

## Kapitel 2

# Zufallsvariable und Verteilungen

Bei manchen Zufallsvorgängen interessiert man sich weniger für das konkrete Ergebnis  $\omega \in \Omega$  als für eine reelle Zahl, die von  $\omega$  abhängt. So wird sich ein Roulettespieler, der auf *Colonne* setzt, nicht so sehr für die ausgespielte Zahl  $\omega$  interessieren, sondern eher für den von  $\omega$  abhängenden Nettogewinn aus dem Spiel. Ein Aktienbesitzer wird sich weniger für das sehr komplexe Ergebnis  $\omega$  des Zufallsvorganges „Entwicklung an der Börse“ interessieren als für den Kurs seiner Aktie an einem Stichtag. Bei der Untersuchung von Haushalten interessiert sich ein Marktforscher meist nicht für alle Spezifika eines beobachteten Haushalts, sondern nur für bestimmte Merkmale, wie z.B. das verfügbare monatliche Haushaltseinkommen oder die monatlichen Ausgaben für Kleidung.

In der deskriptiven Statistik ist ein **Merkmal** eine prinzipiell beobachtbare und durch eine Zahl beschreibbare Eigenschaft von statistischen Einheiten, beispielsweise Individuen, Haushalten oder Unternehmen. Die Menge der statistischen Einheiten, die Träger eines Merkmals sind und über die etwas ausgesagt werden soll, wird dort **Grundgesamtheit** genannt. In der Wahrscheinlichkeitsrechnung und schließenden Statistik entspricht die **Ergebnismenge** der Grundgesamtheit und die **Zufallsvariable** dem Merkmal.

Im folgenden Abschnitt werden die wichtigsten mit Zufallsvariablen zusammenhängenden Begriffe erläutert und an einfachen Beispielen illustriert.

## 2.1 Grundbegriffe

Wir setzen voraus, dass ein Zufallsvorgang mit Ergebnismenge  $\Omega$  gegeben ist. Die Menge der interessierenden Ereignisse sei durch die Ereignisalgebra  $\mathcal{A}$  gegeben, und für alle Ereignisse  $A \in \mathcal{A}$  sei eine Wahrscheinlichkeit  $P(A)$  definiert.

Unter einer **Zufallsvariablen** versteht man formal eine Funktion

$$\boxed{\begin{array}{l} X : \Omega \longrightarrow \mathbb{R}, \\ \omega \longmapsto X(\omega). \end{array}}$$

Die Zufallsvariable  $X$  ordnet also jedem Ergebnis  $\omega$  des Zufallsexperimentes eine reelle Zahl  $X(\omega) = x$  zu. Die Zahl  $x$  wird **Wert** oder **Realisierung** der Zufallsvariablen  $X$  genannt. Hier ist sorgfältig zwischen den Symbolen  $X$  und  $x$  zu unterscheiden.  $x$  stellt eine reelle Zahl dar,  $X$  dagegen eine Funktion. Zufallsvariable werden allgemein mit Großbuchstaben wie  $X, Y, Z$  oder  $X_1, X_2, \dots$  bezeichnet, ihre Realisationen dann mit den entsprechenden Kleinbuchstaben  $x, y, z$  bzw.  $x_1, x_2, \dots$ .

*Beispiel 2.1 (vgl. Beispiel 1.4): Eine Münze mit den Seiten  $K$  (= Kopf) und  $Z$  (= Zahl) wird dreimal geworfen. Die Ergebnismenge besteht aus acht möglichen Ergebnissen,*

$$\Omega = \{(K, K, K), (K, K, Z), \dots, (Z, Z, Z)\}.$$

Bezeichne  $X$  in Abhängigkeit von  $\omega \in \Omega$  die „Anzahl Kopf“,

$$X(\omega) = \text{Anzahl } K \text{ in } \omega.$$

Offenbar wird mehreren Ergebnissen durch  $X$  die gleiche Zahl zugeordnet; z.B. ist

$$X((Z, K, Z)) = X((K, Z, Z)) = X((Z, Z, K)) = 1.$$

*Beispiel 2.2 (vgl. Beispiel 1.3): Beim Roulette ist die Menge der Ergebnisse  $\Omega = \{0, 1, \dots, 36\}$ . Ein Spieler setzt einen Euro auf die erste Column  $C_1$ ,*

$$C_1 = \{1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34\};$$

*vgl. Abbildung 1.2. Wenn die Roulettekugel auf eine der Zahlen in  $C_1$  fällt, erhält der Spieler das Dreifache seines Einsatzes ausgezahlt. Ihn interessiert in der Regel nicht die ausgespielte Zahl, sondern sein Nettogewinn, den wir mit  $Y$  bezeichnen.  $Y$  ist eine Funktion von  $\omega$ , also eine Zufallsvariable. Es gilt*

$$Y(\omega) = \begin{cases} -1, & \text{falls } \omega \notin C_1, \\ 2, & \text{falls } \omega \in C_1. \end{cases}$$

Bei manchen Zufallsvorgängen kommt es nur darauf an, ob ein Ereignis  $A$  eintritt oder nicht. Dies kann man durch eine so genannte **Indikatorvariable**  $X = \mathbf{1}_A$  zum Ausdruck bringen. Die Zufallsvariable, die lediglich das Eintreten von  $A$  anzeigt, wird durch

$$X(\omega) = \mathbf{1}_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \in \overline{A}, \end{cases}$$

definiert. Wenn  $A$  eintritt, nimmt sie den Wert  $x = 1$  an; wenn  $A$  nicht eintritt, den Wert  $x = 0$ . Komplement, Durchschnitt und Vereinigung von Ereignissen entsprechen einfachen Rechenoperationen der Indikatorvariablen. Offenbar gilt (jeweils für alle  $\omega \in \Omega$ )

$$\begin{aligned} \mathbf{1}_\Omega(\omega) &= 1, \\ \mathbf{1}_\emptyset(\omega) &= 0, \\ \mathbf{1}_{A \cap B}(\omega) &= \mathbf{1}_A(\omega) \wedge \mathbf{1}_B(\omega) = \mathbf{1}_A(\omega) \cdot \mathbf{1}_B(\omega), \\ \mathbf{1}_{A \cup B}(\omega) &= \mathbf{1}_A(\omega) \vee \mathbf{1}_B(\omega) = 1 \wedge (\mathbf{1}_A(\omega) + \mathbf{1}_B(\omega)), \end{aligned}$$

wobei das Symbol  $\wedge$  das Minimum und das Symbol  $\vee$  das Maximum zweier Zahlen bezeichnet, z.B.  $-2 \wedge 17 = -2$  und  $1.2 \vee 0.5 = 1.2$ .

*Beispiel 2.2 (Fortsetzung): Den Nettogewinn beim Setzen auf die erste Colonne kann man mit Hilfe einer Indikatorvariablen schreiben,*

$$\begin{aligned} Y(\omega) &= 2 \cdot \mathbf{1}_{C_1}(\omega) - 1 \cdot \mathbf{1}_{\overline{C_1}}(\omega) = 2 \cdot \mathbf{1}_{C_1}(\omega) - (1 - \mathbf{1}_{C_1}(\omega)) \\ &= 3 \cdot \mathbf{1}_{C_1}(\omega) - 1. \end{aligned}$$

Viele Zufallsvorgänge haben selbst Zahlen als Ergebnisse, dann ist  $\Omega \subset \mathbb{R}$ . Wenn das Ergebnis  $\omega$  selbst als Merkmal interessiert, kann man dies durch die Zufallsvariable

$$X(\omega) = \omega, \quad \omega \in \Omega,$$

also durch die identische Abbildung, zum Ausdruck bringen.

*Beispiel 2.3 (vgl. Beispiel 1.2): Werfen eines Würfels mit  $\Omega = \{1, 2, \dots, 6\}$ . Wenn die geworfene Augenzahl als solche von Interesse ist, betrachtet man die Zufallsvariable*

$$X(i) = i \quad \text{für } i = 1, 2, \dots, 6.$$

Sehr wichtig ist es, die Zufallsvariable  $X$  und ihren Wert  $x$  sorgfältig zu unterscheiden. Formal ist  $X$  eine Funktion und  $x$  eine Zahl. Inhaltlich besteht der Unterschied darin, dass  $X$  die Situation *ex ante* beschreibt, bevor sich das zufällige Geschehen ereignet, während  $x$  sich auf die Situation *ex post* bezieht, wenn sich das zufällige Geschehen ereignet hat und sein Ergebnis feststeht. Über  $X$  können Wahrscheinlichkeitsaussagen getroffen werden, über  $x$  jedoch nicht.

### 2.1.1 Verteilungsfunktion

Eine Zufallsvariable  $X$  nimmt ihre Werte mit bestimmten Wahrscheinlichkeiten an. Das grundlegende Hilfsmittel, um mit diesen Wahrscheinlichkeiten zu rechnen, ist die **Verteilungsfunktion** von  $X$ . Dies ist die Funktion

$$\begin{array}{l} F_X : \mathbb{R} \longrightarrow [0, 1], \\ x \longmapsto F_X(x) = P(\{\omega | X(\omega) \leq x\}) . \end{array}$$

Statt  $F_X(x)$  schreibt man auch  $F(x)$ , wenn keine Verwechslungen zu erwarten sind. Die Verteilungsfunktion der Zufallsvariablen  $X$  ordnet also jeder Zahl  $x$  die Wahrscheinlichkeit eines Ereignisses zu, nämlich des Ereignisses, dass  $X$  den Wert  $x$  oder einen kleineren Wert annimmt. Für dieses Ereignis schreibt man auch kürzer

$$\{X \leq x\} = \{\omega | X(\omega) \leq x\}$$

und für seine Wahrscheinlichkeit unter Weglassung der geschweiften Klammern

$$P(X \leq x) = P(\{\omega | X(\omega) \leq x\}) .$$

Hierbei haben wir stillschweigend vorausgesetzt, dass für jedes  $x \in \mathbb{R}$  diese Menge ein Ereignis ist, also  $\{X \leq x\} \in \mathcal{A}$  gilt. Die Funktion

$$B \longmapsto P(X \in B)$$

heißt **Verteilung** der Zufallsvariablen  $X$ , wobei  $B$  alle Teilmengen von  $\mathbb{R}$  durchläuft, die ein Ereignis beschreiben, d.h. für die  $\{\omega | X(\omega) \in B\} \in \mathcal{A}$  ist.

**Eigenschaften der Verteilungsfunktion** Die Verteilungsfunktion  $F$  einer Zufallsvariablen  $X$  besitzt allgemein die folgenden Eigenschaften:

1.  $F$  ist **monoton wachsend**:  $F(x) \leq F(y)$  für  $x < y$ .

Die Monotonie folgt sofort daraus, dass für  $x < y$  die Mengeninklusion  $\{X \leq x\} \subset \{X \leq y\}$  gilt.

2.  $F$  wächst **von null bis eins**:

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} F(x) = 1 .$$

Dies liegt zum einen daran, dass für  $x \rightarrow -\infty$  die Menge  $\{X \leq x\}$  gegen die leere Menge  $\emptyset$  konvergiert und dass  $P(\emptyset) = 0$  ist, zum anderen daran, dass die Menge  $\{X \leq x\}$  für  $x \rightarrow \infty$  gegen  $\Omega$  konvergiert und  $P(\Omega) = 1$  ist.

3.  $F$  ist **rechtsstetig**, d.h. der Funktionswert  $F(x)$  ist an jeder Stelle  $x$  gleich dem **rechtsseitigen Limes**, d.h.

$$\lim_{\substack{y \rightarrow x \\ y > x}} F(y) = F(x).$$

Auf den Beweis der Rechtsstetigkeit wollen wir verzichten. Man beachte, dass eine Verteilungsfunktion im Allgemeinen nicht stetig ist. Sie kann Sprünge aufweisen, an denen der **linksseitige Limes**

$$\lim_{\substack{y \rightarrow x \\ y < x}} F(x)$$

kleiner als der Funktionswert  $F(x)$  ist.

Die drei Eigenschaften einer Verteilungsfunktion lassen sich kurz so zusammenfassen: Eine Verteilungsfunktion wächst monoton von 0 (bei  $x = -\infty$ ) bis 1 (bei  $x = \infty$ ) und ist rechtsstetig.

Umgekehrt kann man zu jeder gegebenen Funktion  $F$ , die diese drei Eigenschaften besitzt, eine Zufallsvariable  $X$  so konstruieren, dass  $F$  die Verteilungsfunktion von  $X$  ist.

*Beispiel 2.2 (Fortsetzung): Die Zufallsvariable  $X$  „Nettogewinn“ bei Setzen von 1 € auf die erste Colonie hat die Verteilungsfunktion*

$$F(x) = \begin{cases} 0 & \text{für } x < -1, \\ \frac{25}{37} & \text{für } -1 \leq x < 2, \\ 1 & \text{für } x \geq 2. \end{cases}$$

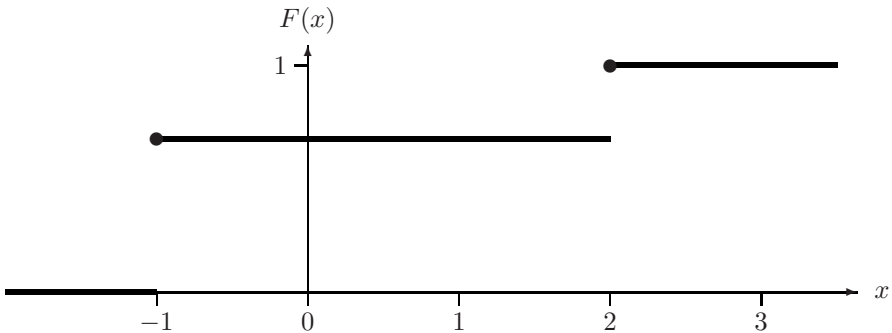


Abbildung 2.1: Verteilungsfunktion der Zufallsvariablen „Nettogewinn“ in Beispiel 2.2

Man sieht in Abbildung 2.1, dass der Graph von  $F$  an den Stellen  $x = -1$  und  $x = 2$  springt,  $F$  aber rechtsstetig ist.

*Beispiel 2.1 (Fortsetzung):* Beim dreimaligen Werfen der Münze nehmen wir an, dass es sich um ein Laplace-Experiment handelt. Dies bedeutet, dass jede der Ergebnisfolgen aus  $K$  (= Kopf) und  $Z$  (= Zahl) die gleiche Wahrscheinlichkeit besitzt. Für die Zufallsvariable  $X =$  „Anzahl Kopf“ gilt dann

$$\begin{aligned} P(X = 0) &= \frac{1}{8}, & P(X = 1) &= \frac{3}{8}, \\ P(X = 2) &= \frac{3}{8}, & P(X = 3) &= \frac{1}{8}, \end{aligned}$$

und für die Verteilungsfunktion

$$F(x) = \begin{cases} 0 & \text{für } x < 0, \\ 0.125 & \text{für } 0 \leq x < 1, \\ 0.5 & \text{für } 1 \leq x < 2, \\ 0.875 & \text{für } 2 \leq x < 3, \\ 1 & \text{für } x \geq 3. \end{cases}$$

In vielen Anwendungen der Wahrscheinlichkeitsrechnung genügt es, statt einer Zufallsvariablen  $X$  lediglich ihre Verteilungsfunktion zu untersuchen. Oft ist es kompliziert oder überhaupt nicht möglich,  $\Omega$  und die Abbildungsvorschrift  $X : \Omega \rightarrow \mathbb{R}$  explizit anzugeben, jedoch lassen sich die Wahrscheinlichkeiten, mit denen eine Zufallsvariable bestimmte Werte annimmt oder der Verlauf ihrer Verteilungsfunktion aus sachlogischen Erwägungen unmittelbar bestimmen. Eine explizite Angabe der Ergebnismenge  $\Omega$  und punktweise Definition einer Zufallsvariablen  $X$  als  $X(\omega)$ ,  $\omega \in \Omega$ , ist in der Regel nicht nötig und aus Sicht des Anwenders ohne Belang.

*Beispiel 2.4:* Die Lebensdauer eines elektronischen Bauteils wird durch die Bedingungen seiner Herstellung, die Umstände seiner Nutzung und durch zahlreiche weitere Einflüsse bestimmt. Wir fassen die Lebensdauer deshalb als Zufallsvariable auf und bezeichnen sie mit  $X$ . Die im konkreten Fall realisierte Dauer  $X(\omega)$  hängt dann vom Ergebnis  $\omega$  des ihr zugrundeliegenden Zufallsvorgangs ab. Um  $\omega \in \Omega$  und die Zuordnung  $\omega \mapsto X(\omega)$  zu beschreiben, müsste man sämtliche Einflussfaktoren spezifizieren, was praktisch kaum möglich ist. Besser ist es, ganz darauf zu verzichten und sich durch sachbezogene Überlegungen eine Vorstellung vom Verlauf der Verteilungsfunktion zu verschaffen. Man kann zeigen (siehe unten Abschnitt 2.4.2), dass, wenn das Bauteil in einem bestimmten Sinn nicht „altert“, die folgende Verteilungsfunktion (mit geeignetem Parameter  $\lambda$ ) angemessen ist:

$$F(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1 - e^{-\lambda x}, & \text{falls } x \geq 0. \end{cases}$$

Zahlreiche weitere Beispiele finden sich in den folgenden Abschnitten 2.3 und 2.4 über spezielle Verteilungen.

Die Verteilungsfunktion  $F$  einer Zufallsvariablen gibt die Wahrscheinlichkeit für alle Ereignisse der Form  $\{X \leq x\}$ ,  $x \in \mathbb{R}$ , an. Hieraus kann man die Wahrscheinlichkeiten für alle anderen interessierenden Ereignisse ableiten. Die Wahrscheinlichkeit, dass  $X$  in ein – beschränktes oder unbeschränktes – **Intervall**<sup>1</sup> fällt, lässt sich in besonders einfacher Weise aus der Verteilungsfunktion berechnen:

$$\begin{aligned} P(X > x) &= 1 - P(X \leq x) = 1 - F(x), \\ P(a < X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a), \end{aligned}$$

wenn  $a < b$  ist. Für das Ereignis  $\{X = x\}$  gilt

$$P(X = x) = P(X \leq x) - P(X < x) = F(x) - \lim_{\substack{y \rightarrow x \\ y < x}} F(y). \quad (2.1)$$

Hieraus folgert man

$$P(X \geq a) = 1 - P(X < a) = 1 - \lim_{\substack{y \rightarrow a \\ y < a}} F(y)$$

und analog

$$P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - \lim_{\substack{y \rightarrow a \\ y < a}} F(y),$$

wenn  $a \leq b$  ist.

### 2.1.2 Quantilfunktion

Für eine gegebene Zahl  $x$  ist  $F_X(x)$  die Wahrscheinlichkeit, dass  $X$  höchstens den Wert  $x$  annimmt. Umgekehrt kann man, wenn eine Wahrscheinlichkeit  $p$  gegeben ist, fragen, welchen Wert  $x$  die Zufallsvariable mit Wahrscheinlichkeit  $p$  nicht überschreitet. Hierbei ist allerdings Vorsicht geboten. Wenn die Verteilungsfunktion stetig ist und streng monoton wächst, besitzt sie eine eindeutige Umkehrfunktion  $Q_X : p \mapsto x_p$ ,  $0 < p < 1$ , genannt **Quantilfunktion**. Dies ist im folgenden Beispiel der Fall.

*Beispiel 2.4 (Fortsetzung): Die Lebensdauer eines elektronischen Bauteils werde als Zufallsvariable  $X$  mit der Verteilungsfunktion*

$$F_X(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1 - e^{-x}, & \text{falls } x \geq 0, \end{cases}$$

<sup>1</sup>Dazu gehören die abgeschlossenen Intervalle  $[a, b]$ , die offenen Intervalle  $]a, b[$ , die halb-offenen Intervalle  $[a, b[$  und  $]a, b]$ , sowie die Halbgeraden  $[c, \infty[$ ,  $]c, \infty[$ ,  $] - \infty, c]$  und  $] - \infty, c[$ , wobei  $a$ ,  $b$  und  $c$  reelle Zahlen sind mit  $a < b$ .

angesehen. Offenbar ist  $F_X$  stetig und für  $x \geq 0$  streng monoton wachsend. Die Quantilfunktion ist in diesem Fall gleich der gewöhnlichen Umkehrfunktion von  $F_X$ . Für jedes  $p \in ]0, 1[$  gilt nämlich

$$F_X(x) = p \quad \Leftrightarrow \quad 1 - e^{-x} = p \quad \Leftrightarrow \quad x = -\ln(1 - p),$$

also

$$Q_X(p) = x_p = -\ln(1 - p) = \ln\left(\frac{1}{1 - p}\right).$$

Diese Quantilfunktion ist zusammen mit der Verteilungsfunktion in Abbildung 2.2.a abgebildet.

Wenn die Verteilungsfunktion jedoch Sprünge macht, gibt es nicht für jedes  $p$  ein solches  $x_p$  und, wenn – wie im obigen Beispiel 2.1 – die Verteilungsfunktion zum Teil waagrecht verläuft, ist dieses  $x_p$  nicht immer eindeutig bestimmt. Die folgende allgemeine Definition trägt dem Rechnung:

Für eine Zufallsvariable  $X$  mit Verteilungsfunktion  $F_X$  heißt die Funktion

$$\begin{aligned} Q_X : ]0, 1[ &\longrightarrow \mathbb{R}, \\ p &\longmapsto Q_X(p) = \min\{x \mid F_X(x) \geq p\}, \end{aligned}$$

**Quantilfunktion** von  $X$ . Der Wert der Quantilfunktion  $x_p = Q_X(p)$  heißt  $p$ -**Quantil** von  $X$ . Das  $p$ -Quantil  $x_p$  ist demnach die kleinste Zahl  $x \in \mathbb{R}$  mit der Eigenschaft, dass  $F_X(x)$  den Wert  $p$  erreicht oder überspringt. Es wird auch als  $p \cdot 100\%$ -**Punkt** bezeichnet. Statt  $Q_X$  und  $F_X$  schreiben wir im Folgenden kürzer  $Q$  und  $F$ , wenn nur eine Zufallsvariable  $X$  in Frage kommt.

Der Graph der Quantilfunktion  $Q$  hängt eng mit dem der Verteilungsfunktion  $F$  zusammen (siehe Abbildung 2.2.a bis c): Er ist nichts anderes als die Spiegelung des Graphen von  $F$  an der 45-Grad-Linie. Dabei gehen etwaige Sprünge von  $F$  (in der Abbildung als senkrechte Linien gestrichelt) in waagrechte Abschnitte des Graphen von  $Q$  über, und etwaige waagrechte Abschnitte des Graphen von  $F$  werden zu senkrechten Linien (das sind die Sprünge) des Graphen von  $Q$ .

*Beispiel 2.5:* Eine Zufallsvariable  $X$  nehme die drei Werte 0.3, 1.2 und 2.4 an, und zwar mit den Wahrscheinlichkeiten

$$P(X = 0.3) = 0.2, \quad P(X = 1.2) = 0.4, \quad P(X = 2.4) = 0.4.$$

Die Verteilungsfunktion und die Quantilfunktion von  $X$  sind in Abbildung 2.2.b dargestellt.



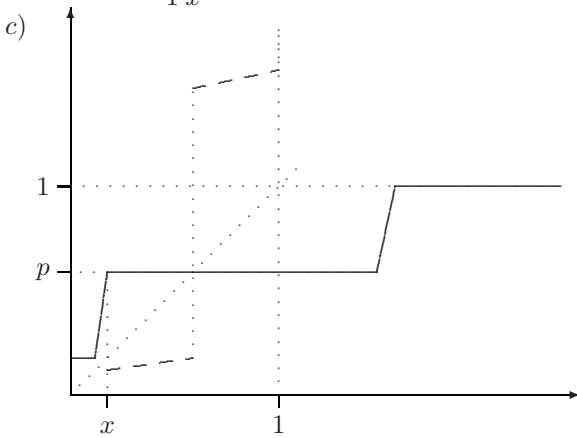
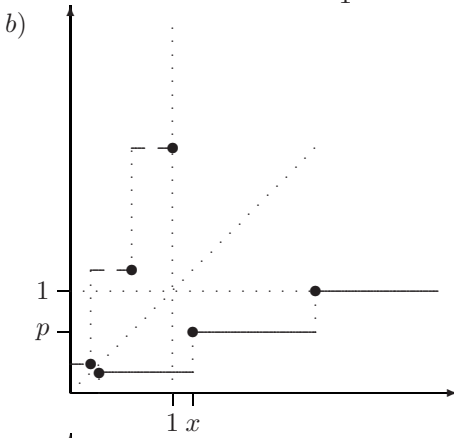
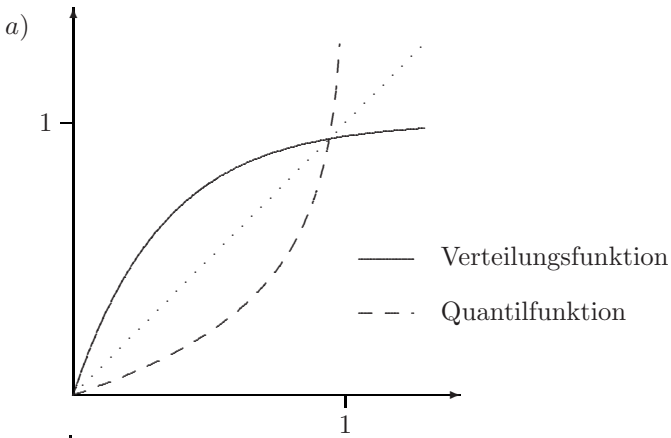


Abbildung 2.2: Drei Verteilungsfunktionen und ihre Quantilfunktionen

Für einige Quantile verwendet man spezielle Namen:

				$x_{0.5}$						<b>Median</b>
			$x_{0.25}$	$x_{0.5}$	$x_{0.75}$					<b>Quartile</b>
		$x_{0.2}$	$x_{0.4}$	$x_{0.6}$	$x_{0.8}$					<b>Quintile</b>
$x_{0.1}$	$x_{0.2}$	$x_{0.3}$	$x_{0.4}$	$x_{0.5}$	$x_{0.6}$	$x_{0.7}$	$x_{0.8}$	$x_{0.9}$		<b>Dezile</b>

Wenn die Zufallsvariable  $X$  genügend viele verschiedene Werte annimmt, lassen sich diese Quantile wie folgt interpretieren:

- Der Median  $x_{0.5}$  teilt die reelle Achse in zwei Teile, die jeweils ungefähr 50% der Wahrscheinlichkeitsmasse tragen.
- Die Quartile  $x_{0.25}$ ,  $x_{0.5}$  und  $x_{0.75}$  teilen die reelle Achse in vier Teile ein, auf denen jeweils ungefähr ein Viertel der Wahrscheinlichkeitsmasse liegt. Analog interpretiert man Quintile und Dezile.

Im Beispiel 2.4 sind Median und oberes Dezil durch

$$x_{0.5} = \ln\left(\frac{1}{1-0.5}\right) = \ln 2 = 0.69 \quad \text{und}$$

$$x_{0.9} = \ln 10 = 2.30$$

gegeben, im Beispiel 2.5 durch  $y_{0.5} = 1.2$  und  $y_{0.9} = 2.4$ .

Im Beispiel 2.4 liegt links und rechts vom Median jeweils exakt die Hälfte der Wahrscheinlichkeitsmasse. Es ist nämlich

$$P(X < x_{0.5}) = F(\ln 2) = 1 - e^{-\ln 2} = 0.5 \quad \text{und}$$

$$P(X > x_{0.5}) = 1 - \lim_{\substack{x \rightarrow \ln 2 \\ x < \ln 2}} F(x) = 1 - (1 - e^{-\ln 2}) = 0.5.$$

Im Beispiel 2.5 hingegen haben wir

$$P(X < x_{0.5}) = P(X < 1.2) = 0.2 \quad \text{und}$$

$$P(X > x_{0.5}) = P(X > 1.2) = 0.4.$$

Wie man sieht, treffen bei Zufallsvariablen, die wie die im Beispiel 2.5 mit relativ großer Wahrscheinlichkeit einzelne Werte annehmen, die obigen Interpretationen nicht genau zu. Weitere Hinweise zur Berechnung von Quantilen finden sich im Abschnitt 2.4. In der statistischen Literatur und Software kommen auch andere Definitionen von Quantilen vor.

