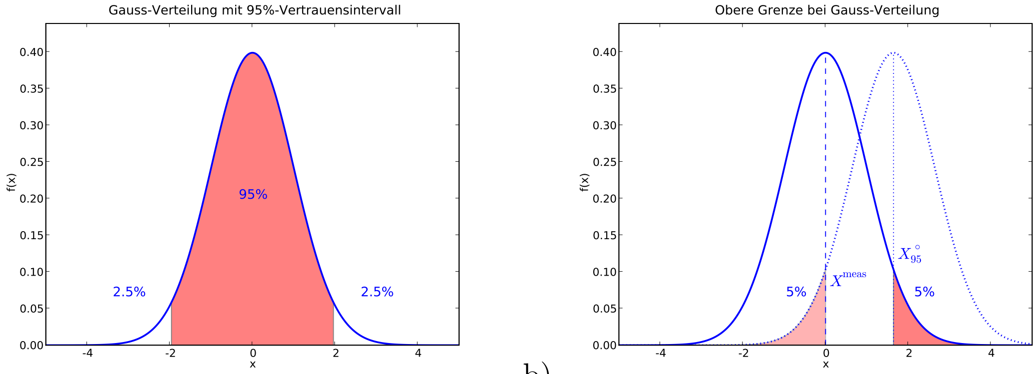
Häufig gibt man auch die Wahrscheinlichkeit, das 'Vertrauensniveau' (confidence level, c. l.), vor und fragt nach den entsprechenden Grenzen (Tab. 2.1b).

Innerhalb von 2 Standardabweichungen, $\pm 1\sigma$, um den Mittelwert liegen also 68.27 % aller Ereignisse. Häufig werden Fehler so definiert, dass 68.27 % innerhalb der Fehlergrenzen liegen, auch wenn die zugrundeliegende Verteilung nicht die Normalverteilung ist ('Standardfehler'). Bei asymmetrischen Verteilungen können die Fehler auch asymmetrisch um den Mittelwert definiert werden, zum Beispiel so, dass jeweils 16 % oberhalb und unterhalb des Fehlerbereichs liegen.

Welches Vertrauensniveau man für eine Aussage verlangt, hängt von der Problemstellung ab. Während man standardmäßig bei Messergebnissen das $1\sigma$-Niveau angibt, verlangt man zur Festlegung von Toleranzgrenzen für Risiken, die das Leben von Menschen gefährden, viel höhere Vertrauensniveaus. Ob man nun 90 % oder 99, 9 % oder 99, 9999 % verlangt, hängt unter anderem von der 'a priori' Wahrscheinlichkeit für das Risiko, also zum Beispiel die Größe der gefährdeten Gruppe, ab ('Bayesischer Ansatz'). Wenn ein Fahrstuhl zum Beispiel im Mittel 1 Million mal während seiner Lebensdauer benutzt wird, sollte die Wahrscheinlichkeit für das Reißen des Seils kleiner als $10^{-6}$ sein.

**Ausschließungsgrenzen:** Häufig möchte man ein bestimmtes Vertrauensniveau angeben, dass bei einem gegebenen Messwert $x^{mess}$ der wahre Wert $x^{wahr}$ oberhalb oder unterhalb einer Grenze liegt.

**Beispiel:** Um in der Elementarteilchenphysik die Entdeckung eines neuen Teilchens zu etablieren, wird ein Vertrauensniveau von mindestens 5 Standardabweichungen verlangt, weil jeder Physiker, der mal 1000 Histogramme mit je etwa 100 Bins angeschaut hat, eine gute Chance hat, wenigstens einen




a) b)

Abbildung 2.5: a) Fläche unter einer Gauss-Kurve, die einem Vertrauensintervall von 95% entspricht. b) Bestimmung einer oberen Grenze bei normalverteilten Fehlern, hier mit einem Vertrauensniveau von 95 %. Links ist die Verteilung um den Messwert, rechts die Verteilung um den Wert der oberen Grenze. Die schattierten Bereiche entsprechen jeweils 5 % Wahrscheinlichkeit. Siehe weitere Erläuterungen im Text.

