

---

## Zufallsvariablen

### Zufallsvariablen. Merkmale

Der Begriff **Zufallsvariable** (Zv.) wird hier sinngleich mit "Beobachtungsgröße" verwendet (vgl. den Abschnitt "Beobachtungsgrößen im Beispiel"). Bei einem Messvorgang ist das Ergebnis zunächst nicht bekannt und kann daher als zufällig (in einem weitergefassten Sinn) angesehen werden.

Wir notieren Zvn. mit großen lateinischen Buchstaben:  $X, Y, Z, \dots$ , die Merkmale zumeist mit entsprechenden Kleinbuchstaben. Die Gleichung  $X = x_3$  beschreibt das Ereignis, dass  $X$  den Wert  $x_3$  annimmt.

#### ■ Beispiele

Bsp. 4:  $X$  bezeichne die Zv. "Familienstand":  $X = \textit{verheiratet}$

Bsp. 7:  $Y$  bezeichne die Zv. "Länge":  $Y = 12$  (mm)

Bsp. 10:  $Z$  bezeichne die Zv. "Wartezeit":  $Z = 3$

Die zu betrachtenden Merkmale  $x$  soll in folgendem Sinn als elementar vorausgesetzt werden: Es gibt kein von  $x$  verschiedenes Merkmal  $x'$  derart, dass das Ereignis  $X = x$  das Ereignis  $X = x'$  nach sich zieht.

Schließen sich die Ereignisse  $X = x$  und  $X = x'$  gegenseitig aus, so heißen die Merkmale  $x, x'$  unverträglich (oder: disjunkt).

### Merkmalausprägungen

Die Merkmale einer Zvn.  $X$  können unterschiedlich beschaffen sein:

- Die Werte von  $X$  sind Maßzahlen (i.a. reelle Zahlen). Man spricht von **quantitativen** Merkmalen.  
Beispiele: 2, 3, 7, 10.
- Die Werte von  $X$  lassen eine Rangfolge (Anordnung) zu. Man spricht von **ordinalen** Merkmalen.  
Beispiele: 1, 8, 9.

▸ Die Werte von  $X$  dienen der Unterscheidung von Dingen. Man spricht von **nominalen** Merkmalen.  
Beispiele: 4, 5, 6.

Ein ordinales Merkmal kann auch nominal, ein quantitatives Merkmal auch ordinal (und erst recht nominal) aufgefasst werden.

N.B.: Wird ein Merkmal durch ein Zahlwort bezeichnet, so schließt das keineswegs seine quantitative Natur ein.

## ■ Beispiele

1. Steuerklassen sind nominale Merkmale. Hausnummern sind ordinale Merkmale.

2. Die Schulnoten ('1' bis '6') sind ordinale Werte; man kann mit ihnen eine Rangfolge herstellen, aber nicht ohne weiteres auch rechnen. Nichtsdestoweniger addiert man Noten und bildet ihre arithmetischen Mittel, die dann Noten "neuer Art" bilden. Allerdings dürfte eine Gleichung wie  $3 - 2 = 5 - 4$  in der Praxis mit Fug und Recht bezweifelt werden.

3. Die Augenzahlen beim Würfelwurf sind zunächst nur nominal. Wenn der Sachzusammenhang (z.B. ein Spiel wie "Mensch ärgere dich nicht") eine Wertung nahelegt, werden sie zu ordinalen Merkmalen ('6' > '5' > ... > '1'). Erlaubt man zusätzlich z.B. ihre Addition, so kommt auch der quantitative Aspekt ins Spiel.

## Merkmalräume

Eine Menge  $\Omega$ , die alle (elementaren, untereinander unverträglichen) Merkmale einer Zvn.  $X$  enthält, nennen wir **Merkmalraum** von  $X$  (bzw. des zu  $X$  gehörigen Versuchs). Ebenfalls geläufige Bezeichnungen sind u.a.: "Stichprobenraum", "Ergebnismenge" oder "Ausfallmenge".

Bei der Modellierung eines Versuchs wird man anstreben, in den Merkmalraum nur tatsächlich auftretende Werte der Zvn. aufzunehmen, z.B. bei einem Spielwürfel:  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . In einigen Fällen kann dies jedoch zu künstlichen Einschränkungen führen, etwa im Beispiel 2 ("Hühnerei"). Es ist hier bequemer (und sinnvoll wegen des im Grunde stetigen Wertverlaufs),  $\Omega$  als reelles Intervall ( $\Omega = [0; 100]$  oder  $\Omega = \mathbb{R}$ ) zu definieren, auch wenn dabei Werte einbezogen werden, die von der Zvn. "Gewicht eines Hühnereis" niemals angenommen werden.

Unendliche Merkmalräume treten nicht nur bei stetigen Wertverläufen der Zvn. auf. Die Zv. "Wartezeit" in Beispiel 10 ist eine diskrete Größe mit möglichen Werten 1, 2, 3, ... Große Werte sind zwar unwahrscheinlich, man kann sie jedoch nicht von vornherein ausschließen. Daher ist der unendliche Merkmalraum  $\Omega = \{1, 2, 3, \dots\}$  zu Grunde zu legen.

Sprechweise: Bei einem endlichen Merkmalraum  $\Omega$  heißt die Anzahl  $m = |\Omega|$  "Anzahl der möglichen Fälle" (für den Ausgang des Versuchs).

## ■ Weitere Beispiele

Bsp. 3:  $\Omega = \mathbb{R}^+$

Bsp. 4:  $\Omega = \{ \text{'ledig'}, \text{'verheiratet'}, \text{'verwitwet'} \}$

Bsp. 8:  $\Omega = \{ \text{'sehr gut'}, \text{'gut'}, \text{'befriedigend'}, \text{'ausreichend'}, \text{'mangelhaft'}, \text{'ungenügend'} \}$

## Ereignisse und ihre Beschreibung

Mit den bisher entwickelten Sprachmitteln lassen sich auf einfache Weise Ereignisse beschreiben, die bei der Durchführung eines Versuchs realisiert werden.

Im Folgenden werden die diesbezüglichen Erklärungen am Beispiel 5 (Spielwürfel) illustriert: Zv.  $X$  (Augenzahl), Merkmalraum  $\Omega = \{1, 2, \dots, 6\}$ .

### ■ Elementarereignisse

Das Ereignis "4 wird geworfen" ist elementar, sofern bei der Zv. nur zu prüfen ist, ob das einzelne Merkmal '4' angenommen wird:  $X = 4$ .

Allgemein sind Elementarereignisse von der Form  $X = x_i$ .

### ■ Zusammengesetzte Ereignisse

Das Ereignis "Eine gerade Augenzahl wird geworfen" ist zusammengesetzt, sofern zu prüfen ist, ob das Merkmal '2' oder das Merkmal '4' oder das Merkmal '6' angenommen wird:  $X = 2 \vee X = 4 \vee X = 6$ .

Auch die Negation eines Elementarereignisses ist (i.a.) zusammengesetzt, z.B. "Es wird keine 4 geworfen" ( $X \neq 4$ ):  $X = 1 \vee X = 2 \vee X = 3 \vee X = 5 \vee X = 6$ .

Mit aussagenlogischen Verknüpfungen (Negation:  $\neg$ , "nicht"; Adjunktion:  $\vee$ , "oder"; Konjunktion:  $\wedge$ , "und"; etc.) lassen sich aus bereits definierten Ereignissen beliebige weitere zusammensetzen.

Liegen ordinale oder quantitative Merkmale vor, so bieten die Relationen  $\leq$ ,  $\geq$ ,  $<$ ,  $>$  bzw. Operationen wie  $+$ ,  $-$ ,  $\cdot$ , etc. zusätzliche Möglichkeiten der Zusammensetzung, z.B.  $X \geq 3$  ("mindestens drei Augen") oder:  $X \bmod 2 = 0$  ("gerade Anzahl von Augen").

### ■ Die Merkalmenge eines Ereignisses

Ein Ereignis lässt sich durch die Menge der Merkmale charakterisieren, die  $X$  zu seiner Realisierung annehmen kann. Z.B. gehört zum Ereignis "gerade Anzahl von Augen" die Merkalmenge  $A = \{2, 4, 6\}$ . Die **Merkalmenge** (oder: **Extension**) eines Ereignisses ist eine Teilmenge von  $\Omega$ .

#### Sprechweisen:

Ereignisse mit der Extension  $\emptyset$  heißen "unmöglich".

Ereignisse mit der Extension  $\Omega$  heißen "sicher".

Ein Merkmal aus der Extension  $A$  eines Ereignisses heißt "günstiger Fall" (für  $A$ ), entsprechend  $|A|$  "Anzahl der günstigen Fälle" des Ereignisses.

## Versuchsreihen. Relative Häufigkeit

### Versuchsreihen

Ein Versuch zu einer Zvn.  $X$  werde  $s$ -mal durchgeführt (Versuchsreihe). Wir notieren jedes auftretende Merkmal in einer Liste  $S$  (Urliste, Protokoll, **Stichprobe**):

$$S = (x_1, x_2, \dots, x_s)$$

Die natürliche Zahl  $s$  heißt **Länge** der Versuchsreihe (auch: **Umfang** der Stichprobe  $S$ ); sie kann stark variieren.

N.B.: In einer Stichprobe  $S$  müssen keinesfalls alle Elemente des zugehörigen Merkmalraums  $\Omega$  vorkommen. Andererseits können Merkmale in  $S$  mehrfach auftreten.

### Relative Häufigkeit

#### ■ Relative Häufigkeit von Merkmalen

Sei  $x \in \Omega$  irgendein Merkmal,  $S$  die Urliste einer Versuchsreihe. Wir durchmustern die in  $S$  auftretenden Merkmale der Reihe nach und prüfen, ob sie mit  $x$  übereinstimmen. Gibt es  $a$  Übereinstimmungen, so heißt  $a$  **absolute Häufigkeit** und der Quotient  $\frac{a}{s}$  **relative Häufigkeit des Merkmals  $x$  in  $S$**  (bzw. in der Versuchsreihe). Die relative Häufigkeit eines Merkmals gibt den Anteil seines Vorkommens innerhalb einer Stichprobe an.

$$\text{Relative Häufigkeit von Merkmal } x = \frac{\text{Anzahl der Vorkommen von } x \text{ in der Stichprobe}}{\text{Umfang der Stichprobe}}$$

Um diesen Ausdruck formal zu präzisieren, setzen wir  $\delta(x, y) = 1$ , falls  $x = y$ , und sonst:  $\delta(x, y) = 0$ . Damit lässt sich die relative Häufigkeit  $H_S(x)$  von Merkmal  $x$  in  $S$  wie folgt schreiben:

$$H_S(x) = \frac{1}{s} \sum_{i=1}^s \delta(x, x_i)$$

Die Bezugnahme auf  $S$  (insbesondere die Indizierung von  $H$ ) kann fortgelassen werden, wenn sie aus dem Zusammenhang hervorgeht.

Offenkundig sind folgende Eigenschaften von  $H_S$ :

1.  $0 \leq H_S(x) \leq 1$  für alle  $x \in \Omega$
2.  $\sum_{x \in \Omega} H_S(x) = 1$

Ein Merkmal  $x$  kommt in der Stichprobe  $S$  genau dann nicht vor, wenn seine relative Häufigkeit 0 beträgt. Relative Häufigkeit 1 bedeutet, dass  $x$  das einzige in  $S$  auftretende Merkmal ist. In der Summe (unter 2.) addieren sich die Anteile aller tatsächlich in  $S$  vorkommenden Merkmale zu 1 auf.

### ■ Relative Häufigkeit von Ereignissen

Es ist naheliegend, den Begriff der relativen Häufigkeit auf Ereignisse auszudehnen. Dies lässt sich auf natürliche Weise durchführen:

Die relative Häufigkeit eines Elementarereignisses  $X = x$  (dessen Merkmalmenge aus  $x$  allein besteht:  $\{x\}$ ) identifizieren wir direkt mit  $H_S(x)$ . Setzt sich das Ereignis aus mehreren Merkmalen  $x_1, \dots, x_m$  zusammen, d.h. besitzt es die Merkmalmenge bzw. Extension  $A = \{x_1, \dots, x_m\}$ , so wird definiert:

$$H_S(A) = \sum_{i=1}^m H_S(x_i)$$

Das hier verwendete Definitionsverfahren heißt **Summenregel**. Es garantiert, dass die relative Häufigkeit eines Ereignisses in einer Stichprobe  $S$  gerade der Anteil ist, den sämtliche Merkmale bilden, deren Auftreten das betreffende Ereignis nach sich zieht.

