Stochastik

Barth, Friedrich

München, [20]03

11. 1. 2. Definitionen und grundlegende Eigenschaften

urn:nbn:de:hbz:466:1-83580

z. B. entnehmen, daß die Wahrscheinlichkeit für einen Verlust etwas größer als 50% ist. Bei einem einzigen Spiel muß man also damit rechnen, 1 DM zu verlieren! Es wäre jedoch vorschnell, daraus zu schließen, daß das Spiel auch auf lange Sicht zu Verlusten für den Spieler führen müsse. Man kann ja nur 1 DM verlieren, die Gewinne können jedoch 1 DM, 2 DM oder sogar 3 DM betragen. Das könnte auf Dauer einen Ausgleich schaffen. Ein Spieler, der dieses Spiel sehr oft spielt, wird sich dafür interessieren, wieviel Geld er im Mittel pro Spiel gewinnen (oder verlieren) wird. Dazu kann er folgende Überlegung anstellen.
Bei n Spielen erwartet er in $\frac{1}{216} \cdot n$ Fällen 3 DM Gewinn, in $\frac{15}{216} \cdot n$ Fällen 2 DM Gewinn, in $\frac{75}{216} \cdot n$ Fällen 1 DM Gewinn und schließlich in $\frac{125}{216} \cdot n$ Fällen 1 DM Verlust. Der zu erwartende Gesamtgewinn ergibt sich also zu

$$3\,\text{DM} \cdot \tfrac{1}{216} \cdot n + 2\,\text{DM} \cdot \tfrac{15}{216} \cdot n + 1\,\text{DM} \cdot \tfrac{75}{216} \cdot n - 1\,\text{DM} \cdot \tfrac{125}{216} \cdot n.$$

Der durchschnittliche Gewinn pro Spiel ergibt sich, indem man diesen Wert durch die Anzahl n der Spiele teilt, zu

$$3\,\text{DM} \cdot \tfrac{1}{216} + 2\,\text{DM} \cdot \tfrac{15}{216} + 1\,\text{DM} \cdot \tfrac{75}{216} - 1\,\text{DM} \cdot \tfrac{125}{216} = -\tfrac{17}{216}\,\text{DM} \approx -0{,}08\,\text{DM}.$$

Der Spieler hat also im Schnitt pro Spiel einen Verlust von knapp 8 Pf zu erwarten. (Er kann sich überlegen, ob ihm die Lust am Spiel diesen Verlust wert ist.)
Man erkennt, daß sich dieser »mittlere Spielwert« als das mit den Wahrscheinlichkeiten gewichtete Mittel der Einzelgewinne ergibt.
Die Zahl $-\frac{17}{216}$ heißt nach *Christiaan Huygens* (1629–1695) **Erwartungswert** der Zufallsgröße »Gewinn beim chuck-a-luck«.
Was sagt uns dieser Wert über die Zufallsgröße?
Er bedeutet, daß man auf lange Sicht damit rechnen muß, etwa 0,08 DM pro Spiel zu verlieren. Die vorschnelle Vermutung, daß das chuck-a-luck ein Verlustspiel für den Spieler ist, hat sich also doch bestätigt, was auch zu erwarten war, weil das Spiel wirklich angeboten wird.

11.1.2. Definitionen und grundlegende Eigenschaften

Wir wollen nun die im letzten Abschnitt eingeführten Begriffe allgemein definieren.

> **Definition 168.1:** Es sei (Ω, P) ein Wahrscheinlichkeitsraum mit endlichem Ergebnisraum Ω. Jede Funktion X, die den Ergebnisraum Ω in die Menge $\mathbb{R}$ der reellen Zahlen abbildet, heißt **Zufallsgröße** X (auf Ω). Es gilt also:
>
> $$X\colon \omega \mapsto X(\omega) \quad \text{mit} \quad D_X = \Omega \quad \text{und} \quad X(\omega) \in \mathbb{R}.$$