**Eigenschaften der Quantilfunktion** Die Quantilfunktion  $Q$  hat ähnliche formale Eigenschaften wie die Verteilungsfunktion: Sie ist

1. **monoton wachsend**, d.h. für  $p < p'$  ist  $Q(p) \leq Q(p')$ ,

2. **linkstetig**, aber im Allgemeinen nicht stetig.

Verteilungsfunktion  $F$  und Quantilfunktion  $Q$  sind wie folgt miteinander verknüpft:

$$\begin{aligned} F(Q(p)) &\geq p && \text{für } p \in ]0, 1[, \\ Q(F(x)) &\leq x && \text{für } x \in \mathbb{R}. \end{aligned}$$

Man beachte, dass die beiden Ungleichheitszeichen im Allgemeinen nicht durch Gleichheitszeichen ersetzt werden dürfen.

### 2.1.3 Diskrete Zufallsvariable

In der elementaren Wahrscheinlichkeitsrechnung unterscheidet man zwei Typen von Zufallsvariablen, diskrete und stetige Variable. Im Umgang mit diesen beiden Typen von Zufallsvariablen werden unterschiedliche mathematische Hilfsmittel benutzt; zum Einen das Rechnen mit endlichen und unendlichen Summen, zum Anderen die Differential- und Integralrechnung.

Eine Zufallsvariable  $X$  heißt **diskret**, wenn es entweder

- endlich viele Punkte  $x_1, x_2, \dots, x_J$  oder
- abzählbar unendlich viele Punkte  $x_1, x_2, x_3, \dots$

so gibt, dass

$$p_j = P(X = x_j) > 0 \quad \text{für alle } j \quad \text{sowie} \quad \sum_j p_j = 1$$

gilt. Die Menge

$$T_X = \{x_1, x_2, \dots, x_J\} \quad \text{bzw.} \quad T_X = \{x_1, x_2, \dots\}$$

heißt **Träger** von  $X$ . Der Träger einer diskreten Zufallsvariablen ist die Menge aller Werte, die  $X$  mit einer positiven Wahrscheinlichkeit annimmt. Es gilt  $P(X \in T_X) = 1$ . Die Funktion

$$f(x) = f_X(x) = \begin{cases} p_j, & \text{falls } x = x_j \quad \text{für ein } j, \\ 0 & \text{sonst,} \end{cases}$$

heißt **Wahrscheinlichkeitsfunktion** von  $X$ . Sie kann etwa mittels eines Stabdiagramms graphisch dargestellt werden.

Die **Verteilung** der diskreten Variablen  $X$  ist durch die Paare  $(x_j, p_j)$ ,  $j = 1, 2, \dots$ , eindeutig bestimmt, denn für jede Menge  $B$  gilt

$$P(X \in B) = \sum_{j|x_j \in B} p_j.$$

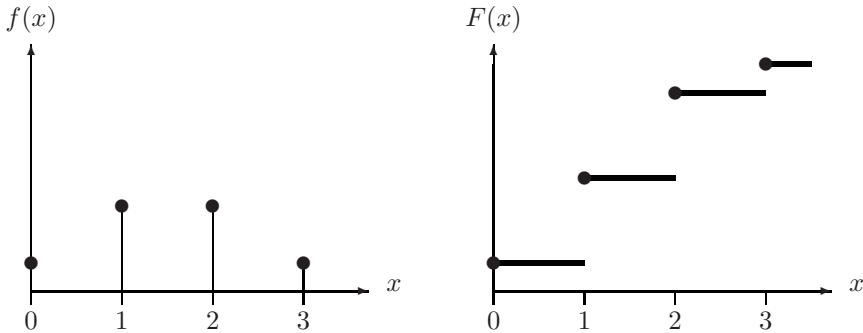


Abbildung 2.3: Wahrscheinlichkeitsfunktion und Verteilungsfunktion von „Anzahl Kopf“ im Beispiel 2.1

Dabei erstreckt sich die Summe über alle  $p_j$ , für die  $x_j$  in  $B$  liegt. Insbesondere ist

$$F(x) = P(X \leq x) = \sum_{j|x_j \leq x} p_j$$

die Verteilungsfunktion an der Stelle  $x \in \mathbb{R}$ . Die **Verteilungsfunktion** einer diskreten Zufallsvariablen ist offenbar eine Treppenfunktion, deren Sprünge an den Stellen  $x_j \in T_X$  liegen. Die Sprunghöhe an der Stelle  $x_j$  ist gleich

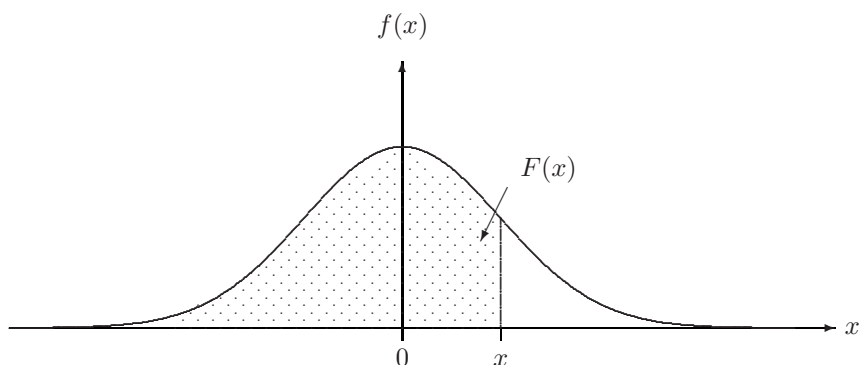
$$F(x_j) - \lim_{\substack{x \rightarrow x_j \\ x < x_j}} F(x) = p_j,$$

also gleich der Wahrscheinlichkeit für  $x_j$ .

*Beispiel 2.1 (Fortsetzung): Beim dreifachen Wurf einer Münze werden alle möglichen Ergebnisse als gleich wahrscheinlich angenommen (Laplace-Experiment). Dann ist  $X$  (= Anzahl Kopf) eine diskrete Zufallsvariable mit Wahrscheinlichkeitsfunktion*

$$f(x) = \begin{cases} 0.125, & \text{falls } x = 0, \\ 0.375, & \text{falls } x = 1, \\ 0.375, & \text{falls } x = 2, \\ 0.125, & \text{falls } x = 3, \\ 0 & \text{sonst.} \end{cases}$$

Wahrscheinlichkeitsfunktion und Verteilungsfunktion sind in Abbildung 2.3 dargestellt.

Abbildung 2.4: Dichte  $f(x)$  und Verteilungsfunktion  $F(x)$ 

## 2.1.4 Stetige Zufallsvariable

Eine Zufallsvariable  $X$  nennt man **stetig**, wenn sich ihre Verteilungsfunktion  $F_X$  als Integral einer Funktion  $f_X : \mathbb{R} \rightarrow [0, \infty[$  schreiben lässt, d.h. wenn

$$F_X(x) = \int_{-\infty}^x f_X(t) dt \quad \text{für alle } x \in \mathbb{R}. \quad (2.2)$$

Die Funktion  $f_X(x)$  heißt dann **Dichtefunktion**, kurz: **Dichte**, von  $X$ , und man sagt,  $X$  sei **stetig verteilt**. Statt  $F_X(x)$  und  $f_X(x)$  schreibt man auch  $F(x)$  bzw.  $f(x)$ .

Die Verteilungsfunktion einer stetigen Zufallsvariablen ist also eine Stammfunktion ihrer Dichtefunktion; in Abbildung 2.4 entspricht der Wert  $F(x)$  dem Inhalt der punktierten Fläche unterhalb des Graphen der Dichte  $f$ . Durch Differenzieren der Verteilungsfunktion erhält man die Dichte,

$$f(x) = F'(x),$$

vorausgesetzt die Verteilungsfunktion  $F$  ist an der Stelle  $x$  tatsächlich differenzierbar. Bei den stetigen Zufallsvariablen, die im Folgenden untersucht werden, ist die Verteilungsfunktion an höchstens ein oder zwei Stellen nicht differenzierbar. An diesen Stellen macht die Dichte einen Sprung.

**Eigenschaften der Dichtefunktion** Die wesentlichen Eigenschaften sind:

1. **Nichtnegativität**  $f(x) \geq 0$  für alle  $x \in \mathbb{R}$ ,

2. **Normiertheit**  $\int_{-\infty}^{\infty} f(x)dx = 1.$

Die beiden Eigenschaften folgen aus der Monotonie von  $F$  und der Tatsache, dass  $\lim_{x \rightarrow \infty} F(x) = 1$  ist. Umgekehrt lässt sich zu jeder Funktion  $f$ , die diese beiden Eigenschaften hat, eine stetige Zufallsvariable konstruieren, die  $f$  als Dichte besitzt.

Eine Besonderheit stetiger Zufallsvariablen besteht darin, dass für jede gegebene Zahl  $x$  die Wahrscheinlichkeit, dass  $X$  den Wert  $x$  annimmt, gleich null ist. Da  $F$  stetig ist, gilt nämlich für jedes  $x \in \mathbb{R}$

$$\lim_{\substack{y \rightarrow x \\ y < x}} F(y) = F(x),$$

und deshalb wegen (2.1)

$$P(X = x) = 0.$$

Dennoch gibt es Mengen  $B$ , für die  $P(B) > 0$  ist, insbesondere gilt offenbar  $P(X \in \mathbb{R}) = 1$ .

**Wahrscheinlichkeiten für Intervalle** Speziell bestimmen wir nun die Wahrscheinlichkeit eines Intervalls, das durch  $a$  und  $b$  begrenzt wird,  $a < b$ . Es gilt:

$$\begin{aligned} P(a < X \leq b) &= F(b) - F(a) \\ &= \int_{-\infty}^b f(t)dt - \int_{-\infty}^a f(t)dt = \int_a^b f(t)dt. \end{aligned}$$

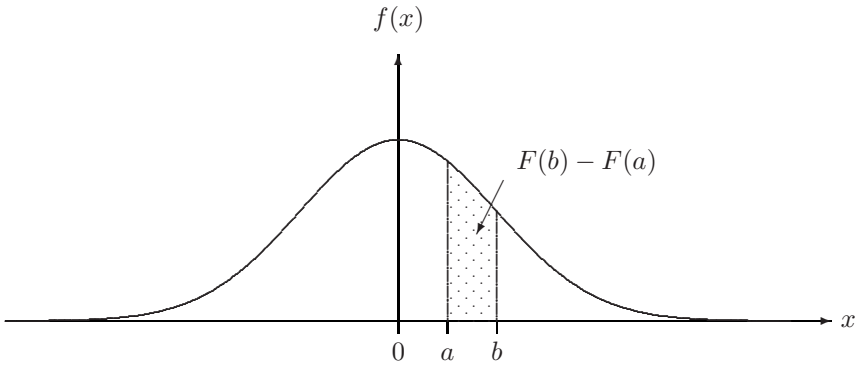
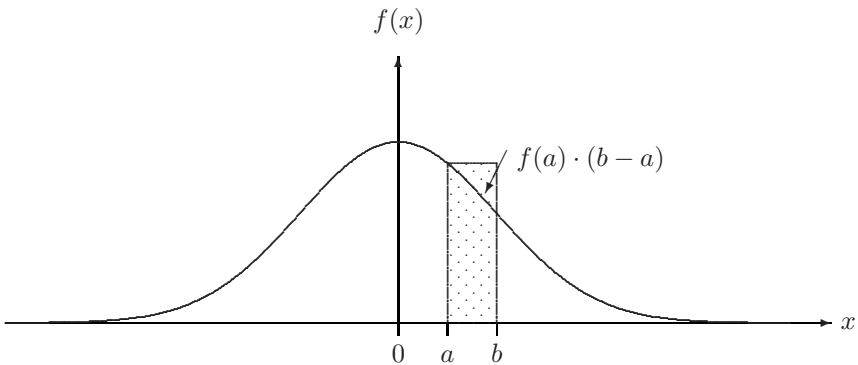
Da  $P(X = a) = P(X = b) = 0$  ist, gilt auch

$$\boxed{P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b)}.$$

Die Wahrscheinlichkeit, dass eine stetige Zufallsvariable  $X$  in ein Intervall mit den Grenzen  $a$  und  $b$  fällt, ergibt sich also als Integral über die Dichte von  $X$  in den Grenzen  $a$  und  $b$ ; siehe Abbildung 2.5. Ob die Grenzen zum Intervall gehören oder nicht, ist dabei gleichgültig.

Man beachte, dass eine Dichte auch Werte größer als 1 annehmen kann. Der Wert  $f(x)$  gibt also nicht etwa die Wahrscheinlichkeit an, dass  $X$  den Wert  $x$  annimmt (diese ist gleich null), vielmehr ist

$$f(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} P(x \leq X \leq x + \varepsilon),$$

Abbildung 2.5: Wahrscheinlichkeit für ein Intervall mit den Grenzen  $a$  und  $b$ Abbildung 2.6: Approximation der Wahrscheinlichkeit für ein Intervall mit den Grenzen  $a$  und  $b$ 

d.h.  $f(x)$  ist näherungsweise gleich  $1/\varepsilon$  mal der Wahrscheinlichkeit, dass  $X$  in ein kleines Intervall bei  $x$  fällt, das die Länge  $\varepsilon$  besitzt. Die Wahrscheinlichkeit, dass  $X$  in ein kleines Intervall  $[a, b]$  fällt, kann deshalb in erster Näherung durch

$$P(a \leq X \leq b) \approx f(a)(b - a)$$

approximiert werden; vgl. Abbildung 2.6.

**Träger** Wir kommen nun zum allgemeinen Begriff des Trägers einer Zufallsvariablen  $X$ . Unter dem **Träger** von  $X$  verstehen wir die kleinste abge-

geschlossene<sup>2</sup> Menge  $T_X \subset \mathbb{R}$  mit der Eigenschaft

$$P(X \in T_X) = 1.$$

$T_X$  ist also die kleinste abgeschlossene Menge, außerhalb derer sich keine Wahrscheinlichkeitsmasse befindet. Falls  $X$  endlich diskret verteilt ist, bilden die Punkte  $x_j$ , die mit positiver Wahrscheinlichkeit  $p_j$  angenommen werden, eine endliche Menge; diese ist abgeschlossen und daher der Träger von  $X$ . Falls  $X$  stetig verteilt ist, gibt es keine einzelnen Punkte, die mit positiver Wahrscheinlichkeit angenommen werden; hier benötigen wir die allgemeinere Form der Definition des Trägers. Alle Beispiele stetiger Verteilungen in diesem Lehrbuch haben ein abgeschlossenes Intervall (beschränkt oder unbeschränkt) als Träger.

**Quantilfunktion** Da die Verteilungsfunktion einer stetigen Zufallsvariablen stetig ist, nimmt ihre Quantilfunktion die folgende einfache Form an:

$$Q(p) = \min \{x \mid F(x) = p\}, \quad 0 < p < 1.$$

Zur Bestimmung eines Quantils  $x_p$  muss man lediglich die kleinste Lösung  $x$  der Gleichung  $F(x) = p$  aufsuchen. In allen Beispielen des Lehrbuchs wird die Lösung der Gleichung eindeutig sein.

*Beispiel 2.6: In einer Fernsehshow wird ein Glücksrad in Bewegung gesetzt und der Stillstand abgewartet. Es sei*

$$X = \text{Position des Glücksrads in Grad bei Stillstand.}$$

*Der Mechanismus ist so konstruiert, dass keine Winkel bevorzugt auftreten. Man kann deshalb davon ausgehen, dass  $X$  eine stetige Zufallsvariable mit der folgenden Dichte ist:*

$$f(x) = \begin{cases} \frac{1}{360}, & \text{falls } 0 \leq x \leq 360, \\ 0 & \text{sonst.} \end{cases}$$

*Für die Verteilungsfunktion gilt dann*

$$F(x) = \begin{cases} 0, & \text{falls } x < 0, \\ \frac{1}{360}x, & \text{falls } 0 \leq x < 360, \\ 1, & \text{falls } x \geq 360. \end{cases}$$

---

<sup>2</sup>Eine Menge  $T \subset \mathbb{R}$  heißt *abgeschlossen*, wenn alle Randpunkte von  $T$  als Elemente zu  $T$  gehören. Jedes abgeschlossene Intervall, jede endliche Vereinigung abgeschlossener Intervalle und jede endliche Menge sind in diesem Sinne abgeschlossen. Auch  $\mathbb{N}$ ,  $\mathbb{R}$  und  $[0, \infty[$  sind abgeschlossene Mengen.



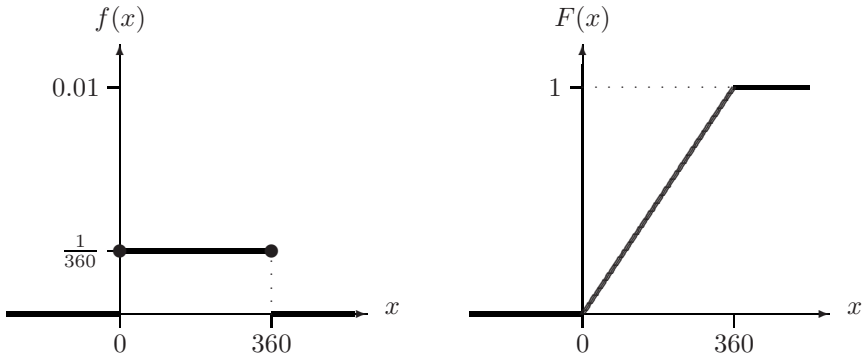


Abbildung 2.7: Dichte und Verteilungsfunktion des „angezeigten Winkels“ in Beispiel 2.6

Träger der Verteilung ist das Intervall  $[0, 360]$ ; vgl. Abbildung 2.7. Die Quantilfunktion ermittelt man durch Auflösung der Gleichung  $p = F(x)$  nach  $x = Q(p)$ , also

$$Q(p) = 360p \quad \text{für } 0 < p < 1.$$

Das untere Quartil z.B. ist  $x_{0.25} = Q(0.25) = 90$ .

*Beispiel 2.4 (Fortsetzung):* Wir betrachten wieder die Lebensdauer  $X$  eines elektronischen Bauteils mit der Verteilungsfunktion

$$F(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1 - e^{-x}, & \text{falls } x \geq 0. \end{cases}$$

Da  $F$  für alle  $x \neq 0$  differenzierbar ist, erhalten wir die Dichte

$$f(x) = \begin{cases} 0, & \text{falls } x < 0, \\ e^{-x}, & \text{falls } x \geq 0. \end{cases}$$

In diesem Beispiel ist der Träger gleich  $[0, \infty[$ . Die Dichte ist in Abbildung 2.8 dargestellt.

**Intervallwahrscheinlichkeiten** Zum Schluss dieses Abschnitts fassen wir die Formeln für die Wahrscheinlichkeit, dass  $X$  in ein Intervall fällt, noch einmal tabellarisch zusammen. Je nach Verteilungstyp gilt Folgendes:

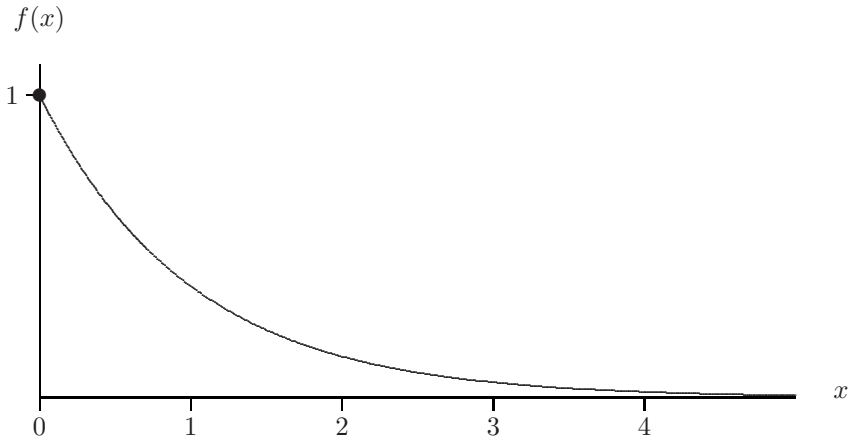


Abbildung 2.8: Dichte der Lebensdauervertelung im Beispiel 2.4

	allgemein	stetige Verteilung	diskrete Verteilung
$P(a < X \leq b)$	$F(b) - F(a)$	$\int_a^b f(x) dx$	$\sum_{j a < x_j \leq b} p_j$
$P(X \leq b)$	$F(b)$	$\int_{-\infty}^b f(x) dx$	$\sum_{j x_j \leq b} p_j$
$P(X > a)$	$1 - F(a)$	$\int_a^{\infty} f(x) dx$	$\sum_{j x_j > a} p_j$

### 2.1.5 Affin-lineare Transformation von Zufallsvariablen

Häufig kennt man die Verteilung einer Zufallsvariablen  $X$  und ist an der Verteilung einer anderen Zufallsvariablen  $Y$  interessiert, die mit  $X$  durch eine monoton wachsende **affin-lineare Transformation** – d.h. eine Nullpunktverschiebung und/oder Skalenänderung – verknüpft ist. Im Folgenden geben wir die Verteilungsfunktion, die Quantilfunktion und – im Fall einer stetigen Variablen – die Dichte von  $Y$  an.

Gegeben sei die Verteilungsfunktion  $F_X$  einer Zufallsvariablen  $X$ . Mit  $a, b \in \mathbb{R}$  und  $b > 0$  wird durch  $Y = a + bX$  eine weitere Zufallsvariable definiert. Dann gilt:

1. Die Verteilungsfunktion  $F_Y$  von  $Y$  ist

$$\boxed{F_Y(y) = F_X\left(\frac{y-a}{b}\right), \quad y \in \mathbb{R},} \quad (2.3)$$

denn für  $b > 0$  gilt

$$F_Y(y) = P(a + bX \leq y) = P\left(X \leq \frac{y-a}{b}\right) = F_X\left(\frac{y-a}{b}\right).$$

2. Wenn  $x_p = Q_X(p)$  das  $p$ -Quantil von  $X$  ist, so ist

$$\boxed{y_p = a + b x_p} \quad (2.4)$$

das  $p$ -Quantil von  $Y = a + bX$ ,  $0 < p < 1$ . Dies folgt direkt aus der Definition des Quantils.

3. Falls  $X$  stetig verteilt ist mit Dichte  $f_X$ , ist auch  $Y$  stetig verteilt.  $Y$  besitzt die Dichte

$$\boxed{f_Y(y) = \frac{1}{b} f_X\left(\frac{y-a}{b}\right).} \quad (2.5)$$

Man erhält sie durch Differenzieren der Verteilungsfunktion (2.3). Soweit  $F_X$  differenzierbar ist, gilt für  $y \in \mathbb{R}$

$$\begin{aligned} F_Y(y) &= F_X\left(\frac{y-a}{b}\right), \\ \frac{d}{dy} F_Y(y) &= \frac{d}{dy} F_X\left(\frac{y-a}{b}\right), \\ f_Y(y) &= F'_X\left(\frac{y-a}{b}\right) \cdot \frac{1}{b} = \frac{1}{b} \cdot f_X\left(\frac{y-a}{b}\right). \end{aligned}$$

*Beispiel 2.6 (Fortsetzung):* Mit dem rotierenden Glücksrad wird in der Fernsehshow folgendes Gewinnspiel verbunden. Ein zuvor bestimmter Kandidat erhält als Gewinn in jedem Fall 999 € ausgezahlt. Der Gewinn erhöht sich pro Grad des angezeigten Winkels um 9.99 €. Wie lautet die Wahrscheinlichkeitsverteilung des Gewinns?

Die Gewinnfunktion (in €) lautet  $g(x) = 999 + 9.99x$ , wobei  $x$  der angezeigte Winkel ist,  $0 \leq x \leq 360$ . Das Minimum ist  $g(0) = 999$ , das Maximum  $g(360) = 999 + 9.99 \cdot 360 = 4595.40$ . Als Zufallsvariable  $Y$  geschrieben, ist der Gewinn eine affin-lineare Transformation der Zufallsvariablen  $X$ , nämlich

$$Y = 999 + 9.99X.$$

Da  $X$  stetig verteilt ist mit der Dichte

$$f_X(x) = \begin{cases} \frac{1}{360}, & \text{falls } 0 \leq x \leq 360, \\ 0 & \text{sonst,} \end{cases}$$

folgt, dass  $Y$  ebenfalls stetig verteilt ist und wegen Formel (2.5) die Dichte

$$f_Y(y) = \frac{1}{9.99} f_X\left(\frac{y-999}{9.99}\right)$$

besitzt, also (unter Anpassung der Intervallgrenzen für  $y$ )

$$f_Y(y) = \begin{cases} \frac{1}{9.99 \cdot 360}, & \text{falls } 999 \leq y \leq 4595.40, \\ 0 & \text{sonst.} \end{cases}$$

## 2.1.6 Unimodalität

Unter den Wahrscheinlichkeitsverteilungen sind diejenigen ausgezeichnet, deren Dichte bzw. Wahrscheinlichkeitsfunktion  $f(x)$  in Abhängigkeit von  $x$  zunächst anwächst und dann wieder abfällt. Solche Verteilungen sind in den obigen Beispielen 2.1 und 2.6 aufgetreten; siehe Abbildung 2.9. Sie werden als **unimodal** oder **eingipflig** bezeichnet. Zur genauen Definition müssen wir zwischen stetigen und diskreten Zufallsvariablen unterscheiden.

**Unimodale stetige Verteilung** Eine stetige Wahrscheinlichkeitsverteilung mit der Dichte  $f$  heißt **unimodal**, wenn es mindestens einen Wert  $x_M$  gibt, für den

$$\begin{aligned} f(x) \leq f(y) \leq f(x_M) & \quad \text{für alle } x, y \text{ mit } x < y < x_M \\ \text{und } f(x_M) \geq f(y) \geq f(x) & \quad \text{für alle } x, y \text{ mit } x_M < y < x \end{aligned}$$

gilt. Jeder solche Wert  $x_M$  wird **Modus** genannt. Offenbar kann eine unimodale stetige Verteilung mehrere Modi besitzen, diese bilden ein Intervall.

*Beispiel 2.7: Die Verteilung mit der Dichte (siehe Abbildung 2.10 links)*

$$f(x) = \begin{cases} 1 - |x|, & \text{falls } -1 \leq x \leq 1, \\ 0 & \text{sonst,} \end{cases}$$

ist unimodal mit einem eindeutigen Modus, nämlich  $x_M = 0$ .

*Beispiel 2.6 (Fortsetzung): Offenbar ist die Verteilung aus Beispiel 2.6 unimodal und jeder Punkt des Intervalls  $[0, 360]$  ist ein Modus; vgl. Abbildung 2.9 rechts.*

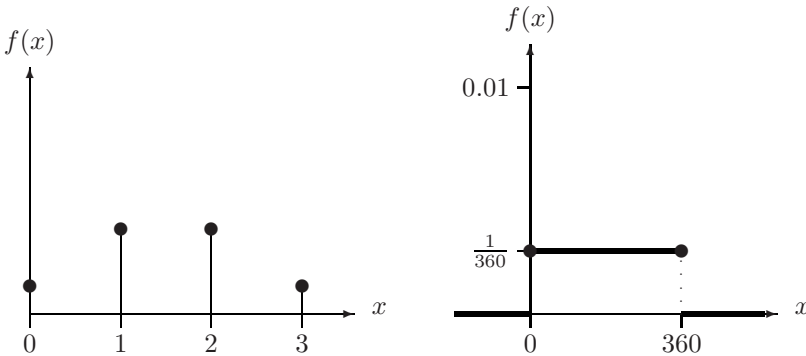


Abbildung 2.9: Wahrscheinlichkeitsfunktion aus Beispiel 2.1 und Dichtefunktion aus Beispiel 2.6

Beispiel 2.8: Abbildung 2.10 rechts zeigt die Dichte einer stetigen Verteilung, die nicht unimodal ist.

**Unimodale diskrete Verteilung** Wir betrachten nun die Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariablen  $X$ . Der Träger von  $X$  sei aufsteigend geordnet,

$$x_j \leq x_k, \quad \text{wenn } j < k,$$

und es sei  $p_j = P(X = x_j), \sum_j p_j = 1$ . Die Verteilung von  $X$  heißt **unimodal**, wenn es mindestens einen Wert  $x_M = x_{j^*}$  mit  $P(X = x_M) = p_{j^*}$  gibt, für den

$p_j \leq p_k \leq p_{j^*} \quad \text{für alle } j, k \text{ mit } j < k < j^*$ $\text{und } p_{j^*} \geq p_k \geq p_j \quad \text{für alle } j, k \text{ mit } j^* < k < j$	(2.5)
---	-------

gilt. Jeder solche Wert  $x_M$  wird **Modus** genannt.