Wo notationstechnisch bequem, wollen wir Ereignisse durch ihre Merkmalmengen  $A, B, C, \dots (\subseteq \Omega)$  wiedergeben.

Es folgt eine Zusammenstellung einfacher Grundtatsachen (wobei eine feste Stichprobe  $S$  zu Grunde liegt, auf die nicht jedesmal Bezug genommen wird):

1.  $H(\emptyset) = 0, H(\Omega) = 1$
2.  $0 \leq H(A) \leq 1$  für alle Ereignisse  $A \subseteq \Omega$
3.  $A \subseteq B \implies H(A) \leq H(B)$
4.  $H(A \setminus B) + H(A \cap B) = H(A) + H(B)$

Beweis:

1.) und 2.) sind sofort ersichtlich, 3.) eine unmittelbare Konsequenz der Summenregel. 4.) ergibt sich daraus, dass die Merkmale im Durchschnitt von  $A$  und  $B$  links und rechts vom Gleichheitszeichen jeweils doppelt gezählt werden. ■

Die Eigenschaft 3 heißt Monotonie. Für disjunkte (unverträgliche) Ereignisse  $A, B$  geht Eigenschaft 4 über in:  $H(A \setminus B) = H(A) + H(B)$  (Additivität).

## ■ Erweiterung einer Stichprobe

In der Praxis kommt es häufig vor, dass eine vorliegende Stichprobe  $S = (x_1, x_2, \dots, x_s)$  durch eine zusätzliche Anzahl  $k$  von Versuchen erweitert wird:  $S^* = (x_1, \dots, x_s, x_{s+1}, \dots, x_{s+k})$ .

Wir wollen untersuchen, wie sich die relative Häufigkeit eines Merkmals bei einer solchen Erweiterung ändert.

Offenbar gilt für die absolute Häufigkeit  $a$  von  $x$  in der erweiterten Stichprobe  $S^*$ :

$$a = s \cdot H_S(x) + \delta(x, x_{s+1}) + \dots + \delta(x, x_{s+k})$$

Die relative Häufigkeit von  $x$  ergibt sich nach Division der absoluten Häufigkeit durch den Umfang  $(s + k)$  der erweiterten Stichprobe  $S^*$ :

$$H_{S^*}(x) = \frac{a}{s+k} = \frac{s}{s+k} H_S(x) + \frac{1}{s+k} \sum_{i=1}^k \delta(x, x_{s+i})$$

Hieraus folgt für die Häufigkeitendifferenz in den Stichproben  $S$  und  $S^*$ :

$$H_{S^*}(x) - H_S(x) = \frac{k}{s+k} \left( -H_S(x) + \frac{1}{k} \sum_{i=1}^k \delta(x, x_{s+i}) \right)$$

Der zweite Summand in der Klammer ist die relative Häufigkeit von  $x$  in der Reststichprobe  $(x_{s+1}, \dots, x_{s+k})$ . Da sie  $\geq 0$  und  $\leq 1$  ist, lässt sich der Klammerausdruck dem Betrage nach mit dem größeren der beiden Werte  $H_S(x)$  und  $1 - H_S(x)$  abschätzen.

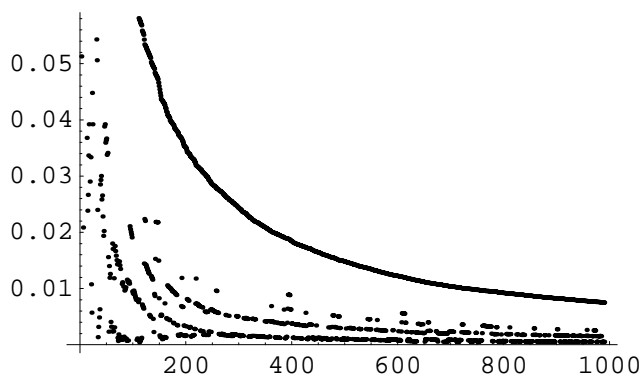
Damit ergibt sich die Ungleichung:

$$|H_{S^*}(x) - H_S(x)| \leq \frac{k}{s+k} \max\{H_S(x), 1 - H_S(x)\}$$

**Folgerung:** Der Absolutbetrag der Änderung, welche die relative Häufigkeit eines Merkmals erleidet, wenn man die zu Grunde liegende Stichprobe vom Umfang  $s$  um  $k$  Elemente erweitert, ist durch den Bruch  $\frac{k}{s+k}$  beschränkt. Bei festem  $k$  strebt dieser Betrag daher gegen Null für  $s \rightarrow \infty$ , d.h. in Abschnitten konstanter Größe werden die Schwankungen der relativen Häufigkeit beliebig klein.

**Warnung:** Dies ist *nicht* gleichbedeutend damit, dass die Häufigkeitenfolge konvergiert!

Die folgende Grafik zeigt (bei einer größeren Stichprobe) deutlich den Verlauf der Schranke (rechte Seite der obigen Ungleichung) oberhalb der Punkte, welche die Differenzbeträge der relativen Häufigkeiten (linke Seite der Ungleichung) darstellen:



## Verteilungen und ihre Darstellung

### Häufigkeitsverteilungen

Sei  $\Omega = \{x_1, x_2, \dots, x_m\}$  der Merkmalraum irgend einer Zufallsvariablen. Liefert sie in einer Versuchsreihe die Stichprobe  $S$ , so nennen wir die Folge der relativen Häufigkeiten  $h_i = H_S(x_i)$  ( $i = 1, 2, \dots, m$ ) **Häufigkeitsverteilung von  $S$**  (über dem Merkmalraum  $\Omega$ ); wir notieren sie auch als  $m$ -Tupel:  $(h_1, h_2, \dots, h_m)$ .

Der folgende Satz kennzeichnet die  $m$ -Tupel (endlichen Folgen) reeller Zahlen, die Häufigkeitsverteilungen (H-Verteilungen) einer Stichprobe sind:

#### Satz

$(h_1, h_2, \dots, h_m) \in \mathbb{R}^m$  ist genau dann eine Häufigkeitsverteilung einer Stichprobe  $S$ , wenn gilt:

- (1)  $h_i \geq 0$  für alle  $i \in \{1, 2, \dots, m\}$
- (2)  $h_1 + h_2 + \dots + h_m = 1$
- (3) Es gibt eine ganze Zahl  $s$  derart, dass alle Produkte  $s \cdot h_i$  ganz sind ( $1 \leq i \leq m$ ).

#### Beweis:

1. Es ist unmittelbar einleuchtend, dass eine H-Verteilung die Eigenschaften (1) und (2) erfüllt. Da die  $h_i$  rationale Zahlen sind, kann  $s$  als das kleinste gemeinschaftliche Vielfache der Nenner in einer Darstellung sämtlicher  $h_i$  als Brüche gewählt werden, womit auch (3) gezeigt ist.

2. Seien nun umgekehrt die Eigenschaften (1)-(3) vorausgesetzt. Wir setzen  $\Omega = \{1, 2, \dots, m\}$  und denken uns eine Stichprobe  $S$  gegeben, in welcher das Merkmal  $i$  genau  $s \cdot h_i$ -mal vorkommt (leicht zu erreichen). Dann gilt:

$|S| = \sum_{i=1}^m s \cdot h_i = s(h_1 + \dots + h_m) = s$  sowie  $H_S(i) = \frac{s \cdot h_i}{s} = h_i \geq 0$  für alle  $i \in \{1, 2, \dots, m\}$ . Das vorgegebene  $m$ -Tupel  $(h_1, h_2, \dots, h_m)$  ist somit die H-Verteilung der so konstruierten Stichprobe  $S$ . ■

### ■ Verallgemeinerung

Der Begriff der H-Verteilung lässt sich noch etwas allgemeiner fassen, wenn die im obigen Satz genannte Bedingung (3) entfällt sowie zusätzlich erlaubt wird, dass die Zahlenfolge unendlich ist.



Eine Folge reeller Zahlen  $(p_1, p_2, \dots, p_n, \dots)$  heie (diskrete) **Verteilung**, wenn gilt:

1.  $p_i \geq 0$  fr alle  $i \geq 1$
2.  $\sum_{i \geq 1} p_i = 1$ .

### Verteilung wovon?

Lassen sich die Zahlen  $p_i$  als relative Hufigkeiten elementarer Merkmale deuten, so haben wir es mit einer H-Verteilung zu tun.

Lassen die  $p_i$  sich als Wahrscheinlichkeiten von Elementarereignissen deuten, so sprechen wir von einer Wahrscheinlichkeitsverteilung (W-Verteilung).

Z.B. bentigt man bei Entscheidungen unter Risiko eine W-Verteilung ber der Menge  $B$  der Bedingungen.

## Darstellungsformen

Fr Stichproben beobachteter Merkmale und fr ihre H-Verteilungen gibt es eine Reihe unterschiedlicher Darstellungsformen. Welcher Typ von Diagramm in einer bestimmten Situation angemessen ist, hngt vom Zweck der Visualisierung und von der jeweils vorliegenden Ausprgung der Merkmale ab.

Wir behandeln hier nur kurz die wichtigsten Diagrammformen fr Stichproben und fr H-Verteilungen.

### **Darstellung von Stichproben:**

Punktediagramme

Liniendiagramme

### **Darstellung von Verteilungen:**

Stabdiagramme

Kreisdiagramme

Hufigkeitspolygone

## ■ Computergesttzte Visualisierung

Die Darstellungsformen werden im Folgenden mit Hilfe eines *Mathematica*-Pakets "Verteilungen.m" demonstriert, das es u.a. erlaubt, Zufallsstichproben ber beliebigen endlichen Merkmalrumen zu erzeugen und die H-Verteilungen numerisch und grafisch auszugeben.

Zunchst muss das Paket geladen werden:

```
<< Graphics`Graphics`
<< Modellbildung`Verteilungen`
```

Wir legen als Merkmalraum die Menge der Augenzahlen eines Spielwrfels zu Grunde und erzeugen ber ihm eine Zufallsstichprobe vom Umfang 30:

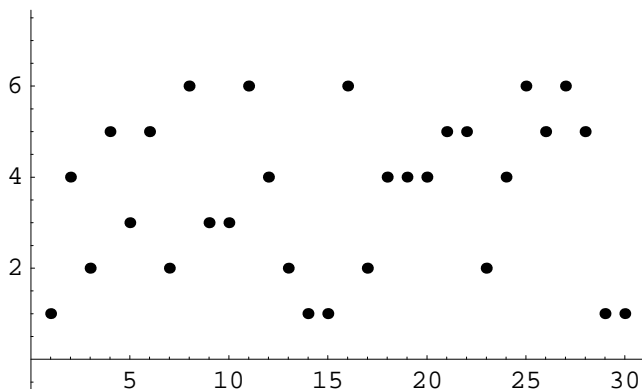
```
m = {1, 2, 3, 4, 5, 6};  
sprobe = Stichprobe[m, 30]
```

```
{1, 4, 2, 5, 3, 5, 2, 6, 3, 3, 6, 4, 2,  
1, 1, 6, 2, 4, 4, 4, 5, 5, 2, 4, 6, 5, 6, 5, 1, 1}
```

## ■ Punktediagramme

In einem rechtwinkligen Koordinatensystem gezeichnete Punktmengen. Auf der Ordinate werden die Merkmale abgetragen. Ausprägungen: quantitativ, ordinal (unter Umständen auch nominale Merkmale möglich). Die Abszisse dient i.a. als Zeitachse, d.h. die Merkmale treten in ihrer zeitlichen Reihenfolge auf:

```
PunkteDiagramm[sprobe, 2]
```

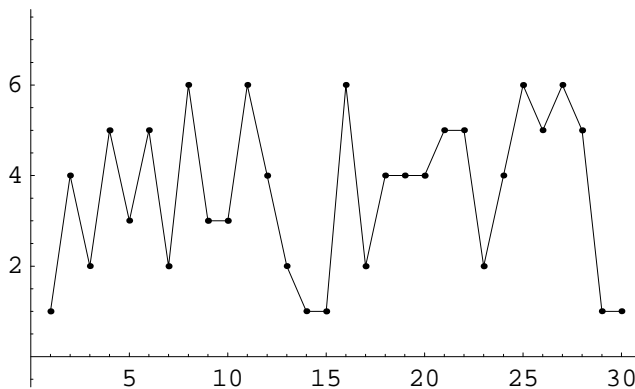


Hier ist die Stichprobe mit der Punktgröße 2 dargestellt (mögliche Punktgrößen: 1, 2, 3).