Wir haben schon viele Zufallsgrößen* kennengelernt, ohne sie »Zufallsgröße« genannt zu haben. Wir erinnern an Augenzahl und Augensumme beim Würfeln oder an die Anzahl der gezogenen schwarzen Kugeln beim Ziehen aus einer Urne. Ein weiteres Beispiel einer Zufallsgröße liefert jede Wahrscheinlichkeitsverteilung P auf Ω vermöge der Zuordnung: $\omega \mapsto X(\omega)$ mit $X(\omega) = P(\{\omega\})$. Auch die relative Häufigkeit ist eine Zufallsgröße. Ein triviales Beispiel einer Zufallsgröße ist eine konstante Zufallsgröße, die für alle $\omega \in \Omega$ den konstanten Wert a

* Statt »Zufallsgröße« sagt man auch »Stochastik«, »zufällige Größe«, »Zufallsvariable« oder »aleatorische Größe«. Siehe auch Fußnote ** auf Seite 165.

annimmt. Man verwendet dann sowohl für die Zufallsgröße als auch für ihren einzigen Funktionswert denselben Buchstaben a.
Jede Zufallsgröße X erzeugt auf natürliche Weise eine Zerlegung von Ω. Eine Komponente dieser Zerlegung ist dabei die Menge derjenigen Ergebnisse ω, denen vermöge X derselbe Funktionswert x zugeordnet wird. Es gilt also:

$$\Omega = \bigcup_{i=1}^{n} \{\omega \mid X(\omega) = x_i\}.$$

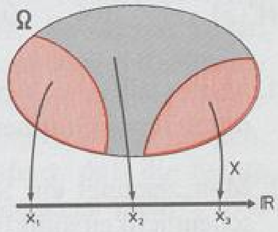

Fig. 169.1 Zerlegung von Ω durch eine Zufallsgröße X

Ω wird demnach in so viele Ereignisse zerlegt, wie es verschiedene Funktionswerte von X gibt (siehe Figur 169.1).
Betrachtet man nur eine einzige Zufallsgröße X auf Ω, so kann man Ω vergröbern, indem man die Wertemenge von X als neuen Ergebnisraum Ω' auffaßt. Das bedeutet, daß man in Ω die Ergebnisse ω in den einzelnen Komponenten der durch X erzeugten Zerlegung nicht mehr unterscheidet.
Jeder reellen Zahl x, die Wert der Zufallsgröße X ist, läßt sich damit die Wahrscheinlichkeit des Ereignisses $\{\omega \mid X(\omega) = x\}$ zuordnen. Ordnet man allen anderen reellen Zahlen x, also gerade denen, die nicht Funktionswerte von X sind, auch $P(\{\omega \mid X(\omega) = x\})$ als Wert zu – was bedeutet, ihnen den Wert 0 zuzuordnen –, so können wir damit auf ganz $\mathbb{R}$ eine Funktion definieren, die **Wahrscheinlichkeitsfunktion** W. Für den recht schwerfälligen Term $P(\{\omega \mid X(\omega) = x\})$ führen wir die Abkürzung $P(X = x)$ ein, also

$$\boxed{P(X = x) := P(\{\omega \mid X(\omega) = x\})}$$

Damit gewinnen wir

> **Definition 169.1:** Die Funktion $W\colon x \mapsto P(X = x)$, $D_W = \mathbb{R}$, heißt **Wahrscheinlichkeitsfunktion** der Zufallsgröße X. Man sagt, die Zufallsgröße X ist nach W verteilt.

Auf Grund der Definition von W erkennen wir, daß W eine Wahrscheinlichkeitsverteilung P' auf dem vergröberten Ergebnisraum Ω' ist vermöge $P'(\{x\}) := P(X = x) = W(x)$ für alle $x \in \Omega'$.
Man nennt daher W auch **Wahrscheinlichkeitsverteilung der Zufallsgröße** X, kurz auch **Verteilung der Zufallsgröße** X.
Figur 169.2 zeigt den Zusammenhang zwischen einer Zufallsgröße X und ihrer Wahrscheinlichkeitsfunktion W.

