

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 der Übungsaufgabe 1: 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 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).

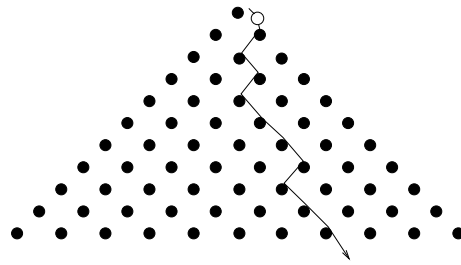


Abbildung 2.1: Galton-Brett.

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:

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

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 Beispiel 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 $3p^2(1-p)$). Da jede dieser Wahrscheinlichkeiten gleich ist, muss man also nur die Anzahl der möglichen Kombinationen bestimmen.

Um diese Anzahl der Kombinationen zu bestimmen, beginnt man damit, k unterscheidbare Ereignisse A_1, \dots, A_k auf n Stellen zu verteilen. Man hat für

$$\begin{array}{lll}
 A_1 & n & \text{Möglichkeiten,} \\
 A_2 & n - 1 & \text{Möglichkeiten} \\
 & \dots & \\
 A_k & n - (k - 1) & \text{Möglichkeiten.}
 \end{array}$$

Das sind insgesamt

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

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:

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

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

$$\begin{array}{ccccccc} & & & & & & n \\ & & & & & & 0 \\ & & & & & & 1 \\ & & & & & & 1 & 1 \\ & & & & & & 2 & 2 & 1 \\ & & & & & & 3 & 3 & 3 & 1 \\ & & & & & & 4 & 6 & 4 & 1 \\ & & & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & & k \rightarrow \end{array}$$

Damit ergibt sich die Binomial-Verteilung:

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

Es ist einfach zu sehen, dass die Normierung

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

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

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

Mittelwert:

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

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

$$\implies \langle k \rangle = np \quad (2.9)$$

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

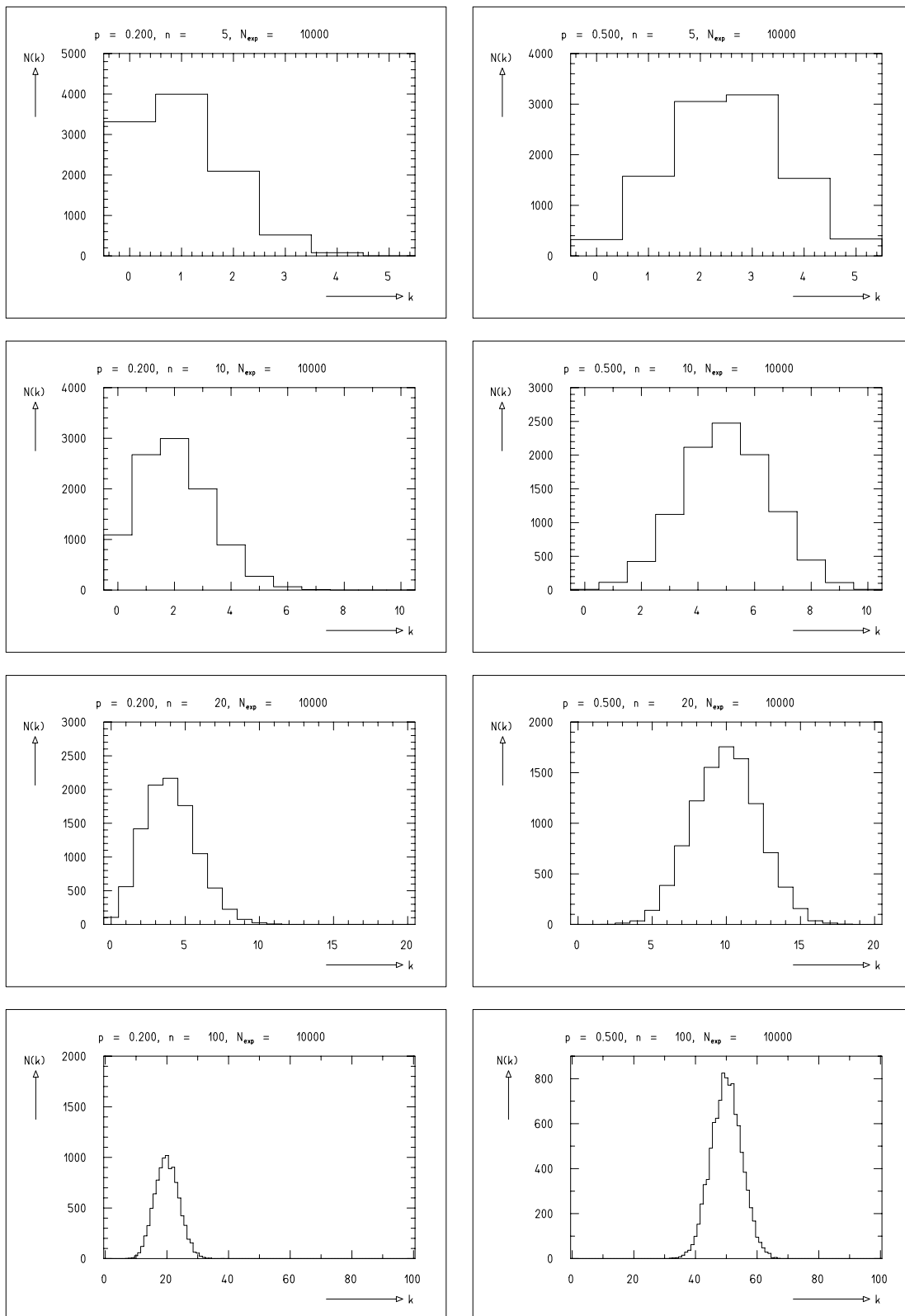


Abbildung 2.2: Beispiele von Binomial-Verteilungen mit verschiedenen Parametern n und p (erstellt mit dem Programm e3ds [1]).

Varianz: Die Varianz ist die mittlere quadratische Abweichung vom Mittelwert, die sich nach (1.39) zerlegen läßt:

$$\sigma^2 = \langle (k - \langle k \rangle)^2 \rangle = \langle k^2 \rangle - \langle k \rangle^2 \quad (2.10)$$

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

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

Damit ergibt sich für die Varianz:

$$\sigma^2 = np(1-p). \quad (2.12)$$

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

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

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

2. Die relative Breite wird mit wachsendem n kleiner:

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

3. Für große n und np (p nicht zu klein) nähert sich die Binomial-Verteilung der Normalverteilung mit $\mu = np$ und $\sigma^2 = np(1-p)$ an:

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

2.2 Multinomial-Verteilung

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

$$E = \sum_{j=1}^l A_j; \quad A_i \cap A_j = \emptyset \quad \forall i \neq j. \quad (2.16)$$

Daraus folgt für die Summe der Wahrscheinlichkeiten aller Klassen:

$$\sum_{j=1}^l p_j = 1 \quad (2.17)$$

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

$$W_{k_1, k_2, \dots, k_l}^n = n! \prod_{j=1}^l \frac{p_j^{k_j}}{k_j!} \quad (2.18)$$

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

$$W_{k_1, k_2}^n = n! \frac{p_1^{k_1}}{k_1!} \cdot \frac{p_2^{k_2}}{k_2!} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = W_k^n \quad (2.19)$$

Der vollständige Beweis der Formel (2.18) kann durch Induktion von $l - 1$ auf l durchgeführt werden.

Beispiele:

- Die Häufigkeit der Buchstaben in Texten, im allgemeinen $p_i \neq p_j$, wird zur Analyse von Texten, 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:

$$\lim_{\substack{n \rightarrow \infty \\ p \rightarrow 0}} W_k^n = P_k^\lambda \quad (n \cdot p = \lambda \text{ endlich}) \quad (2.20)$$