Beispiel 2.1 (Fortsetzung): Die Verteilung beim dreifachen Münzwurf (vgl. Abbildung 2.9 links) ist unimodal. Sie besitzt zwei Modi, nämlich  $x_{M1} = 1$  und  $x_{M2} = 2$ .

Wenn eine Verteilung unimodal ist, lässt sich ihre Wölbung (vgl. Abschnitt 2.2.4) anschaulich interpretieren. Unimodale Verteilungen besitzen – gegenüber allgemeinen Verteilungen – zahlreiche günstige Eigenschaften. So lassen sich beispielsweise Abschätzungen von Wahrscheinlichkeiten durch die **Tschebyscheff-Ungleichung** (Abschnitt 2.2.3) verbessern, wenn die zugrundeliegende Verteilung unimodal ist. Eine umfassende Darstellung der

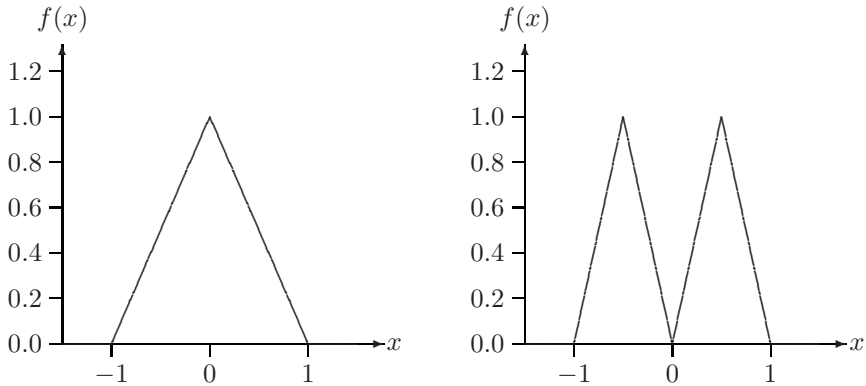


Abbildung 2.10: Unimodale (Beispiel 2.7) und nicht unimodale (Beispiel 2.8) Verteilung

Eigenschaften unimodaler Verteilungen findet man in Dharmadhikari und Joag-dev (1988).

### 2.1.7 Symmetrie

In vielen Fällen schwankt eine Zufallsvariable um einen zentralen Wert. Wenn die Abweichungen nach oben und unten mit den gleichen Wahrscheinlichkeiten auftreten, spricht man von einer symmetrischen Wahrscheinlichkeitsverteilung.

Sei  $X$  eine Zufallsvariable und  $c \in \mathbb{R}$  eine Zahl. Wenn

$$P(X \leq c - y) = P(X \geq c + y) \quad \text{für alle } y \in \mathbb{R}$$

gilt, sagt man,  $X$  sei **symmetrisch zu  $c$  verteilt**. Gleichbedeutend gilt  $P(X - c \leq -y) = P(-(X - c) \leq -y)$  für alle  $y$ , mit anderen Worten, die beiden Zufallsvariablen

$$X - c \quad \text{und} \quad -(X - c) \quad \text{haben die gleiche Verteilung.}$$

Wenn  $X$  stetig verteilt ist mit Dichte  $f(x)$  und Verteilungsfunktion  $F(x)$ , ist die Symmetrie der Verteilung gleichbedeutend mit

$$F(c - y) = 1 - F(c + y) \quad \text{für alle } y \geq 0. \quad (2.6)$$

Durch Differenzieren von (2.6) folgt

$$f(c - y) = f(c + y) \quad \text{für alle } y \geq 0. \quad (2.7)$$

Umgekehrt kann man leicht durch Integration zeigen, dass aus (2.7) wiederum (2.6) folgt, beide Bedingungen also in äquivalenter Weise die Symmetrie einer stetigen Zufallsvariablen charakterisieren.

*Beispiel 2.9:* Zwei Beispiele für Dichten von symmetrischen Verteilungen findet man in Abbildung 2.4 und 2.7. Die in Abbildung 2.4 dargestellte Dichte ist offenbar symmetrisch zu  $c = 0$ , während die Dichte in Abbildung 2.7 symmetrisch zu  $c = 180$  ist; vgl. auch Beispiel 2.6. Die erste Dichte gehört zur Klasse der Normalverteilungen, die zweite zur Klasse der Rechteckverteilungen. Beides sind wichtige Klassen von symmetrischen Verteilungen, die wir in den Abschnitten 2.4.1 und 2.4.4 ausführlich behandeln werden.

## 2.2 Verteilungsparameter

In diesem Abschnitt wollen wir einige Größen einführen, die die Verteilung einer Zufallsvariablen im Hinblick auf ihre Lage, Streuung und Schiefe beschreiben. Sie werden **Verteilungsparameter** genannt, da sie einerseits von der Verteilung abhängen, andererseits die Verteilung unter bestimmten Aspekten charakterisieren.

### 2.2.1 Erwartungswert

Als Erstes soll die allgemeine Lage einer Zufallsvariablen  $X$  beschrieben werden. Ist  $X$  diskret, so ist die Verteilung von  $X$  durch die Gesamtheit der möglichen Werte  $x_j$  und ihrer Wahrscheinlichkeiten  $p_j$  gegeben. Ist  $X$  stetig, so ist die Verteilung durch die Dichte  $f(x)$  bestimmt. Der **Erwartungswert** der Verteilung von  $X$  ist folgendermaßen definiert<sup>3</sup>:

$$E[X] = \begin{cases} \sum_j x_j p_j, & \text{falls } X \text{ diskret,} \\ \int_{-\infty}^{\infty} x f(x) dx, & \text{falls } X \text{ stetig.} \end{cases}$$

Man schreibt auch  $E[X] = \mu_X$ .

Bei einer diskreten Zufallsvariablen  $X$  werden also alle Werte  $x_j$  aus dem Träger von  $X$  mit den zugehörigen Wahrscheinlichkeiten  $p_j = P(X = x_j)$  gewichtet und aufsummiert, bei einer stetigen Zufallsvariablen tritt an die Stelle der Summe ein Integral.

<sup>3</sup>Es gibt Zufallsvariable (mit unendlichem Träger), bei denen die den Erwartungswert definierende Summe bzw. das Integral einen unendlichen Wert ergeben oder überhaupt nicht konvergieren. Zum Beispiel besitzt die stetige Zufallsvariable mit der Dichte  $f(x) = \frac{1}{\pi(1+x^2)}$ ,  $x \in \mathbb{R}$ , keinen Erwartungswert, da das uneigentliche Integral  $\int_{-\infty}^{\infty} x \frac{1}{\pi(1+x^2)} dx$  nicht konvergiert. Weitere Beispiele liefern die Pareto-Verteilungen in Abschnitt 2.4.3.

*Beispiel 2.2 (Fortsetzung): Beim Roulette erhält man*

- den dreifachen Einsatz, wenn man auf *Colonne* gesetzt hat (und die gesetzte *Colonne* erscheint),
- den 36-fachen Einsatz, wenn man auf *Zahl* gesetzt hat (und die gesetzte *Zahl* erscheint).

*Die Nettoauszahlung ist in beiden Fällen eine diskrete Zufallsvariable,*

$$X_{\text{Colonne}} = \begin{cases} 2, & \text{falls die Colonne erscheint,} \\ -1, & \text{falls nicht,} \end{cases}$$

*bzw.*

$$X_{\text{Zahl}} = \begin{cases} 35, & \text{falls die Zahl erscheint,} \\ -1, & \text{falls nicht.} \end{cases}$$

*Für die Erwartungswerte gilt*

$$E[X_{\text{Colonne}}] = (-1) \cdot \frac{25}{37} + 2 \cdot \frac{12}{37} = -\frac{1}{37},$$

$$E[X_{\text{Zahl}}] = (-1) \cdot \frac{36}{37} + 35 \cdot \frac{1}{37} = -\frac{1}{37}.$$

Offenbar stellt der Erwartungswert einer Zufallsvariablen  $X$  nicht immer einen möglichen Wert von  $X$  dar. So nimmt im Beispiel die Zufallsvariable  $X_{\text{Colonne}}$  nur die beiden Werte  $-1$  und  $2$  an. Ihr Erwartungswert ist jedoch  $E[X_{\text{Colonne}}] = -\frac{1}{37}$ ; er gehört nicht zum Träger von  $X_{\text{Colonne}}$ .

Ist  $X$  eine stetige Zufallsvariable, so gilt  $P(X = x) = 0$  für alle  $x \in \mathbb{R}$ . Die Berechnung des Erwartungswerts kann deshalb nicht wie bei diskreten Zufallsvariablen erfolgen. Vielmehr wird bei stetigen Zufallsvariablen nach der obigen Formel jeder Wert  $x$  aus dem Träger mit der Dichte  $f(x)$  gewichtet und dann integriert. Dabei kann man den Integrationsbereich auf den Träger beschränken, da außerhalb des Trägers die Dichte gleich null ist.

*Beispiel 2.6 (Fortsetzung): Die Zufallsvariable  $X$  hat die Dichte  $f(x) = \frac{1}{360}$ , falls  $0 \leq x \leq 360$ ,  $f(x) = 0$  sonst. Ihr Erwartungswert ist*

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_0^{360} x \frac{1}{360} dx = 180.$$

**Erwartungswert einer transformierten Zufallsvariablen** Häufig benötigt man nicht den Erwartungswert von  $X$  selbst, sondern den einer transformierten Zufallsvariablen  $Y = g(X)$ . Die Transformation kann beispielsweise eine affin-lineare Funktion  $Y = a + bX$ , eine Potenz wie  $Y = X^2$  oder



$Y = X^3$  oder eine Exponentialfunktion  $Y = \exp X$  sein. Für eine beliebige Funktion  $g: \mathbb{R} \rightarrow \mathbb{R}$  berechnet man den Erwartungswert der transformierten Zufallsvariablen  $Y = g(X)$  mit der Formel

$$E[g(X)] = \begin{cases} \sum_j g(x_j) p_j, & \text{falls } X \text{ diskret,} \\ \int_{-\infty}^{\infty} g(x) f_X(x) dx, & \text{falls } X \text{ stetig.} \end{cases}$$

*Beispiel 2.2 (Fortsetzung):* Wir betrachten die Zufallsvariablen  $X_{\text{Colonne}}$  und  $X_{\text{Zahl}}$  sowie die Funktion  $g(x) = x^2$ . Dann ist

$$\begin{aligned} E[X_{\text{Colonne}}^2] &= (-1)^2 \cdot \frac{25}{37} + 2^2 \cdot \frac{12}{37} = \frac{73}{37} = 1.97, \\ E[X_{\text{Zahl}}^2] &= (-1)^2 \cdot \frac{36}{37} + 35^2 \cdot \frac{1}{37} = \frac{1261}{37} = 34.08. \end{aligned}$$

*Beispiel 2.6 (Fortsetzung):*

a) Für die Zufallsvariable  $X$  mit Dichte  $f_X(x) = \frac{1}{360}$ , falls  $0 \leq x \leq 360$ , gilt

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_0^{360} x^2 \frac{1}{360} dx = \frac{1}{360} \frac{1}{3} x^3 \Big|_0^{360} = 43200.$$

b) Nachdem der Kandidat eine weitere Runde erreicht hat, bietet ihm der Showmaster das folgende Spiel an: Wenn das Glücksrad in der Stellung  $x$  stehen bleibt, erhält der Kandidat

$$1000 \cdot \exp\left(\frac{x}{360}\right) \text{ €}$$

ausgezahlt. Wie hoch ist der Erwartungswert der Auszahlung?

Der Erwartungswert beträgt

$$\begin{aligned} E\left[1000 \cdot \exp\left(\frac{X}{360}\right)\right] &= \int_0^{360} 1000 \cdot \exp\left(\frac{x}{360}\right) \frac{1}{360} dx \\ &= 1000 \frac{1}{360} 360 \cdot \exp\left(\frac{x}{360}\right) \Big|_0^{360} \\ &= 1000 e^1 - 1000 = 1718.28 \text{ €}. \end{aligned}$$

**Rechnen mit dem Erwartungswert** Transformiert man die Werte der Zufallsvariablen  $X$  mit einer affin-linearen Abbildung  $x \mapsto a + bx$ , so erhält

man eine neue Zufallsvariable  $Y$  mit den Werten  $Y(\omega) = a + bX(\omega)$ ,  $\omega \in \Omega$ . Der Erwartungswert der transformierten Zufallsvariablen  $Y = a + bX$  ist dann gleich

$$\boxed{E[a + bX] = a + bE[X]}. \quad (2.8)$$

Anschaulich besagt die Formel (2.8), dass die Erwartungswertbildung mit der Addition einer Konstanten und der Multiplikation einer Zahl „vertauscht“ werden kann.

**BEWEIS** Im Fall einer stetigen Variablen  $X$  mit Dichte  $f$  gilt

$$\begin{aligned} E[a + bX] &= \int_{-\infty}^{\infty} (a + bx) f(x) dx = \int_{-\infty}^{\infty} a f(x) dx + \int_{-\infty}^{\infty} bx f(x) dx \\ &= a \int_{-\infty}^{\infty} f(x) dx + b \int_{-\infty}^{\infty} x f(x) dx = a \cdot 1 + b E[X]. \end{aligned}$$

Falls  $X$  diskret ist, ist auch  $a + bX$  eine diskrete Zufallsvariable und es gilt

$$\begin{aligned} E[a + bX] &= \sum_j (a + bx_j) p_j = \sum_j a p_j + \sum_j bx_j p_j \\ &= a \sum_j p_j + b \sum_j x_j p_j = a + b \sum_j x_j p_j \\ &= a + b E[X]. \quad \square \end{aligned}$$

*Beispiel 2.10:* Die Zufallsvariable  $X$  gebe das Bruttomonatsgehalt eines zufällig ausgewählten Angestellten einer Unternehmensholding wieder. Von  $X$  sei nur der Erwartungswert  $\mu_X = 4700$  € bekannt. Tarifverhandlungen ergaben eine Lohnsteigerung um einen Sockelbetrag von 300 € und eine zusätzliche lineare Erhöhung von 3%. Das Bruttogehalt nach der Tarifierhöhung beträgt  $Y = 300 + 1.03 X$ . Sein Erwartungswert beträgt dann

$$\begin{aligned} E[Y] &= E[300 + 1.03 X] = 300 + 1.03 E[X] \\ &= 300 + 1.03 \cdot 4700 = 5141 \text{ €}. \end{aligned}$$

Es ist hier nicht erforderlich, die Wahrscheinlichkeitsfunktion der Zufallsvariablen  $Y$  zu kennen.

**Zentrierung** Zieht man von einer Zufallsvariablen  $X$  ihren Erwartungswert ab, so erhält man die **zentrierte Zufallsvariable**  $X - \mu_X$ . Sie hat wegen (2.8) den Erwartungswert

$$E[X - \mu_X] = \mu_X - \mu_X = 0.$$

Wenn  $X$  **symmetrisch** zu einem bekannten Wert  $c$  verteilt ist und der Erwartungswert existiert, gilt

$$E[X] = c.$$

**BEWEIS** Die Symmetrie der Verteilung besagt (s.o. Abschnitt 2.1.7), dass  $X - c$  wie  $-X + c$  verteilt ist, also auch den gleichen Erwartungswert besitzt. Folglich ist

$$\begin{aligned} E[X - c] &= E[-X + c], \\ E[X] - c &= -E[X] + c, \\ 2E[X] &= 2c, \end{aligned}$$

woraus die Behauptung  $E[X] = c$  folgt.  $\square$

**Erwartungswert als Lageparameter** Nicht nur für symmetrische Verteilungen charakterisiert der Erwartungswert die Lage einer Zufallsvariablen. Wegen (2.8) gilt nämlich: Wird die Zufallsvariable  $X$  um die Konstante  $a$  auf der reellen Achse verschoben, so verschiebt sich der Erwartungswert um die gleiche Konstante,  $E[a + X] = a + E[X]$ . Ein Verteilungsparameter, der diese Eigenschaft besitzt, heißt **lageäquvariant**. Wird  $X$  mit einem Skalenfaktor  $b > 0$  multipliziert, so ist  $E[bX] = bE[X]$ . Ein Verteilungsparameter, der sich um den gleichen Skalenfaktor ändert, wird **skalenäquvariant** genannt.

Einen Verteilungsparameter, der lage- und skalenäquvariant ist, nennt man einen **Lageparameter**. Demnach ist der Erwartungswert ein Lageparameter. Wegen Gleichung (2.4) ist auch jedes Quantil  $x_p$  von  $X$  ein Lageparameter, insbesondere der Median  $x_{0.5}$ .

Erwartungswert und Median sind nicht nur Lageparameter, sie beschreiben auch – in unterschiedlicher Weise – ein „Zentrum“ der Verteilung. Wenn die Verteilung symmetrisch zu einem Wert  $c$  ist, sind beide gleich diesem Wert  $c$ . Im Allgemeinen sind Erwartungswert und Median jedoch verschieden.

Weitere Zentralitätseigenschaften des Erwartungswerts hängen mit der Varianz der Zufallsvariablen zusammen. Sie werden im folgenden Abschnitt behandelt.

## 2.2.2 Varianz

Neben der Lage einer Zufallsvariablen, die durch den Erwartungswert charakterisiert wird, ist ihre Streuung von besonderem Interesse. Sie wird durch die Varianz beschrieben. Die **Varianz** von  $X$ ,

$$V[X] = E[(X - \mu_X)^2],$$

ist der Erwartungswert der Zufallsvariablen  $g(X) = (X - \mu_X)^2$ , also der Erwartungswert der quadrierten Abweichung der Zufallsvariablen  $X$  von ihrem

Erwartungswert  $\mu_X$ . Für diskrete bzw. stetige Verteilungen ist die Varianz durch die Formeln

$$V[X] = \begin{cases} \sum_j (x_j - \mu_X)^2 p_j, & \text{falls } X \text{ diskret,} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx, & \text{falls } X \text{ stetig,} \end{cases}$$

gegeben. Offenbar kann die Varianz keinen negativen Wert annehmen.<sup>4</sup> Die positive Wurzel aus der Varianz,

$$\sigma[X] = +\sqrt{V[X]},$$

heißt **Standardabweichung**. Weitere Notationen für Standardabweichung und Varianz sind

$$\sigma_X = \sigma[X], \quad \sigma_X^2 = \sigma^2[X] = \text{Var}[X] = V[X].$$

Zur Berechnung der Varianz ist die folgende Formel nützlich:

$$V[X] = E[X^2] - \mu_X^2 \quad (2.9)$$

In (2.9) genügt es,  $\mu_X$  sowie  $E[X^2]$  zu berechnen, was meistens weniger Aufwand als die Berechnung mit Hilfe der obigen Definitionsformel erfordert.

Der Beweis von (2.9) folgt aus

$$\begin{aligned} E[(X - \mu_X)^2] &= E[X^2] - 2E[X \cdot \mu_X] + \mu_X^2 \\ &= E[X^2] - 2E[X] \cdot \mu_X + \mu_X^2 \\ &= E[X^2] - \mu_X^2. \end{aligned} \quad \square$$

*Beispiel 2.2 (Fortsetzung):* Wir betrachten wieder beim Roulette die Zufallsvariablen  $X_{\text{Colonne}}$  und  $X_{\text{Zahl}}$ . Beide haben den gleichen Erwartungswert  $-\frac{1}{37}$ . Für die Varianzen ergibt sich

$$\begin{aligned} V[X_{\text{Colonne}}] &= E[X_{\text{Colonne}}^2] - (E[X_{\text{Colonne}}])^2 \\ &= \frac{73}{37} - \left(-\frac{1}{37}\right)^2 = 1.9722, \\ \sigma[X_{\text{Colonne}}] &= 1.40, \\ V[X_{\text{Zahl}}] &= E[X_{\text{Zahl}}^2] - (E[X_{\text{Zahl}}])^2 \\ &= \frac{1261}{37} - \left(-\frac{1}{37}\right)^2 = 34.0804, \\ \sigma[X_{\text{Zahl}}] &= 5.84. \end{aligned}$$

<sup>4</sup>Es gibt allerdings Zufallsvariable, bei denen diese Summe bzw. dieses Integral unendlich groß werden, also keine endliche Varianz existiert. Beispiele sind die  $t$ -verteilten Variablen mit  $\nu \leq 2$  Freiheitsgraden (Abschnitt 4.3.2).

In diesem Beispiel kann man die Varianzen (und Standardabweichungen) als Risikomaße interpretieren. Offensichtlich birgt „Setzen auf Zahl“ mehr Risiko (im Sinne möglicher Schwankungen der Auszahlung nach oben wie nach unten!) als „Setzen auf Colonne“.

*Beispiel 2.6 (Fortsetzung): Die Varianz des Ergebnisses  $X$  berechnet man mit der Formel (2.9) als*

$$V[X] = E[X^2] - \mu_X^2 = 43\,200 - 180^2 = 10\,800.$$

Die Standardabweichung ist dann  $\sigma[X] = \sqrt{10\,800} = 103.92$ .

Transformiert man  $X$  zur Zufallsvariablen  $Y = a + bX$ ,  $a, b \in \mathbb{R}$ , so ändern sich Varianz und Standardabweichung wie folgt:

$$\boxed{\begin{array}{l} V[a + bX] = b^2 V[X], \\ \sigma[a + bX] = |b| \sigma[X]. \end{array}} \quad (2.10)$$

Dies folgt sofort aus den Definitionen und der Lage- und Skalenäquivarianz des Erwartungswerts:

$$\begin{aligned} V[Y] &= E[(Y - \mu_Y)^2] = E[(b(X - \mu_X))^2] \\ &= b^2 E[(X - \mu_X)^2] = b^2 V[X] \end{aligned}$$

Varianz und Standardabweichung einer Zufallsvariablen werden also durch die Addition einer Konstanten nicht beeinflusst; beide Verteilungsparameter sind **lageinvariant**. Sie ändern sich, wenn  $X$  mit einem positiven Skalenfaktor multipliziert wird. Die Varianz multipliziert sich dann mit dem Quadrat des Faktors, während sich die Standardabweichung mit dem Faktor selbst multipliziert. Beide Verteilungsparameter messen demzufolge die Streuung von  $X$ . Die Standardabweichung ist **skalenäquvariant**, die Varianz ist es jedoch nicht.

*Beispiel 2.10 (Fortsetzung): Das Bruttomonatsgehalt  $X$  habe die Standardabweichung  $\sigma_X = 950$  €. Nach einer allgemeinen Gehaltserhöhung um einen Sockelbetrag von 300 € und 3% linear beträgt das Gehalt  $Y = 300 + 1.03 X$ . Für die Standardabweichung von  $Y$  gilt dann*

$$\begin{aligned} \sigma_Y &= \sqrt{V[300 + 1.03 X]} = 1.03 \sigma_X \\ &= 1.03 \cdot 950 = 978.5 \text{ €}. \end{aligned}$$

Generell sind Varianz und Standardabweichung größer oder gleich null. Von Interesse ist der Fall der kleinsten Streuung, wenn also die Varianz (und damit auch die Standardabweichung) gleich null ist. Es lässt sich zeigen, dass

$$V[X] = 0 \quad \iff \quad P(X = \mu_X) = 1$$

gilt, die Varianz von  $X$  also genau dann verschwindet, wenn  $X$  (mit Wahrscheinlichkeit 1) konstant gleich seinem Erwartungswert ist.

Die nachfolgende Minimaleigenschaft des Erwartungswerts rechtfertigt seine Interpretation als **Schwerpunkt** der Verteilung. Es ist

$$\min_{a \in \mathbb{R}} E[(X - a)^2] = E[(X - \mu_X)^2], \quad (2.11)$$

d.h. der Erwartungswert ist der Punkt, von dem die mittlere quadratische Abweichung der Zufallsvariablen minimal ist. Die Streuung der Verteilung um einen Punkt  $a$  ist also am kleinsten für  $a = \mu_X$ . Der Begriff „Schwerpunkt“ kommt aus der Physik; er erklärt sich wie folgt: Wenn man die Wahrscheinlichkeiten  $p_j$  einer diskreten Verteilung als Massen an den zugehörigen Punkten  $x_j$  der reellen Achse befestigt, so entspricht der Erwartungswert der Verteilung gerade dem physikalischen Schwerpunkt dieser Punktmassen.

**BEWEIS** Der Beweis von (2.11) folgt aus

$$\begin{aligned} E[(X - a)^2] &= E[((X - \mu_X) + (\mu_X - a))^2] \\ &= E[(X - \mu_X)^2] + \\ &\quad \underbrace{2 E[(X - \mu_X)(\mu_X - a)]}_{=0} + (\mu_X - a)^2. \end{aligned}$$

Die rechte Seite (und damit auch die linke Seite) ist offensichtlich minimal, wenn man  $a = \mu_X$  wählt.  $\square$

### 2.2.3 Ungleichung von Tschebyscheff

Erwartungswert  $\mu_X$  und Varianz  $\sigma_X^2$  sind Parameter der Lage bzw. der Streuung einer Zufallsvariablen  $X$ . Wir geben eine von  $\sigma_X^2$  abhängige Mindestwahrscheinlichkeit dafür an, dass  $X$  in ein Intervall vorgegebener Breite fällt, das seinen Erwartungswert symmetrisch umgibt.

Die **Ungleichung von Tschebyscheff** besagt, dass für jede positive Zahl  $\varepsilon$

$$P(|X - \mu_X| < \varepsilon) \geq 1 - \frac{\sigma_X^2}{\varepsilon^2} \quad (2.12)$$

gilt. Dies heißt, die Wahrscheinlichkeit für das Ereignis

$$\{\mu_X - \varepsilon < X < \mu_X + \varepsilon\}$$

lässt sich durch die Schranke  $1 - (\sigma_X^2/\varepsilon^2)$  nach unten abschätzen. Die Wahrscheinlichkeit, dass  $X$  in das Intervall  $]\mu_X - \varepsilon, \mu_X + \varepsilon[$  fällt, ist also umso

größer, je kleiner  $\sigma_X^2$  ist. Offenbar ist die Ungleichung nur für  $\varepsilon^2 \geq \sigma_X^2$  interessant, da sonst die rechte Seite negativ wird.

Auch die Wahrscheinlichkeit des Komplementärereignisses, dass  $X$  nicht in das angegebene Intervall fällt, lässt sich mit der Tschebyscheff-Ungleichung abschätzen, und zwar nach oben wie folgt:

$$P(|X - \mu_X| \geq \varepsilon) \leq \frac{\sigma_X^2}{\varepsilon^2}. \quad (2.13)$$

**BEWEIS** Die Ungleichung (2.13) ist zur Ungleichung (2.12) äquivalent. Um Ungleichung (2.13) einzusehen, definieren wir eine neue Zufallsvariable  $Z$ , die durch

$$Z(\omega) = \begin{cases} \varepsilon^2, & \text{falls } |X(\omega) - \mu_X| \geq \varepsilon, \\ 0 & \text{sonst,} \end{cases}$$

gegeben ist. Offenbar gilt

$$\begin{aligned} Z(\omega) &\leq (X(\omega) - \mu_X)^2, \\ E[Z] &\leq E[(X - \mu_X)^2] = \sigma_X^2, \\ E[Z] &= \varepsilon^2 \cdot P(|X - \mu_X| \geq \varepsilon). \end{aligned}$$

Also ist

$$P(|X - \mu_X| \geq \varepsilon) \leq \frac{\sigma_X^2}{\varepsilon^2},$$

wie behauptet. □

*Beispiel 2.11:* Sei  $X$  die Verweildauer (in Tagen) eines zufällig ausgewählten Patienten in einem Großklinikum. Erwartungswert  $\mu_X = 10$  und Standardabweichung  $\sigma_X = 4$  seien bekannt. Wie groß ist mindestens die Wahrscheinlichkeit, dass der Patient mehr als fünf, aber weniger als fünfzehn Tage im Krankenhaus verbringt? Es ist

$$P(5 < X < 15) = P(|X - 10| < 5) \geq 1 - \frac{4^2}{5^2} = \frac{9}{25}.$$

Setzt man  $\varepsilon = \lambda\sigma_X$  (mit einem  $\lambda > 1$ ), so ergeben sich zwei äquivalente Versionen der Tschebyscheff-Ungleichung,

$$\begin{aligned} P(|X - \mu_X| < \lambda\sigma_X) &\geq 1 - \frac{1}{\lambda^2} && \text{bzw.} \\ P(|X - \mu_X| \geq \lambda\sigma_X) &\leq \frac{1}{\lambda^2}. \end{aligned}$$

Die zweite Ungleichung erhält man aus der ersten, indem man zum Komplementärereignis übergeht. Offenbar kann man in der ersten Ungleichung das „<“-Zeichen durch „ $\leq$ “ ersetzen, da sich hierdurch die Wahrscheinlichkeit nicht verringert.