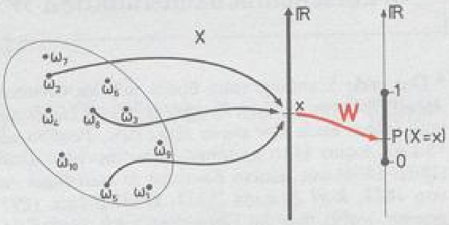

Fig. 169.2 Zusammenhang zwischen einer Zufallsgröße X und ihrer Wahrscheinlichkeitsverteilung W

Zur Veranschaulichung von Wahrscheinlichkeitsverteilungen verwendet man üblicherweise 3 Darstellungen. Wir zeigen sie an Hand einer Zufallsgröße X mit der Wahrscheinlichkeitsverteilung

x	$-1{,}5$	$-0{,}5$	$0{,}5$	2
$W(x)$	$0{,}4$	$0{,}2$	$0{,}1$	$0{,}3$

1. Man stellt die Wahrscheinlichkeitsfunktion durch ihren Graphen dar wie in Figur 170.1a).

2. Man zeichnet ein Stabdiagramm der Wahrscheinlichkeitsfunktion wie in Figur 170.1b).*

3. Man verwendet Histogramme (Figur 170.1c), die allgemein wie folgt definiert werden.**

Definition 170.1:
Auf der Zahlengeraden $\mathbb{R}$ werden $m+1$ Punkte $a_0 < a_1 < \ldots < a_m$ so gewählt, daß alle Werte $x_1, x_2, \ldots, x_n$ der Zufallsgröße X zwischen a_0 und a_m liegen. Es entstehen die $m+2$ Intervalle $]-\infty, a_0]$, $]a_0, a_1]$, $\ldots$, $]a_{m-1}, a_m]$, $]a_m, +\infty[$.
Über jedem Intervall $]a_i, a_{i+1}]$ wird ein Rechteck errichtet, dessen Flächenmaßzahl gleich der Wahrscheinlichkeit dafür ist, daß die Zufallsgröße X Werte aus diesem Intervall $]a_i, a_{i+1}]$ annimmt. Die Rechtecke auf dem ersten Intervall $]-\infty, a_0]$ und auf dem letzten Intervall $]a_m, +\infty[$ haben natürlich die Höhe 0. Rechtecke der Höhe 0 werden nicht gezeichnet. Die sich so ergebende Anordnung von Rechtecken heißt ein **Histogramm** der Zufallsgröße X bzw. der zugehörigen Wahrscheinlichkeitsfunktion W.

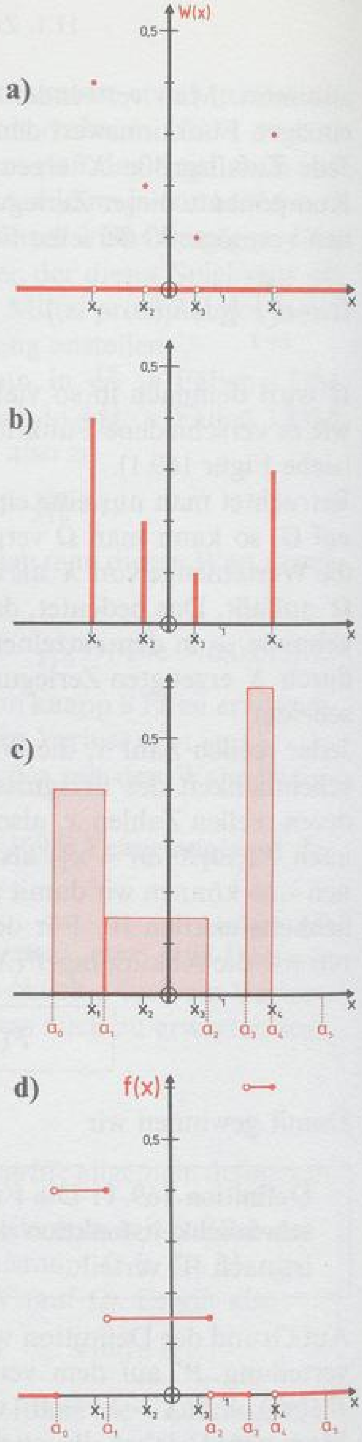

Fig. 170.1
a) Graph einer Wahrscheinlichkeitsfunktion W
b) Stabdiagramm von W
c) Ein Histogramm von W
d) Graph der zum Histogramm von **c)** gehörenden Dichtefunktion f