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

$$\langle k \rangle = \lambda = n \cdot p, \quad (2.21)$$

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 Δt zerfällt, ist dagegen sehr klein, aber die mittlere Zerfallsrate λ ist endlich.
- Die Anzahl der Sterne, die man in einem gegebenen Ausschnitt eines Teleskops bei einer bestimmten Auflösung beobachtet, hat einen bestimmten Mittelwert λ , 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 (2.20) aus der Binomialverteilung abgeleitet werden. Mit $\lambda = n \cdot p$ beziehungsweise $p = \lambda/n$ gilt:

$$\begin{aligned}
 W_k^n &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\
 &= \frac{n!}{(n-k)!k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
 &= \frac{\lambda^k}{k!} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\rightarrow e^{-\lambda} \text{ für } n \rightarrow \infty} \underbrace{\frac{n(n-1)\dots(n-k-1)}{n^k \left(1 - \frac{\lambda}{n}\right)^k}}_{\rightarrow 1 \text{ für } n \rightarrow \infty}
 \end{aligned} \tag{2.22}$$

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

$$P_k^\lambda = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \tag{2.23}$$

Ausgehend von

$$P_0^\lambda = e^{-\lambda} \tag{2.24}$$

ist vor allem zum Programmieren folgende Rekursionsformel nützlich:

$$P_{k+1}^\lambda = P_k^\lambda \cdot \frac{\lambda}{k+1} \tag{2.25}$$

Normierung: Die Poisson-Verteilung (2.23) ist richtig normiert:

$$\sum_{k=0}^{\infty} P_k^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^\lambda} = e^{-\lambda} \cdot e^\lambda = 1 \tag{2.26}$$

Mittelwert: Nach Konstruktion ist der Erwartungswert von k gleich λ :

$$\langle k \rangle = \lambda, \quad (2.27)$$

was sich durch explizite Berechnung bestätigen läßt:

$$\langle k \rangle = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \cdot e^{-\lambda} = \lambda. \quad (2.28)$$

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

$$\sigma^2 = np = \lambda. \quad (2.29)$$

Die Standardabweichung ist dann

$$\sigma = \sqrt{\lambda}. \quad (2.30)$$

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 λ und als Fehler wird der Schätzwert für die Standardabweichung benutzt:

$$\hat{\sigma} = \sqrt{N}. \quad (2.31)$$

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 λ (nahe 1) ergibt sich eine asymmetrische Verteilung.
3. Für wachsende λ wird die Verteilung immer symmetrischer und nähert sich einer Gauss-Verteilung mit Mittelwert und Varianz λ (Laplace-Verteilung):

$$P_k^\lambda \rightarrow P(k; \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{(k-\lambda)^2}{2\lambda}\right) \quad (2.32)$$