## ■ Liniendiagramme

In einem rechtwinkligen Koordinatensystem gezeichnete Kurven (Linien). Die Abszisse ist Zeitachse, auf der Ordinate werden die Merkmale abgetragen (nur quantitative Merkmale sinnvoll!). Im Prinzip wird der Merkmalraum als kontinuierlich angenommen. Die Verbindung der Datenpunkte durch Kurvenstücke (häufig: Strecken) ermöglicht es, Zwischenwerte zu schätzen bzw. Trends zu erkennen. Liniendiagramme betonen den Gesamtverlauf bzw. den Zusammenhang, der (eventuell) zwischen den Beobachtungsdaten besteht.

```
LinienDiagramm[sprobe, 1]
```



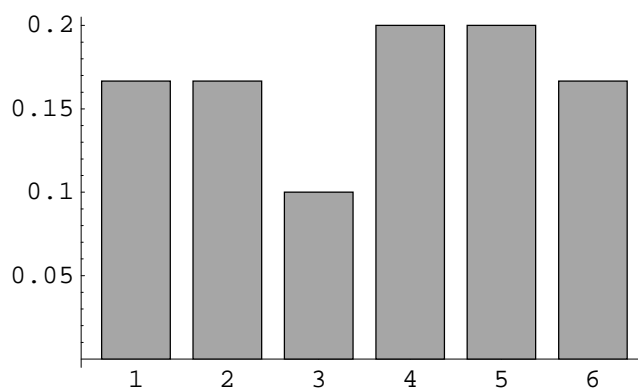
Hier ist die Stichprobe mit der Liniendicke 1 dargestellt (mögliche Liniendicken: 1, 2, 3).

### ■ Stabdiagramme

Den Begriff "Stabdiagramm" verwenden wir hier synonym zu "Histogramm". Stabdiagramme visualisieren eine Verteilung aus relativen Häufigkeiten. (Eine Vorform des Stabdiagramms, das sog. Säulendiagramm, besteht aus rechteckigen Säulen, deren Höhen die absoluten Häufigkeiten von Merkmalen in einer Stichprobe wiedergeben.)

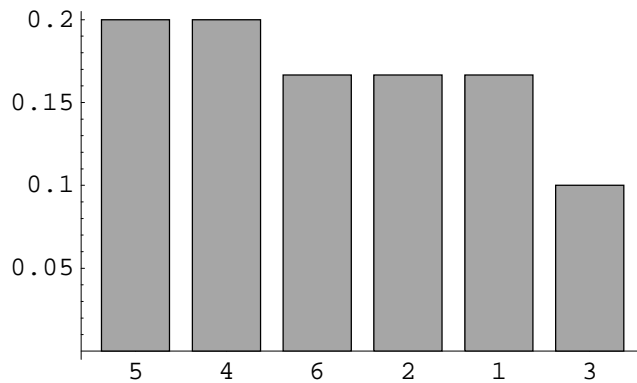
Es liegt ein rechtwinkliges Koordinatensystem zugrunde. Auf der Abszisse werden die Elemente des Merkmalraums abgetragen (alle Ausprägungen möglich!) und über ihnen je ein schmales Rechteck gezeichnet, dessen Höhe die relative Häufigkeit des Merkmals darstellt.

```
StabDiagramm[m, sprobe]
```



Nützlich ist (besonders im Fall nominaler Merkmale) das **Staffeldiagramm** (angeordnetes Stabdiagramm), das die Merkmale nach abnehmender (bzw. zunehmender) Häufigkeit darstellt:

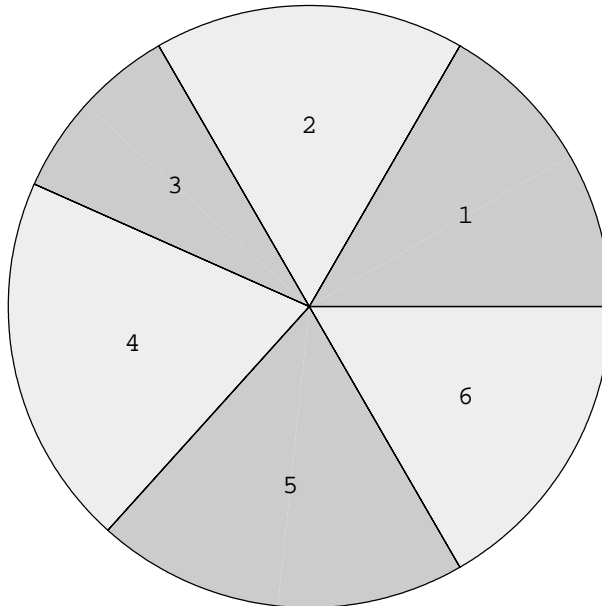
```
StaffelDiagramm[m, sprobe]
```



### ■ Kreisdiagramme

Kreisdiagramme (auch "Tortendiagramme" genannt) stellen eine Verteilung dar, indem sie jedem Merkmal einen Kreissektor ("Kuchenstück") zuordnen, dessen Anteil am Kreis der relativen Häufigkeit des betreffenden Merkmals entspricht. Diese Darstellungsart eignet sich besonders für nominale Merkmale. Grundsätzlich bietet sie sich aber immer dann an, wenn die Aufteilung eines Ganzen in Anteile gezeigt bzw. betont werden soll.

```
KreisDiagramm[m, sprobe]
```

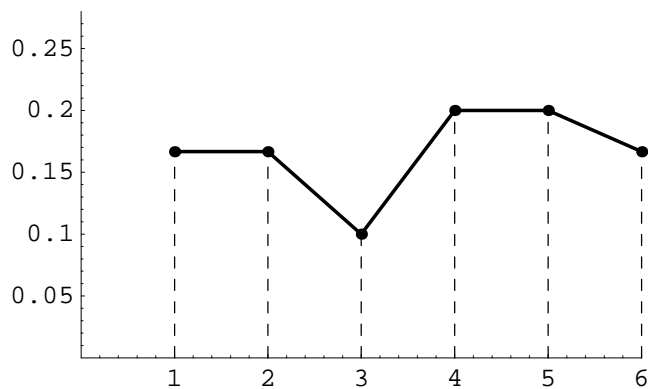


## ■ Häufigkeitspolygone

Es handelt sich um Streckenzüge in einem rechtwinkligen Koordinatensystem. Die Strecken verbinden Punkte, deren Abszissenwerte quantitative Merkmale (häufig Messwerte) und deren Ordinatenwerte die zugehörigen relativen Häufigkeiten (in der Stichprobe bzw. Messreihe) sind.

Der Merkmalraum ist bevorzugt als kontinuierlich vorzustellen, sodass die Punkte auf den Strecken als Zwischenwerte gedeutet werden können.

```
HaeufigkeitsPolygon[m, sprobe]
```



---

# Wahrscheinlichkeiten

## Elementare Wahrscheinlichkeiten

### ■ Allgemeine Vorbemerkungen

Der *Begriff der Wahrscheinlichkeit* bezieht sich auf unsichere Ereignisse bzw. Aussagen. Dazu gehören u.a.

- Ausfälle (Ergebnisse) von wiederholbaren Zufallsversuchen (z.B. Würfelwurf, Münzwurf)
- einmalige Ereignisse in der Zukunft (z.B. bei der Wettervorhersage)
- Hypothesen über Sachverhalte, die sich nicht unmittelbar überprüfen lassen (z.B. Ursache einer Krankheit).

Der Begriff "Zufallsversuch" (Z-Versuch) wird im Folgenden synonym mit dem Begriff "Versuch" gebraucht, betont dabei aber die Tatsache, dass es unsicher ist, welches Ergebnis (Merkmal) nach der Versuchsdurchführung zu beobachten sein wird.

Eine Wahrscheinlichkeit  $p$  (lat. *probabilitas*) wird durch eine reelle Zahl zwischen 0 und 1 wiedergegeben. Für Anwendungen ist zu erklären, was es bedeutet, dass ein bestimmtes Ereignis  $A$  die Wahrscheinlichkeit  $P(A) = p$  erhält.

Die Wahrscheinlichkeitsrechnung behandelt die Frage, nach welchen Regeln (bzw. mit welchen Modellen) sich Wahrscheinlichkeiten aus anderen (bereits bekannten!) Wahrscheinlichkeiten berechnen lassen. Es handelt sich im Kern um ein innermathematisches Problem.

Eine überwiegend außermathematische Frage ist die nach den "ersten" Wahrscheinlichkeiten; das sind diejenigen, die sich nicht (jedenfalls nicht ohne weiteres) auf grundlegendere Wahrscheinlichkeiten zurückführen lassen. Diese Problematik fällt in die Zuständigkeit von Physik, Statistik oder Entscheidungstheorie.

*Stochastik* ist eine zusammenfassende Bezeichnung für Wahrscheinlichkeitstheorie und Statistik.

## ■ Wahrscheinlichkeiten für elementare Ereignisse

Vom Standpunkt der Anwendung aus verdient das folgende fundamentale Problem besondere Aufmerksamkeit:

Ein Zufallsversuch habe die möglichen Ausgänge  $x_1, x_2, x_3, \dots \in \Omega$ .

Wie gelangt man zu Wahrscheinlichkeiten  $p_i$  für die zugehörigen Elementarereignisse  $X = x_i$ ?

Wir betrachten (der Einfachheit halber und ohne Beschränkung der Allgemeinheit) Z-Versuche mit zwei Ausgängen:  $\Omega = \{x_1, x_2\}$ . Es genügt, die Wahrscheinlichkeit  $p_1 = P(X = x_1)$  für das Auftreten von  $x_1$  zu bestimmen (da sich beide Wahrscheinlichkeiten zu 1 ergänzen:  $p_1 + p_2 = 1$ ).

Jemand behauptet:  $p_1 = 0.5$

Was kann eine solche Wahrscheinlichkeitsaussage bedeuten?

Es gibt mehrere (ihrer Natur nach grundsätzlich verschiedene) Möglichkeiten, einer solchen Aussage einen Sinn beizulegen.

### 1. Physikalische Deutung

Man betrachtet den Versuch als einen (im Prinzip) physikalisch beschreibbaren Vorgang. Zwar lässt sich ein individueller Ausgang nicht vorhersagen, doch aufgrund bekannter (oder angenommener) Symmetrien werden Wahrscheinlichkeiten als objektive, der Versuchsvorrichtung innewohnende Eigenschaften ("Neigungen") aufgefasst.

Beim Wurf einer Münze ist keine der beiden Seiten physikalisch ausgezeichnet. Die Annahme, die Ergebnisse "Wappen" und "Zahl" seien gleichwahrscheinlich, führt auf die Wahrscheinlichkeit  $\frac{1}{2}$ .

Versuche mit gleichwahrscheinlichen Ausgängen heißen Laplace-Versuche. Sie beruhen in den meisten Fällen auf Geräten mit eingebauter Symmetrie: Münze, Spielwürfel, Roulette, etc. Wo solche Symmetrien nicht gezielt hergestellt (und technisch kontrolliert) werden, ist die Annahme von Gleichwahrscheinlichkeit vielfach problematisch (z.B. bei der Laplace-Regel, vgl. "Entscheidungen bei Unsicherheit"). Z.B. ist das Geschlecht eines Neugeborenen ist keine gleichverteilte Zufallsvariable.

### 2. Frequentistische Deutung

Auch in dieser Deutung ist die Wahrscheinlichkeit eine objektive Eigenschaft. Sie wird jedoch gemessen (geschätzt), indem man eine längere Versuchsreihe durchführt und die relative Häufigkeit des in Frage stehenden Merkmals  $x_1$  in der Stichprobe bestimmt. Ihr Wert liegt in der Nähe der Wahrscheinlichkeit.

Eine solche empirische Messung ist auch bei Laplace-Versuchen sinnvoll. Nach einer großen Zahl von Münzwürfen wird sich eine ungefähre Gleichverteilung einstellen, und die relativen Häufigkeiten erscheinen so als Näherungswerte der theoretischen Apriori-Wahrscheinlichkeit 0.5.