* Das erste Stabdiagramm findet sich im *Commercial and Political Atlas* (1786) von *William Playfair* (1759–1823). Stabdiagramme nannte man früher auch *pipe organ chart* bzw. *Panflötendiagramm*.
** Die ersten Histogramme stammen von *André Michel Guerry* (1802–1866) aus seinem *Essai sur la statistique morale de la France* von 1833. *Karl Pearson* (1857–1933) prägte 1895 das Wort *Histogramm*, wofür man im Deutschen auch *Staffelbild* sagt.

Bemerkungen:

1. Die Flächenmaßzahl eines Rechtecks hat den Wert $P(\{\omega | a_i < X(\omega) \leqq a_{i+1}\})$, kurz $P(a_i < X \leqq a_{i+1})$.
2. Die Gestalt des Histogramms hängt von der Auswahl der Intervalle ab. Es ist üblich, Intervalle von gleicher Länge (meist 1) zu wählen und sie symmetrisch um die Werte von X anzuordnen. (Vergleiche Figur 167.3 und 167.4.)

Histogramme sind keine Funktionen; man kann aus jedem von ihnen aber einen Funktionsgraphen auf folgende Art gewinnen. Man nimmt die oberen Seiten der Histogrammrechtecke als Stücke des Graphen. Nicht gezeichnete Rechtecke der Höhe 0 ergeben als Graphenstücke dann Teile der x-Achse. Weil aber eine Funktion eindeutig sein muß, müssen wir an den Rechtecksecken noch festlegen, welche Ecke Punkt des Graphen sein soll. Es wird vereinbart, daß die rechte obere Ecke Punkt des Graphen ist, die linke obere Ecke hingegen nicht*. Die so durch diesen Graphen definierte Funktion f ist also linksseitig stetig. Sie heißt **Wahrscheinlichkeitsdichtefunktion** der Zufallsgröße X, kurz **Dichtefunktion****. Ihr Term ist sehr kompliziert; es gilt nämlich

$$f(x) := \begin{cases} \dfrac{P(a_i < X \leqq a_{i+1})}{a_{i+1} - a_i} & \text{für} \quad x \in \,]a_i; a_{i+1}] \\ 0 \quad \text{sonst} \end{cases}$$

Zu jedem Histogramm gehört also genau eine Dichtefunktion; siehe Figur 170.1c) und Figur 170.1d).
Die Wahrscheinlichkeit dafür, daß die Zufallsgröße X Werte im Intervall $]a_j; a_k]$ annimmt, ist gleich der Summe der Maßzahlen der Rechtecksflächen über $]a_j; a_k]$. Diese Summe läßt sich mittels der Dichtefunktion f auch als Integral schreiben.

Es gilt nämlich: $P(a_j < X \leqq a_k) = \int\limits_{a_j}^{a_k} f(x)\,\mathrm{d}x$.

Diese Beziehung kann dazu dienen, den Namen »Dichtefunktion« plausibel zu machen. Denkt man sich nämlich einen Stab mit der eindimensionalen Dichteverteilung f, so erhält man als Masse zwischen den Punkten a_j und a_k den Wert

$m = \int\limits_{a_j}^{a_k} f(x)\,\mathrm{d}x$.