Für  $\lambda = 2$  und  $\lambda = 3$  erhalten wir Abschätzungen für die Wahrscheinlichkeit, dass  $X$  in den so genannten **Zwei-** bzw. **Drei-Sigma-Bereich** fällt, nämlich

$$P\left(\underbrace{\mu_X - 2\sigma_X < X < \mu_X + 2\sigma_X}_{\text{Zwei-Sigma-Bereich}}\right) \geq \frac{3}{4},$$

$$P\left(\underbrace{\mu_X - 3\sigma_X < X < \mu_X + 3\sigma_X}_{\text{Drei-Sigma-Bereich}}\right) \geq \frac{8}{9}.$$

Dies sind sehr grobe Abschätzungen, die für beliebige Verteilungen gelten. Sie lassen sich verfeinern, wenn weitere Informationen über die Verteilung vorliegen, etwa wenn man voraussetzen kann, dass  $X$  normalverteilt ist (vgl. Abschnitt 2.4.4).

*Beispiel 2.12:* Die Verteilung von  $X$  beschreibe die Verteilung des Bruttomonatseinkommens in einer Gemeinde (in €). Es sei  $\mu_X = 2500$  und  $\sigma_X = 750$  bekannt. Dann liegen mindestens 75% der Einkommen zwischen 1000 € und 4000 €. Mindestens  $8/9 \cdot 100\%$  der Einkommen liegen zwischen 250 € und 4750 €.

## 2.2.4 Schiefe und Wölbung

Neben Maßzahlen zur Charakterisierung der Lage und Streuung der Verteilung einer Zufallsvariablen  $X$  gibt es noch solche, die die Form der Verteilung charakterisieren, so genannte **Formparameter**, auch **Gestaltparameter** genannt.

**Schiefe** Von Bedeutung ist hier zunächst die Abweichung der Verteilung von einer symmetrischen Verteilung gleicher Lage und Streuung. Im Folgenden betrachten wir zwei Maße für diese Abweichung. Als **Momentenschiefe**, kurz: **Schiefe**, definiert man

$$\gamma_1[X] = E\left[\left(\frac{X - \mu_X}{\sigma_X}\right)^3\right] = \frac{E[(X - \mu_X)^3]}{\sigma_X^3}.$$

Es gilt

$$\begin{aligned} \gamma_1[X] &= \frac{E[X^3] - 3E[X^2]\mu_X + 3E[X]\mu_X^2 - \mu_X^3}{\sigma_X^3} \\ &= \frac{E[X^3] - 3E[X^2]\mu_X + 2\mu_X^3}{\sigma_X^3}. \end{aligned}$$



Ist die Verteilung von  $X$  symmetrisch zu  $c$ , so gilt  $\mu_X = c$  und man sieht leicht, dass die Verteilung von  $(X - \mu_X)^3$  und  $-(X - \mu_X)^3$  übereinstimmt. Daraus folgt, dass bei einer symmetrischen Verteilung  $E[(X - \mu_X)^3] = 0$  und damit  $\gamma_1[X] = 0$  ist.

Man beachte, dass die Umkehrung nicht gilt: Aus  $\gamma_1[X] = 0$  kann nicht die Symmetrie der Verteilung von  $X$  gefolgert werden. Ist  $\gamma_1[X] \neq 0$ , so ist die Verteilung asymmetrisch, wobei man

bei  $\gamma_1[X] > 0$  von **Rechtsschiefe** (d.h. **Linkssteilheit**),  
bei  $\gamma_1[X] < 0$  von **Linksschiefe** (d.h. **Rechtssteilheit**)

spricht. Im ersten Fall überwiegen größere positive Werte von  $X - \mu_X$ , im zweiten Fall größere negative Werte; vgl. Abbildung 2.11 und 2.12.  $\gamma_1[X]$  ist nicht normiert, sondern kann prinzipiell beliebige Werte annehmen.

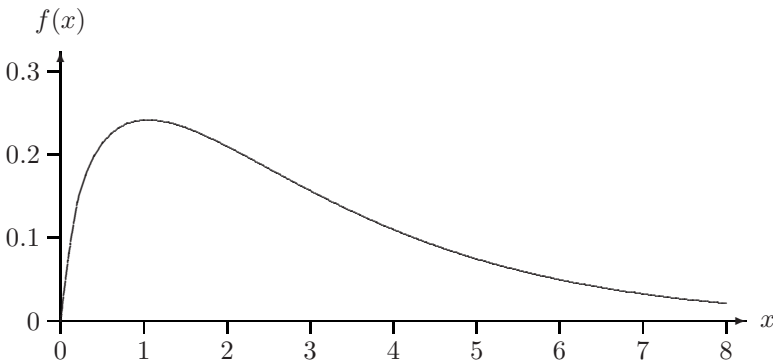


Abbildung 2.11: Dichtefunktion einer stetigen rechtsschiefen Verteilung

Im Gegensatz zur Momentenschiefe ist die **Quartilsschiefe**  $\gamma_1^Q[X]$  beschränkt. Ihre Definition lautet

$$\begin{aligned} \gamma_1^Q[X] &= \frac{[Q(0.75) - Q(0.5)] - [Q(0.5) - Q(0.25)]}{[Q(0.75) - Q(0.5)] + [Q(0.5) - Q(0.25)]} \\ &= \frac{Q(0.75) + Q(0.25) - 2 \cdot Q(0.5)}{Q(0.75) - Q(0.25)}, \end{aligned} \quad (2.13)$$

wobei  $Q(p)$  das  $p$ -Quantil der Verteilung von  $X$  bezeichnet. Für beliebige  $X$  gilt

$$-1 \leq \gamma_1^Q[X] \leq 1.$$

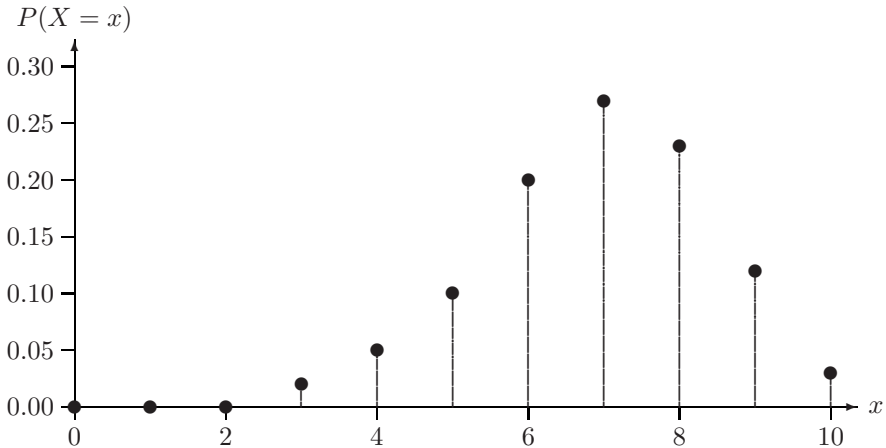


Abbildung 2.12: Wahrscheinlichkeitsfunktion einer diskreten linksschiefen Verteilung

Im Fall einer symmetrischen Verteilung ist  $Q(0.75) - Q(0.5) = Q(0.5) - Q(0.25)$ , also  $\gamma_1^Q[X]$  ebenfalls gleich null. Die Umkehrung gilt auch hier nicht: Aus  $\gamma_1^Q[X] = 0$  kann nicht auf die Symmetrie der Verteilung von  $X$  geschlossen werden. Wenn die Quartilsschiefe ungleich null ist, bezeichnet man die Verteilung als **rechtsschief** (wenn  $\gamma_1^Q[X] > 0$ ) bzw. **linksschief** (wenn  $\gamma_1^Q[X] < 0$ ) im Sinne der Quartilsschiefe.

Wird  $X$  affin-linear transformiert,  $Y = a + bX$ , so ändert sich  $\gamma_1$  offenbar nicht,  $\gamma_1[X] = \gamma_1[Y]$ ; ebenso gilt  $\gamma_1^Q[X] = \gamma_1^Q[Y]$ . Beide Maßzahlen der Schiefe sind daher sowohl lage- als auch skaleninvariant.

**Wölbung** Ein weiterer, häufig untersuchter Gestaltparameter ist die **Wölbung**, die auch als **Kurtosis** bezeichnet wird. Sie ist durch

$$\gamma_2[X] = E \left[ \left( \frac{X - \mu_X}{\sigma_X} \right)^4 \right] = \frac{E[(X - \mu_X)^4]}{\sigma_X^4}$$

definiert. Die Wölbung ist offenbar lage- und skaleninvariant. Für beliebige  $X$  gilt  $\gamma_2[X] \geq 1$ . Der Wert  $\gamma_2[X] = 1$  ergibt sich für eine Zufallsvariable  $X$  mit Zweipunktverteilung mit  $P(X = -1) = P(X = 1) = \frac{1}{2}$ . Bei der Normalverteilung (siehe Abschnitt 2.4.4) gilt  $\gamma_2[X] = 3$ . Die Wölbung einer symmetrischen unimodalen Verteilung misst, wie spitz die Dichte bzw. die Wahrscheinlichkeitsfunktion um den Erwartungswert  $\mu_X$  herum verläuft und wie viel Wahrscheinlichkeitsmasse auf den Flanken der Verteilung liegt.

Symmetrische unimodale Verteilungen mit einer Wölbung größer als 3 sind also um  $\mu_X$  herum spitzer und haben mehr Masse auf den Flanken als die Normalverteilung. Die Maßzahl  $\gamma_2 - 3$  wird als **Exzess** bezeichnet.

Der Parameter  $\gamma_2$  sollte nur für solche Verteilungen berechnet werden, die symmetrisch oder angenähert symmetrisch und unimodal sind, da die obige Interpretation bei anderen Verteilungen nicht zutrifft.

Beispiele zur Berechnung der Momentenschiefe, der Quartilsschiefe und der Wölbung finden sich in den Abschnitten 2.4.2 und 2.4.4 über spezielle stetige Verteilungen. Die Übersicht in Abschnitt 2.4.6 enthält die Momentenschiefe und die Wölbung von einigen weiteren, häufig angewandten Wahrscheinlichkeitsverteilungen.

## 2.3 Spezielle diskrete Verteilungen

In diesem Abschnitt behandeln wir einige diskrete Verteilungen, die zum Kernbestand der Wahrscheinlichkeitsrechnung gehören, da sie bei den unterschiedlichsten Anwendungen auftreten.

Die drei zunächst beschriebenen Verteilungen treten im Zusammenhang mit der Bernoulli-Versuchsreihe auf. Ein Zufallsexperiment, bei dem man sich nur dafür interessiert, ob ein Ereignis  $A$  eintritt oder nicht, nennt man ein **Bernoulli-Experiment**. Das Eintreten von  $A$  wird auch als **Erfolg**, das Eintreten von  $\bar{A}$  als **Misserfolg** bezeichnet. Eine **Bernoulli-Versuchsreihe** ist die  $n$ -malige unabhängige Durchführung eines Bernoulli-Experiments; vgl. Abschnitt 1.3.4. Sei  $A_i$  das Ereignis, dass beim  $i$ -ten Experiment ein Erfolg eintritt, und sei  $\pi = P(A_i)$  für  $i = 1, 2, \dots, n$ . Die Ereignisse  $A_1, A_2, \dots, A_n$  werden als global unabhängig angenommen. Die Indikatorvariable

$$\mathbf{1}_{A_i} = \begin{cases} 1, & \text{falls } A_i \text{ eintritt („Erfolg“),} \\ 0, & \text{falls } A_i \text{ nicht eintritt („Misserfolg“),} \end{cases}$$

ist eine Zufallsvariable, die anzeigt, ob  $A_i$  eintritt oder nicht.

*Beispiel 2.13:* Ein fairer Würfel werde  $n$ -mal geworfen. Sei  $A$  das Ereignis, eine „6“ zu würfeln,  $A_i$  der Erfolg beim  $i$ -ten Wurf. Wenn die Ereignisse  $A_i$  global unabhängig sind, handelt es sich um eine Bernoulli-Versuchsreihe. Man kann zeigen, dass die  $A_i$  genau dann global unabhängig sind, wenn keine der möglichen Folgen von Augenzahlen bevorzugt auftritt, d.h. jede die gleiche Wahrscheinlichkeit besitzt; vgl. Beispiel 1.26.

Ein Grundmodell, an dem viele Begriffe der Wahrscheinlichkeitsrechnung erläutert werden können, ist das folgende **Urnenmodell**: Eine Urne ist ein

mit Kugeln gefüllter Behälter, aus dem einzelne Kugeln in zufälliger Weise gezogen werden. Die Ziehungswahrscheinlichkeit ist für jede Kugel die gleiche.

Speziell enthalte eine Urne  $N$  Kugeln, von denen  $M$  rot und  $N - M$  weiß sind. Nacheinander werden insgesamt  $n$  Kugeln aus der Urne gezogen. Sei  $A_i$  das Ereignis, dass die  $i$ -te Ziehung eine rote Kugel erbringt. Jede einzelne solche Ziehung ist offenbar ein Bernoulli-Experiment.

Beim **Ziehen mit Zurücklegen** wird die gezogene Kugel jedes Mal wieder in die Urne gelegt und mit den übrigen Kugeln gemischt, so dass jede weitere Ziehung das gleiche Zufallsexperiment wie die erste Ziehung darstellt. Daher können die Ereignisse  $A_1, \dots, A_n$  als global unabhängig angesehen werden. Das Ziehen mit Zurücklegen ist deshalb eine Bernoulli-Versuchsreihe mit

$$\pi = P(A_i) = \frac{M}{N} \quad \text{für alle } i = 1, 2, \dots, n.$$

Im Unterschied dazu besteht das **Ziehen ohne Zurücklegen** aus  $n$  Ziehungen, bei denen eine Kugel gezogen, aber nicht wieder in die Urne zurückgelegt wird. Da der Anteil der roten Kugeln in der Urne bei einer Ziehung davon abhängt, wie viele rote Kugeln in den vorhergehenden Ziehungen gezogen wurden, sind die Ereignisse  $A_1, \dots, A_n$  *nicht* unabhängig, und das Ziehen ohne Zurücklegen ist *keine* Bernoulli-Versuchsreihe.

### 2.3.1 Binomialverteilung

Wir gehen von einer Bernoulli-Versuchsreihe aus und betrachten die Zufallsvariable

$$X = \sum_{i=1}^n \mathbf{1}_{A_i}.$$

Offenbar kann  $X$  die Werte  $0, 1, 2, \dots, n$  annehmen. Die Zufallsvariable  $X$  zählt die eingetretenen Ereignisse  $A_i$ , sie gibt an, wie oft bei der  $n$ -maligen unabhängigen Durchführung des Bernoulli-Experiments ein Erfolg aufgetreten ist.

*Beispiel 2.14:* Für eine Bernoulli-Versuchsreihe der Länge  $n = 4$  gebe man alle möglichen Folgen von Ergebnissen an, bei denen genau zwei Erfolge eintreten. Dies sind offensichtlich die Folgen:

$$\begin{aligned}
 &A_1, A_2, \overline{A_3}, \overline{A_4} \\
 &A_1, \overline{A_2}, A_3, \overline{A_4} \\
 &A_1, \overline{A_2}, \overline{A_3}, A_4 \\
 &\overline{A_1}, A_2, A_3, \overline{A_4} \\
 &\overline{A_1}, A_2, \overline{A_3}, A_4 \\
 &\overline{A_1}, \overline{A_2}, A_3, A_4
 \end{aligned}$$

Es sind  $\binom{4}{2} = 6$  verschiedene Folgen.

Bei einer Bernoulli-Versuchsreihe der Länge  $n$  gibt es  $\binom{n}{k}$  Folgen, bei denen genau  $k$ -mal ein Erfolg eintritt. Da zu  $k$  Erfolgen jeweils  $n - k$  Misserfolge gehören, ist die Wahrscheinlichkeit jeder dieser Folgen (wegen der Unabhängigkeit der Wiederholungen) gleich

$$\pi^k (1 - \pi)^{n-k}.$$

Daraus folgt

$$P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \quad \text{für } x = 0, 1, \dots, n. \quad (2.14)$$

Eine diskrete Zufallsvariable  $X$  mit dieser Wahrscheinlichkeitsfunktion heißt **binomialverteilt** mit den Parametern  $n$  und  $\pi$ , in Zeichen:  $X \sim B(n, \pi)$ . Die Verteilung  $B(n, \pi)$  von  $X$  heißt **Binomialverteilung**, ihr Träger ist  $T_X = \{0, 1, \dots, n\}$ . Abbildung 2.12 zeigt die Wahrscheinlichkeitsfunktion der Binomialverteilung mit  $n = 10$  und  $\pi = 0.7$ .

Die Gesamtheit der Binomialverteilungen  $B(n, \pi)$  mit  $n \in \mathbb{N}$  und  $\pi \in ]0, 1[$  bildet eine **Verteilungsfamilie**, in der die einzelne Verteilung durch die beiden **Parameter**  $n$  und  $\pi$  charakterisiert ist.

Im Spezialfall  $n = 1$ , d.h., wenn  $X \sim B(1, \pi)$ , heißt  $X$  **Bernoulli-verteilt**. In diesem Fall gilt

$$\begin{aligned}
 P(X = 1) &= \pi, \\
 P(X = 0) &= 1 - \pi.
 \end{aligned}$$

*Beispiel 2.15 (Ziehen mit Zurücklegen): Eine Urne enthält zehn Kugeln, davon drei rote. Es werden drei Kugeln mit Zurücklegen gezogen. Wie groß ist die Wahrscheinlichkeit, mindestens zwei rote Kugeln zu ziehen?*

*Bezeichne  $X$  die Anzahl der roten Kugeln unter den drei gezogenen Kugeln. Dann ist ( $\hookrightarrow$  Calc/SPSS)*

$$X \sim B\left(n = 3, \pi = \frac{3}{10} = 0.3\right)$$

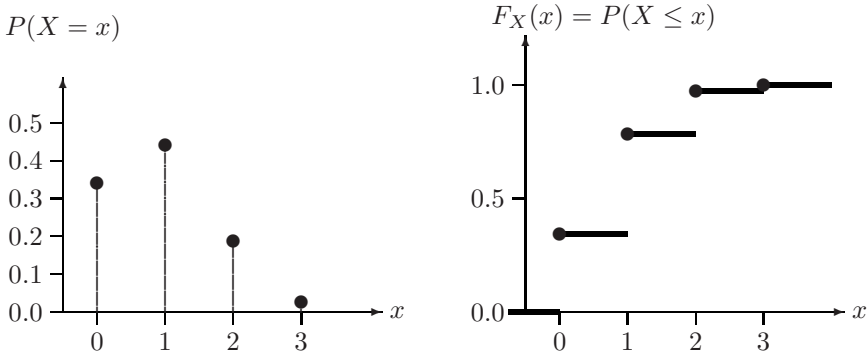


Abbildung 2.13: Wahrscheinlichkeitsfunktion und Verteilungsfunktion von  $B(n=3, \pi=0.3)$  im Beispiel 2.15 (Ziehen mit Zurücklegen)

und

$$\begin{aligned} P(X \geq 2) &= 1 - P(X=0) - P(X=1) \\ &= 1 - \binom{3}{0} \cdot 0.3^0 \cdot 0.7^3 - \binom{3}{1} \cdot 0.3^1 \cdot 0.7^2 = 0.216. \end{aligned}$$

**Erwartungswert und Varianz** einer  $B(1, \pi)$ -verteilten Zufallsvariablen lassen sich leicht direkt ausrechnen. Es ist

$$\begin{aligned} E[X] &= 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi, \\ E[X^2] &= 1^2 \cdot \pi + 0^2(1 - \pi) = \pi, \end{aligned}$$

und deshalb

$$V[X] = E[X^2] - (E[X])^2 = \pi - \pi^2 = \pi(1 - \pi).$$

Für eine  $B(n, \pi)$ -verteilte Zufallsvariable  $X$  gelten die Formeln

$$\boxed{E[X] = n\pi \quad \text{und} \quad V[X] = n\pi(1 - \pi)}. \quad (2.15)$$

Wir beweisen sie hier nicht, da sie in einfacher Weise mit Hilfsmitteln des Kapitels 3 hergeleitet werden können.

**Tabellen** Da sich für größere  $n$  die Wahrscheinlichkeiten der Binomialverteilung nur umständlich „von Hand“ berechnen lassen, sind in Tabelle 1 im Anhang dieses Buchs die Wahrscheinlichkeiten  $P(X=x)$  für  $X \sim B(n, \pi)$  und  $n = 1, 2, \dots, 15, 20, 30$  sowie  $\pi = 0.05, 0.10, \dots, 0.50$  tabelliert. Für  $\pi > 0.5$  benutzt man die Beziehung

$$X \sim B(n, \pi) \quad \Longleftrightarrow \quad n - X \sim B(n, \pi' = 1 - \pi).$$

BEWEIS Wenn  $X$  die Zahl der „Erfolge“  $A_i$  einer Bernoulli-Versuchsreihe beschreibt, dann beschreibt die Zufallsvariable  $Y = n - X$  nämlich die Zahl der „Misserfolge“  $\bar{A}_i$ . Die Unabhängigkeit der Folge  $A_1, \dots, A_n$  zieht die Unabhängigkeit der Folge  $\bar{A}_1, \dots, \bar{A}_n$  nach sich. Es gilt  $P(\bar{A}_i) = 1 - \pi$  für  $i = 1, \dots, n$ . Also handelt es sich bei den  $\bar{A}_i$  ebenfalls um eine Bernoulli-Versuchsreihe, und  $Y = n - X$  ist binomialverteilt mit den Parametern  $n$  und  $\pi' = 1 - \pi$ .  $\square$

*Beispiel 2.16:* Es sei  $X \sim B(n = 10, \pi = 0.8)$ . Man bestimme  $P(X \geq 9)$ . Die gesuchte Wahrscheinlichkeit beträgt

$$\begin{aligned} P(X \geq 9) &= P(X = 9) + P(X = 10) \\ &= P(Y = 1) + P(Y = 0) \\ &= 0.2684 + 0.1074 = 0.3758, \end{aligned}$$

wobei für  $Y = n - X \sim B(n = 10, \pi' = 0.2)$  die Werte der Tabelle 1 verwendet wurden.

### 2.3.2 Poisson-Verteilung

Wir gehen wieder von einer Bernoulli-Versuchsreihe aus, treffen aber spezielle Annahmen über die Parameter  $n$  und  $\pi$ . Wir nehmen an, dass das Bernoulli-Experiment sehr häufig wiederholt wird,  $n$  also groß ist. Andererseits sei die Wahrscheinlichkeit  $\pi = P(A)$  für einen Erfolg sehr klein.  $A$  wird deshalb ein „seltenes Ereignis“ genannt. Die Anzahl  $X$  der Erfolge bei  $n$  Versuchen ist dann binomialverteilt,  $X \sim B(n, \pi)$ , mit

$$P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \quad \text{für } x = 0, 1, 2, \dots, n. \quad (2.16)$$

In diesem Ausdruck lässt man nun

- $n$  gegen unendlich und zugleich
- $\pi$  gegen null gehen, derart, dass  $n\pi$  gegen eine positive Zahl  $\mu$  konvergiert.

Die Binomialwahrscheinlichkeit (2.16) konvergiert bei diesem Grenzübergang, und zwar gegen

$$\lim_{\substack{\pi \rightarrow 0, n \rightarrow \infty \\ n\pi \rightarrow \mu}} \binom{n}{x} \pi^x (1 - \pi)^{n-x} = e^{-\mu} \frac{\mu^x}{x!}. \quad (2.17)$$

Man kann zeigen, dass der Limes für jede nichtnegative ganze Zahl  $x$  definiert und größer als null ist. Für die Exponentialfunktion gilt die Reihendarstellung

$$e^\mu = \exp(\mu) = \sum_{k=0}^{\infty} \frac{\mu^k}{k!}, \quad \mu \in \mathbb{R}.$$

Summiert man die rechte Seite der Gleichung (2.17) über alle  $x = 0, 1, 2, \dots$ , ergibt sich deshalb (mit Summationsindex  $x$  statt  $k$ )

$$\begin{aligned} \sum_{x=0}^{\infty} e^{-\mu} \frac{\mu^x}{x!} &= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \\ &= e^{-\mu} e^{\mu} = e^0 = 1. \end{aligned}$$

Der Limes definiert also eine Wahrscheinlichkeitsfunktion. Eine diskrete Zufallsvariable  $X$  mit der Wahrscheinlichkeitsfunktion

$$P(X = x) = e^{-\mu} \frac{\mu^x}{x!} \quad \text{für } x = 0, 1, 2, \dots,$$

heißt **Poisson-verteilt**<sup>5</sup> mit Parameter  $\mu > 0$ , in Zeichen:  $X \sim Po(\mu)$ . Der Träger von  $X$  ist  $T_X = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$ .

**Tabellen** Die Wahrscheinlichkeiten  $P(X = x)$  lassen sich leicht berechnen. Für Werte von  $\mu$  zwischen 0.1 und 10 sind sie außerdem in Tabelle 2 des Anhangs tabelliert.

Für kleine  $\mu$  zeigt die Poisson-Verteilung eine starke Asymmetrie (Rechtschiefe), während die Verteilung für größere  $\mu$  fast symmetrisch ist. Abbildung 2.14 illustriert dies anhand zweier Poisson-Verteilungen mit  $\mu = 2$  bzw.  $\mu = 12$ .

Viele ökonomische und technische Sachverhalte können mit Poisson-verteilten Zufallsvariablen beschrieben werden. Beispielsweise lässt sich die Anzahl der in einem bestimmten Zeitraum ankommenden Kunden in einem Bedienungssystem häufig auf diese Weise modellieren; vgl. Abschnitt 3.3.2.

Die Poisson-Verteilung  $Po(\mu)$  kann man wegen der erwähnten Limesbeziehung als **Approximation der Binomialverteilung**  $B(n, \pi)$  für kleines  $\pi$  und großes  $n$  auffassen. In diesem Fall gilt die Näherung

$$\binom{n}{x} \pi^x (1 - \pi)^{n-x} \approx e^{-\mu} \frac{\mu^x}{x!} \quad \text{mit } \mu = n\pi.$$

Eine gebräuchliche **Faustregel** lautet, dass die Approximation im konkreten Fall als zulässig angesehen werden kann, wenn die drei Bedingungen

$$\pi \leq 0.1, \quad n \geq 50 \quad \text{und} \quad n\pi \leq 9$$

erfüllt sind.

*Beispiel 2.17: 50 Säuglinge erhalten eine bestimmte Impfung. Die Wahrscheinlichkeit, dass ein Säugling die Impfung nicht verträgt, ist  $\pi = 0.05$ .*

<sup>5</sup>Siméon Denis Poisson (1781–1840)



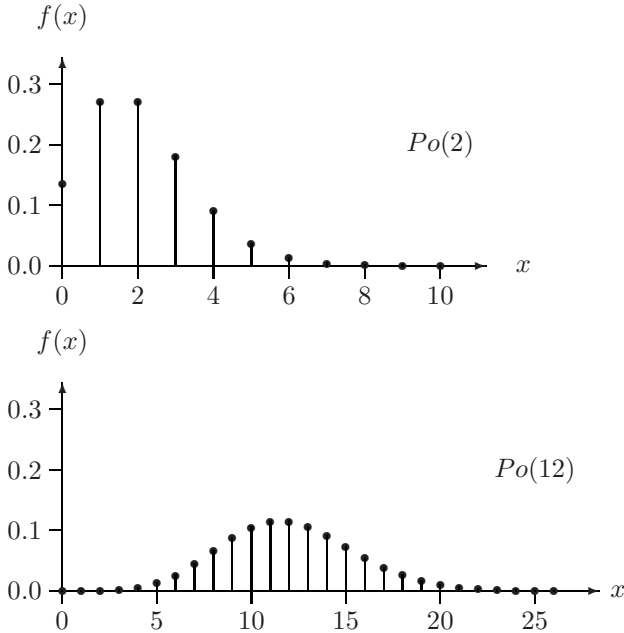


Abbildung 2.14: Wahrscheinlichkeitsfunktion der Poisson-Verteilung  $Po(\mu)$  mit  $\mu = 2$  bzw.  $\mu = 12$

Die Reaktionen der einzelnen Säuglinge auf die Impfung sind unabhängig voneinander. Sei

$X =$  Anzahl der Säuglinge, die die Impfung nicht vertragen.