Werfen wir statt einer Münze eine Reiszwecke, so liefert die physikalische Deutung keinen plausiblen theoretischen Wert, und es bleibt allein die Möglichkeit eines Massenexperiments.

Die frequentistische Deutung ist natürlich nicht anwendbar, wenn sich die Wahrscheinlichkeitsaussage auf einen einmaligen, nicht wiederholbaren Versuch bezieht (z.B. Stichwahl zweier Staatspräsidenten).

### 3. Epistemische Deutung

Die epistemische Deutung macht auch bei einmaligen Ereignissen Sinn. Ihr zufolge drückt nämlich die Wahrscheinlichkeit den Grad aus, in dem eine Person an etwas glaubt, wodurch sie als eine Eigenschaft des Wissens (gr. *episteme*) erscheint, das die betreffende Person über die Welt hat.

Die Wahrscheinlichkeit, die zu einer Person gehört, kann durch ihr Wettverhalten gemessen werden. Ist z.B. jemand bereit, 1 € auf  $x_1$  ("Wappen" einer Laplace-Münze) zu setzen, um im Erfolgsfall 2 € zurück zu erhalten, so ist diese 1:1-Wette ein Äquivalent zur Wahrscheinlichkeitsaussage  $p_1 = 0.5$ . – Ebenso gut kann man einer Wette darüber, wie die Stichwahl zweier Kandidaten ausgeht, die subjektiven Glaubensgrade der Wettenden entnehmen.

Wird statt der Münze eine Reiszwecke geworfen, so könnte man z.B. mit einer 1:2-Wette beginnen. Nach einigen Versuchen mag es ratsam erscheinen, das Wettverhältnis (und damit die Wahrscheinlichkeiten) an den Informationsstand anzupassen und auch weiterhin zu revidieren. Auf diese Weise – das kann hier nur angedeutet werden – lässt sich die frequentistische Auffassung mit der epistemischen Deutung der Wahrscheinlichkeit verbinden (sog. Bayes-Statistik).

### ■ Wahrscheinlichkeitsverteilung einer Zufallsvariablen

Sei  $X$  eine Zufallsvariable, welche die Werte (Merkmale)  $x_1, x_2, x_3, \dots \in \Omega$  annehmen kann. Die relativen Häufigkeiten  $h_i$  der Ereignisse  $X = x_i$  bilden die Häufigkeitsverteilung von  $X$  (in einer Stichprobe). In Analogie dazu wird die Folge der Wahrscheinlichkeiten  $p_i = P(X = x_i)$  **Wahrscheinlichkeitsverteilung** (kurz: **W-Verteilung**) von  $X$  genannt. Der Bezug auf eine Stichprobe entfällt hier.

## Wahrscheinlichkeiten von Ereignissen

### ■ Definition (Summenregel)

In derselben Weise wie der Begriff der relativen Häufigkeit lässt sich auch der Begriff der Wahrscheinlichkeit von Elementarereignissen (Merkmalen) auf beliebige Ereignisse (Mengen von Merkmalen) erweitern. Die naheliegende Definition erfolgt ebenfalls mit Hilfe der Summenregel:

Besteht ein Ereignis  $A (\subseteq \Omega)$  aus genau den Merkmalen  $x_1, \dots, x_n$  mit bekannten Elementarwahrscheinlichkeiten  $p_1, \dots, p_n$ , so ist die Wahrscheinlichkeit  $p$  von  $A$  definiert durch:  $p = P(A) = p_1 + \dots + p_n$ .

D.h.: Ist eine W-Verteilung über dem Merkmalraum  $\Omega$  festgelegt, so sind die Wahrscheinlichkeiten aller Ereignisse  $A$  (Teilmengen von  $\Omega$ ) bestimmt.

Beispiel:

In einem Kasten liegen vier Kugeln; drei davon sind farbig (rot, grün, gelb), eine ist weiß. Das verdeckte Ziehen einer Kugel ist ein Laplace-Versuch, d.h. eine bestimmte Kugel zu ziehen, besitzt die Wahrscheinlichkeit  $\frac{1}{4}$ . Bezeichne  $A$  das Ereignis, eine farbigte Kugel zu ziehen:  $A = \{\text{rot, grün, gelb}\}$ .

Dann gilt nach Definition:  $P(A) = P(\text{rot}) + P(\text{grün}) + P(\text{gelb}) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

### ■ Einfache Folgerungen

Das Ereignis, dass  $A$  nicht eintritt, wird **Komplementär-** oder **Gegenereignis** genannt und mit  $\bar{A}$  bezeichnet. Es gilt stets:  $P(A) + P(\bar{A}) = 1$ . Die Wahrscheinlichkeit  $P(\bar{A}) = q = 1 - p$  heißt **Gegenwahrscheinlichkeit**.



Sind  $A$  und  $B$  zwei unverträgliche Ereignisse, so bezeichnet  $A + B$  das Ereignis, dass sich (entweder)  $A$  oder  $B$  realisiert. Aus der Summenregel folgt sofort:  $P(A + B) = P(A) + P(B)$ .

Das Ereignis  $A + \bar{A}$  (im obigen Beispiel: Ziehen einer farbigen Kugel oder einer weißen Kugel) hat die Wahrscheinlichkeit 1 (**sicheres** Ereignis). Das Ereignis, eine blaue Kugel zu ziehen, hat die Wahrscheinlichkeit 0 (**unmögliches** Ereignis).

Für ein Wahrscheinlichkeitsmaß  $P$  gelten dieselben Grundaussagen wie für die relative Häufigkeit  $H_S$ :

1.  $P(\emptyset) = 0, P(\Omega) = 1$
2.  $0 \leq P(A) \leq 1$  für alle Ereignisse  $A \subseteq \Omega$
3.  $A \subseteq B \implies P(A) \leq P(B)$
4.  $P(A \cup B) + P(A \cap B) = P(A) + P(B)$

Beweis: Wörtlich wie für  $H$  (mittels Summenregel). ■

---

## Unabhängige Versuche. Produktregel

### Abhängigkeit zwischen Zufallsvariablen

In der Praxis beobachtet man häufig Zvn.  $X, Y, \dots$ , die in einem bestimmten Sinn voneinander abhängig sind.

Beispiele:

- $X = \text{Monat im Jahr}, \quad Y = \text{Niederschlagsmenge}$
- $X = \text{Nikotinkonsum}, \quad Y = \text{Lebenserwartung}$
- $X = \text{Altersgruppe}, \quad Y = \text{Gefährdung im Straßenverkehr}$

Tritt ein bestimmtes Ereignis  $X = x_1$  ein, so hat dies i.a. Einfluss darauf, welches Merkmal in der Folge bei  $Y$  beobachtet wird, d.h. die Wahrscheinlichkeit etwa für  $Y = y_1$  ist unter der Bedingung  $X = x_1$  (möglicherweise) eine andere als unter der Bedingung  $X = x_2$ .

Abhängigkeit dieser Art tritt häufig dann auf, wenn die Beobachtungen in Systemen durchgeführt werden, die physikalisch nicht gegeneinander abgeschlossen sind (an den obigen Beispielen offensichtlich).

#### ■ Ein Farbwürfel

Wir werfen einen guten Spielwürfel mit gefärbten Seiten (grün bei den Augenzahlen 3, 4 und 6, sonst rot). Zwei Zvn. werden beobachtet:  $X = \text{Farbe}, Y = \text{Augenzahl}$ .

Der Würfel werde geworfen. Wie wahrscheinlich ist eine 3? Da der Würfel gut ist (Laplace-Versuch), lautet die Antwort:  $P(Y = 3) = \frac{1}{6}$ . Wird beim selben Wurf das Ereignis  $X = \text{grün}$  beobachtet, so ändert sich die Wahrscheinlichkeit zu  $\frac{1}{3}$ . Die Bedingung wird kenntlich gemacht in der Schreibweise:

$$P(Y = 3 \mid X = \text{grün}) = \frac{1}{3}$$

Im Übrigen gilt:  $P(Y = 3 \mid X = \text{rot}) = 0$ .

## Unabhängige Versuche

Dass zwei Zvn.  $X, Y$  unabhängig sind, lässt sich durch die Forderung ausdrücken:  $P(X = x | Y = y) = P(X = x)$  für alle Merkmale  $x, y$  (die jeweils von  $X$  bzw.  $Y$  angenommen werden können).

Eine genauere Definition der sog. bedingten Wahrscheinlichkeiten  $P(X = x | Y = y)$  soll an dieser Stelle unterbleiben. Sie muss in der Anwendungspraxis auch keineswegs immer überprüft werden. Gehören die Zvn. zu Systemen, die gegeneinander abgeschlossen sind, so heißt dies: kein Ereignis  $Y = y$  übt eine Wirkung auf das System aus, in dem  $X$  beobachtet wird. Mithin stimmen die bedingten Wahrscheinlichkeiten  $P(X = x | Y = y)$  bei variierendem  $y$  sämtlich überein.

Wir klären den Sachverhalt am Beispiel von Laplace-Versuchen.

### ■ Beispiel 1

Wir betrachten zwei Versuche:

$Z_1$ : Werfen eines normalen Spielwürfels (mit  $X =$  Augenzahl)

$Z_2$ : Werfen einer Münze (mit  $Y =$  Münzseite [Wappen, Zahl])

$Z_1$  und  $Z_2$  denke man sich zu einem Gesamtversuch kombiniert. Es ist völlig klar, dass Würfelwurf und Münzwurf sich gegenseitig nicht beeinflussen. Die Zvn.  $X, Y$  sind unabhängig, weil die Teilversuche, zu denen sie gehören, unabhängig (im physikalischen Sinn) sind. Dabei spielt es keine Rolle, ob  $Z_1$  und  $Z_2$  gleichzeitig oder nacheinander ausgeführt werden.

### ■ Beispiel 2

Wir führen den Versuch  $Z_1$  zweimal hintereinander durch. Es sei  $X$  die Augenzahl im ersten,  $Y$  die Augenzahl im zweiten Wurf.

Auch in diesem 2-stufigen Versuch sind die Zvn. unabhängig, obwohl sie am selben Gerät beobachtet werden. Abhängigkeit würde hier ja bedeuten, dass ein realisiertes Ergebnis ein künftiges Ergebnis beeinflusst. Bei einem guten Spielwürfel ist das nicht der Fall. Würde er sich aber z.B. beim Werfen in nennenswertem Ausmaß verformen, so hinge ein Wurfresultat durchaus von den bisherigen Verformungen ab.

### ■ Anmerkungen

Manchen Menschen leuchtet die Unabhängigkeit wiederholter Versuche nicht in letzter Konsequenz ein, z.B. Roulette-Spielern, die glauben, dass nach einer längeren Serie Schwarz die Farbe Rot immer wahrscheinlicher werde.

Angenommen, bei 20-maliger Durchführung von  $Z_1$  erscheint ausnahmslos eine Sechs. Anstatt zu glauben, es sei nun Zeit für eine andere Augenzahl, ist es in diesem Fall vernünftiger anzunehmen, dass der Spielwürfel manipuliert ist (und sich weiterhin nicht wie ein Laplace-Gerät verhält). Mit der Prüfung solcher Hypothesen beschäftigt sich die Statistik.

## Produktregel

Wir betrachten den kombinierten Versuch aus **Beispiel 1** und fragen nach der Wahrscheinlichkeit des Ereignisses:

$Z_1$  hat den Ausgang "4" und  $Z_2$  hat den Ausgang "Wappen"

Dasselbe lässt sich mit Hilfe der Zvn.  $X, Y$  formal wie folgt wiedergeben:

$$X = 4 \wedge Y = \text{Wappen}$$

Bereits bekannt ist  $P(X = 4) = \frac{1}{6}$ ,  $P(Y = \text{Wappen}) = \frac{1}{2}$ .

Hieraus errechnet sich die gesuchte Wahrscheinlichkeit durch Produktbildung:

$$P(X = 4 \wedge Y = \text{Wappen}) = P(X = 4) \cdot P(Y = \text{Wappen}) = \frac{1}{12}$$

Die hier benutzte Produktregel gilt für unabhängige Ereignisse.