> $4\sigma$-Effekt zu beobachten. Ist dagegen ein Teilchen vorhergesagt und man findet oberhalb eines Untergrundes kein Signal, gibt man in der Regel untere Grenzen für die Häufigkeit der Erzeugung des Teilchens mit 90% oder 95% Vertrauensniveau an.

Will man zum Beispiel mit 95 % Vertrauensniveau (95 % c. l.) bei gegebenem Messwert $x^{mess}$ eine obere Grenze für $x^{wahr}$ angeben, stellt man die Frage: Was ist der Wert $x^o_{95}$, für den die Wahrscheinlichkeit, einen Messwert $x^{mess}$ oder kleiner zu erhalten, 5 % beträgt. Die Grenze $x^o_{95}$ wird also als Mittelwert einer Gauss-Verteilung (mit bekannter, gemessener oder geschätzter Standardabweichung) gesucht, deren Integral von $-\infty$ bis $x^{mess}$ 5 % beträgt (Abb. 2.5b). Wegen der Symmetrie der Gauss-Verteilung kann man aber auch von einer entsprechenden Gaussverteilung um den gemessenen Wert ausgehen und $x^o_{95}$ als denjenigen Wert bestimmen, für den das Integral über $x > x^o_{95}$ die geforderten 5 % bzw. das Komplement 95 % ergibt:

$$F(x^o_{95}) = 0.95 \tag{2.57}$$

Entsprechend ergibt sich für eine untere Grenze mit 95 % Vertrauensniveau:

$$F(x^u_{95}) = 0.05 \tag{2.58}$$

Man schreibt dann zum Beispiel:

$$x < x^u_{95}, \quad 95\% \text{ c. l.} \tag{2.59}$$

Bei angenommenen gauss-verteilten Fehlern sind also die Grenzen einfach aus der Verteilungsfunktion zu bestimmen. Im allgemeinen Fall muss man aber auf die oben angegebene Definition zurückgreifen. Zum Beispiel kommt es häufig vor, dass man auf der Suche nach einem Ereignis nichts findet, also ein Nullergebnis hat. Wenn es sich um ein Zählratenexperiment handelt, ergibt sich bekanntlich für eine Poisson-Verteilung eine endliche Wahrscheinlichkeit auch bei einem nicht-verschwindenden

Tabelle 2.2: Untere und obere Grenze der Vertrauensintervalle von 90 % und 95 % für den Erwartungswert einer Posison-Verteilung gegeben, dass $n$ Ereignisse (frei von Untergrund) gemessen wurden.

| | $\epsilon = 90\,\%$ | | $\epsilon = 90\,\%$ | |
|---|---|---|---|---|
| n | $\lambda^u$ | $\lambda^o$ | $\lambda^u$ | $\lambda^o$ |
| 0 | - | 2.30 | - | 3.00 |
| 1 | 0.105 | 3.89 | 0.051 | 4.74 |
| 2 | 0.532 | 5.32 | 0.355 | 6.30 |
| 3 | 1.10 | 6.68 | 0.818 | 7.75 |
| 4 | 1.74 | 7.99 | 1.37 | 9.15 |
| 5 | 2.43 | 9.27 | 1.97 | 10.51 |

Mittelwert ($\lambda \neq 0$) ein Nullergebnis zu erhalten. Man kann dann nur eine obere Grenze für den wahren Wert von $\lambda$ geben. Entsprechend der oben angegebene Definition fragt man für ein gefordertes Vertrauensniveau $\epsilon$: für welchen Mittelwert $\lambda^o_\epsilon$ ist die Wahrscheinlichkeit die Zählrate 0 (oder kleiner) zu erhalten gerade $1 - \epsilon$:

$$p(n, \lambda) = p(0, \lambda^o_\epsilon) = \frac{(\lambda^o_\epsilon)^0}{0!} e^{-\lambda^o_\epsilon} = e^{-\lambda^o_\epsilon} \stackrel{!}{=} 1 - \epsilon \tag{2.60}$$

$$\Longrightarrow \quad \lambda^o_\epsilon = -\ln(1 - \epsilon) \tag{2.61}$$

Die Grenzen für 90 und 95 % Vertrauensniveau sind bei 0 beobachteten Ereignissen:

$$\begin{aligned} \lambda^o_{90} &= 2.30 \\ \lambda^o_{95} &= 3.00 \end{aligned} \tag{2.62}$$

Für eine beobachtete Anzahl $n > 0$ ergeben sich obere und untere Grenzen $\lambda^o$ und $\lambda^u$, die in Tab. 2.2 für $\epsilon = 90\,\%$ und $95\,\%$ zusammengestellt sind.