Man berechne  $P(X = 0), P(X = 1), P(X = 2), P(X \geq 3)$  exakt und approximativ und vergleiche die Werte der Wahrscheinlichkeiten ( $\leftrightarrow$  Calc/SPSS).

Es gilt  $X \sim B(n = 50, \pi = 0.05)$ . Wegen  $\pi \leq 0.1, n \geq 50$  und  $n\pi = 2.5 \leq 9$  können wir die Verteilung von  $X$  durch die Poisson-Verteilung  $Po(\mu = 2.5)$  approximieren. Die exakten und genäherten Wahrscheinlichkeiten sind:

	$B(n = 50, \pi = 0.05)$	$Po(\mu = 2.5)$
$P(X = 0)$	0.0769	0.0821
$P(X = 1)$	0.2025	0.2052
$P(X = 2)$	0.2611	0.2565
$P(X \geq 3)$	0.4595	0.4562

Erwartungswert und Varianz einer Zufallsvariablen  $X \sim Po(\mu)$  sind

$$E[X] = \mu, \quad V[X] = \mu.$$

Sowohl der Erwartungswert einer Poisson-Verteilung als auch ihre Varianz sind also gleich dem Parameter  $\mu$  der Verteilung. Dies macht man sich anschaulich daran klar, dass die  $Po(\mu)$ -Wahrscheinlichkeiten die Grenzwerte von  $B(n, \pi)$ -Wahrscheinlichkeiten sind. Erwartungswert und Varianz einer jeden  $B(n, \pi)$ -Verteilung betragen  $n\pi$  bzw.  $n\pi(1 - \pi)$ . Wegen  $n\pi \rightarrow \mu$  muss der Erwartungswert im Limes gleich  $\mu$  sein. Ebenso gilt wegen  $\pi \rightarrow 0$  für die Varianz die Limesbeziehung  $n\pi(1 - \pi) \rightarrow \mu$ . Dies lässt sich auch exakt zeigen:

BEWEIS Es gilt

$$\begin{aligned} E[X] &= \sum_{x=0}^{\infty} x e^{-\mu} \frac{\mu^x}{x!} = e^{-\mu} \sum_{x=1}^{\infty} x \frac{\mu^x}{x!} \\ &= e^{-\mu} \sum_{x=1}^{\infty} \mu \frac{\mu^{x-1}}{(x-1)!} = \mu e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = \mu. \end{aligned}$$

Die zweite dieser Gleichungen ergibt sich durch Weglassen des ersten Summanden (der gleich null ist), die vierte durch Übergang vom Summationsindex  $x$  (beginnend mit eins) zum Summationsindex  $x - 1$  (beginnend mit null).

Um die Varianz von  $X \sim Po(\mu)$  zu bestimmen, berechnen wir zunächst  $E[X^2]$ ,

$$\begin{aligned} E[X^2] &= \sum_{x=0}^{\infty} x^2 e^{-\mu} \frac{\mu^x}{x!} = e^{-\mu} \sum_{x=1}^{\infty} x^2 \frac{\mu^x}{x!} \\ &= e^{-\mu} \sum_{x=1}^{\infty} \mu x \frac{\mu^{x-1}}{(x-1)!} = \mu e^{-\mu} \sum_{x=0}^{\infty} (x+1) \frac{\mu^x}{(x)!} \\ &= \mu \left[ e^{-\mu} \sum_{x=0}^{\infty} x \frac{\mu^x}{x!} + e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \right] \\ &= \mu[\mu + 1] = \mu^2 + \mu \end{aligned}$$

Hierbei erhält man wie oben die zweite Gleichung durch Weglassen des ersten Summanden (der gleich null ist) und die vierte durch die gleiche Umindizierung. Es folgt  $V[X] = E[X^2] - \mu^2_X = \mu^2 + \mu - \mu^2 = \mu$ .  $\square$

### 2.3.3 Geometrische Verteilung

Wir kehren zurück zur Bernoulli-Versuchsreihe und nehmen an, dass das zugrunde liegende Bernoulli-Experiment im Prinzip beliebig oft wiederholt werden kann. Es interessiert nun der Zeitpunkt des ersten „Erfolges“, d.h. des erstmaligen Eintretens des relevanten Ereignisses. Die Zufallsvariable  $X$

gebe an, beim wievielten Versuch das Ereignis erstmals eintritt. Man kann  $X$  in der Form

$$X = \min \left\{ n \mid \sum_{i=1}^n \mathbf{1}_{A_i} = 1 \right\}$$

schreiben. Der Träger von  $X$  umfasst offensichtlich die natürlichen Zahlen,  $T_X = \mathbb{N}$ . Wenn der erste Erfolg beim  $x$ -ten Versuch eintritt, gehen ihm genau  $x - 1$  Misserfolge voraus. Es ist

$$\begin{aligned} P(X = 1) &= \pi, \\ P(X = 2) &= \pi(1 - \pi), \\ P(X = 3) &= \pi(1 - \pi)^2, \text{ usw.} \end{aligned}$$

Eine Zufallsvariable  $X$  mit der Wahrscheinlichkeitsfunktion

$$P(X = x) = \pi(1 - \pi)^{x-1} \text{ für } x \in \mathbb{N}$$

heißt **geometrisch verteilt** mit Parameter  $\pi$ , in Zeichen:  $X \sim G(\pi)$ . Der Träger von  $X$  ist  $T_X = \mathbb{N}$ .

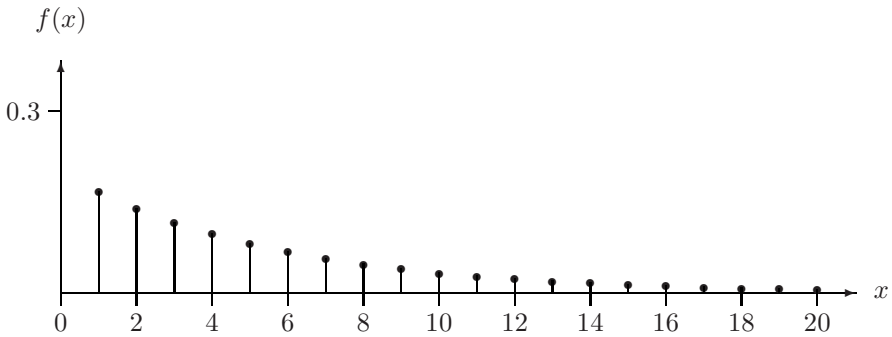


Abbildung 2.15: Wahrscheinlichkeitsfunktion der geometrischen Verteilung  $G(\pi)$ ,  $\pi = \frac{1}{6}$

Man rechnet mit Hilfe der unendlichen geometrischen Reihe leicht nach, dass sich die Wahrscheinlichkeiten insgesamt zu eins addieren:

$$\sum_{x=1}^{\infty} P(X = x) = \sum_{x=1}^{\infty} \pi(1 - \pi)^{x-1} = \pi \cdot \frac{1}{1 - (1 - \pi)} = 1$$

Aus den Einzelwahrscheinlichkeiten von  $X$  lässt sich mit der endlichen geo-

metrischen Reihe die **Verteilungsfunktion** bestimmen. An jeder der Stellen  $k = 1, 2, 3, \dots$  gilt

$$F(k) = \sum_{i=1}^k \pi(1-\pi)^{i-1} = 1 - (1-\pi)^k,$$

und allgemein für beliebige  $x \in \mathbb{R}$

$$F(x) = \begin{cases} 0, & \text{falls } x < 1, \\ 1 - (1-\pi)^k, & \text{falls } k \leq x < k+1. \end{cases}$$

*Beispiel 2.18:* Aus einer Urne mit 10 Kugeln (davon 4 rot und 6 weiß) wird mit Zurücklegen zufällig gezogen. Wie groß ist die Wahrscheinlichkeit, dass

- bei der dritten Ziehung zum ersten Mal eine rote Kugel gezogen wird,
- frühestens bei der dritten Ziehung zum ersten Mal eine rote Kugel gezogen wird?

Mit

$X$  = Nummer der Ziehung, bei der zum ersten Mal eine rote Kugel gezogen wird,

gilt

$$X \sim G(\pi = 0.4).$$

zu a) Es ist

$$P(X = 3) = 0.4 \cdot 0.6^2 = 0.144.$$

zu b) Es ist

$$\begin{aligned} \sum_{x=3}^{\infty} P(X = x) &= 1 - P(X = 1) - P(X = 2) \\ &= 1 - 0.4 - 0.4 \cdot 0.6 = 0.36. \end{aligned}$$

**Erwartungswert** und **Varianz** einer  $G(\pi)$ -verteilten Zufallsvariablen  $X$  sind

$$E[X] = \frac{1}{\pi}, \quad V[X] = \frac{1-\pi}{\pi^2}.$$

**BEWEIS** Der Erwartungswert von  $X$  ist die Summe einer unendlichen Reihe,

$$E[X] = \sum_{k=1}^{\infty} k\pi(1-\pi)^{k-1}.$$

Es gilt

$$\begin{aligned} \frac{E[X]}{\pi} &= \sum_{k=1}^{\infty} k(1-\pi)^{k-1} \\ &= \underbrace{\sum_{k=1}^{\infty} (1-\pi)^{k-1}}_{=\frac{1}{\pi}} + \sum_{k=1}^{\infty} (k-1)(1-\pi)^{k-1}. \end{aligned}$$

Die erste Summe der letzten Gleichung ist eine geometrische Reihe; sie hat den Wert  $\frac{1}{1-(1-\pi)} = \frac{1}{\pi}$ . Wir untersuchen nun die zweite Summe:

$$\begin{aligned} \sum_{k=1}^{\infty} (k-1)(1-\pi)^{k-1} &= (1-\pi) \sum_{k=1}^{\infty} (k-1)(1-\pi)^{k-2} \\ &= (1-\pi) \sum_{k=2}^{\infty} (k-1)(1-\pi)^{k-2} \\ &= (1-\pi) \sum_{k=1}^{\infty} k(1-\pi)^{k-1} \\ &= (1-\pi) \frac{E[X]}{\pi} \end{aligned}$$

Man erhält also

$$\frac{E[X]}{\pi} = \frac{1}{\pi} + (1-\pi) \cdot \frac{E[X]}{\pi}$$

und daraus  $E[X] = \frac{1}{\pi}$ . Die Varianz wird auf ähnliche Weise berechnet.  $\square$

Man beachte, dass der Erwartungswert umso größer ist, je kleiner  $\pi$  ist. Das leuchtet unmittelbar ein, denn bei kleinem  $\pi$  ist zu erwarten, dass es länger als bei großem  $\pi$  dauert, bis  $A$  erstmals eintritt. Ähnliches gilt für die Varianz; sie ist umso größer, je kleiner  $\pi$  ist.

*Beispiel 2.19:* Beim „Mensch-Ärgere-Dich-Nicht“-Spiel wird so lange gewürfelt, bis die erste „6“ erscheint. Die Zufallsvariable  $X$  sei definiert als die Nummer des Wurfes, bei dem erstmals eine Sechs auftritt. Dann ist  $X$  geometrisch verteilt,  $X \sim G(\pi)$  mit  $\pi = \frac{1}{6}$  und es gilt  $E[X] = 6$ . Man benötigt also im Durchschnitt sechs Versuche, um erstmals eine „6“ zu würfeln.

In Teilen der Literatur wird die geometrische Verteilung etwas anders definiert, nämlich als Zahl  $Y$  der „Misserfolge“ bis zum ersten Erfolg. Offenbar ist dann  $Y = X - 1$ , wobei  $X$  im hier definierten Sinne geometrisch verteilt ist. Es gilt

$$P(Y = y) = \pi(1-\pi)^y.$$

Der Erwartungswert von  $Y$  ist

$$E[Y] = E[X - 1] = E[X] - 1 = \frac{1}{\pi} - 1 = \frac{1 - \pi}{\pi},$$

die Varianz ist

$$V[Y] = V[X] = \frac{1 - \pi}{\pi^2}.$$

### 2.3.4 Hypergeometrische Verteilung

Zufälliges Ziehen mit Zurücklegen ist ein Beispiel für eine Bernoulli-Versuchsreihe, da bei jeder Wiederholung die gleiche Ausgangssituation vorliegt und daher die „Erfolge“ der einzelnen Versuche global unabhängig sind. Beim Ziehen ohne Zurücklegen ist dies, wie wir jetzt sehen werden, nicht der Fall.

Wir betrachten wiederum Ziehungen aus einer Urne mit  $N$  Kugeln, davon seien  $M$  rot und  $N - M$  weiß. Es werden  $n$  Kugeln nacheinander **ohne Zurücklegen** und ohne Berücksichtigung der Reihenfolge gezogen. Sei

$A_i =$  bei der  $i$ -ten Ziehung wird eine rote Kugel gezogen.

Dann ist

$$\begin{aligned} P(A_1) &= \frac{M}{N} = \pi, \\ P(A_2) &= P(A_2 | A_1) \cdot P(A_1) + P(A_2 | \bar{A}_1) \cdot P(\bar{A}_1) \\ &= \frac{M-1}{N-1} \cdot \frac{M}{N} + \frac{M}{N-1} \cdot \frac{N-M}{N} = \frac{M^2 - M + MN - M^2}{(N-1)N} \\ &= \frac{M(N-1)}{(N-1)N} = \frac{M}{N} = \pi. \end{aligned}$$

$A_1$  und  $A_2$  sind jedoch nicht unabhängig, da

$$P(A_1 \cap A_2) = \frac{M}{N} \cdot \frac{M-1}{N-1} \neq \frac{M}{N} \cdot \frac{M}{N} = P(A_1) \cdot P(A_2).$$

Wegen der fehlenden Unabhängigkeit der Ziehungen ist das Ziehen ohne Zurücklegen keine Bernoulli-Versuchsreihe. Wie beim Ziehen mit Zurücklegen kann man jedoch auch hier die Verteilung der Zufallsvariablen

$X =$  Anzahl der bei  $n$  Versuchen gezogenen roten Kugeln

herleiten. Da es nicht darauf ankommt, in welcher Reihenfolge rote und weiße Kugeln gezogen werden, gibt es (vgl. Abschnitt 1.2.3)

$\binom{N}{n}$  Möglichkeiten,  $n$  Kugeln aus  $N$  Kugeln ohne Zurücklegen auszuwählen,

$\binom{M}{x}$  Möglichkeiten,  $x$  rote Kugeln aus  $M$  roten Kugeln ohne Zurücklegen auszuwählen,

$\binom{N-M}{n-x}$  Möglichkeiten,  $n-x$  weiße Kugeln aus  $N-M$  weißen Kugeln ohne Zurücklegen auszuwählen.

Nach dem klassischen Wahrscheinlichkeitsbegriff ergibt sich daraus die Wahrscheinlichkeitsfunktion:

$$P(X = x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} \quad \text{für } \begin{array}{l} x \in \{0, 1, \dots, n\} \\ \text{mit } x \leq M \text{ und} \\ n-x \leq N-M. \end{array}$$

Eine solche diskrete Zufallsvariable  $X$  nennt man **hypergeometrisch verteilt**, in Zeichen:  $X \sim H(n, N, M)$ . Die hypergeometrische Verteilung hat drei Parameter,  $n, N$  und  $M$  mit  $n, N \in \mathbb{N}$ ,  $n \leq N$ ,  $M \in \{0, 1, \dots, N\}$ .

Dabei kommt die Einschränkung  $x \leq M$  dadurch zustande, dass nicht mehr rote Kugeln gezogen werden können als in der Urne sind, und die Einschränkung  $n-x \leq N-M$  dadurch, dass nicht mehr weiße Kugeln, als in der Urne vorhanden sind, gezogen werden können. Die Anzahl der gezogenen roten Kugeln ist demnach mindestens  $n-(N-M)$  und höchstens  $M$ , außerdem ist sie mindestens 0 und höchstens  $n$ . Der Träger von  $X$  ist die Menge

$$T_X = \{\max\{0, n-N+M\}, \dots, \min\{n, M\}\}.$$

*Beispiel 2.20: Eine Warensendung enthält zehn elektronische Bauteile, von denen drei defekt sind. Aus der Sendung werden zwei Bauteile zufällig und ohne Zurücklegen entnommen und geprüft. Wie groß ist die Wahrscheinlichkeit, dass darunter mindestens ein defektes Bauteil ist?*

Sei

$X = \text{Anzahl der defekten Bauteile unter den zwei entnommenen.}$

Dann gilt ( $\hookrightarrow$  Calc/SPSS)  $X \sim H(n=2, N=10, M=3)$  und damit

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{3}{0} \binom{7}{2}}{\binom{10}{2}} = \frac{8}{15} = 0.5333.$$

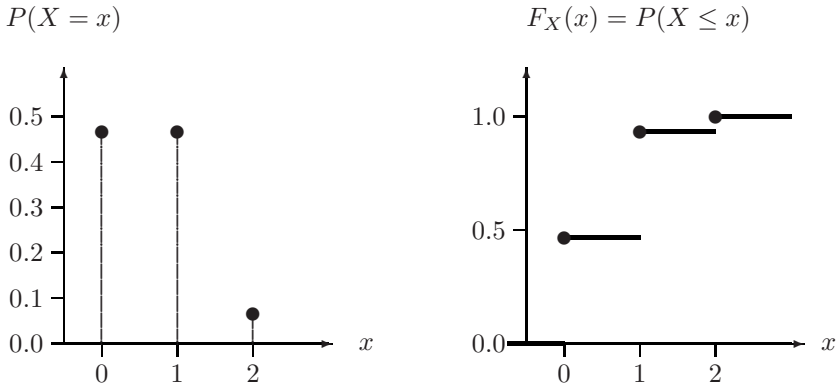


Abbildung 2.16: Wahrscheinlichkeitsfunktion und Verteilungsfunktion von  $H(n=2, N=10, M=3)$  im Beispiel 2.20 (Ziehen ohne Zurücklegen)

Für  $n=1$  stimmt die hypergeometrische Verteilung  $H(n=1, N, M)$  mit der Bernoulli-Verteilung  $B(n=1, \pi = \frac{M}{N})$  überein. Für  $n > 1$  gibt es Unterschiede zur Binomialverteilung, die sich durch den unterschiedlichen Ziehungsmodus erklären.

**Approximation** Andererseits ist unmittelbar klar, dass dann, wenn sich viele Kugeln in der Urne befinden und nur wenige Kugeln gezogen werden, kein großer Unterschied zwischen dem Ziehen mit und ohne Zurücklegen besteht. Für jedes feste  $x$  und  $n$  gilt die Grenzbeziehung

$$\frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \xrightarrow[\substack{N \rightarrow \infty \\ M \rightarrow \infty \\ \frac{M}{N} \rightarrow \pi}]{\quad} \binom{n}{x} \pi^x (1-\pi)^{n-x}.$$

Die Wahrscheinlichkeiten der hypergeometrischen Verteilung  $H(n, N, M)$  können deshalb durch die einer Binomialverteilung  $B(n, \pi = \frac{M}{N})$  approximiert werden. Für die Zulässigkeit der Approximation im konkreten Fall ist die folgende **Faustregel** üblich:  $\frac{n}{N} \leq 0.05$ . Der Quotient  $\frac{n}{N}$  heißt **Auswahlsatz**.

*Beispiel 2.21: Aus einer Urne mit 100 Kugeln (davon 50 roten) werden drei Kugeln ohne Zurücklegen gezogen. Wie groß ist die Wahrscheinlichkeit, dass genau eine rote unter den drei gezogenen Kugeln ist?*

Mit

$X = \text{Anzahl der roten Kugeln unter den drei gezogenen}$



gilt  $X \sim H(n = 3, N = 100, M = 50)$  und deshalb

$$P(X = 1) = \frac{\binom{50}{1} \binom{50}{2}}{\binom{100}{3}} = 0.3788.$$

Da der Auswahlatz  $\frac{n}{N} = \frac{3}{100}$  klein genug ist, dürfen wir diese Wahrscheinlichkeit auch näherungsweise mithilfe der Binomialverteilung bestimmen. Es gilt  $\frac{M}{N} = \frac{50}{100} = 0.5$  und approximativ  $X \sim B(n = 3, \pi = 0.5)$ . Aus der Binomialverteilungstabelle (Tabelle 1 im Anhang) erhalten wir  $P(X = 1) \approx 0.3750$ .

**Erwartungswert** und **Varianz** von  $X \sim H(n, N, M)$  sind

$$\begin{array}{l} E[X] = n \frac{M}{N}, \\ V[X] = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}. \end{array}$$

Mit  $\pi = \frac{M}{N}$  erhalten wir

$$\begin{array}{l} E[X] = n\pi, \\ V[X] = n\pi(1 - \pi) \frac{N-n}{N-1}. \end{array}$$

Während es beim Erwartungswert keinen Unterschied zwischen dem Ziehen mit und ohne Zurücklegen gibt, verhält es sich mit der Varianz anders: Wegen

$$\frac{N-n}{N-1} < 1, \quad \text{falls } n > 1,$$

ist die Varianz von  $X$  beim Ziehen ohne Zurücklegen kleiner als beim Ziehen mit Zurücklegen. Im Extremfall  $n = N$ , d.h. wenn alle Kugeln aus der Urne entnommen werden, gilt sogar  $V[X] = 0$ , denn dann ist  $P(X = M) = 1$ .

## 2.4 Spezielle stetige Verteilungen

In diesem Abschnitt behandeln wir verschiedene stetige Verteilungen, die zum Kernbestand der Wahrscheinlichkeitsrechnung gehören und in vielen Anwendungen eine Rolle spielen. Anders als bei den diskreten Verteilungen des Abschnitts 2.3, die sämtlich aus dem Urnenmodell mit bzw. ohne Zurücklegen entwickelt werden konnten, liegen diesen stetigen Verteilungen sehr unterschiedliche Modellansätze zugrunde. Einige von ihnen werden wir in späteren Anwendungen erläutern.

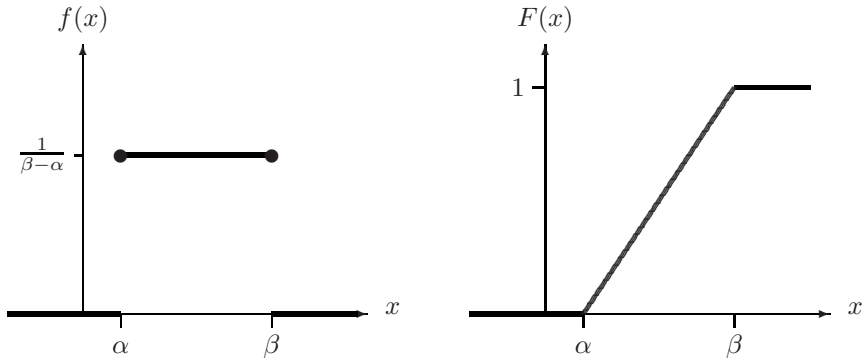


Abbildung 2.17: Dichte und Verteilungsfunktion der Rechteckverteilung  $R(\alpha, \beta)$

### 2.4.1 Rechteckverteilung

Eine Zufallsvariable  $X$  heißt **rechteckverteilt** mit Parametern  $\alpha, \beta \in \mathbb{R}$ ,  $\alpha < \beta$ , in Zeichen:  $X \sim R(\alpha, \beta)$ , wenn sie die Dichte

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & \text{falls } \alpha \leq x \leq \beta, \\ 0 & \text{sonst,} \end{cases}$$

besitzt. Die Rechteckverteilung wird auch als **uniforme Verteilung** oder **stetige Gleichverteilung** bezeichnet.

Der Träger der Rechteckverteilung ist  $T_X = [\alpha, \beta]$ . Auf dem Intervall  $[\alpha, \beta]$  ist die Verteilungsfunktion linear mit konstanter Steigung  $1/(\beta - \alpha)$ , ansonsten ist sie konstant, also (vgl. Abbildung 2.17)

$$F(x) = \begin{cases} 0, & \text{falls } x < \alpha, \\ \frac{x - \alpha}{\beta - \alpha}, & \text{falls } \alpha \leq x \leq \beta, \\ 1, & \text{falls } x > \beta. \end{cases}$$

Die Rechteckverteilung  $R(\alpha, \beta)$  ist dann ein sinnvolles Modell für eine Zufallsvariable  $X$ , wenn man weiß, dass  $X$  nur Werte zwischen  $\alpha$  und  $\beta$  annimmt und innerhalb des Intervalls  $[\alpha, \beta]$  keine Werte „bevorzugt“ auftreten. Dann hängt die Wahrscheinlichkeit dafür, dass  $X$  in ein Teilintervall von  $[\alpha, \beta]$  fällt, nur von der Länge des Teilintervalls ab und ist proportional dieser Länge.

*Beispiel 2.22: Ein Autofahrer gibt seinen Wagen zur Inspektion und muss deshalb mit der Straßenbahn fahren. Er weiß nur, dass die Züge im Zehn-Minuten-Rhythmus verkehren und geht zur Haltestelle, um auf den nächsten*

Zug zu warten. Sei  $X$  seine Wartezeit an der Haltestelle (in Minuten). Welche Wahrscheinlichkeitsverteilung für  $X$  bietet sich an?

Offenbar ist  $X \sim R(\alpha = 0, \beta = 10)$  eine plausible Annahme.

*Beispiel 2.6 (Fortsetzung):* In dem Spiel des Fernseh-Showmasters bezeichne  $X$  den angezeigten Winkel in Grad bei Stillstand des Glücksrads. Dann ist  $X \sim R(\alpha = 0, \beta = 360)$  eine sinnvolle Verteilungsannahme.

Unter den Rechteckverteilungen ist die mit den Parametern  $\alpha = 0$  und  $\beta = 1$  von besonderer Bedeutung. Eine Zufallsvariable  $Z \sim R(0, 1)$  hat die Dichte

$$f_Z(x) = \begin{cases} 1, & \text{falls } 0 \leq x \leq 1, \\ 0 & \text{sonst,} \end{cases}$$

und die Verteilungsfunktion

$$F_Z(x) = \begin{cases} 0, & \text{falls } x < 0, \\ x, & \text{falls } 0 \leq x \leq 1, \\ 1, & \text{falls } x > 1. \end{cases}$$

Man nennt ein solches  $Z \sim R(0, 1)$  auch **standard-rechteckverteilt**. Wir zeigen nun, dass eine affin-lineare Transformation der Zufallsvariablen  $Z$  wieder rechteckverteilt ist. Sei  $\alpha, \beta \in \mathbb{R}, \alpha < \beta$ , und

$$X = \alpha + (\beta - \alpha)Z.$$

Dann ist  $X \sim R(\alpha, \beta)$ .

**BEWEIS** Wegen der Formel (2.5) für die Dichte einer affin transformierten Variablen gilt

$$f(x) = \frac{1}{\beta - \alpha} f_Z\left(\frac{x - \alpha}{\beta - \alpha}\right), \quad x \in \mathbb{R}.$$

Es folgt

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & \text{falls } 0 \leq \frac{x - \alpha}{\beta - \alpha} \leq 1, \text{ d.h. falls } \alpha \leq x \leq \beta, \\ 0 & \text{sonst.} \end{cases}$$

Also ist  $X \sim R(\alpha, \beta)$ . □

Wenn eine Zufallsvariable  $X$  allgemein rechteckverteilt ist, d.h.  $X \sim R(\alpha, \beta)$ , zeigt man entsprechend, dass die Zufallsvariable

$$\frac{X - \alpha}{\beta - \alpha} \sim R(0, 1)$$

standard-rechteckverteilt ist.