Um das einzusehen, verschaffen wir uns zunächst einen Merkmalraum  $\Omega$  für den Gesamtversuch. Dessen Merkmale bestehen aus allen geordneten Paaren  $(x, y)$  mit  $x \in \{1, 2, 3, 4, 5, 6\}$  und  $y \in \{\text{Wappen}, \text{Zahl}\}$ . Es ist also

$$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{\text{Wappen}, \text{Zahl}\}$$

Insgesamt erhalten wir 12 "kombinierte" Merkmale. Da sie gleichwahrscheinlich sind, ergibt sich die Wahrscheinlichkeit für  $(4, \text{Wappen})$  zu  $\frac{1}{12}$ .

### ■ Definition

In Anlehnung an diese Überlegung bei Laplace-Versuchen definieren allgemeiner wir die **Unabhängigkeit zweier Elementarereignisse**  $X = x$  und  $Y = y$  mit den Wahrscheinlichkeiten  $p = P(X = x)$  und  $q = P(Y = y)$  wie folgt:

$$P(X = x \wedge Y = y) = p \cdot q$$

Der Sachverhalt lässt sich auf Ereignisse allgemein übertragen (und auch hier nicht beschränkt auf Laplace-Versuche). Sind  $A, B$  irgendwelche Ereignisse, so bezeichnen wir mit  $A B$  das Ereignis " $A$  und  $B$ " (Produktereignis).

#### Produktregel

Sind  $A, B$  unabhängige Ereignisse, so gilt  $P(A B) = P(A) \cdot P(B)$ .

#### Beweis:

Seien  $A = \{x_1, \dots, x_m\}$ ,  $B = \{y_1, \dots, y_n\}$  die Ereignisse (als Merkmalmenge hingeschrieben) und  $p_i, q_j$  die Wahrscheinlichkeiten für das Auftreten der Merkmale  $x_i$  bzw.  $y_j$ . Das Produktereignis  $A B$  tritt genau dann ein, wenn ein Merkmal  $(x_i, y_j)$  ( $1 \leq i \leq m$ ,  $1 \leq j \leq n$ ) beobachtet wird. Letzteres geschieht wegen der Unabhängigkeit mit der Wahrscheinlichkeit  $p_i q_j$ . Nach zweimaliger Anwendung der Summenregel ergibt sich damit:

$$P(A B) = \sum_{1 \leq i \leq m, 1 \leq j \leq n} p_i q_j = \left( \sum_{i=1}^m p_i \right) \left( \sum_{j=1}^n q_j \right) = P(A) \cdot P(B)$$

■

---

## Einfache Laplace-Versuche

### Zufallszahlen erzeugen

#### ■ Zufall oder Pseudozufall

Echte Zufallszahlen erzeugt man am besten mit dafür vorgesehenen Laplace-Geräten (Münze, Würfel, Roulette, etc.) oder mit authentischem Ziffernmaterial [vgl. das 1955 in The Free Press (New York) erschienene Buch *A Million Random Digits* bzw. den Auszug daraus in Arthur Engel: *Wahrscheinlichkeitsrechnung und Statistik*, Band 1, Klett Verlag: Stuttgart 1973, S. 154-155].

Ein Computer ist eine deterministische Maschine. Die Werte einer berechenbaren Funktion (bzw. die Ausgaben eines Computer-Programms) können nicht wirklich zufällig sein. Mit speziellen Algorithmen lassen sich aber sog. Pseudozufallszahlen erzeugen. Diese entstehen aus einer Startzahl durch wiederholte Ausführung (Iteration) einer geeigneten Funktion. Bei der Wahl des Iterationsalgorithmus kommt es darauf an, dass die erzeugten Zahlen "regellos" aufeinander zu folgen scheinen (und sich insbesondere nicht in zu kurzen Perioden wiederholen). Man vgl. zu dieser interessanten Problemstellung Kapitel 6 des oben erwähnten Buchs von A. Engel sowie [A. Engel: *Mathematisches Experimentieren mit dem PC*: Klett Verlag: Stuttgart 1991, S. 91-114]. Eine eingehende Erörterung des Themas liefert [Donald E. Knuth: *The Art of Computer Programming*, Vol. 2, Chapter 3, das den Umfang eines Buches besitzt].

#### ■ Rückgriff auf eingebaute Funktionen

In den meisten Programmiersprachen sind brauchbare Zufallszahlen-Generatoren eingebaut. Am weitesten verbreitet ist eine Funktion (zumeist **Random** genannt), die eine reelle Pseudozufallszahl im Intervall  $[0,1)$  ausgibt. Die Ergebnisse sind gleichverteilt, d.h. jede Zahl wird mit derselben Wahrscheinlichkeit gewählt.

Aufruf von **Random** in *Mathematica*:

```
Random[ ]
```

```
0.550926
```

In JavaScript ist **random** eine Methode des **Math**-Objekts. Mit ihr lässt sich leicht eine selbstdefinierte Funktion hinschreiben, die gleichverteilte ganze Zufallszahlen aus  $\{0, 1, \dots, m - 1\}$  erzeugt:

```
function RandomInt(m)
{
  return parseInt( m * Math.random() )
}
```

## Elementare Zufallsgeneratoren

Im Folgenden wollen wir einige Funktionen bereitstellen, die einfache (diskrete) Laplace-Versuche und zugehörige Versuchsreihen simulieren, dabei aber von den komfortableren Möglichkeiten von *Mathematica* Gebrauch machen.

### ■ Ganze Zufallszahl zwischen zwei Grenzen

Die Funktion **Random[Integer, {m,n}]** erzeugt eine ganze Pseudozufallszahl zwischen den ganzen Zahlen **m** und **n** (einschließlich):

```
Random[Integer, {100, 120}]
```

```
115
```

Denselben Effekt erzielt man mit folgender Zeile:

```
100 + Floor[21 * Random[]]
```

```
113
```

Die dementsprechende allgemeine Funktion:

```
RandomInteger[m_, n_] := m + Floor[(n - m + 1) * Random[]]
```

Eine Versuchsreihe dazu:

```
Table[RandomInteger[100, 120], {30}]
```

```
{117, 106, 114, 115, 107, 103, 101, 100, 101, 109, 101, 114, 118, 105, 105,
 108, 100, 102, 115, 101, 104, 102, 105, 108, 107, 116, 112, 113, 100, 113}
```

## ■ Urnenmodell I: Ziehen mit Zurücklegen

Die Funktion **ZieheMitZuruecklegen** (aus dem Paket **Zufallsversuche.m**) erzeugt eine gewünschte Anzahl zufällig einer beliebigen Menge (Merkmalraum) entnommener Elemente, genauer (im Sinne der Kombinatorik): eine zufällige Permutation mit Wiederholung.

```
ZieheMitZuruecklegen[m_, anzahl_] := Module[{mr = Union[m]},
  Table[Part[mr, Random[Integer, {1, Length[mr]}]], {anzahl}]
]
```

```
ZieheMitZuruecklegen[{"Otto", 7, "€"}, 10]
```

```
{€, Otto, €, Otto, Otto, 7, Otto, Otto, €, Otto}
```

Die Funktion simuliert das Urnenmodell der Stochastik: Aus einem Kasten mit irgendwelchen Objekten (Kugeln) wird (verdeckt) eines gewählt, jedoch nicht entfernt ("Ziehen mit Zurücklegen"). Die als erstes Argument übergebene Liste steht für den Kasten bzw. die Urne; das zweite Argument gibt die Anzahl der Ziehungen an.

Anmerkung: **ZieheMitZuruecklegen** ist funktionell identisch mit **Stichprobe** aus dem *Mathematica*-Paket **Verteilungen.m**.

## ■ Urnenmodell II: Ziehen ohne Zurücklegen

Das Gegenstück **ZiehenOhneZuruecklegen** (vom bekannten Zahlenlotto "6 aus 49" her bekannt) erfordert, dass die Urne das gezogene Element anschließend nicht mehr enthält. Dank leistungsfähiger Listenoperationen lässt sich auch diese Funktion in *Mathematica* auf natürliche Weise formulieren:

```
ZieheOhneZuruecklegen[m_, anzahl_] :=
Module[{mr = Union[m], sprobe = {}, elem, j = 0},
  While[j < anzahl && Length[mr] > 0,
    elem = Part[mr, Random[Integer, {1, Length[mr]}]];
    AppendTo[sprobe, elem];
    mr = Drop[mr, Position[mr, elem][[1]]];
    j++;
  ]
sprobe
```

Die While-Schleife erledigt nacheinander 1) die Entnahme einer Kugel (Element), 2) ihre Hinzügung zur Stichprobe und 3) ihre Entfernung aus der Urne. Dies geschieht höchstens so oft wie im zweiten Argument (**anzahl**) angegeben; sobald die Urne keine Kugel mehr enthält, wird abgebrochen.

```
urne = Table[k, {k, 1, 49}];  
ZieheOhneZuruecklegen[urne, 6]
```

```
{38, 22, 42, 47, 5, 20}
```

In dieser Form realisiert der Zufallsversuch eine zufällige 6-Permutation von 49 Elementen ohne Wiederholung. Wenn wir die Stichprobenelemente anordnen (weil es, wie beim Zahlenlotto, auf ihre Reihenfolge nicht ankommt), so erhalten wir 6-Kombinationen:

```
Sort[%]
```

```
{5, 20, 22, 38, 42, 47}
```

Beschränkt man sich auf herkömmliche Programmiersprachen (Pascal, C, etc.), sind die algorithmischen Lösungen dieses Problems ein wenig schwieriger zu verstehen. Vgl. das in HTML und JavaScript entwickelte Beispiel-Formular *Simulierte Lottoziehung "6 aus 49"*.



## Bernoulli-Räder

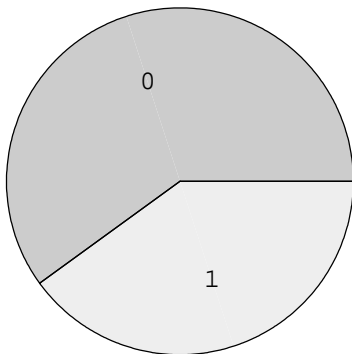
Bernoulli-Räder (im Folgenden auch "Glückräder" genannt) sind Simulationsmodelle für Zufallsversuche, deren Ausgänge nicht notwendig gleichverteilt sind.

Im Folgenden benötigte Packages:

```
<< Graphics`Graphics`  
<< Modellbildung`Verteilungen`
```

### Bernoulli-Rad mit 2 Sektoren

Das einfachste Glücksrad besitzt zwei Sektoren: 0 und 1, die im allgemein verschieden groß sind.



Wir stellen uns einen Zeiger vor, der sich um das Zentrum dreht. Es bezeichne  $p_i$  die Wahrscheinlichkeit dafür, dass er in Sektor  $i$  stehen bleibt. Da in der Beispiel-Figur Sektor 0 doppelt so groß ist wie Sektor 1, ergibt sich  $p_0 = \frac{2}{3}$ ,  $p_1 = \frac{1}{3}$ .

Häufig benutzte Bezeichnungen: Der Versuchsausgang 1 heißt **Treffer** (Erfolg), entsprechend 0 **Fehlschlag** (Misserfolg);  $p$  ist die Treffer-Wahrscheinlichkeit,  $q (= 1 - p)$  die Wahrscheinlichkeit für einen Fehlschlag.

Einmaliges Drehen des Bernoulli-Rads entspricht einem Aufruf von **Random[ ]**, der als Treffer gewertet wird, wenn das Ergebnis  $< p$  ist, sonst als Fehlschlag:

```
p = 1 / 3;
If[Random[] < p, 1, 0]
```

```
0
```

Wir simulieren 1000 Drehungen und vergleichen die Häufigkeitsverteilung in der entstehenden Stichprobe mit der W-Verteilung  $(\frac{2}{3}, \frac{1}{3})$ :

```
br = Table[If[Random[] < p, 1, 0], {1000}];
```

```
RelHVerteilung[{0, 1}, br] // N
```

```
{0.675, 0.325}
```

Es liegt nahe, das mehrfache Drehen eines Bernoulli-Rades durch folgende Funktion wiederzugeben:

```
BernoulliRad[p_, anzahl_] := Table[If[Random[] < p, 1, 0], {anzahl}]
```

## Bernoulli-Rad mit n Sektoren

Eine Verallgemeinerung des obigen Bernoulli-Rads ist das **Bernoulli-Rad ("Glücksrad") mit  $n$  Sektoren** 1, 2, ...,  $n$  und zugehöriger W-Verteilung  $(p_1, p_2, \dots, p_n)$ . Wir bezeichnen es kurz als  $(p_1, p_2, \dots, p_n)$ -Glücksrad.