Man beachte also, daß der Funktionswert $f(x)$ der Dichtefunktion f keine Wahrscheinlichkeit darstellt, so wie die Dichte auch keine Masse ist. Nur eine *Fläche* unter der Dichtefunktion f zwischen den Grenzen a_j und a_k liefert uns eine Wahrscheinlichkeit! Es kann sogar vorkommen, daß die Funktionswerte der Dichtefunktion größer als 1 sind – wenn nämlich die Intervallbreiten beim zugehörigen Histogramm klein genug sind. Man vergleiche dazu das Histogramm von Figur 167.4. Dagegen gilt immer $\int\limits_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$.

Beim chuck-a-luck betrachteten wir den mittleren Gewinn pro Spiel auf lange

* Fällt die linke obere Ecke mit der rechten oberen Ecke des vorhergehenden Rechtecks zusammen, so gehöre sie dem Graphen an.

** Merke: D**i**chtefunktion – **l**i nksseitig stetig.

Sicht und nannten ihn den Erwartungswert der Zufallsgröße Gewinn. In Verallgemeinerung definiert man für eine beliebige Zufallsgröße X ihren mittleren Wert pro Versuch auf lange Sicht als ihren Erwartungswert gemäß

Definition 172.1: Die Zufallsgröße X habe die Wertemenge $\{x_1, x_2, \ldots, x_n\}$. Die zugehörigen Wahrscheinlichkeiten seien $W(x_1)$, $W(x_2), \ldots, W(x_n)$. Dann heißt die Zahl

$$\mu := \mathscr{E}X := \sum_{i=1}^{n} x_i W(x_i)$$

Erwartungswert der Zufallsgröße X.

Bemerkungen:

1. Statt $\mathscr{E}X$ schreibt man auch $\mathscr{E}(X)$, vor allem dann, wenn Mißverständnisse zu befürchten sind. $\mathscr{E}$ ist nämlich eine Funktion, die »Erwartung«, durch die der Zufallsgröße X eine reelle Zahl $\mathscr{E}X$ zugeordnet wird. Man unterscheidet also zwischen der Funktion $\mathscr{E}$ und dem Funktionswert $\mathscr{E}X$.
2. Der Erwartungswert einer Zufallsgröße ist das mit ihren Wahrscheinlichkeiten gewichtete arithmetische Mittel der Werte der Zufallsgröße,* also **der mittlere Wert der Zufallsgröße pro Versuch auf lange Sicht.** Das Symbol μ soll an diese Bedeutung des Erwartungswerts als Mittelwert erinnern.
3. Der Erwartungswert $\mathscr{E}X$ einer Zufallsgröße X wird im allgemeinen *nicht* dem Wertebereich der Zufallsgröße X angehören! (Beim chuck-a-luck ist $-\frac{17}{216}$ keine Gewinnzahl.)
4. Der Erwartungswert $\mathscr{E}X$ muß nicht einmal in der Nähe des wahrscheinlichsten Wertes der Zufallsgröße X liegen. (Vgl. Aufgabe 188/**15**.)
5. Die Zahl $\mathscr{E}X$ kann auch physikalisch gedeutet werden. Denkt man sich auf einem masselosen Stab an den Stellen x_i die Massen $W(x_i)$ angebracht, so stellt $\mathscr{E}X$ die Koordinate des Massenschwerpunkts dar.*
6. Wegen $W(x_i) = \sum\limits_{X(\omega) = x_i} P(\{\omega\})$ läßt sich $\mathscr{E}X$ folgendermaßen schreiben:

$$\boxed{\mathscr{E}X = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\})}$$

Mit dem Erwartungswert des Gewinns als »Wert des Spiels« schuf *Christiaan Huygens* (1629–1695) den ersten wahrscheinlichkeitstheoretischen Begriff, der lange Zeit auch der einzige blieb, mit dem Aufgaben über Spiele gelöst werden konnten. *Huygens* beginnt seine Abhandlung *van reeckening in speelen van geluck* (1657) mit 3 Sätzen über den Erwartungswert, für den er noch keinen Fachausdruck hatte. Er umschreibt ihn mit

"Het is my soo veel weerdt" – »Das ist mir soviel wert«,

was *Frans van Schooten* (um 1615–1660) kurz mit *expectatio mea* und *valor expectationis meae* ins Lateinische übersetzt.