2.4 Gleichverteilung

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

$$f(x) = c = \text{const} \quad (2.33)$$

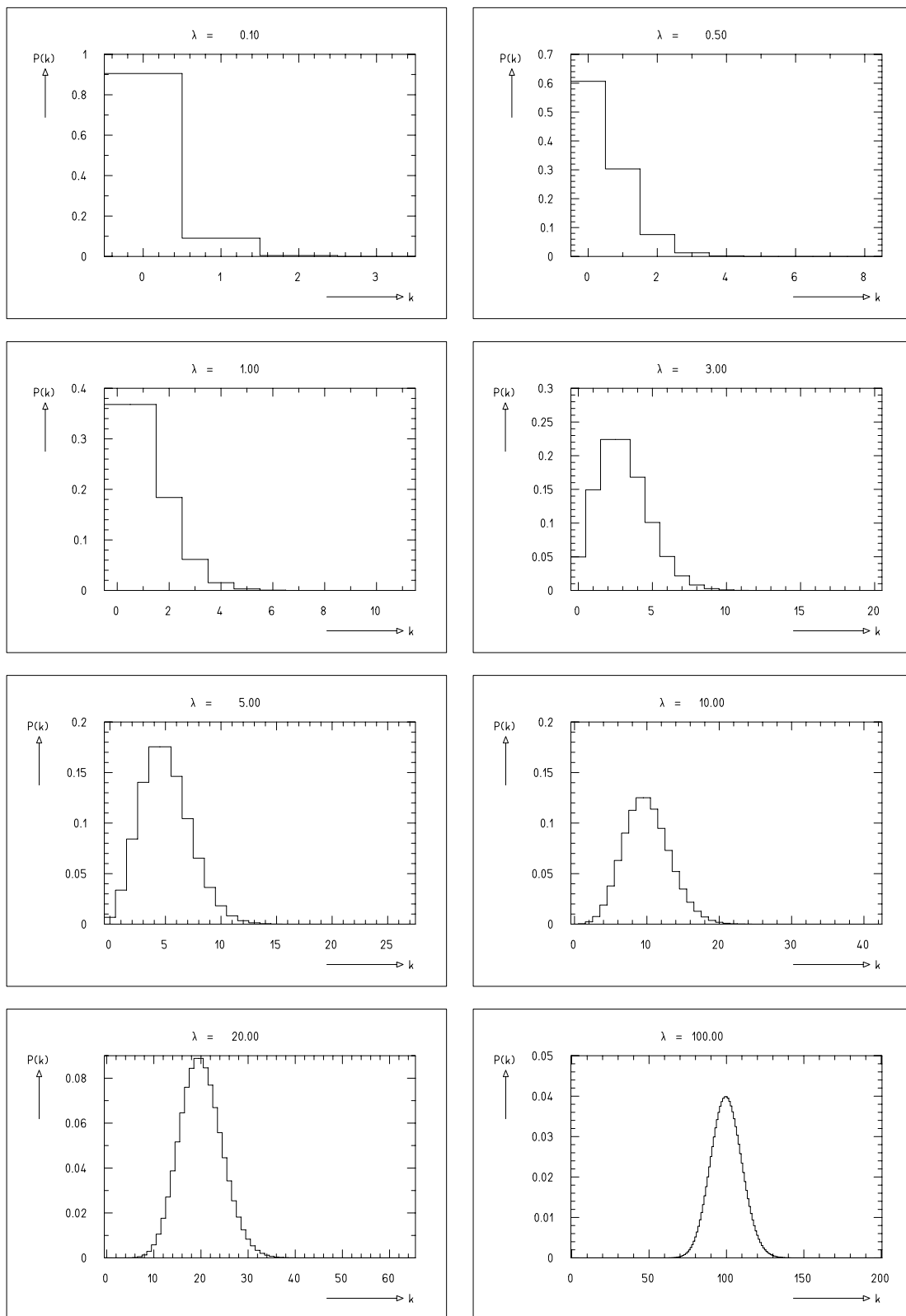
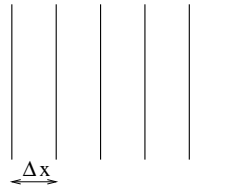


Abbildung 2.3: Beispiele von Poisson-Verteilungen mit verschiedenen Parametern λ (erstellt mit dem Programm s1sd [1]).

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:

$$1 = \int_{x_1}^{x_2} f(x) dx = \int_{x_1}^{x_2} c dx = c(x_2 - x_1) = c \Delta x \implies c = \frac{1}{\Delta x} \quad (2.34)$$

Zum Beispiel ergibt sich für eine Uhr:

$$f(\varphi) = \frac{1}{360^\circ} \quad (2.35)$$

Mittelwert:

$$\bar{x} = \langle x \rangle = \frac{1}{\Delta x} \int_{x_1}^{x_2} x dx = \frac{1}{2} \frac{x_2^2 - x_1^2}{x_2 - x_1} = \frac{x_1 + x_2}{2} \quad (2.36)$$

Varianz:

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{3} \frac{x_2^3 - x_1^3}{x_2 - x_1} - \left(\frac{x_1 + x_2}{2} \right)^2 = \frac{(\Delta x)^2}{12} \quad (2.37)$$

Die Standardabweichung ist dann

$$\sigma = \frac{\Delta x}{\sqrt{12}}. \quad (2.38)$$

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

Verteilungsfunktion: Die Verteilungsfunktion steigt linear mit x an:

$$F(x) = \frac{1}{\Delta x} \int_{x_1}^x d\xi = \frac{x - x_1}{\Delta x} \quad (2.39)$$