Wie beim Glücksrad mit zwei Sektoren stellen wir die Wahrscheinlichkeit  $p_1$  durch das Intervall  $[0, p_1)$  dar, die Wahrscheinlichkeit  $p_2$  durch das sich daran anschließende Intervall  $[p_1, p_1 + p_2)$ , ferner  $p_3$  durch  $[p_1 + p_2, p_1 + p_2 + p_3)$ , usw. Die so definierten  $n$  Intervalle haben beziehentlich die Längen  $p_1, p_2, p_3, \dots$  und bilden eine Zerlegung von  $[0,1)$ . Damit ist die sektorale Zerlegung des Kreises auf das Einheitsintervall abgebildet. Wir können eine Drehung des Glücksrads daher wie folgt modellieren: Liefert **Random[]** den Wert  $r$ , so suchen wir das (eindeutig bestimmte) Intervall, in dem  $r$  liegt:

$$p_1 + \dots + p_{i-1} \leq r < p_1 + \dots + p_{i-1} + p_i$$

Es ist dann  $i$  die Nummer des angezeigten Sektors.

In der folgenden Hilfsfunktion bezeichnen **a** und **b** die linke bzw. rechte Intervallgrenze für  $r$  beim Durchlaufen der Intervallfolge; **pvert** ist die zu Grunde liegende W-Verteilung:

```
Sektor[r_, pvert_] := Module[{n = Length[pvert], i = 1, a = 0, b = pvert[[1]]},
  While[Not[a ≤ r < b],
    i++;
    a = a + pvert[[i - 1]];
    b = b + pvert[[i]];
  ]
  i
]
```

Ein Testaufruf mit 4 Sektoren:

```
r = Random[]  
Sektor[r, {0.1, 0.3, 0.2, 0.4}]
```

```
0.283578
```

```
2
```

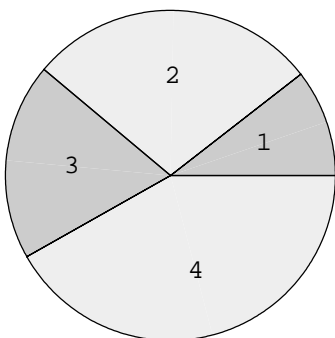
Damit lässt sich das mehrmalige Drehen eines Glücksrads mit vorgegebener W-Verteilung simulieren:

```
Gluecksrad[pvert_, anzahl_] := Table[Sektor[Random[], pvert], {anzahl}]
```

Wir drehen 1000-mal:

```
gr = Gluecksrad[{0.1, 0.3, 0.2, 0.4}, 1000];  
RelHVerteilung[{1, 2, 3, 4}, gr] // N  
KreisDiagramm[{1, 2, 3, 4}, gr]
```

```
{0.105, 0.284, 0.193, 0.418}
```



## Mehrstufige Bernoulli-Versuche

Das Drehen eines Bernoulli-Rads (gleichgültig mit wie vielen Sektoren und gleichgültig, ob simuliert oder real durchgeführt) nennen wir einen (einstufigen) Bernoulli-Versuch. Bernoulli-Versuche, die gleichzeitig mit verschiedenen Glücksrädern oder hintereinander mit demselben Glücksrad durchgeführt werden, betrachten wir als unabhängige Zufallsversuche.

Im Folgenden legen wir ein  $(p, q)$ -Bernoulli-Rad mit den Sektoren "Treffer" (1) bzw. "Fehlschlag" (0) zu Grunde und realisieren einen  $N$ -**stufigen Versuch**, indem das Rad  $N$ -mal hintereinander gedreht wird (oder dazu gleichwertig:  $N$  baugleiche Räder gleichzeitig gedreht werden). Die Versuchsergebnisse sind  $N$ -gliedrige Folgen bestehend aus den Ziffern 0 und 1.

Wir interessieren uns für die Zufallsvariable  $X = \text{Anzahl der Treffer}$ . Sie ist als Summe unabhängiger Zvn. darstellbar:  $X = X_1 + X_2 + \dots + X_N$ . Dabei ist  $X_i$  das Ergebnis (0 oder 1) des  $i$ -ten Teilversuchs.

Es soll die Wahrscheinlichkeit einiger Ereignisse bestimmt werden, die sich mittels  $X$  beschreiben lassen.

### ■ Lauter Treffer

Eine Trefferserie besteht aus  $N$  Treffern: 1, 1, ..., 1, d.h.  $X = N$ . Nach der Produktregel hat dieses Ereignis die Wahrscheinlichkeit  $p^N$ :

$$P(X = N) = p^N$$

### ■ Mindestens ein Treffer

Eine Serie von  $N$  Fehlschlägen hat (analog zur Trefferserie) die Wahrscheinlichkeit  $q^N$ . Es ist also  $1 - q^N = 1 - (1 - p)^N$  die Wahrscheinlichkeit für  $X \neq 0$ , d.h. für das Ereignis, dass nicht lauter Fehlschläge auftreten. Dies ist gleichbedeutend mit dem Ereignis  $X \geq 1$ , dass mindestens ein Treffer erzielt wird:

$$P(X \geq 1) = 1 - (1 - p)^N$$

### ■ Genau $k$ Treffer: $X = k$

Wir nehmen zunächst an, dass  $k$  Treffer vorneweg und danach sämtliche  $N - k$  Fehlschläge hintereinander auftreten:

$$\underbrace{1 \dots 1}_k \underbrace{0 \dots 0}_{N-k}$$

Die Wahrscheinlichkeit dafür ist (wiederum nach der Produktregel)  $p^k \cdot q^{N-k}$ . Für jede andere Platzierung der  $k$  Treffer erhalten wir dieselbe Wahrscheinlichkeit.

Nun lassen sich von den  $N$  in der Folge vorhandenen Plätzen die  $k$  Plätze für die Treffer auf  $\binom{N}{k}$  Weisen festlegen.

Nach der Summenregel ergibt sich damit als Wahrscheinlichkeit für das Ereignis " $k$  Treffer":

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

## Die Binomialverteilung

### Treffer-Verteilung in mehrstufigen Bernoulli-Versuchen

Beim Studium von Bernoulli-Rädern hat sich gezeigt: Bei  $N$ -maligem Drehen eines  $(p, q)$ -Bernoulli-Rades ergibt sich für die Zv.  $X = \text{Trefferanzahl}$  die Verteilung

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

Eine solche Verteilung heißt (wegen der hier in Erscheinung tretenden Koeffizienten) **Binomialverteilung**.

Es ist praktisch, für diesen Ausdruck eine Funktion parat zu haben:

```
BW[p_, n_, k_] := Binomial[n, k] * p^k * (1 - p)^(n - k)
```

#### ■ Beispiel

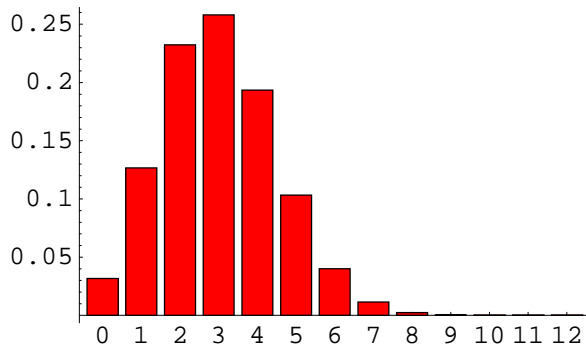
Wir erzeugen eine Wertetabelle und betrachten ein Stabdiagramm dieser W-Verteilung für  $p = \frac{1}{4}$  und  $N = 12$ :

```
<< Graphics`Graphics`  
<< Statistics`DescriptiveStatistics`  
<< Modellbildung`Verteilungen`  
<< Modellbildung`Zufallsversuche`
```

```
bv = BinomialVerteilung[0.25, 12]
```

```
{0.0316764, 0.126705, 0.232293, 0.258104,  
0.193578, 0.103241, 0.0401495, 0.0114713, 0.00238985,  
0.000354052, 0.0000354052, 2.14577 × 10-6, 5.96046 × 10-8}
```

```
BarChart[bv, BarLabels -> {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}];
```



Die wahrscheinlichste Trefferanzahl ist  $X = 3$ . Dies stimmt mit der Idee überein, der zufolge die Wahrscheinlichkeit  $p$  den (theoretischen) Anteil eines Merkmals in der Stichprobe angibt (vgl. die frequentistische Deutung des W-Begriffs), hier also:  $p \cdot N = \frac{12}{4} = 3$ .

## Anwendungen

Die Binomialverteilung besitzt zahlreiche Anwendungen in vielfältigen Lebensbereichen. Wir betrachten zwei einfache Situationen.

### ■ Beispiel 1

Die Wahrscheinlichkeit, dass unter Zwillingen beide Kinder Knaben sind, wird auf  $p = 0.32$  geschätzt. Wie wahrscheinlich ist es, dass von 6 Zwillingspaaren die Hälfte Knabenpaare sind?

Lösung:  $P(X = 3) = \binom{6}{3} p^3 (1-p)^{6-3} = 0.206 \dots$

```
BW[0.32, 6, 3]
```

```
0.206066
```

### ■ Beispiel 2

Ein Expertenteam, das regelmäßig mit Prognosen zur wirtschaftlichen Entwicklung beauftragt wird, hat eine Trefferquote von 65 %. In nächster Zeit sind 10 neue Gutachten zu erstellen. Wie groß ist die Wahrscheinlichkeit, dass die Experten dabei eine Trefferquote von mindestens 80 % erzielen werden?

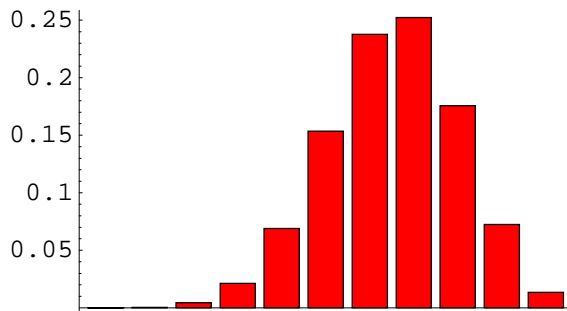
Lösung: Es ist  $p = 0.65$ ,  $N = 10$ . Die gesuchte Wahrscheinlichkeit ist nach der Summenregel zu berechnen:  
 $P(X \geq 8) = P(X = 8) + P(X = 9) + P(X = 10) = 0.26 \dots$

```
Sum[BW[0.65, 10, k], {k, 8, 10}]
```

```
0.261607
```

Der niedrige Wert entspricht dem kleinen Maximum (bei  $X = 7$ ) und der dahinter steil abfallenden Verteilung:

```
BarChart[BinomialVerteilung[0.65, 10], BarLabels -> None];
```



## Simulation von Massenversuchen

Die von der Binomialverteilung gelieferten Wahrscheinlichkeiten zeichnen sich als Stabilisierungsniveaus relativer Häufigkeiten ab, wenn der betreffende (mehrstufige) Bernoulli-Versuch oft wiederholt wird. Im Folgenden wollen wir dies am eingangs dargestellten Beispiel eines 12-mal gedrehten  $(\frac{1}{4}, \frac{3}{4})$ -Glücksrades nachvollziehen.

Die 1-malige Ausführung des 12-stufigen Versuchs simuliert die Funktion **BernoulliRad** (aus dem Paket **Zufallsversuche.m**):

```
p = 0.25;
n = 12;
BernoulliRad[p, n]
```

```
{1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0}
```

Die Summe aller in dem  $N$ -Tupel realisierten Ausfälle ergibt die Trefferzahl. Es liegt also nahe, eine Funktion dieses Namens zu definieren, welche die Anzahl der Treffer auf diese Weise berechnet:

```
Trefferzahl[p_, n_] := Plus @@ BernoulliRad[p, n]
```

Der zugehörige Merkmalraum enthält alle möglichen Werte von **Trefferzahl**:  $0, 1, \dots, N$ .