* Diese Interpretation gab *Nikolaus Bernoulli* (1687–1759) in *De usu artis conjectandi in jure* 1709.

Wie sehr diese Wortschöpfung damals und auch heute noch der Umgangssprache widerspricht, zeigt sich in der Erklärung, die *Jakob Bernoulli* (1655–1705) in seiner *Ars Conjectandi* (1713) der *Huygens*schen Definition beifügt:

»Aus dem Gesagten ist zu entnehmen, daß das Wort *Erwartung* nicht nur in dem gewöhnlichen Sinne zu nehmen ist, in dem wir üblicherweise das erwarten oder hoffen, was für uns das Allerbeste ist; es kann vielmehr auch den Sinn haben, daß Schlimmeres uns zufällt. Erwartung bedeutet also unsere Hoffnung, das Beste zu gewinnen, soweit diese nicht durch die Furcht, Ungünstigeres zu erzielen, gemäßigt und verkleinert wird, und zwar in dem Maße, daß mit Wert der Erwartung immer das Mittel aus dem Besten, das wir erhoffen, und dem Schlimmsten, das wir befürchten, bezeichnet wird.«

In der französischen Literatur hat sich daher neben *espérance mathématique* auch *valeur moyenne* und in der englischen neben *expected value* auch *mean value* für Erwartungswert eingebürgert.

Zur Illustration des Begriffs *Erwartungswert* führen wir nun 3 Beispiele vor.

Beispiel 1: Die Zufallsgröße »Augenzahl« beim einfachen Wurf eines Laplace-Würfels hat die Wahrscheinlichkeitsfunktion

x	1	2	3	4	5	6
$W(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Damit ergibt sich

$$\mathscr{E}X = \mu = 1\cdot\tfrac{1}{6} + 2\cdot\tfrac{1}{6} + 3\cdot\tfrac{1}{6} + 4\cdot\tfrac{1}{6} + 5\cdot\tfrac{1}{6} + 6\cdot\tfrac{1}{6} = \tfrac{21}{6} = 3{,}5.$$

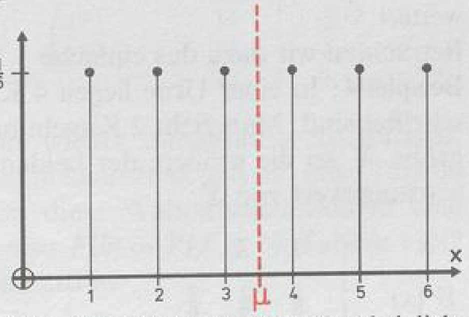

Fig. 173.1 Stabdiagramm der Wahrscheinlichkeitsfunktion der Zufallsgröße »Augenzahl eines Laplace-Würfels« mit Erwartungswert μ

Beispiel 2: Die Zufallsgröße »Augensumme« beim Wurf zweier Laplace-Würfel hat, wie Bild 26.2 veranschaulicht, die Wahrscheinlichkeitsverteilung

x	2	3	4	5	6	7	8	9	10	11	12
$W(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Damit ergibt sich

$$\mathscr{E}X = \mu = 2\cdot\tfrac{1}{36} + 3\cdot\tfrac{2}{36} + 4\cdot\tfrac{3}{36} + 5\cdot\tfrac{4}{36} + 6\cdot\tfrac{5}{36} + 7\cdot\tfrac{6}{36} + 8\cdot\tfrac{5}{36} + 9\cdot\tfrac{4}{36} + 10\cdot\tfrac{3}{36} + 11\cdot\tfrac{2}{36} + 12\cdot\tfrac{1}{36} = \tfrac{252}{36} = 7.$$