## 2.6 Zentraler Grenzwertsatz

Die Gauss-Verteilung hat unter allen Verteilungen eine besondere Bedeutung, weil sie für viele Verteilungen ein Grenzfall für große Zahlen darstellt. Wir hatten das bereits für die Binomial- und die Poisson-Verteilung gesehen, die beide im Grenzfall großer Mittelwerte in die Gauss-Verteilung übergehen.

Die Gauss-Verteilung kann interpretiert werden als Verteilung von Abweichungen um einen Mittelwert, die sich als Überlagerung vieler kleiner Störungen ergeben. Tatsächlich findet man, dass die Summe von $n$ beliebigen Zufallsvariablen für große $n$ einer Gauss-Verteilung zustrebt. In Übungsaufgabe 8 wurde das für die Summe von gleichverteilten Zufallszahlen gezeigt, wobei sich zeigte, dass die Verteilung der Summe von 12 solchen Zufallszahlen bereits sehr gut eine Gauss-Verteilung approximiert (Abb. 2.6).

Diese Eigenschaft der Gauss-Verteilung wird mathematisch im Zentralen Grenzwertsatz formuliert: Gegeben seinen $n$ unabhängige Variablen $x_i,\ i = 1, \ldots, n$, die

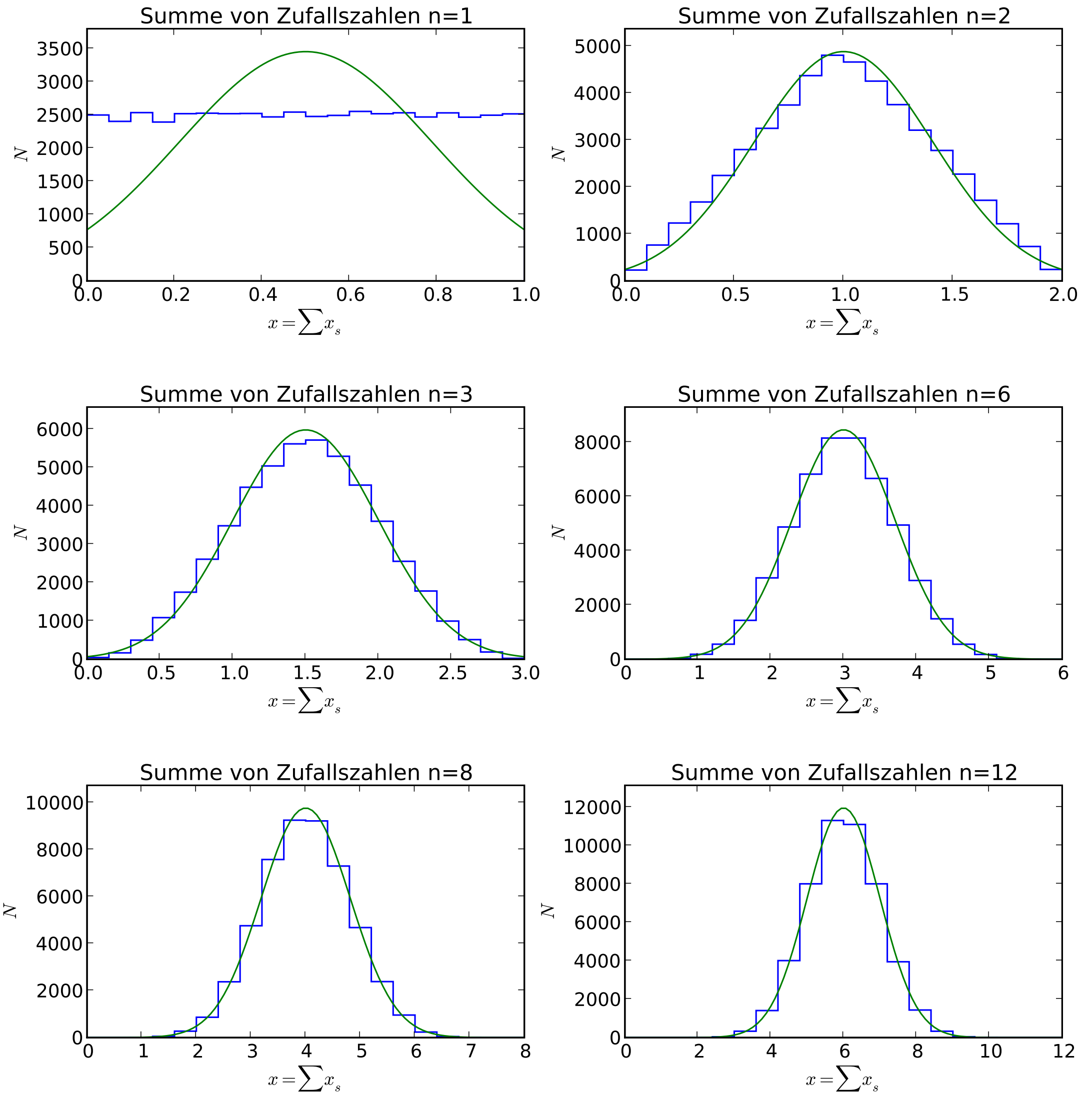



Abbildung 2.6: Beispiele von Verteilungen der Summen von $n$ zwischen 0 und 1 gleichverteilten Zufallszahlen. Die Verteilungen werden mit Gauss-Verteilungen mit Mittelwert $\mu = n/2$ und Varianz $\sigma^2 = n/12$ verglichen.

jeweils einer Verteilung mit Mittelwert $\mu_i$ und Varianz $\sigma_i$ entnommen sind (die Verteilungen sind ansonsten beliebig). Dann hat die Verteilung der Summe