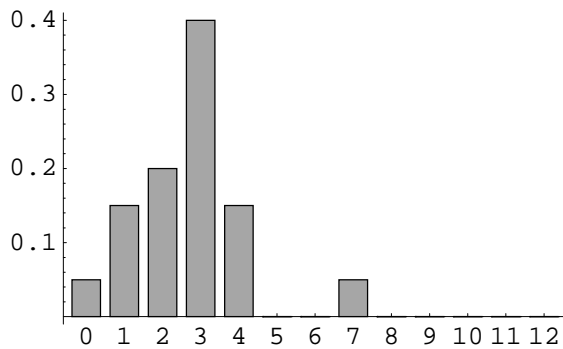
Als nächstes ist eine Stichprobe von Trefferzahlen herzustellen, z.B. vom Umfang 20:

```
sprobe = Table[Trefferzahl[p, n], {20}]
```

```
{3, 2, 1, 4, 3, 2, 4, 2, 3, 0, 4, 3, 7, 2, 3, 3, 1, 3, 3, 1}
```

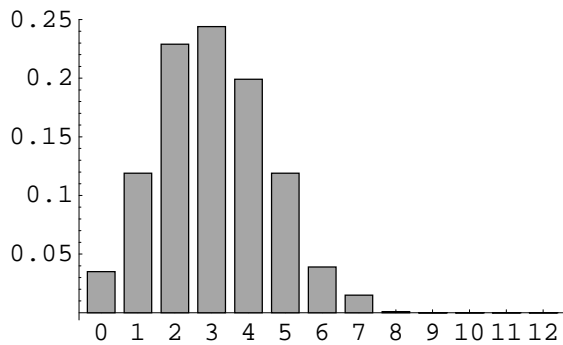
Die hierdurch gegebene Häufigkeitsverteilung über dem Merkmalraum  $\Omega = \{0, 1, \dots, 12\}$  ist die empirische Vorform der Binomialverteilung:

```
omega = Table[k, {k, 0, 12}];  
StabDiagramm[omega, sprobe]
```



Die Annäherung an den theoretischen Verlauf wird umso besser, je länger die realisierte Stichprobe ausfällt, hier z.B. bei 1000 Versuchen:

```
StabDiagramm[omega, Table[Trefferzahl[p, n], {1000}]]
```

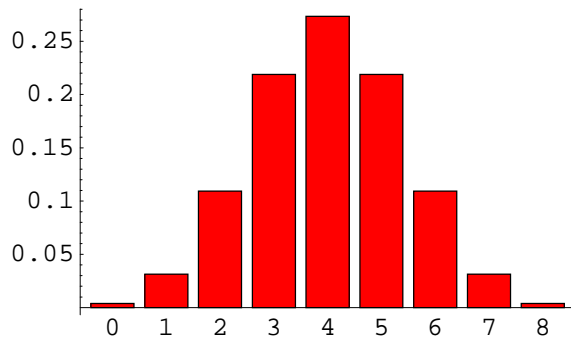


## Normalverteilung

Binomialverteilungen (bzw. simulierte Trefferzahl-Verteilungen) sind eingipflig. Die beobachteten Merkmale verteilen sich mehr oder weniger gleichmäßig um den wahrscheinlichsten Wert herum. Bei  $p = \frac{1}{2}$  erweist sich eine Binomialverteilung als vollkommen symmetrisch:

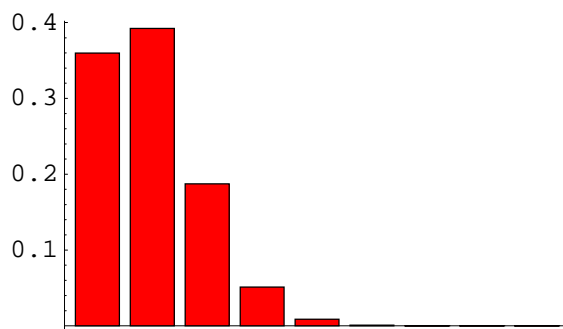


```
bv = BinomialVerteilung[0.5, 8];  
BarChart[bv, BarLabels -> {0, 1, 2, 3, 4, 5, 6, 7, 8}];
```



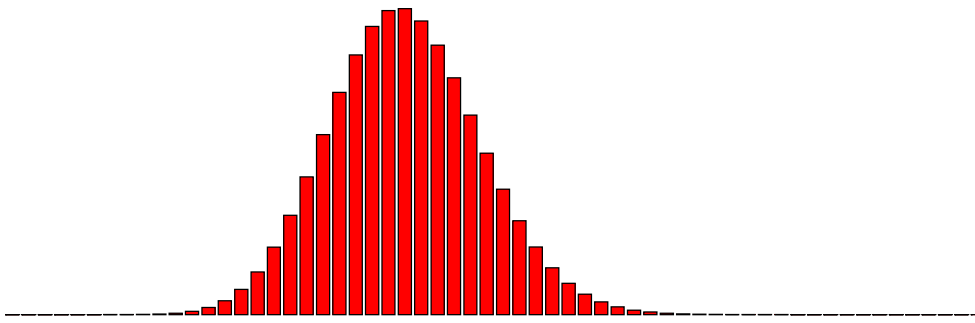
Je mehr  $p$  von diesem Wert abweicht, desto "schiefer" gerät das Verteilungsbild, hier z.B. für  $p = 0.12$ :

```
bv = BinomialVerteilung[0.12, 8];  
BarChart[bv, BarLabels -> None];
```



In diesem Fall liegt ein 8-stufiger Bernoulli-Versuch vor. Wenn wir nun die Stufenzahl  $N$  vergrößern, gibt es entsprechend mehr konkurrierende Merkmale und die resultierende Verteilung gewinnt an Symmetrie.

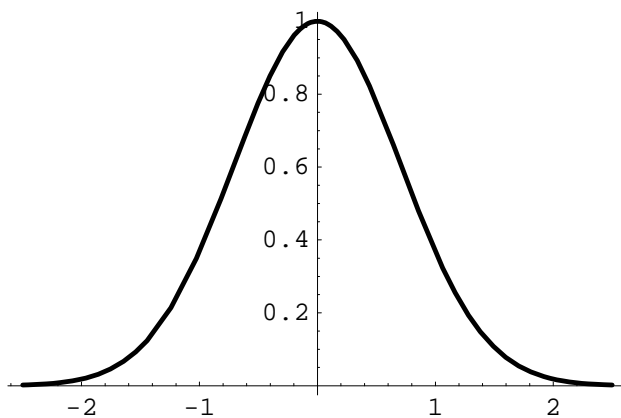
Hier das Schaubild für  $p = 0.12$ ,  $N = 200$ :



Die sich allmählich abzeichnende Grenzkurve (*Gaußsche Glockenkurve*) ist eine universale Gestalt. Sie repräsentiert die sog. **Normalverteilung**. Die Normalverteilung stellt sich als gleichsam "natürliche" Verteilung solcher Beobachtungsgrößen ein, die sich aus einer großen Zahl unabhängiger Zufallsvariablen aufsummieren.

Die Normalverteilungskurve wird im wesentlichen – d.h. bis auf Verschiebungen und Stauchungen – durch die Funktion  $y = e^{-x^2}$  beschrieben:

```
Plot[Exp[-x^2], {x, -2.5, 2.5}, PlotStyle -> Thickness[0.008]];
```



## Wartezeiten-Verteilung (geometrische Verteilung)

### Die Zufallsvariable "Wartezeit"

Ein  $(p, q)$ -Glücksrad wird solange gedreht bis ein Treffer erscheint. Geschieht dies beim  $n$ -ten Versuch, so nennt man  $n$  die **Wartezeit** für einen Treffer. Wir verwenden eine Zufallsvariable  $Z$  und schreiben  $Z = n$ .

Das Ereignis  $Z = n$  lässt sich so beschreiben: Zunächst werden  $n - 1$  Fehlschläge realisiert, anschließend ein Treffer:  $\underbrace{0 \dots 0}_{n-1} 1$ . Nach der Produktregel ergibt sich für die Wahrscheinlichkeit:

$$P(Z = n) = (1 - p)^{n-1} p$$

Welche Werte  $n$  kann  $Z$  annehmen? Natürlich  $n = 1, 2, 3, \dots$ . Dabei lässt sich (jedenfalls theoretisch) kein größter Wert angeben, der diese Folge beschränkt. Zum einen resultieren für kleines  $p$  große Wartezeiten. Zum anderen ist, etwa für  $p = 0.5$ , kein Grund ersichtlich, der eine Versuchsvorrichtung zwingen könnte, innerhalb einer vorgeschriebenen Versuchsanzahl einen Treffer zu realisieren. Bei unabhängigen Versuchen ereignen sich Treffer und Fehlschlag jedesmal mit derselben Wahrscheinlichkeit. – Somit ist der Merkmalraum von  $Z$  als Menge der positiven ganzen Zahlen anzusetzen.

### Die geometrische Verteilung

Wir verschaffen uns einen Eindruck vom Verlauf der sog. **geometrischen Verteilung** (wie die W-Verteilung der Wartezeiten auf einen Treffer auch genannt wird).

Die folgende Funktion berechnet die Wahrscheinlichkeit für  $Z = n$  nach obiger Formel:

$$GW[p\_ , n\_ ] := (1 - p) ^ (n - 1) * p$$

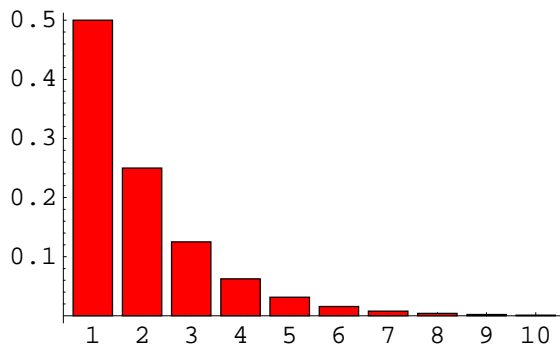
Wir legen eine Tabelle an für  $p = 0.5$ , wobei  $n$  von 1 bis 10 läuft:

```
gwtab = Table[GW[0.5, n], {n, 1, 10}]
```

```
{0.5, 0.25, 0.125, 0.0625, 0.03125, 0.015625,
 0.0078125, 0.00390625, 0.00195313, 0.000976563}
```

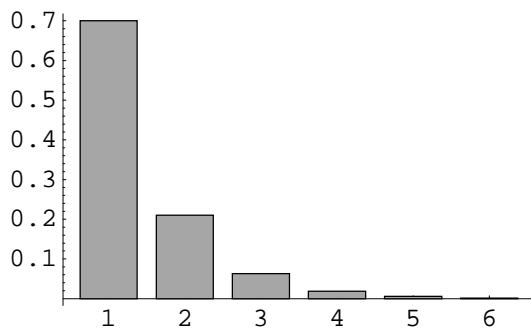
Das zugehörige Schaubild:

```
BarChart[gwtab, BarLabels -> {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}];
```



Das Paket **Verteilungen.m** enthält die Funktion **GeomVerteilung**, die zu gegebener Treffer-Wahrscheinlichkeit  $p$  die theoretische  $W$ -Verteilung zugleich berechnet und als Stabdiagramm ausgibt.

```
GeomVerteilung[0.7]
```



```
{0.7, 0.21, 0.063, 0.0189, 0.00567, 0.001701}
```

## Simulation eines Wartezeiten-Versuchs

Ein einfacher Weg zur empirischen Untersuchung der Wartezeiten-Verteilung: Man dreht viele Male ein Glücksrad.

```
brtab = BernoulliRad[0.4, 100]
```

```
{1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1,
 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0,
 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1,
 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1}
```

Aus der Ergebnisliste stellen wir eine Liste (als Stichprobe für  $Z$ ) her, in der die aufgetretenen Wartezeiten verzeichnet sind. Dazu bestimmen wir zunächst die Positionen der Treffer "1":

```
posliste = Flatten[Position[brtab, 1]]
```

```
{1, 2, 10, 12, 19, 20, 22, 23, 24, 25, 26, 28, 30,
 31, 34, 35, 36, 37, 45, 48, 53, 56, 60, 61, 63, 68, 69, 70,
 71, 75, 76, 77, 78, 79, 85, 88, 91, 95, 96, 97, 98, 100}
```

und berechnen die Wartezeiten als die Differenzen benachbarter Positionen:

```
zliste = Table[posliste[[i + 1]] - posliste[[i]], {i, 1, Length[posliste] - 1}]
```

```
{1, 8, 2, 7, 1, 2, 1, 1, 1, 1, 2, 2, 1, 3, 1, 1, 1, 8, 3,
 5, 3, 4, 1, 2, 5, 1, 1, 1, 4, 1, 1, 1, 1, 6, 3, 3, 4, 1, 1, 1, 2}
```