**Erwartungswert und Varianz** Da  $X \sim R(\alpha, \beta)$  symmetrisch zur Mitte des Intervalls  $[\alpha, \beta]$  verteilt ist, ist der Erwartungswert gleich dem Mittelpunkt des Intervalls,

$$E[X] = \frac{\alpha + \beta}{2}.$$

Die Varianz beträgt

$$V[X] = \frac{(\beta - \alpha)^2}{12}.$$

**BEWEIS** Um die Varianz von  $X$  zu bestimmen, berechnen wir zunächst die Varianz einer standard-rechteckverteilten Zufallsvariablen  $Z \sim R(0, 1)$ . Es ist  $E[Z] = \frac{1}{2}$ ,

$$E[Z^2] = \int_0^1 z^2 \cdot 1 \, dz = \frac{z^3}{3} \Big|_0^1 = \frac{1}{3},$$

und daraus folgt

$$\begin{aligned} V[Z] &= E[Z^2] - (E[Z])^2 \\ &= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}. \end{aligned}$$

Nach den Rechenregeln (2.8) und (2.10) für Erwartungswert und Varianz einer affin-linear transformierten Zufallsvariablen erhalten wir

$$\begin{aligned} V[X] &= V[\alpha + (\beta - \alpha)Z] = V[(\beta - \alpha)Z] \\ &= (\beta - \alpha)^2 V[Z] = \frac{(\beta - \alpha)^2}{12}. \quad \square \end{aligned}$$

*Beispiel 2.23: Bei der Produktion von Automobilen muss ein bestimmter Arbeitsgang per Hand ausgeführt werden. Gehen Sie davon aus, dass die Dauer  $X$  des Arbeitsganges mindestens sieben Minuten und höchstens zwölf Minuten beträgt.*

- a) Berechnen Sie den Erwartungswert und die Standardabweichung der Dauer des Arbeitsganges  $X$ , indem Sie eine Rechteckverteilung unterstellen.
- b) Bestimmen Sie die Quantilfunktion von  $X$  und berechnen Sie den 90%-Punkt. Wie ist dieser zu interpretieren?

zu a) Es ist  $\alpha = 7$ ,  $\beta = 12$ ,

$$\begin{aligned} E[X] &= \frac{12+7}{2} = 9.5 \quad [\text{Minuten}], \\ \sqrt{V[X]} &= \sqrt{\frac{(12-7)^2}{12}} = 1.44 \quad [\text{Minuten}]. \end{aligned}$$

zu b) Um die Quantilfunktion zu bestimmen, müssen wir

$$F(x) = \frac{x-7}{12-7} = p, \quad 7 \leq x \leq 12,$$

für  $0 < p < 1$  nach  $x$  auflösen. Es ist

$$Q(p) = x_p = 7 + 5p.$$

Der 90%-Punkt ist  $Q(0.9) = 7 + 5 \cdot 0.9 = 11.5$  [Minuten], d.h. mit Wahrscheinlichkeit 0.9 ist der Arbeitsgang nach 11.5 Minuten beendet.

Wir zeigen nun eine wichtige Eigenschaft der Rechteckverteilung, die man für die Monte-Carlo-Simulation (siehe Abschnitt 3.3.3) benötigt:

Sei  $F$  eine vorgegebene Verteilungsfunktion und  $Q$  die zugehörige Quantilfunktion,

$$Q(p) = \min \{x | F(x) \geq p\}.$$

Ist  $Z$  eine  $R(0,1)$ -verteilte Zufallsvariable und  $X = Q(Z)$ , dann hat  $X$  die Verteilungsfunktion  $F$ .

BEWEIS Es gilt nämlich für jedes  $x \in \mathbb{R}$

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(Q(Z) \leq x) \\ &= P(F(Q(Z)) \leq F(x)) \\ &= P(Z \leq F(x)) = F(x). \end{aligned} \quad \square$$

Wenn  $F$  eine stetige Funktion ist, gilt darüber hinaus, dass

$$F(X) \sim R(0,1),$$

d.h. dass  $F(X)$  eine Rechteckverteilung in  $[0,1]$  besitzt.

BEWEIS Für  $z \in [0,1]$  gilt

$$\begin{aligned} F_{F(X)}(z) &= P(F(X) \leq z) = P(Q(F(X)) \leq Q(z)) \\ &= P(X \leq Q(z)) = F(Q(z)) = z. \end{aligned} \quad \square$$

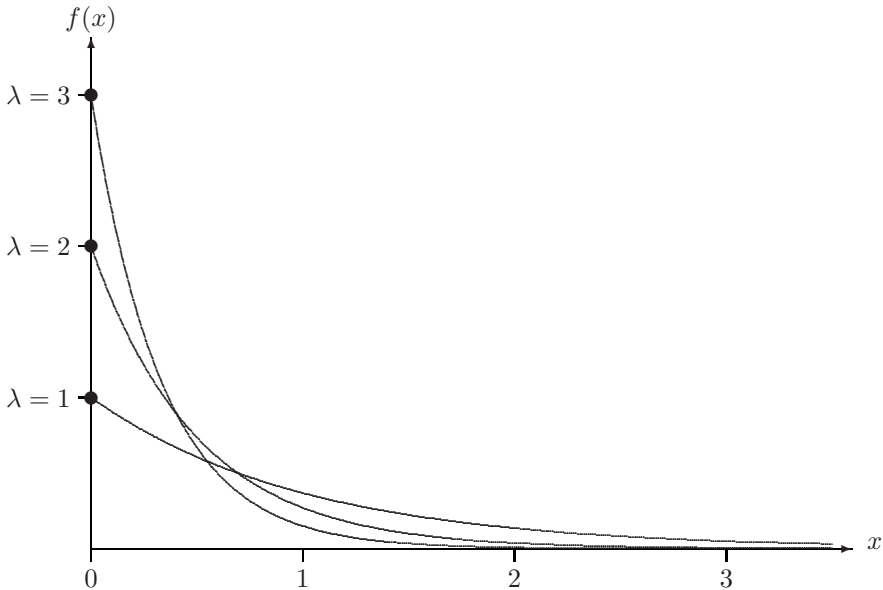


Abbildung 2.18: Dichtefunktion der Exponentialverteilung  $Exp(\lambda)$  mit Parameter  $\lambda = 1, 2$  und  $3$

## 2.4.2 Exponentialverteilung

Eine stetige Zufallsvariable  $X$  nennt man **exponentialverteilt** mit Parameter  $\lambda > 0$ , in Zeichen:  $X \sim Exp(\lambda)$ , falls ihre Dichte durch

$$f(x) = \begin{cases} 0, & \text{falls } x < 0, \\ \lambda e^{-\lambda x}, & \text{falls } x \geq 0, \end{cases}$$

gegeben ist. Der Träger von  $X$  ist demnach  $T_X = [0, \infty[$ .

Offenbar ist  $f$  die Ableitung der folgenden Verteilungsfunktion  $F$ ,

$$F(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1 - e^{-\lambda x}, & \text{falls } x \geq 0. \end{cases}$$

Die Abbildungen 2.18 und 2.19 zeigen die Dichte bzw. die Verteilungsfunktion von Exponentialverteilungen mit verschiedenen Parametern  $\lambda$ . Wenn  $\lambda$  groß ist, befindet sich relativ viel Wahrscheinlichkeitsmasse in der Nähe von 0 und die Dichte fällt mit wachsendem  $x$  rasch ab. Wenn  $\lambda$  klein ist, ist es

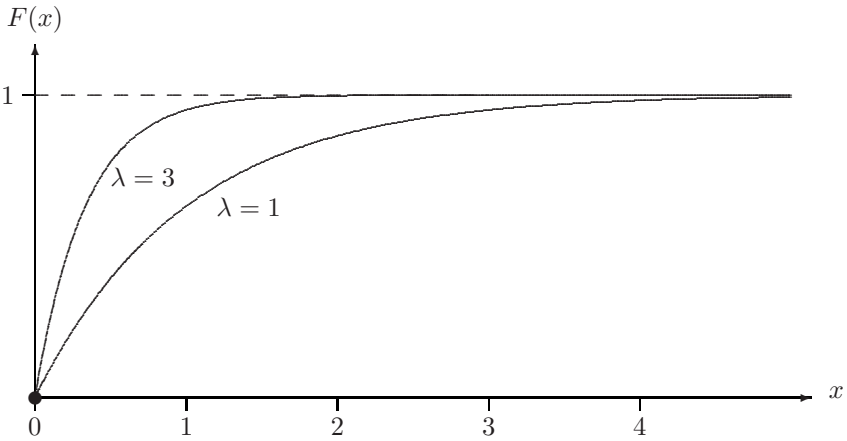


Abbildung 2.19: Verteilungsfunktion der Exponentialverteilung  $Exp(\lambda)$  für  $\lambda = 1$  und 3

umgekehrt: Es liegt relativ wenig Wahrscheinlichkeitsmasse bei null und die Dichte fällt mit wachsendem  $x$  nur langsam. Die Rolle von  $\lambda$  kann auch an der Quantilfunktion demonstriert werden. Es ist (vgl. Abschnitt 2.1.2)

$$Q(p) = x_p = -\frac{1}{\lambda} \ln(1-p).$$

Für jedes feste  $p$  ist das  $p$ -Quantil  $x_p$  also umso kleiner, je größer  $\lambda$  ist. Der Median  $x_{0.5}$  beträgt

$$x_{0.5} = -\frac{\ln 0.5}{\lambda} = \frac{0.6931}{\lambda}.$$

**Erwartungswert** und **Varianz** einer Zufallsvariablen  $X \sim Exp(\lambda)$  sind

$$E[X] = \frac{1}{\lambda}, \quad V[X] = \frac{1}{\lambda^2}.$$

Der Erwartungswert und die Varianz sind also umso kleiner, je größer  $\lambda$  ist.

BEWEIS Den Erwartungswert berechnet man mittels partieller Integration,

$$\begin{aligned}
 E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \cdot (\lambda e^{-\lambda x}) dx \\
 &= -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} 1 \cdot e^{-\lambda x} dx \\
 &= -x e^{-\lambda x} \Big|_0^{\infty} - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} \\
 &= -0 + 0 - 0 + \frac{1}{\lambda},
 \end{aligned}$$

also  $\mu = E[X] = \frac{1}{\lambda}$ . Ebenso erhält man mit zweifacher partieller Integration

$$E[X^2] = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}.$$

Hieraus ergibt sich  $V[X] = E[X^2] - \mu^2 = \frac{1}{\lambda^2}$ . □

**Schiefe** Wir berechnen nun die Momentenschiefe  $\gamma_1[X]$  und die Quartilsschiefe  $\gamma_1^Q[X]$  einer Zufallsvariablen  $X \sim Exp(\lambda)$ . Sie betragen

$$\gamma_1[X] = 2, \quad \gamma_1^Q[X] = \frac{\ln 4 - \ln 3}{\ln 3} = 0.2619,$$

hängen also beide nicht von  $\lambda$  ab. Sowohl die Momentenschiefe als auch die Quartilsschiefe einer Exponentialverteilung sind positiv und zeigen Rechtsschiefe (= Linksteilheit) an.

BEWEIS Laut Definition ist die Momentenschiefe gleich

$$\gamma_1[X] = \frac{E[X^3] - 3E[X^2]\mu_X + 2\mu_X^3}{\sigma_X^3}.$$

Mithilfe dreimaliger partieller Integration erhält man  $E[X^3] = \frac{6}{\lambda^3}$  und daher

$$\begin{aligned}
 \gamma_1[X] &= \frac{E[X^3] - 3E[X^2]\mu_X + 2\mu_X^3}{\sigma_X^3} \\
 &= \frac{\frac{6}{\lambda^3} - 3 \cdot \frac{2}{\lambda^2} \cdot \frac{1}{\lambda} + 2 \left(\frac{1}{\lambda}\right)^3}{\left(\frac{1}{\lambda}\right)^3} = 2.
 \end{aligned}$$

Zur Berechnung der Quartilsschiefe

$$\gamma_1^Q = \frac{Q(0.75) + Q(0.25) - 2 \cdot Q(0.5)}{Q(0.75) - Q(0.25)}$$



benötigen wir die Quantile

$$\begin{aligned} Q(0.25) &= -\frac{1}{\lambda} \ln(1 - 0.25) = \frac{1}{\lambda} \ln\left(\frac{4}{3}\right), \\ Q(0.5) &= -\frac{1}{\lambda} \ln(1 - 0.50) = \frac{1}{\lambda} \ln 2, \\ Q(0.75) &= -\frac{1}{\lambda} \ln(1 - 0.75) = \frac{1}{\lambda} \ln 4. \end{aligned}$$

Es ergibt sich

$$\begin{aligned} \gamma_1^Q[X] &= \frac{\ln 4 + \ln\left(\frac{4}{3}\right) - 2 \cdot \ln 2}{\ln 4 - \ln\left(\frac{4}{3}\right)} = \frac{\ln\left(\frac{4 \cdot 4}{3 \cdot 2^2}\right)}{\ln 3} \\ &= \frac{\ln\left(\frac{4}{3}\right)}{\ln 3} = \frac{\ln 4 - \ln 3}{\ln 3} = 0.2619. \quad \square \end{aligned}$$

Die Exponentialverteilung ist geeignet zur Modellierung von Lebensdauern von technischen Geräten und von Wartezeiten in Bedienungssystemen. Sie hat die folgende interessante Eigenschaft, die im Folgenden als **Gedächtnislosigkeit** interpretiert wird: Für  $X \sim \text{Exp}(\lambda)$  und  $t, s > 0$  gilt

$$\begin{aligned} P(X > t + s \mid X > t) &= \frac{P(X > t + s)}{P(X > t)} = \frac{1 - F(t + s)}{1 - F(t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s), \end{aligned}$$

also

$$\boxed{P(X > t + s \mid X > t) = P(X > s)}. \quad (2.18)$$

Die bedingte Wahrscheinlichkeit, dass die Lebensdauer  $X$  den Wert  $t + s$  überschreitet unter der Bedingung, dass sie bereits den Wert  $t$  überschritten hat, ist also gleich der (nicht bedingten) Wahrscheinlichkeit, dass die Lebensdauer den Wert  $s$  überschreitet. Umgekehrt lässt sich zeigen, dass eine stetige Zufallsvariable  $X$ , die die Eigenschaft (2.18) besitzt, eine Exponentialverteilung haben muss. Wenn für eine diskrete Zufallsvariable (2.18) zutrifft, ist sie geometrisch verteilt.

Die Eigenschaft (2.18) lässt sich so als „Gedächtnislosigkeit“ interpretieren: Beträgt das Alter eines Geräts bereits  $t$  Zeiteinheiten, so ist die Wahrscheinlichkeit, dass seine Funktionstüchtigkeit noch mindestens  $s$  weitere Zeiteinheiten anhält, ebenso groß wie die Wahrscheinlichkeit, dass ein gleichartiges neues Gerät mindestens  $s$  Zeiteinheiten lang funktionstüchtig ist. Das Gerät „vergisst“ also in jedem Zeitpunkt  $t$  sein Alter.

Ob für ein bestimmtes Objekt die Eigenschaft der Gedächtnislosigkeit zutrifft oder nicht, kann natürlich nur durch inhaltliche Überlegungen festgestellt

werden. Falls sie zutrifft, ist die Lebensdauer exponentialverteilt, wobei der Parameter  $\lambda$  noch zu bestimmen ist.

*Beispiel 2.24:* Ein Ingenieur unterstellt, dass die Lebensdauer  $X$  eines bestimmten Typs von Glühlampen (in Stunden) durch eine Exponentialverteilung mit  $E[X] = 800$  beschrieben werden kann ( $\hookrightarrow$  Calc/SPSS).

- Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Glühlampe länger als 1000 Stunden brennt?
- Welche Brenndauer wird nur mit einer Wahrscheinlichkeit von 10% überschritten?

zu a) Es ist  $E[X] = \frac{1}{\lambda} = 800$ , also  $\lambda = \frac{1}{800}$ . Deshalb ist

$$P(X > 1000) = e^{-\frac{1}{800} \cdot 1000} = e^{-\frac{10}{8}} = 0.29.$$

zu b) Gesucht ist der 90%-Punkt  $x_{0.9}$ . Es ist

$$x_{0.9} = -\frac{1}{\lambda} \ln(1 - 0.9) = 1842.07.$$

Im Folgenden treten mehrfach Wahrscheinlichkeiten der Form  $P(X > x)$  auf. Allgemein wird für eine nichtnegative Zufallsvariable  $X$  die Funktion

$$S(x) = 1 - F(x) = P(X > x), \quad x \geq 0,$$

**Überlebensfunktion** (englisch: **survival function**) von  $X$  genannt. Ist  $X$  eine Lebensdauer, so gibt  $S(x)$  an, wie groß die Wahrscheinlichkeit ist, den Zeitpunkt  $x$  zu überleben. Ein weiterer wichtiger Begriff in diesem Zusammenhang ist die **Ausfallrate** (englisch: **hazard rate**)

$$h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)}, \quad x \geq 0.$$

Offenbar gilt

$$h(x) = -\frac{d}{dx} \ln(S(x)), \quad (2.19)$$

d.h. die Ausfallrate ist gleich dem negativen Wert der logarithmischen Ableitung der Überlebensfunktion. Umgekehrt lässt sich durch Integration aus der Ausfallrate die Überlebensfunktion und damit die Verteilungsfunktion bestimmen; es gilt

$$\begin{aligned} S(x) &= \exp\left(-\int_0^x h(t) dt\right), \\ F(x) &= 1 - S(x) = 1 - \exp\left(-\int_0^x h(t) dt\right). \end{aligned}$$

Für ein kleines Zeitintervall  $\Delta x$  ist

$$\begin{aligned} h(x)\Delta x &= \frac{f(x)\Delta x}{S(x)} \approx \frac{P(x < X \leq x + \Delta x)}{P(X > x)} \\ &= P(x < X \leq x + \Delta x | X > x). \end{aligned}$$

$h(x)\Delta x$  gibt also näherungsweise die Wahrscheinlichkeit an, dass ein Objekt im Intervall  $]x, x + \Delta x]$  ausfällt unter der Bedingung, dass es den Zeitpunkt  $x$  erlebt.

Speziell sei  $X$  exponentialverteilt,  $X \sim \text{Exp}(\lambda)$ . Dann ist  $S(x) = e^{-\lambda x}$  und  $f(x) = \lambda e^{-\lambda x}$ , also

$$h(x) = \frac{f(x)}{S(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda \quad \text{für } x \geq 0.$$

Bei der Exponentialverteilung ist demnach die Ausfallrate konstant und hängt nicht vom Alter  $x$  des Objektes ab. Man kann leicht (nämlich durch Integration der konstanten Ausfallrate  $\lambda$ ) mit Hilfe von Formel (2.19) zeigen, dass die Exponentialverteilung sogar die einzige Lebensdauerverteilung ist, die diese Eigenschaft besitzt. Die Konstanz der Ausfallrate entspricht offenbar der Gedächtnislosigkeit der Exponentialverteilung.

### 2.4.3 Pareto-Verteilung

Wir beginnen mit einem Anwendungsbeispiel. Aus einer Grundgesamtheit von Haushalten werde auf zufällige Weise ein Haushalt ausgewählt und sein Einkommen mit  $X$  bezeichnet. Dann ist  $X$  eine Zufallsvariable. Für jedes  $x$  entspricht der Wert der Verteilungsfunktion  $F(x)$  von  $X$  dem Anteil der Haushalte, die ein Einkommen von  $x$  oder weniger beziehen.  $S(x) = 1 - F(x)$  entspricht dem Anteil der Haushalte mit einem Einkommen größer als  $x$ .

Pareto<sup>6</sup> untersuchte die Einkommensverteilungen wohlhabender Haushalte und stellte dabei fest, dass die empirischen Verteilungen – besonders für große Einkommen  $x$  – der approximativen Beziehung

$$\ln(S(x)) \approx A - \alpha \ln x \tag{2.20}$$

genügen. Dies ist eine (approximative) lineare Gleichung in den logarithmierten Größen  $\ln(S(x))$  und  $\ln x$ . Trägt man in einem doppelt logarithmischen Koordinatensystem den Anteil  $S(x)$  gegen  $x$  ab, so ergibt sich näherungsweise eine Gerade mit Steigung  $-\alpha$ .

---

<sup>6</sup>Vilfredo Pareto (1848–1923)

Beschränken wir uns nun auf die Verteilung der Einkommen, die ein bestimmtes Niveau  $c$  übersteigen, und nehmen an, dass (2.20) mit geeigneten Zahlen  $A$  und  $\alpha$  exakt gilt:  $\ln(S(x)) = A - \alpha \ln x$  für  $x > c$ . Dies ist zur Gleichung  $S(x) = e^A x^{-\alpha}$  äquivalent, also zur Gleichung

$$F(x) = 1 - S(x) = 1 - e^A x^{-\alpha} \quad \text{für } x > c.$$

Wegen der Beschränkung der Grundgesamtheit auf Haushalte mit Mindesteinkommen  $c$  setzen wir  $F(x) = 0$  für  $x \leq c$ . Speziell gilt  $F(c) = 0$  und daher  $e^A c^{-\alpha} = 1$ , d.h.  $e^A = c^\alpha$ . Insgesamt folgt

$$F(x) = \begin{cases} 0, & \text{falls } x < c, \\ 1 - \left(\frac{c}{x}\right)^\alpha, & \text{falls } x \geq c. \end{cases}$$

Eine Zufallsvariable  $X$  mit dieser Verteilungsfunktion heißt **Pareto-verteilt** mit Parametern  $c > 0$  und  $\alpha > 0$ , in Zeichen:  $X \sim \text{Par}(\alpha, c)$ . Der Träger von  $X \sim \text{Par}(\alpha, c)$  ist  $T_X = [c, \infty[$ , die Dichte

$$f(x) = \begin{cases} 0, & \text{falls } x < c, \\ \frac{\alpha}{c} \left(\frac{c}{x}\right)^{\alpha+1} = \alpha c^\alpha x^{-\alpha-1}, & \text{falls } x \geq c. \end{cases}$$

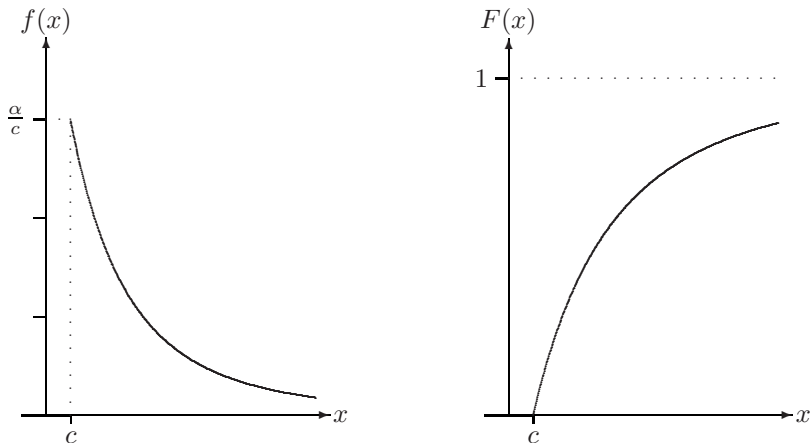


Abbildung 2.20: Dichte und Verteilungsfunktion der Pareto-Verteilung  $\text{Par}(\alpha = 2.5, c = 4)$

*Beispiel 2.25:* Von den Studierenden an einer exklusiven Privatuniversität sei bekannt, dass jeder ein verfügbares Monatseinkommen von mindestens 1500 € hat. Es sei gemäß  $Par(\alpha = 2.1, c = 1500)$  verteilt ( $\hookrightarrow$  SPSS).

- a) Berechnen Sie den Median der Einkommensverteilung.  
 b) Wie viel Prozent der Studierenden haben 2500 € oder mehr zur Verfügung?

zu a) Gesucht ist also  $x_{0.5}$ . Es gilt:

$$\begin{aligned} F(x_{0.5}) &= 1 - \left(\frac{1500}{x_{0.5}}\right)^{2.1} = 0.5 \\ \left(\frac{1500}{x_{0.5}}\right)^{2.1} &= 0.5 \\ \frac{1500}{x_{0.5}} &= 0.5^{\frac{1}{2.1}} \\ x_{0.5} &= 2086.60. \end{aligned}$$

Der Median beträgt 2086.60 €. zu b) Wir berechnen

$$1 - F(2500) = 1 - \left(1 - \left(\frac{1500}{2500}\right)^{2.1}\right) = 0.3421.$$

34.2% der Studierenden verfügen über ein Einkommen von mehr als 2500 €.

**Erwartungswert und Varianz** Wir wollen noch Erwartungswert und Varianz einer  $Par(\alpha, c)$ -Verteilung bestimmen. Dazu berechnen wir  $E[X^n]$  für beliebiges  $n \in \mathbb{N}$ ,

$$\begin{aligned} E[X^n] &= \int_c^\infty x^n \alpha c^\alpha x^{-\alpha-1} dx \\ &= \alpha c^\alpha \int_c^\infty x^{n-\alpha-1} dx \\ &= \alpha c^\alpha \frac{1}{n-\alpha} x^{n-\alpha} \Big|_c^\infty, \quad \text{falls } \alpha > n, \\ &= \alpha c^\alpha \frac{1}{\alpha-n} c^{n-\alpha} = \frac{\alpha c^n}{\alpha-n}. \end{aligned}$$

Die Integration erfordert, dass  $\alpha > n$  ist. Für **Erwartungswert** und **Varianz** einer Pareto-Verteilung erhalten wir

$$\mu = E[X] = \frac{\alpha c}{\alpha - 1}, \quad \text{falls } \alpha > 1,$$

$$V[X] = E[X^2] - \mu^2 = \frac{\alpha c^2}{\alpha - 2} - \left(\frac{\alpha c}{\alpha - 1}\right)^2,$$

$$V[X] = \frac{\alpha c^2}{(\alpha - 2)(\alpha - 1)^2}, \quad \text{falls } \alpha > 2.$$

*Beispiel 2.25 (Fortsetzung): Erwartungswert und Standardabweichung der Pareto-Verteilung mit  $\alpha = 2.1$  und  $c = 1500$  berechnet man wie folgt:*

$$\begin{aligned} E[X] &= \frac{2.1 \cdot 1500}{1.1} = 2863.64 \text{ €}, \\ V[X] &= \frac{2.1 \cdot 1500^2}{0.1 \cdot 1.1^2} = 39\,049\,586.78 \text{ €}^2, \\ \sqrt{V[X]} &= 6248.97 \text{ €}. \end{aligned}$$

Ist  $X \sim \text{Par}(\alpha, c)$  und ist  $c' > c$ , so gilt für  $x > c'$

$$P(X > x \mid X > c') = \frac{P(X > x)}{P(X > c')} = \frac{\left(\frac{c}{x}\right)^\alpha}{\left(\frac{c}{c'}\right)^\alpha} = \left(\frac{c'}{x}\right)^\alpha.$$

Die **bedingte Verteilung** von  $X$  unter der Bedingung  $X > c'$  ist also wieder eine Pareto-Verteilung. Sie hat denselben Parameter  $\alpha$  wie die nicht bedingte Verteilung; ihr anderer Parameter ist  $c'$ .

*Beispiel 2.26: Die Einkommen (in €) der Mitglieder eines noblen Golfklubs seien  $\text{Par}(\alpha, c)$ -verteilt mit  $\alpha = 2.4$  und  $c = 200\,000$ . Wie ist das Einkommen derjenigen Teilgesamtheit der Mitglieder verteilt, deren Einkommen mindestens  $c' = 300\,000$  € beträgt?*

*Nach obiger Überlegung ist dieses Einkommen  $\text{Par}(\alpha = 2.4; c' = 300\,000)$ -verteilt.*

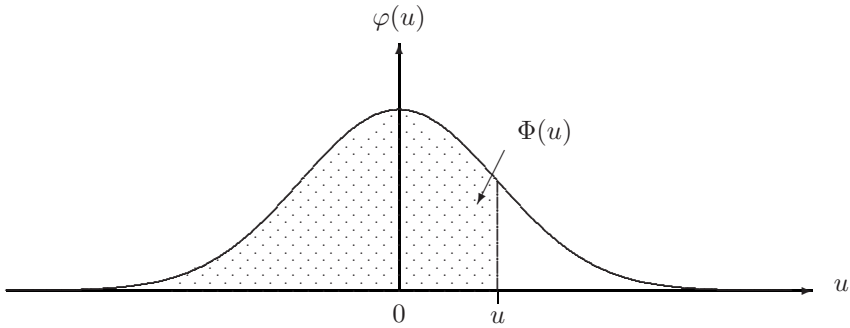
## 2.4.4 Normalverteilung

Eine Zufallsvariable  $X$  heißt **normalverteilt** oder **Gauß-verteilt**<sup>7</sup> mit Parametern  $\mu \in \mathbb{R}$  und  $\sigma^2 > 0$ , in Zeichen:  $X \sim N(\mu, \sigma^2)$ , wenn ihre Dichte durch

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R},$$

gegeben ist.