$$X = \sum_{i=1}^{n} x_i \tag{2.63}$$

folgende Eigenschaften:

(i) Erwartungswert:

$$\langle X \rangle = \sum_{i=1}^{n} \mu_i; \tag{2.64}$$

(ii) Varianz:

$$\sigma_X^2 = \sum_{i=1}^{n} \sigma_i^2; \tag{2.65}$$

(iii) die Verteilung nähert sich einer Gauss-Verteilung für

$$n \to \infty. \tag{2.66}$$

Zum Beweis von (2.64) und (2.65) benutzt man die Linearität der Erwartungswertbildung: der Erwartungswert einer Summe unabhängiger Zufallszahlen ist die Summe der Erwartungswerte. Für den Erwartungswert von $X$ ergibt sich:

$$\langle X \rangle = \left\langle \sum_i x_i \right\rangle = \sum_i \langle x_i \rangle = \sum_i \mu_i. \tag{2.67}$$

Entsprechend ergibt sich für die Varianz:

$$\begin{aligned} \sigma_X^2 &= \langle (X - \langle X \rangle)^2 \rangle = \left\langle \left( \sum_i x_i - \sum_i \mu_i \right)^2 \right\rangle = \left\langle \left( \sum_i (x_i - \mu_i) \right)^2 \right\rangle \\ &= \sum_i \left\langle (x_i - \mu_i)^2 \right\rangle + \sum_i \sum_{j \neq i} \underbrace{\langle (x_i - \mu_i)(x_j - \mu_j) \rangle}_{=0,\ \text{wenn i, j unabhängig}} = \sum_i \sigma_i^2 \end{aligned} \tag{2.68}$$

Der Beweis der wichtigen Aussage (2.66) ist schwieriger und kann in Statistikbüchern nachgelesen werden, zum Beispiel [1, 2]. Abbildung 2.6 zeigt die Summe gleichverteilter Variablen, die sich der Gauss-Verteilung mit wachsender Anzahl Variabler annähert.