In der Liste fehlt das erste Element, das bei dem Differenzenverfahren verloren geht. Wir fügen es nachträglich hinzu:

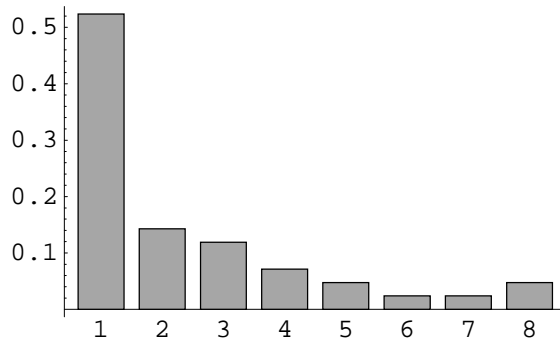
```
zliste = Join[{posliste[[1]]}, zliste]
```

```
{1, 1, 8, 2, 7, 1, 2, 1, 1, 1, 1, 2, 2, 1, 3, 1, 1, 1, 8, 3,
 5, 3, 4, 1, 2, 5, 1, 1, 1, 4, 1, 1, 1, 1, 6, 3, 3, 4, 1, 1, 1, 2}
```

Damit haben wir die Stichprobe hergestellt, in der die Häufigkeitsverteilung der Wartezeiten zu berechnen ist.

```
RelHVerteilung[zliste, zliste] // N  
StabDiagramm[zliste, zliste]
```

```
{0.52381, 0.142857, 0.119048, 0.0714286,  
0.047619, 0.0238095, 0.0238095, 0.047619}
```



Das Paket **Zufallsversuche.m** enthält eine Funktion, welche diese manuell durchgeführten Operationen zusammenfasst und automatisiert:

```
WartezeitenListe[0.4, 100]
```

```
{2, 1, 2, 3, 1, 1, 2, 1, 1, 3, 1, 1, 1, 2, 1, 1, 2, 4, 2, 2, 2, 3,  
1, 1, 2, 1, 2, 2, 2, 1, 5, 4, 1, 1, 3, 11, 1, 4, 2, 8, 2, 1, 2, 1, 3}
```

## Erwartungswert einer Zufallsvariablen

### Definition des Erwartungswerts

Die Definition des Erwartungswerts (einer Zv.) lässt sich auf einfache Weise plausibel machen, wenn man vom arithmetischen Mittel ausgeht.

#### ■ Arithmetisches Mittel einer Stichprobe

Wir denken uns eine Stichprobe aus lauter quantitativen Merkmalen gegeben. Ihr arithmetisches Mittel errechnet sich dadurch, dass man die Summe ihrer Elemente durch die Stichprobenlänge dividiert.

Das arithmetische Mittel einer Stichprobe lässt sich auch auffassen als das mit den relativen Häufigkeiten ihrer Merkmale gewichtete Mittel aller Merkmale.

Man fasse in der Stichprobensumme gleiche Merkmale zusammen. Jedes Merkmal wird dann mit seiner absoluten Häufigkeit multipliziert. Aus dem jeweiligen Faktor wird nach Division durch die Stichprobenlänge die relative Häufigkeit des Merkmals.

Der Sachverhalt soll noch formal hingeschrieben werden: Sei  $X$  eine quantitative Zufallsvariable mit dem Merkmalraum  $\Omega = \{x_1, \dots, x_m\}$  und  $S$  eine Stichprobe der Länge  $s$ . Wir betrachten die relativen Häufigkeiten  $h_i$  aller in Betracht kommenden Merkmale:

$$h_i := H_S(x_i) \quad (i = 1, 2, \dots, m)$$

Merkmal  $x_i$  kommt in  $S$  offenbar  $s \cdot h_i$ -mal vor ( $i = 1, 2, \dots, m$ ). Daher ist  $s h_1 x_1 + \dots + s h_m x_m$  gleich der Summe aller Merkmale, die in  $S$  aufgelistet sind. Dividiert man diese Summe durch die Länge  $s$  der Stichprobe, so erhält man ihr arithmetisches Mittel  $\mu$ :

$$\mu = h_1 x_1 + \dots + h_m x_m$$

#### ■ Der Erwartungswert als Verallgemeinerung des arithmetischen Mittels

Ersetzen wir in der zuletzt gewonnenen Darstellung von  $\mu$  die relativen Häufigkeiten durch die Wahrscheinlichkeiten der W-Verteilung von  $X$  – genauer: durch  $p_i = P(X = x_i)$  – so erhalten wir den **Erwartungswert von  $X$** :

$$E(X) = p_1 x_1 + \dots + p_m x_m$$

Auch für den Erwartungswert  $E(X)$  ist die Abkürzung  $\mu$  gebräuchlich.

N.B.: Bei abzählbar-unendlichem Merkmalraum  $\Omega = \{x_1, x_2, x_3, \dots\}$  ist eine unendliche Summe (Reihe)  $E(X) = \sum_{i=1}^{\infty} p_i x_i$  auszuwerten (vgl. mittlere Wartezeit).

## Zwei einfache Beispiele

### ■ Mittlere Augenzahl beim Spielwürfel

Die Augenzahl  $X$  eines guten Spielwürfels hat den Merkmalraum  $\Omega = \{1, 2, \dots, 6\}$  und die W-Verteilung  $p_i = \frac{1}{6}$  ( $i = 1, 2, \dots, 6$ ). Für den Erwartungswert ergibt sich:

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = 3.5$$

Offenkundig ist der Erwartungswert einer Zvn. nicht notwendig auch ein Wert, den die Zv. tatsächlich annehmen kann.

Ein Massenversuch (500 Würfe) möge das Ergebnis von der frequentistischen Seite her beleuchten:

```
<< Statistics`DescriptiveStatistics`
<< Modellbildung`Zufallsversuche`
```

```
sprobe = ZieheMitZuruecklegen[{1, 2, 3, 4, 5, 6}, 500];
Mean[sprobe] // N
```

```
3.436
```

### ■ Mittlerer Gewinn beim Münzwurf

Beim Wurf einer Laplace-Münze werde 1 € gesetzt. Erscheint "Wappen", erhält man den doppelten Einsatz zurück (andernfalls ist der Einsatz verloren).

Aus der Sicht eines Spielers ist die Zv.  $X = \text{Gewinn}$  von Interesse. Wir berechnen ihren Erwartungswert. "Gewinn" meint hier Reingewinn; bei "Wappen" erhält der Spieler 2 € nach Abzug des Einsatzes bleibt ihm daher 1 € als Gewinn. Daher gilt:

$$E(X) = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$$

Für Spieler (und ebenso Gegenspieler) entsteht bei dieser Wette keine positive (oder negative) Gewinnerwartung. Man spricht von einem fairen Spiel.



Da in einer Versuchsreihe das aus den tatsächlich realisierten Häufigkeiten gebildete arithmetische Mittel um den (theoretischen) Erwartungswert herum schwankt, können für einzelne Spielteilnehmer sehr wohl Gewinn- oder Verlustsituationen eintreten (auch bei nicht-fairen Spielen). Darin liegt der Reiz von Glücksspielen.

```
sprobe = ZieheMitZuruecklegen[{-1, 1}, 500];
Mean[sprobe] // N
```

```
-0.028
```

## Grundlegende Eigenschaften

Wir beweisen im Folgenden zwei grundlegende Eigenschaften des Erwartungswerts von reellwertigen Zufallsvariablen über demselben Merkmalraum.

### ■ $E$ ist homogen

Für alle reellen  $c$  gilt:  $E(c X) = c E(X)$

Beweis: Sei  $p_i = P(X = x_i)$  für  $i = 1, 2, \dots, m$ . Auch das Ereignis  $c X = c x_i$  hat die Wahrscheinlichkeit  $p_i$ . Mithin gilt:  $E(c X) = p_1(c x_1) + \dots + p_m(c x_m) = c(p_1 x_1 + \dots + p_m x_m) = c E(X)$ . ■

### ■ $E$ ist additiv

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

Beweis: Als elementare Merkmale von  $X + Y$  wählen wir sämtliche Summen  $x_i + y_i$  von Merkmalen von  $X$  und von  $Y$ , die einen Wert von  $X + Y$  darstellen ( $i = 1, 2, \dots, n$ ). Sei  $p_i$  die Wahrscheinlichkeit, mit der die  $i$ -te Summe realisiert wird. Dann ergibt sich:

$$E(X + Y) = p_1(x_1 + y_1) + \dots + p_n(x_n + y_n) = (p_1 x_1 + \dots + p_n x_n) + (p_1 y_1 + \dots + p_n y_n)$$

In dem Ausdruck  $p_1 x_1 + \dots + p_n x_n$  sind nun eventuell noch Summanden mit gleichen Merkmalwerten  $x_i$  zusammenzufassen, um einzusehen, dass  $E(X) = p_1 x_1 + \dots + p_n x_n$ . Entsprechendes gilt für die zweite Klammer:  $E(Y) = p_1 y_1 + \dots + p_n y_n$ . ■

Eine homogene und additive Operation heißt **linear**. Der Erwartungswert ist ein linearer Operator.

## Mittlere Trefferzahl im Bernoulli-Versuch

Ein  $N$ -stufiger Versuch mit einem  $(p, q)$ -Glücksrad werde durchgeführt. Die Zv.  $X_i$  stehe für das Ergebnis beim  $i$ -ten Teilversuch ( $i = 1, 2, \dots, N$ ). Die Trefferzahl wird dann durch  $X = X_1 + \dots + X_N$  dargestellt. Ihre W-Verteilung ist die Binomialverteilung. Wir berechnen den Erwartungswert von  $X$  (die "mittlere" Trefferzahl).

Wegen der Additivität von  $E$  hat man zunächst

$$E(X) = E(X_1 + \dots + X_N) = E(X_1) + \dots + E(X_N)$$

In allen Teilversuchen haben Treffer dieselbe Wahrscheinlichkeit  $p$ , d.h. es gilt  $P(X_i = 1) = p$  und damit für die Erwartungswerte:

$$E(X_i) = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = 0 \cdot q + 1 \cdot p = p$$

Insgesamt erhalten wir daraus:

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

Zugleich ist dies der Erwartungswert einer binomialverteilten Zvn.

Die mittlere Trefferzahl  $\mu$  liegt in der Nähe des wahrscheinlichsten Wertes  $x_{\max}$  von  $X$ . Genauer gilt:  
 $\mu - q \leq x_{\max} \leq \mu + p$ . – (Beweis als Übung)

## Mittlere Wartezeit

Beim Wartezeiten-Versuch wird ein  $(p, q)$ -Glücksrad mit  $p > 0$  solange gedreht, bis ein Treffer erzielt ist. Für die Wartezeit  $Z$  gilt:  $P(Z = n) = q^{n-1} p$ . Wir berechnen die mittlere Wartezeit  $E(Z)$ :

$$E(Z) = \sum_{n=1}^{\infty} n \cdot P(Z = n) = \sum_{n=1}^{\infty} n q^{n-1} p = p \sum_{n=1}^{\infty} n q^{n-1}$$

Die rechterhand verbliebene Summe  $\sum_{n=1}^{\infty} n q^{n-1}$  ist noch auszuwerten:

$$1 + 2q + 3q^2 + \dots = \frac{1}{(1-q)^2}$$

Man sieht dies ein, indem man die Gleichung mit  $1 - q$  multipliziert. Links vom Gleichheitszeichen bleibt dann die unendliche geometrische Reihe  $1 + q + q^2 + \dots$  stehen, die wegen  $q < 1$  zum Wert  $\frac{1}{1-q}$  konvergiert. Beachtet man  $1 - q = p$ , so ergibt sich insgesamt:

$$\mu = E(Z) = \frac{1}{p}$$

Für eine Münze (Warten auf "Wappen",  $p = \frac{1}{2}$ ) gilt:  $\mu = 2$ ; für einen Spielwürfel (Warten auf eine Sechs,  $p = \frac{1}{6}$ ) gilt:  $\mu = 6$ .