<sup>7</sup>Carl Friedrich Gauß (1777–1855)

Abbildung 2.21: Dichte  $\varphi(u)$  der Standard-Normalverteilung

Im Spezialfall  $\mu = 0$  und  $\sigma^2 = 1$  heißt die Verteilung **Standard-Normalverteilung**. In diesem Buch bezeichnen wir eine standard-normalverteilte Zufallsvariable mit  $U$ ,  $U \sim N(0, 1)$ , und ihre Dichte mit  $\varphi(u)$ . Es ist

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}, \quad u \in \mathbb{R}.$$

Zwischen der Normalverteilungsdichte  $f$  mit Parametern  $\mu$  und  $\sigma^2$  und der Standard-Normalverteilungsdichte  $\varphi$  besteht offenbar die folgende Beziehung:

$$f(x) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right), \quad x \in \mathbb{R}.$$

Die Formel entspricht der Formel (2.5) für die Dichte einer affin-linearen transformierten Variablen. Auf die affin-lineare Transformationsbeziehung zwischen  $U$  und  $X$  werden wir unten genauer eingehen.

**Eigenschaften der Dichtefunktion** Die Dichtefunktion  $\varphi(u)$ ,  $u \in \mathbb{R}$ , der Standard-Normalverteilung hat folgende wesentliche Eigenschaften (vgl. Abbildung 2.21):

1.  $\varphi(u)$  hat sein Maximum an der Stelle  $u = 0$  mit  $\varphi(0) = \frac{1}{\sqrt{2\pi}} = 0.39894$ .
2.  $\varphi(u)$  ist symmetrisch zu  $c = 0$ , d.h.  $\varphi(-u) = \varphi(u)$  für alle  $u \in \mathbb{R}$ .
3.  $\varphi(u)$  ist positiv für alle  $u \in \mathbb{R}$  und geht sowohl für  $u \rightarrow \infty$  als auch für  $u \rightarrow -\infty$  gegen 0.
4.  $\varphi(u)$  besitzt Wendepunkte an den Stellen  $u = 1$  und  $u = -1$ .

Insbesondere ist eine standard-normalverteilte Zufallsvariable  $U \sim N(0, 1)$  unimodal mit eindeutigem Modus 0 und nimmt Werte auf der ganzen reellen Achse an. Aus den Eigenschaften von  $\varphi$  folgen unmittelbar die Eigenschaften der Dichte  $f(x)$  einer **beliebigen Normalverteilung**  $N(\mu, \sigma^2)$ , nämlich:

1.  $f(x)$  hat sein Maximum an der Stelle  $x = \mu$  mit

$$f(\mu) = \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} = \frac{1}{\sigma} \cdot 0.3989.$$

2.  $f(x)$  ist symmetrisch zu  $\mu$ , d.h.  $f(\mu - x) = f(\mu + x)$  für alle  $x \in \mathbb{R}$ .
3.  $f(x)$  ist positiv für alle  $x \in \mathbb{R}$  und geht sowohl für  $x \rightarrow \infty$  als auch für  $x \rightarrow -\infty$  gegen 0.
4.  $f(x)$  besitzt Wendepunkte an den Stellen  $x = \mu + \sigma$  und  $x = \mu - \sigma$ .

Eine normalverteilte Zufallsvariable  $X \sim N(\mu, \sigma^2)$  ist unimodal mit eindeutigem Modus  $\mu$  und nimmt beliebig große und kleine Werte an.

**Verteilungsfunktion** Die zur Standard-Normalverteilung gehörende Verteilungsfunktion wird mit  $\Phi$  bezeichnet,

$$\Phi(u) = \int_{-\infty}^u \varphi(t) dt,$$

vgl. Abbildung 2.23. Dieses Integral lässt sich nicht durch bekannte Funktionen wie Polynome oder Exponentialfunktionen ausdrücken oder sonst auf einfache Weise berechnen. Eine Wertetabelle für  $\Phi(u)$  findet man als Tabelle 4 im Tabellenanhang. Manche Taschenrechner geben ebenfalls die Werte von  $\Phi(u)$  aus.

Dennoch lassen sich einige **Eigenschaften** der **Verteilungsfunktion**  $\Phi$  direkt ableiten:

1.  $\Phi$  ist **streng monoton wachsend**, da

$$\Phi'(u) = \varphi(u) > 0.$$

2. Da die Dichtefunktion  $\varphi$  symmetrisch zu 0 ist, folgt für die Verteilungsfunktion

$$\Phi(u) = 1 - \Phi(-u), \quad u \in \mathbb{R}.$$

Insbesondere ist  $\Phi(0) = \frac{1}{2}$ .



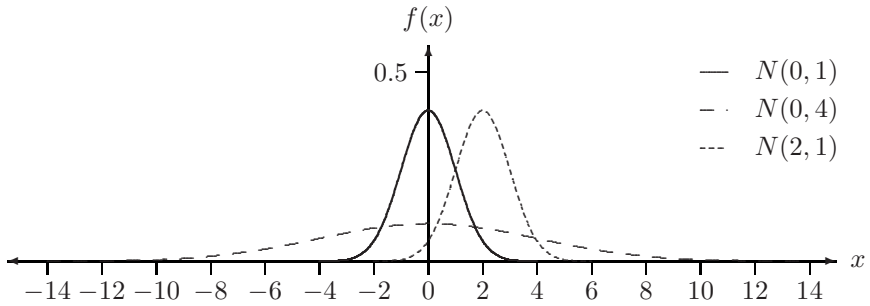


Abbildung 2.22: Dichten verschiedener Normalverteilungen

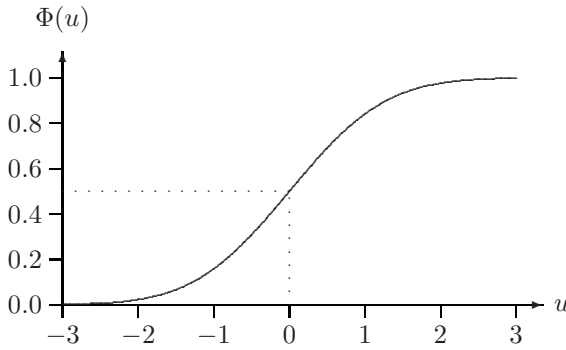


Abbildung 2.23: Verteilungsfunktion  $\Phi(u)$  der Standardnormalverteilung

Da die Verteilungsfunktion  $\Phi$  streng monoton wachsend ist, besitzt sie eine Umkehrfunktion. Für jedes  $p \in ]0, 1[$  erhält man deshalb das  $p$ -Quantil der Standard-Normalverteilung

$$u_p = Q(p) = \Phi^{-1}(p)$$

als eindeutige Lösung der Gleichung  $\Phi(u) = p$ . Es gilt hier

$$\begin{aligned} \Phi(\Phi^{-1}(p)) &= p && \text{für } 0 < p < 1, \\ \Phi^{-1}(\Phi(u)) &= u && \text{für } -\infty < u < \infty. \end{aligned}$$

Aus der Eigenschaft  $\Phi(u) = 1 - \Phi(-u)$  folgt für die Quantilfunktion die entsprechende Eigenschaft

$$\boxed{\Phi^{-1}(p) = -\Phi^{-1}(1 - p), \quad \text{d.h.} \quad u_p = -u_{1-p}.}$$

Die Quantile der Standard-Normalverteilung erhält man aus Tabelle 4 im Anhang, indem man sie „rückwärts“ liest. Einige besonders häufig verwendete Quantile  $u_p$  sind in Tabelle 3 zusammengestellt.

**Erwartungswert und Varianz** einer standard-normalverteilten Zufallsvariablen  $U \sim N(0, 1)$  lauten

$$\boxed{E[U] = 0, \quad V[U] = 1.}$$

BEWEIS  $E[U] = 0$  folgt sofort aus der Symmetrie der Dichte zu 0. Die Varianz von  $U$  berechnet man mit partieller Integration wie folgt:

$$\begin{aligned} V[U] &= E[U^2] - (E[U])^2 = E[U^2] \\ &= \int_{-\infty}^{\infty} u^2 \varphi(u) du = \int_{-\infty}^{\infty} u \cdot [u \cdot \varphi(u)] du \\ &= -u \cdot \varphi(u) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \varphi(u) du \\ &= (-0 + 0) + 1 = 1 \end{aligned}$$

Hierbei wurde verwendet, dass

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

die Stammfunktion  $-\varphi(u)$  besitzt und dass

$$\lim_{u \rightarrow \infty} u \cdot \varphi(u) = \lim_{u \rightarrow -\infty} u \cdot \varphi(u) = 0.$$

□

*Beispiel 2.27:* Es sei  $U \sim N(0, 1)$ -verteilt ( $\leftrightarrow$  Calc/SPSS).

- Man bestimme  $P(U > 0.5)$  und  $P(-1 \leq U \leq 2)$ .
- Man bestimme den 20%- und den 70%-Punkt von  $U$ .
- Man bestimme  $b$  so, dass  $P(-b \leq U \leq b) = 0.9$ .

zu a) Wir berechnen

$$\begin{aligned} P(U > 0.5) &= 1 - P(U \leq 0.5) = 1 - \Phi(0.5) \\ &= 1 - 0.6915 = 0.3085, \\ P(-1 \leq U \leq 2) &= P(U \leq 2) - P(U \leq -1) = \Phi(2) - \Phi(-1) \\ &= \Phi(2) - (1 - \Phi(1)) = 0.9772 - 1 + 0.8413 \\ &= 0.8185. \end{aligned}$$

Dabei wurden die Werte von  $\Phi(2)$  und  $\Phi(1)$  der Tabelle 4 entnommen.

zu b) Gesucht wird das 0.2-Quantil (bzw. das 0.7-Quantil). Es muss also gelten:  $P(U \leq u) = \Phi(u) = 0.2$  (bzw. 0.7). Tabelle 3 liefert  $u_{0.2} = -u_{0.8} = -0.8416$ . Aus Tabelle 4 erhält man durch lineare Interpolation  $u_{0.7} = 0.5244$ .

zu c) Es gilt

$$\begin{aligned} P(-b \leq U \leq b) &= P(U \leq b) - P(U \leq -b) = \Phi(b) - \Phi(-b) \\ &= \Phi(b) - (1 - \Phi(b)) = 2 \cdot \Phi(b) - 1. \end{aligned}$$

Daraus folgt  $2 \cdot \Phi(b) - 1 = 0.9$ , also  $\Phi(b) = 0.95$ . Aus Tabelle 3 ergibt sich  $b = 1.6449$ .

Im Folgenden betrachten wir wieder eine Normalverteilung  $N(\mu, \sigma^2)$  und diskutieren die affin-lineare Transformation, durch die sie mit der Standard-Normalverteilung  $N(0, 1)$  verknüpft ist. Grundlegend sind die folgenden Aussagen:

1. Ist  $X \sim N(\mu, \sigma^2)$ , so gilt für die **standardisierte Variable**

$$\boxed{\frac{X - \mu}{\sigma} \sim N(0, 1) .}$$

2. Ist  $U \sim N(0, 1)$  und sind  $\mu \in \mathbb{R}$  und  $\sigma > 0$ , so gilt für die mit  $\mu$  und  $\sigma$  **transformierte Zufallsvariable**

$$\boxed{\sigma U + \mu \sim N(\mu, \sigma^2) .}$$

**BEWEIS** Die erste Aussage weist man so nach: Die Verteilungsfunktion der standardisierten Variablen ist

$$\begin{aligned} F_{\frac{X-\mu}{\sigma}}(u) &= P\left(\frac{X-\mu}{\sigma} \leq u\right) \\ &= P(X \leq \sigma u + \mu) = F_X(\sigma u + \mu) \\ &= \int_{-\infty}^{\sigma u + \mu} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = \Phi(u) \quad \text{für } u \in \mathbb{R}. \end{aligned}$$

Hier haben wir beim Integrieren

$$\frac{x - \mu}{\sigma} = y \quad \text{mit} \quad x = \mu + \sigma y, \quad \frac{dy}{dx} = \frac{1}{\sigma} \quad \text{und} \quad dx = \sigma dy$$

substituiert. Da die Verteilungsfunktion von  $\frac{X-\mu}{\sigma}$  gleich  $\Phi$  ist, ist

$$\frac{X - \mu}{\sigma} \sim N(0, 1).$$

Die zweite Aussage zeigt man auf ähnliche Weise. □

Aus der ersten Aussage folgt

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \quad \text{für } x \in \mathbb{R}.$$

Wenn  $X \sim N(\mu, \sigma^2)$  ist, lässt sich die **Verteilungsfunktion** von  $X$  durch die Standard-Normalverteilungsfunktion  $\Phi$  ausdrücken. Für die Dichte von  $X$  gilt

$$f_X(x) = F'_X(x) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right),$$

und das  **$p$ -Quantil** ist gemäß (2.4) gleich

$$x_p = \mu + \sigma u_p,$$

wobei  $u_p$  das  $p$ -Quantil der Standard-Normalverteilung bezeichnet.

Wir können nun leicht den **Erwartungswert** und die **Varianz** einer normalverteilten Zufallsvariablen  $X \sim N(\mu, \sigma^2)$  bestimmen.  $X$  ist die affin-lineare Transformation einer standard-normalverteilten Zufallsvariablen  $U$ ,

$$X = \mu + \sigma U, \quad U \sim N(0, 1),$$

also gilt

$$E[X] = \mu + \sigma E[U] = \mu.$$

Dies folgt auch unmittelbar aus der Symmetrie der  $N(\mu, \sigma^2)$ -Verteilung zu  $\mu$ . Weiter ist

$$V[X] = V[\mu + \sigma U] = \sigma^2 V[U] = \sigma^2.$$

Die beiden Parameter  $\mu$  und  $\sigma^2$  einer Normalverteilung lassen sich also als **Erwartungswert** bzw. **Varianz** interpretieren.

**Schiefe und Kurtosis** Da eine Normalverteilung symmetrisch zu  $\mu$  ist, sind sowohl die Momentenschiefe  $\gamma_1$  als auch die Quartilsschiefe  $\gamma_1^Q$  gleich null. Für die Wölbung  $\gamma_2$  ergibt sich mittels mehrfacher partieller Integration

$$\gamma_2 = \int_{-\infty}^{\infty} u^4 \varphi(u) du = 3.$$

Beispiel 2.28: Es sei  $X \sim N(\mu = 2, \sigma^2 = 25)$  ( $\hookrightarrow$  Calc/SPSS).

- a) Man berechne  $P(X > 3)$  und  $P(0 \leq X \leq 2)$ .  
 b) Man bestimme den 80%- und 95%-Punkt von  $X$ .  
 c) Man bestimme  $a$  so, dass  $P(2 - a \leq X \leq 2 + a) = 0.9$  ist.

zu a) Es ist

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) = 1 - F_X(3) \\ &= 1 - \Phi\left(\frac{3-2}{5}\right) = 1 - \Phi(0.2) \\ &= 1 - 0.5793 = 0.4207, \\ P(0 \leq X \leq 2) &= P(X \leq 2) - P(X \leq 0) = F_X(2) - F_X(0) \\ &= \Phi\left(\frac{2-2}{5}\right) - \Phi\left(\frac{0-2}{5}\right) = \Phi(0) - \Phi(-0.4) \\ &= \Phi(0) - (1 - \Phi(0.4)) \\ &= 0.5 - 1 + 0.6554 = 0.1554. \end{aligned}$$

zu b) Gesucht sind die Quantile  $x_{0.8}$  und  $x_{0.95}$ . Man berechnet sie als

$$\begin{aligned} x_{0.8} &= \mu + \sigma u_{0.8} = 2 + 5 \cdot 0.8416 = 6.2081, \\ x_{0.95} &= \mu + \sigma u_{0.95} = 2 + 5 \cdot 1.6449 = 10.2245. \end{aligned}$$

zu c) Es ist

$$\begin{aligned} 0.9 &= P(2 - a \leq X \leq 2 + a) \\ &= \Phi\left(\frac{2+a-2}{5}\right) - \Phi\left(\frac{2-a-2}{5}\right) \\ &= \Phi\left(\frac{a}{5}\right) - \Phi\left(\frac{-a}{5}\right) = 2\Phi\left(\frac{a}{5}\right) - 1. \end{aligned}$$

Daraus folgt  $\Phi\left(\frac{a}{5}\right) = 0.95$ ,  $\frac{a}{5} = u_{0.95} = 1.6449$  und damit  $a = 8.2245$ .

Unter dem **zentralen Schwankungsintervall** der Breite  $2a$  versteht man das Intervall

$$[\mu - a, \mu + a].$$

Für  $X \sim N(\mu, \sigma^2)$  gilt

$\begin{aligned} P(X \in [\mu - a, \mu + a]) &= P(\mu - a \leq X \leq \mu + a) \\ &= \Phi\left(\frac{\mu+a-\mu}{\sigma}\right) - \Phi\left(\frac{\mu-a-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{a}{\sigma}\right) - \Phi\left(\frac{-a}{\sigma}\right) = 2\Phi\left(\frac{a}{\sigma}\right) - 1. \end{aligned}$
--

Besonders häufig verwendete zentrale Schwankungsintervalle einer  $N(\mu, \sigma^2)$ -Verteilung sind die so genannten **Ein-, Zwei- und Drei-Sigma-Bereiche**. Sie lauten:

$$\begin{aligned} P(\mu - \sigma \leq X \leq \mu + \sigma) &= 2\Phi(1) - 1 = 0.6826, \\ P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &= 2\Phi(2) - 1 = 0.9544, \\ P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= 2\Phi(3) - 1 = 0.9974. \end{aligned}$$

Man vergleiche die durch die Normalverteilung gegebenen Wahrscheinlichkeiten dieser Intervalle mit den allgemeinen Wahrscheinlichkeitsschranken, die die Tschebyscheff-Ungleichung (Abschnitt 2.2.3) liefert: Der Zwei-Sigma-Bereich einer beliebigen Verteilung hat mindestens die Wahrscheinlichkeit 0.7500 und der Drei-Sigma-Bereich die Wahrscheinlichkeit 0.8889. Man sieht daran, dass die Annahme einer zugrundeliegenden Normalverteilung schärfere Aussagen über die Wahrscheinlichkeiten erlaubt.

Normalverteilungen eignen sich in vielen Anwendungsfällen gut dazu, um so genannte **Fehlerverteilungen** zu beschreiben. Hierbei kann es sich um einen **Messfehler** handeln, d.h. die zufällige Abweichung eines Messwerts vom „wahren Wert“ der zu messenden Größe. Dies spielt insbesondere in den Naturwissenschaften und der Technik eine wichtige Rolle. Auch lassen sich oft **Fertigungs- oder Produktionsfehler** (das sind Abweichungen von vorgegebenen Sollgrößen) gut durch normalverteilte Zufallsvariable beschreiben.

Auch im Bereich der Wirtschafts- und Sozialwissenschaften wird die Normalverteilung gerne verwendet, vor allem als Verteilung einer „Störgröße“, nämlich der Differenz zwischen dem beobachteten Wert eines Merkmals und seinem durch ein theoretisches Modell vorhergesagten Wert. Dies lässt sich dann rechtfertigen, wenn die beobachtete Abweichung vom theoretischen Wert durch eine Vielzahl von unabhängigen kleinen Abweichungen zustandekommt; vgl. den Zentralen Grenzwertsatz (Abschnitt 3.2.4).

*Beispiel 2.29: Das tatsächliche Gewicht  $X$  einer (zufällig ausgewählten) Kaffeepackung in Gramm sei normalverteilt mit  $\mu = 1000$  und  $\sigma = 10$ . Wie groß ist die Wahrscheinlichkeit, dass das tatsächliche Gewicht geringer als 985 Gramm ist?*

*Es ist  $X \sim N(\mu = 1000, \sigma^2 = 100)$ . Für die gesuchte Wahrscheinlichkeit ergibt sich*

$$\begin{aligned} P(X < 985) &= \Phi\left(\frac{985 - 1000}{10}\right) \\ &= \Phi(-1.5) = 1 - \Phi(1.5) \\ &= 1 - 0.9332 = 0.0668. \end{aligned}$$

*Beispiel 2.30:* Sei  $K_0 = 40 \text{ €}$  der Schlusskurs einer bestimmten Aktie an der Frankfurter Börse am heutigen Tag, und bezeichne  $K$  den morgigen Schlusskurs der Aktie. Über die Verteilung der Ein-Tagesrendite (in %)

$$R = \frac{K - K_0}{K_0} \cdot 100$$

trifft ein Anleger die Annahme  $R \sim N(0, \sigma^2 = 16)$ . Wie groß ist dann die Wahrscheinlichkeit, dass der Schlusskurs  $K$  am morgigen Tag  $38 \text{ €}$  oder weniger beträgt?

Aus den Annahmen folgt, dass

$$K = K_0 + \frac{K_0 R}{100} \sim N\left(\mu_K = 40, \sigma_K^2 = \frac{40^2 \cdot 16}{100^2} = 0.4^2 \cdot 16\right),$$

also

$$\begin{aligned} P(K \leq 38) &= \Phi\left(\frac{38 - 40}{0.4 \cdot 4}\right) = \Phi(-1.25) = 1 - \Phi(1.25) \\ &= 1 - 0.8944 = 0.1056. \end{aligned}$$

Die Normalverteilungsannahme erweist sich in vielen wirtschaftswissenschaftlichen Anwendungen als problematisch. Eine Aktienrendite wie im Beispiel 2.30 ist schon deshalb nicht exakt normalverteilt, weil sie nicht kleiner als  $-100\%$  (das entspricht dem Totalverlust) werden kann. Weiter beobachtet man in den meisten Anwendungsfällen, dass empirische Renditeverteilungen sowohl in der Mitte der Verteilung als auch an den Flanken der Verteilung stark von der Normalverteilung abweichen.

Ob und inwiefern die Annahme einer Normalverteilung eine hinreichend genaue Approximation darstellt, ist im konkreten Anwendungsfall sorgfältig zu prüfen. Auf dieses Problem kommen wir im Abschnitt 6.4.2 zurück.

### 2.4.5 Lognormalverteilung

Ist  $X$  normalverteilt mit den Parametern  $\mu \in \mathbb{R}$  und  $\sigma^2 > 0$ , so nennt man  $Y = e^X$  **lognormalverteilt** mit diesen Parametern, in Zeichen:

$$Y \sim LN(\mu, \sigma^2).$$

Offensichtlich ist eine nichtnegative Zufallsvariable  $Y$  genau dann lognormalverteilt, wenn  $\ln Y$  normalverteilt ist. Wir leiten zunächst für gegebene Werte der Parameter  $\mu$  und  $\sigma^2$  die Verteilungsfunktion und die Dichte von  $Y$  her. Für  $y > 0$  gilt

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(e^X \leq y) \\ &= P(X \leq \ln y) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right), \end{aligned}$$

da  $X \sim N(\mu, \sigma^2)$  verteilt ist. Insgesamt erhalten wir

$$F_Y(y) = \begin{cases} 0, & \text{falls } y \leq 0, \\ \Phi\left(\frac{\ln y - \mu}{\sigma}\right), & \text{falls } y > 0. \end{cases}$$

Um den Median  $y_{0.5}$  von  $Y$  zu ermitteln, setzen wir  $F_Y(y_{0.5}) = 0.5$ ; es folgt

$$\frac{\ln y_{0.5} - \mu}{\sigma} = 0, \quad \text{also} \\ y_{0.5} = e^\mu.$$

Durch Differenzieren von  $F_Y(y)$  erhält man die Dichte von  $Y$ . Es ist

$$f_Y(y) = \begin{cases} 0, & \text{falls } y \leq 0, \\ \frac{1}{\sigma y} \varphi\left(\frac{\ln y - \mu}{\sigma}\right), & \text{falls } y > 0. \end{cases}$$

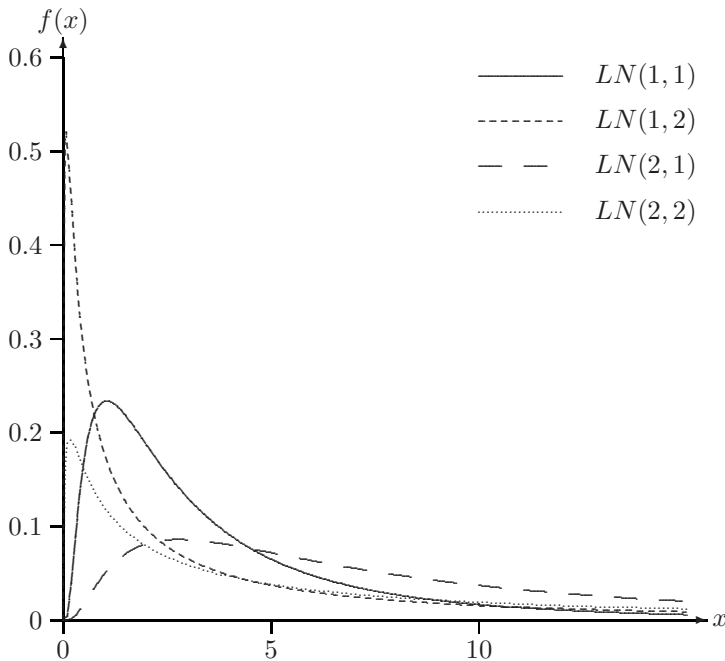


Abbildung 2.24: Dichtefunktionen verschiedener Lognormalverteilungen



Man rechnet leicht aus, dass die Dichte an der Stelle  $y = e^{\mu - \sigma^2}$  ihr einziges Maximum besitzt. Die Lognormalverteilung ist daher unimodal mit eindeutigen **Modus**

$$y_M = e^{\mu - \sigma^2}.$$

**Erwartungswert und Varianz** Bei den Parametern  $\mu$  und  $\sigma^2$  einer lognormalverteilten Zufallsvariablen  $Y$  ist Vorsicht geboten: Sie sind gleich dem Erwartungswert und der Varianz von  $X = \ln Y$ , aber keineswegs von  $Y$  selbst. Vielmehr gilt

$$\begin{aligned} E[Y] &= e^{\mu + \frac{\sigma^2}{2}}, \\ V[Y] &= e^{2\mu + \sigma^2} (e^{\sigma^2} - 1). \end{aligned}$$

Häufig wird die lognormale Verteilung mit Hilfe zweier anderer Größen parametrisiert, nämlich mit ihrem Median  $\xi = y_{0.5}$  und dem Quotienten

$$\eta = \frac{y_{0.5}}{y_M}$$

aus Median und Modus. Dann ist

$$\xi = e^\mu > 0, \quad \eta = e^{\sigma^2} > 1.$$

Es gilt  $\mu = \ln \xi$  und  $\sigma^2 = \ln \eta$ , also

$$F_Y(y) = \begin{cases} 0, & \text{falls } y \leq 0, \\ \Phi\left(\frac{\ln y - \ln \xi}{\sqrt{\ln \eta}}\right), & \text{falls } y > 0. \end{cases}$$

Mit den Parametern  $\xi$  und  $\eta$  lassen sich Erwartungswert und Varianz offenbar so ausdrücken:

$$E[Y] = \xi\sqrt{\eta}, \quad V[Y] = \xi^2\eta(\eta - 1).$$

Jede Lognormalverteilung ist rechtsschief. Als Momentenschiefe ergibt sich

$$\gamma_1 = \sqrt{\eta - 1}(\eta + 2) > 0,$$

da  $\eta > 1$  ist. Für die Wölbung gilt

$$\gamma_2 = \eta^4 + 2\eta^3 + 3\eta^2 - 3 > 3.$$

*Beispiel 2.31:* Die Verteilung der verfügbaren Einkommen (in 1000 €) in einer Gesamtheit von Haushalten sei durch  $LN(\mu = 0.9, \sigma^2 = 0.1)$  gegeben. Man berechne Erwartungswert, Standardabweichung, Median und Modus dieser Einkommensverteilung.