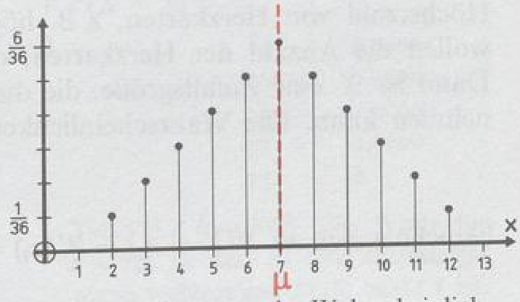

Fig. 173.2 Stabdiagramm der Wahrscheinlichkeitsfunktion der Zufallsgröße »Augensumme beim Doppelwurf« mit Erwartungswert μ

Beispiel 3: Die Zufallsgröße »Quadrat der Augenzahl« beim einfachen Wurf eines Laplace-Würfels hat die Wahrscheinlichkeitsverteilung

x	1	4	9	16	25	36
$W(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Damit ergibt sich

$$\begin{aligned}\mathscr{E}X = \mu &= 1\cdot\tfrac{1}{6} + 4\cdot\tfrac{1}{6} + \\ &\quad + 9\cdot\tfrac{1}{6} + 16\cdot\tfrac{1}{6} + \\ &\quad + 25\cdot\tfrac{1}{6} + 36\cdot\tfrac{1}{6} = \\ &= \tfrac{91}{6} = \\ &= 15\tfrac{1}{6}.\end{aligned}$$

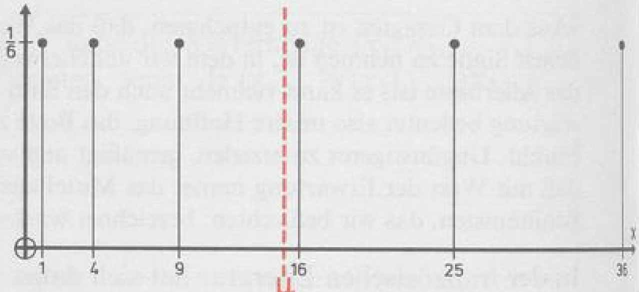

Fig. 174.1 Stabdiagramm der Wahrscheinlichkeitsfunktion der Zufallsgröße »Quadrat der Augenzahl eines Laplace-Würfels« mit Erwartungswert μ

Die Berechnung des Erwartungswerts wird besonders übersichtlich, wenn man die Wertetabelle der Wahrscheinlichkeitsverteilung um die Zeile $xW(x)$ erweitert.
Betrachten wir dazu das einfache
Beispiel 4: In einer Urne liegen 4 Kugeln, die mit den Zahlen 0, 1, 2 und 3 beschriftet sind. Man zieht 2 Kugeln nacheinander ohne Zurücklegen. Die Zufallsgröße X sei die größere der beiden gezogenen Zahlen. Wir berechnen den Erwartungswert von X:

x	1	2	3	
$W(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	
$xW(x)$	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{9}{6}$	$\Rightarrow \mathscr{E}X = \frac{14}{6} = 2\frac{1}{3}$.

11.2. Die kumulative Verteilungsfunktion einer Zufallsgröße

Einführendes Beispiel: Die 52 Karten des Bridgespiels werden auf 4 Spieler verteilt. Theodor hat dabei keine Herzkarte erhalten. Nun interessiert er sich dafür, mit welcher Wahrscheinlichkeit seine Spielgegnerin Dorothea eine bestimmte Höchstzahl von Herzkarten, z. B. höchstens 8 Herzkarten, erhalten hat. Wir wollen die Anzahl der Herzkarten, die Dorothea erhält, mit X bezeichnen. Dann ist X eine Zufallsgröße, die die Werte $x_1 = 0$, $x_2 = 1, \ldots, x_{14} = 13$ annehmen kann. Die Wahrscheinlichkeit dafür, daß sie den Wert x_i annimmt, berechnet sich zu $P(X = x_i) = W(x_i) = \dfrac{\binom{13}{x_i}\binom{26}{13-x_i}}{\binom{39}{13}}$. Die Werte der Wahrscheinlichkeitsfunktion W sind in Tabelle 175.1 wiedergegeben. Theodor interessiert sich also für Wahrscheinlichkeiten der Ereignisse »X ist höchstens so