Bezeichne  $Y$  das Einkommen eines zufällig ausgewählten Haushaltes. Der Modus ist  $y_M = e^{0.9-0.1} = 2.2255$ , der Median  $\xi = e^{0.9} = 2.4596$ , der Gestaltparameter  $\eta = e^{0.1} = 1.1052$ . Weiter erhalten wir

$$\begin{aligned} E[Y] &= \xi\sqrt{\eta} = 2.5857, \\ \sqrt{V[Y]} &= \sqrt{\xi^2\eta(\eta-1)} = 0.8387. \end{aligned}$$

Die Lognormalverteilung hat wichtige Anwendungen in der Analyse von Renditeverteilungen am Aktienmarkt; siehe Beispiel 3.14 im Abschnitt 3.2.4 und das Zahlenbeispiel im Kapitelanhang ( $\leftrightarrow$  Calc/SPSS).

### 2.4.6 Übersicht über einige spezielle Verteilungen

Die folgende Übersicht enthält Erwartungswert, Varianz und Träger aller bisher eingeführten speziellen Verteilungen. Für alle außer der hypergeometrischen Verteilung sind auch die Momentenschiefe und die Wölbung aufgeführt.

$X \sim$	$E[X]$	$V[X]$	$T_X$	$\gamma_1$	$\gamma_2$
$B(n, \pi)$	$n\pi$	$n\pi(1 - \pi)$	$\{0, 1, \dots, n\}$	$\frac{1 - 2\pi}{\sqrt{n\pi(1 - \pi)}}$	$\frac{1 - 6\pi(1 - \pi)}{n\pi(1 - \pi)} + 3$
$H(n, N, M)$	$n \frac{M}{N}$	$\frac{M}{n} \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N - n}{N - 1}$	$\{x   x \in \mathbb{N} \cup \{0\}, 0 \leq x \leq M, 0 \leq n - x \leq N - M\}$		
$Po(\mu)$	$\mu$	$\mu$	$\mathbb{N} \cup \{0\}$	$\frac{1}{\sqrt{\mu}}$	$\frac{1}{\mu} + 3$
$G(\pi)$	$\frac{1}{\pi}$	$\frac{1 - \pi}{\pi^2}$	$\mathbb{N}$	$\frac{2 - \pi}{\sqrt{1 - \pi}}$	$\frac{9 - 9\pi + \pi^2}{1 - \pi}$
$N(0, 1)$	0	1	$\mathbb{R}$	0	3
$N(\mu, \sigma^2)$	$\mu$	$\sigma^2$	$\mathbb{R}$	0	3
$R(\alpha, \beta)$	$\frac{\alpha + \beta}{2}$	$\frac{(\beta - \alpha)^2}{12}$	$[\alpha, \beta]$	0	$\frac{9}{5}$
$Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$[0, \infty[$	2	9
$Par(\alpha, c)$	$\frac{\alpha c}{\alpha - 1}, \alpha > 1$	$\frac{\alpha c^2}{(\alpha - 2)(\alpha - 1)^2}, \alpha > 2$	$[c, \infty[$	$\frac{2(\alpha + 1)\sqrt{\alpha - 2}}{(\alpha - 3)\sqrt{\alpha}}, \alpha > 3$	$\frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)} + 3, \alpha > 4$
$LN(\mu, \sigma^2)$	$e^{\mu + \frac{\sigma^2}{2}}$	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$	$[0, \infty[$	$\sqrt{e^{\sigma^2} - 1} (e^{\sigma^2} + 2)$	$e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3$

## 2.5 Ergänzungen

### 2.5.1 Borel-Mengen, Verteilung einer Zufallsvariablen

Mit Hilfe der Verteilungsfunktion  $F$  einer Zufallsvariablen  $X$  kann man besonders leicht die Wahrscheinlichkeit dafür angeben, dass  $X$  in ein gegebenes Intervall fällt. Aus der Verteilungsfunktion lässt sich jedoch auch für allgemeinere Mengen  $B$  von reellen Zahlen die Wahrscheinlichkeit

$$P(X \in B) = P(\{\omega | X(\omega) \in B\})$$

ableiten. Beispielsweise, wenn  $B$  die Vereinigung von disjunkten Intervallen ist, ergibt sich  $P(X \in B)$  als Summe der Wahrscheinlichkeiten der einzelnen Intervalle. Man kann beweisen, dass durch die Verteilungsfunktion von  $X$  die Wahrscheinlichkeiten  $P(X \in B)$  für alle  $B \in \mathcal{B}$  bestimmt sind. Dabei ist  $\mathcal{B}$  das Mengensystem der so genannten **Borel-Mengen**; dies sind die Mengen, die sich durch wiederholte Ereignisoperationen aus Intervallen bilden lassen. Die Borel-Mengen bilden eine Ereignisalgebra (vgl. Abschnitt 1.1.2). Darüber hinaus besitzen sie folgende Eigenschaft: Die Vereinigung von abzählbar vielen Mengen in  $\mathcal{B}$  ist wieder eine Menge in  $\mathcal{B}$ . Die Borel-Mengen bilden demnach eine **Ereignis- $\sigma$ -Algebra**.

Alle für Anwendungen interessanten Mengen sind Borel-Mengen. (Es ist sogar ausgesprochen schwierig, überhaupt eine Menge zu finden, die keine Borel-Menge ist.) Die Funktion

$$B \mapsto P(X \in B), \quad B \in \mathcal{B},$$

heißt **Wahrscheinlichkeitsverteilung**, kurz: **Verteilung**, von  $X$ . Die Verteilung von  $X$  ist durch die Verteilungsfunktion

$$x \mapsto P(X \in ] - \infty, x]), \quad x \in \mathbb{R},$$

schon vollständig festgelegt.

Die Verteilung einer diskreten Zufallsvariablen  $X$  ist bereits durch die Einzelwahrscheinlichkeiten  $p_j = P(X = x_j)$ ,  $j = 1, 2, 3, \dots$  bestimmt. Die Verteilung einer stetigen Zufallsvariablen  $X$  ist durch ihre Dichte  $f(x)$ ,  $x \in \mathbb{R}$ , eindeutig festgelegt.

### 2.5.2 Erwartungswert einer Wette als subjektive Wahrscheinlichkeit

Im Abschnitt 1.4.2 wurde bereits auf die Begründung subjektiver Wahrscheinlichkeiten durch Wetten hingewiesen.

Eine Wette biete die Auszahlung von 100€, falls ein bestimmtes Ereignis  $A$  eintritt, andernfalls eine Auszahlung von 0€. Die Auszahlung ist dann eine diskrete Zufallsvariable mit zwei Ausprägungen:  $P(X = 100) = \pi$  und  $P(X = 0) = 1 - \pi$ , mit  $x_1 = 100$ ,  $p_1 = \pi$ ,  $x_2 = 0$  und  $p_2 = 1 - \pi$ . Der Erwartungswert beträgt

$$E[X] = x_1 \cdot p_1 + x_2 \cdot p_2 = 100\pi + 0(1 - \pi) = 100\pi.$$

Die subjektive Wahrscheinlichkeit für das Eintreten des Ereignisses  $A$  lässt sich ermitteln, indem der Spieler befragt wird, welchen Betrag er maximal bereit sei, für die Wette einzusetzen.

Man geht davon aus, dass der Spieler „risikoneutral“ ist, das bedeutet, dass sein maximaler Einsatz gleich dem Erwartungswert einer Wette ist. Ist eine Person also bereit, einen Einsatz von  $M$  zu leisten, so kann man aus der Beziehung

$$M = E[X] = 100\pi$$

ableiten, dass ihre subjektive Wahrscheinlichkeit  $\pi$  den Wert  $\pi = \frac{M}{100}$  hat.

Derartige Überlegungen spielen bei vielen wirtschaftlichen Fragestellungen eine Rolle. Beim Abschluss einer Versicherung etwa hängt die Höhe der Versicherungsprämie, die der Versicherungsnehmer maximal zu zahlen bereit ist, von der Wahrscheinlichkeit ab, die er dem Eintritt des Versicherungsfalles zubilligt.

### **Ergänzende Literatur zu Kapitel 2:**

Die in den Literaturhinweisen zu Kapitel 1 aufgeführten Bücher enthalten Ausführungen über Zufallsvariablen und deren Verteilungen. Zahlreiche spezielle diskrete und stetige Verteilungen werden in Johnson et al. (2005, 1994, 1995) sowie in Patil et al. (1975) ausführlich behandelt.

## 2.6 Anhang: Verwendung von Excel/Calc und SPSS

Für konkrete Anwendungen der Wahrscheinlichkeitsrechnung und für statistische Auswertungen stehen leistungsfähige Computerprogramme zur Verfügung. Weit verbreitet sind die Softwarepakete Microsoft® Excel (bzw. sein Freeware-Pendant OpenOffice.org Calc) und SPSS.<sup>8</sup>

In diesem Kapitelanhang sowie in Anhängen zu den Kapiteln 4 bis 7 werden wir deshalb einige knappe Hinweise geben, wie die zuvor dargestellten Berechnungen und statistischen Verfahren am Computer mit Hilfe von Excel/Calc oder SPSS durchgeführt werden können. Dabei beziehen wir uns auf Beispiele des Lehrbuchtextes. Es empfiehlt sich für den Leser, diese Beispiele am Computer nachzurechnen. Daten zu den Beispielen sind im Internet unter [www.wisostat.uni-koeln.de/statblb](http://www.wisostat.uni-koeln.de/statblb) verfügbar.

Eine knappe Einführung in die Verwendung von Excel ist im Lehrbuch „Deskriptive Statistik und Wirtschaftsstatistik“ (Mosler und Schmid, 2009) zu finden. Dort wird der Umgang mit Excel-Tabellen und die Durchführung von Aufgaben der deskriptiven Statistik mit Excel beschrieben.

Da die Vorgehensweisen in Excel und Calc beinahe identisch sind, werden im weiteren Verlauf die Calc-Notationen verwendet. Auf funktionale Unterschiede wird an den entsprechenden Stellen hingewiesen. Insbesondere sei angemerkt, dass Dezimalzahlen in Calc mit einem Komma, in SPSS jedoch mit einem Punkt geschrieben werden. Für Details über die verschiedenen Versionen von Calc und SPSS sei auf deren Online-Hilfefunktionen sowie auf die am Ende dieses Abschnitts genannte Literatur verwiesen.

Im Folgenden bezeichnen Angaben in dieser SCHRIFT Befehle und Buttons aus dem Calc- bzw. SPSS-Menü. Namen von Funktionen einschließlich ihrer Argumente werden in dieser Schriftart notiert. Text, der per Hand eingegeben werden muss, erscheint in dieser Schrift. (Rechen-) Befehle in der Kommandozeile werden zum Beispiel als  $= A1 + A2$  hervorgehoben.

**Calc** Alle Zahlen können vom Benutzer für die Ausgabe „gerundet“ angezeigt werden. Im Folgenden runden wir die Ergebnisse auf vier Dezimalstellen (FORMAT / ZELLEN / ZAHLEN  $\leftrightarrow$  Auswahl von ZAHL bei KATEGORIE  $\leftrightarrow$  Eingabe von 4 in Feld NACHKOMMASTELLEN und OK). Die Zwischenrechnungen führt Calc dabei stets mit allen verfügbaren Dezimalstellen durch.

**SPSS** In der VARIABLENANSICHT kann unter DEZIMALSTELLEN eingestellt werden, wieviele Nachkommastellen die zu einer bestimmten Variablen gehörenden Zahlen aufweisen sollen.

<sup>8</sup>Calc sowie weitere OpenOffice-Programme können unter [www.de.openoffice.org](http://www.de.openoffice.org) heruntergeladen werden.

Für viele Aufgaben der Wahrscheinlichkeitsrechnung und der schließenden Statistik bieten die beiden Softwarepakete Funktionen an, die im Rahmen von „Menüs“ aufgerufen werden können.

**Calc** Die im Folgenden verwendeten Funktionen von Calc sind, solange nicht anders erklärt, im Menü unter EINFÜGEN/FUNKTION (Kategorie STATISTIK) zu finden. Weiterhin weisen die auf Wahrscheinlichkeitsverteilungen bezogenen Funktionen eine bestimmte Struktur auf. Diese sieht beispielsweise bei der Normalverteilung so aus: Verteilungsfunktion und Dichte an der Stelle  $x$  erhält man mithilfe von `NORMVERT(x;Mittelwert;Standabwn;Kumuliert)`, wobei

`NORMVERT(x;Mittelwert;Standabwn;wahr)` die Verteilungsfunktion und `NORMVERT(x;Mittelwert;Standabwn;falsch)` die Dichte aufruft.

`NORMINV(p;Mittelwert;Standabwn)` berechnet die Quantilfunktion an der Stelle  $p$ .

Bei einer diskreten Verteilung wird die Wahrscheinlichkeitsfunktion statt der Dichte berechnet (s.u.). Anstelle von `wahr` kann man auch jeden beliebigen Wert  $\neq 0$  eingeben, anstelle von `falsch` genügt auch eine 0.

**SPSS** Um Daten mit SPSS zu analysieren, beginnt man so: Man liest eine Datenmatrix in die DATENANSICHT ein, definiert sodann eine Variable in der VARIABLENANSICHT und weist ihr Daten zu. Dabei muss zumindest eine metrisch skalierte Variable in der VARIABLENANSICHT definiert und in der Datenansicht initialisiert (z.B. = 0 gesetzt) werden, um diese dann als ZIELVARIABLE der statistischen Analyse verwenden zu können (vgl. Angele (2010)).

Die im Folgenden verwendeten Funktionenfamilien findet man im Menü unter TRANSFORMIEREN/BERECHNEN. Namen von Funktionen besitzen in SPSS einen typischen Aufbau, der am Beispiel der Binomialverteilung verdeutlicht werden soll.

`CDF.BINOM(q,n,p)`  $\cong$  CDF (CUMULATIVE DISTRIBUTION FUNCTION) bedeutet Zugriff auf die Verteilungsfunktion,

`PDF.BINOM(q,n,p)`  $\cong$  PDF (PROBABILITY DENSITY FUNCTION) Zugriff auf die Dichtefunktion bzw. Wahrscheinlichkeitsfunktion,

`IDF.BINOM(q,n,p)`  $\cong$  IDF (INVERSE DISTRIBUTION FUNCTION) Zugriff auf die Quantilfunktion,

`RV.BINOM(q,n,p)`  $\cong$  RV (RANDOM VARIABLE) generiert binomialverteilte Zufallszahlen.

**Zur Binomialverteilung** (Beispiel 2.15, Seite 77):

**Calc** Binomialwahrscheinlichkeiten lassen sich mit der Funktion `BINOMVERT(Anzahl Erfolge; Versuche; Erfolgswahrscheinlichkeit; Kumuliert)` berechnen. Im Beispiel ist  $P(X \geq 2)$  gesucht.

Es gilt  $P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$ . Die Lösung ergibt sich für  $X \sim B(3, 0.3)$  aus

$$\begin{aligned} P(X \geq 2) &= 1 - \text{BINOMVERT}(0;3;0,3;\text{falsch}) \\ &\quad - \text{BINOMVERT}(1;3;0,3;\text{falsch}) \\ &= 0.216, \end{aligned}$$

oder alternativ aus

$$P(X \geq 2) = 1 - \text{BINOMVERT}(1;3;0,3;\text{wahr}) = 0.216.$$

**SPSS** In SPSS lässt sich die gesuchte Wahrscheinlichkeit mit den Funktionen CDF.BINOM(q,n,p) oder PDF.BINOM(q,n,p) bestimmen:

$$\begin{aligned} P(X \geq 2) &= 1 - \text{PDF.BINOM}(0,3,0,3) - \text{PDF.BINOM}(1,3,0,3) \\ &= 1 - \text{CDF.BINOM}(1,3,0,3) \\ &= 0.216 \end{aligned}$$

Zur Berechnung der Bernoulli-Verteilung enthält SPSS auch die Funktionen CDF.BERNOULLI(q,p) bzw. PDF.BERNOULLI(q,p).

**Calc** Für die Berechnung der Binomialverteilung steht in Calc außerdem die Funktion B(Anzahl Versuche; Erfolgswahrscheinlichkeit; a; b) für  $P(a \leq X \leq b)$  zur Verfügung. Mittels  $a = 0$  kann man auch die kumulierte Wahrscheinlichkeit  $P(X \leq b)$  berechnen. Im obigen Beispiel gilt dann:

$$P(X \geq 2) = 1 - \text{B}(3;0,3;0;1) = \text{B}(3;0,3;2;3) = 0.216$$

Außerdem steht in Calc zusätzlich noch die Funktion KRITBINOM(Versuche; Erfolgswahrscheinlichkeit; Alpha) zur Verfügung. Dies ist die Quantilfunktion der Binomialverteilung. Sie gibt das kleinste  $k$  an, für das  $P(X \leq k) \geq \alpha$  gilt. Für  $X \sim B(n = 3, \pi = 0.3)$  und  $\alpha = 0.8$  berechnet man:

$$k = \text{KRITBINOM}(3;0,3;0,8) = 2$$

Beachten Sie, dass Excel die Funktion B nicht unterstützt!

**Zur Poissonverteilung** (Beispiel 2.17, Seite 80):

**Calc** Wahrscheinlichkeiten Poisson-verteilter Zufallsvariablen werden in Calc mit der Funktion POISSON(x;Mittelwert;Kumuliert) berechnet. Dabei entspricht x der Anzahl der Fälle, während Mittelwert für den Erwartungswert  $\mu$  steht. Im Beispiel 2.17 ist  $X \sim Po(2.5)$ . Man erhält

$$\begin{aligned} P(X = 1) &= \text{POISSON}(1;2,5;\text{falsch}) = 0.2052, \\ P(X \geq 3) &= 1 - P(X \leq 2) = 1 - \text{POISSON}(2;2,5;\text{wahr}) = 0.4562. \end{aligned}$$



**SPSS** Zur Berechnung von Poisson-Wahrscheinlichkeiten können die Funktionen CDF.POISSON(q,mittel) und PDF.POISSON(q,mittel) aufgerufen werden. In SPSS entspricht q der Anzahl  $x$  und mittel dem Mittelwert  $\mu$  der Poisson-Verteilung. Für die Beispielaufgabe erhält man

$$P(X = 1) = \text{PDF.POISSON}(1,2.5) = 0.2052,$$

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - \text{CDF.POISSON}(2,2.5) = 0.4562.$$

**Zur hypergeometrischen Verteilung** (Beispiel 2.20, Seite 87):

Gesucht ist  $P(X \geq 1)$ , wenn  $X \sim H(n = 2, N = 10, M = 3)$ .

**Calc** Die Lösung dieser Aufgabe ergibt sich unter Benutzung der Funktion HYPGEOMVERT(X;n;M;N) als

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \text{HYPGEOMVERT}(0;2;3;10) = 0.5333.$$

Bei Calc (und Excel) ist es nicht möglich, kumulierte Wahrscheinlichkeiten der hypergeometrischen Verteilung direkt abzurufen. Man kann sich aber mit einer einfachen Tabelle behelfen; vgl. `Beispiel22.xls`.

**SPSS** Zur Lösung derselben Aufgabe mit SPSS wird die Funktion PDF.HYPER(q,gesamt,stichpr,treffer) benutzt. Anders als Calc erlaubt es SPSS zusätzlich, mit der Funktion CDF.HYPER(q,gesamt,stichpr,treffer) kumulierte Wahrscheinlichkeiten der hypergeometrischen Verteilung direkt zu berechnen. Da  $P(X = 0) = P(X \leq 0)$  gilt, lässt sich die gegebene Aufgabenstellung mit jeder der beiden erwähnten Funktionen lösen:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = 1 - \text{PDF.HYPER}(0,10,2,3) \\ &= 1 - \text{CDF.HYPER}(0,10,2,3) = 0.5333 \end{aligned}$$

**Zur Exponentialverteilung** (Beispiel 2.24, Seite 98):

**Calc** Beispiel 2.24 wird in Calc mit der Funktion EXPONVERT(x;Lambda;Kumuliert) berechnet. Für die in dem Beispiel unter a) gesuchte Wahrscheinlichkeit erhält man

$$P(X > 1000) = 1 - \text{EXPONVERT}(1000; \frac{1}{800}; \text{wahr}) = 0.2865.$$

**SPSS** SPSS stellt die Funktionengruppe \*.EXP zur Verfügung. Man berechnet

$$P(X > 1000) = 1 - \text{CDF.EXP}(1000, \frac{1}{800}) = 0.2865.$$

Mit dieser Funktionengruppe kann auch Aufgabenteil b) leicht gelöst werden:

$$x_{0,9} = \text{IDF.EXP}(0.9, \frac{1}{800}) = 1842.07$$

**Zur Pareto-Verteilung** (Beispiel 2.25, Seite 101):

**Calc** Bei Calc (und Excel) ist keine Funktion für die Berechnung der Pareto-Verteilung verfügbar.

**SPSS** In SPSS ist die Funktionenfamilie \*.PARETO implementiert. Aufgabenteil a) lässt sich mittels der Quantilfunktion IDF.PARETO(p,schwelle,form) lösen:

$$F(x_{0.5}) = 0.5 \implies x_{0.5} = \text{IDF.PARETO}(0.5,1500,2.1) = 2086.60$$

Für Aufgabenteil b) verwendet man die Verteilungsfunktion CDF.PARETO(q,schwelle,form):

$$1 - F(2500) = 1 - \text{CDF.PARETO}(2500,1500,2.1) = 0.3421$$

**Zur Normalverteilung** (Beispiele 2.27 und 2.28, Seiten 106 und 109):

**Calc** Rund um die Normalverteilung stellt Calc mehrere Funktionen zur Verfügung. Zur Lösung der ersten Teilaufgabe aus Beispiel 2.27 a) kann die Verteilungsfunktion STANDNORMVERT(z) eingesetzt werden. Es gilt

$$\begin{aligned} P(U > 0.5) &= 1 - \Phi(0.5) \\ &= 1 - \text{STANDNORMVERT}(0,5) = 0.3085. \end{aligned}$$

Beachten Sie, dass die Funktion STANDNORMVERT(z) immer kumuliert ist, da sie die Verteilungsfunktion  $\Phi$  der Standardnormalverteilung darstellt. Möchte man die Dichtefunktion der Standardnormalverteilung berechnen, so benutzt man die Funktion NORMVERT(z;0;1;falsch).

Zur Ermittlung der Quantile  $x_{0.2}$  bzw.  $x_{0.7}$  mittels Calc wird auf die Quantilfunktion STANDNORMINV(Wahrsch) zurückgegriffen:

$$\begin{aligned} x_{0.2} &= \text{STANDNORMINV}(0,2) = -0.8416 \\ x_{0.7} &= \text{STANDNORMINV}(0,7) = 0.5244 \end{aligned}$$

Neben der Verteilungsfunktion und der Quantilfunktion der Standardnormalverteilung bietet Calc auch die entsprechenden Funktionen für allgemeine Normalverteilungen der Form  $N(\mu, \sigma^2)$  an.

Im Beispiel 2.28 ist  $X \sim N(\mu = 2, \sigma^2 = 25)$ . Die Wahrscheinlichkeit  $P(X > 3)$  lässt sich mit der Funktion NORMVERT(x;Mittelwert;Standabwn;Kumuliert) berechnen:

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \text{NORMVERT}(3;2;5;wahr) = 0.4207$$

Die Quantile  $x_p$  werden durch die Funktion NORMINV(Wahrsch;Mittelwert;Standabwn) berechnet,

$$\begin{aligned} x_{0.8} &= \text{NORMINV}(0,8;2;5) = 6.2081, \\ x_{0.95} &= \text{NORMINV}(0,95;2;5) = 10.2243. \end{aligned}$$

**SPSS** In SPSS lassen sich diese Aufgaben mit Hilfe der Funktionenfamilie \*.NORMAL(...) bearbeiten. Die Bestimmung von  $P(X > 3)$  lässt sich mit der Funktion CDF.NORMAL(q,mittel,stdAbw) wie folgt bewerkstelligen:

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) \\ &= 1 - \text{CDF.NORMAL}(3,2,5) = 0.4207 \end{aligned}$$

Die Berechnung der Quantile  $x_p$  kann mit der Funktion IDF.NORMAL(p,mittel,stdAbw) ähnlich wie in Calc vorgenommen werden:

$$\begin{aligned} x_{0.8} &= \text{IDF.NORMAL}(0.8,2,5) = 6.2081 \\ x_{0.95} &= \text{IDF.NORMAL}(0.95,2,5) = 10.2243 \end{aligned}$$

**Zur Lognormalverteilung** (Abschnitt 2.4.5):

**Calc** Calc bietet die Verteilungsfunktion und die Quantilfunktion der Lognormalverteilung  $LN(\mu, \sigma^2)$ . Häufig wird von Aktienkursen angenommen, dass sie lognormalverteilt sind. Sei  $S_t$  der Kurs einer bestimmten Aktie zur Zeit  $t$  und gelte

$$\begin{aligned} S_t &= \exp(\alpha + \beta \cdot t + \sigma \cdot Z_t), \quad \text{d.h.} \\ \ln(S_t) &= \alpha + \beta \cdot t + \sigma \cdot Z_t \quad \text{mit } Z_t \sim N(0, t). \end{aligned}$$

Dann ist  $S_t \sim LN(\alpha + \beta \cdot t, \sigma^2)$ . Als Beispiel sei  $\alpha = 1$ ,  $t = 2$ ,  $\beta = 0.5$  und  $\sigma = 5$  gewählt. Gesucht sei  $P(S_1) > 1$ . In Calc wird die Aufgabe mit der Funktion LOGNORMVERT(x;Mittelwert;Standabwn) gelöst. Achtung: Hier sind mit Mittelwert und Standabwn die Parameter  $\mu$  und  $\sigma^2$  der  $LN(\mu, \sigma^2)$ -Verteilung bezeichnet! Für die genannten Parameter erhalten wir dann  $\mu = \alpha + \beta \cdot t = 1 + 2 \cdot 0,5 = 2$  und damit

$$P(S_1 > 1) = 1 - P(S_1 \leq 1) = 1 - \text{LOGNORMVERT}(1;2;5) = 0.6554.$$

Die Quantile  $x_p$  der Lognormalverteilung werden mit der Funktion LOGINV(p;Mittelwert;Standabwn) bestimmt. Der Median der Lognormalverteilung ist mit den genannten Parametern

$$x_{0.5} = \text{LOGINV}(0,5;2;5) = 7.3890.$$

**SPSS** Die Lösung der obigen Aufgabe ist in SPSS unter Verwendung der Funktionenfamilie \*.LNORMAL(q,a,b) möglich. Dabei entspricht (anders als in Calc!) a dem Median  $e^\mu$  und b dem Parameter  $\sigma$  der Lognormalverteilung  $LN(\mu, \sigma^2)$  sowie q dem Argument der Verteilungsfunktion. Mit den obigen Parametern erhalten wir

$$\begin{aligned} P(S_1 > 1) &= 1 - P(S_1 \leq 1) \\ &= 1 - \text{CDF.LNORMAL}(1, \exp(2), 5) = 0.6554, \\ x_{0.5} &= \text{IDF.LNORMAL}(0.5, \exp(2), 5) = 7.3890. \end{aligned}$$

**Literatur zur Verwendung von Computerprogrammen:**

Elementare Einführungen in die Durchführung statistischer Verfahren mit Excel und SPSS bieten Monka et al. (2008) und Hafner und Waldl (2001).

Zwerenz (2007) stellt den Einsatz von Excel bei vielfältigen Aufgabenstellungen der Statistik dar und beinhaltet darüber hinaus interaktive Zahlenbeispiele und Simulationen. Eine allgemeine Einführung in Excel bieten die Broschüren RRZN (2006) und RRZN (2007).<sup>9</sup>

Als anwendungsnaher Einstieg in SPSS für Wirtschaftswissenschaftler ist Eckstein (2008) zu empfehlen. Angele (2010) leitet darüber hinaus zur Implementierung von Verfahren in SPSS an. Toutenburg und Heumann (2008) ist ein Lehrbuch der Wahrscheinlichkeitsrechnung und schließenden Statistik, das zahlreiche Beispiele zur Verwendung von SPSS mit Wirtschaftsdaten enthält.

Möglichkeiten der Auswertung von Daten mit dem Computer bieten auch interaktive Lernprogramme wie Teach/Me (Lohninger, 2002) und die Software von Schaich und Münnich (2001).

---

<sup>9</sup>Näheres im Internet unter [www.uni-koeln.de/RRZK/dokumentation/handbuecher/](http://www.uni-koeln.de/RRZK/dokumentation/handbuecher/), „Die Handbücher des RRZ Niedersachsen (RRZN)“.