

## Die Binomialverteilung

### Treffer-Verteilung in mehrstufigen Bernoulli-Versuchen

Beim Studium vom Bernoulli-Rädern hatte 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

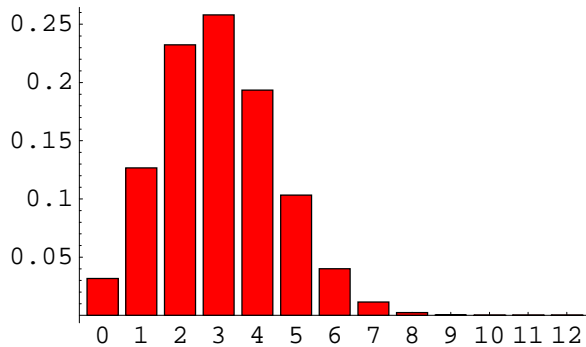
Wir erzeugen eine Wertetabelle und betrachten ein Stabdiagramm dieser W-Verteilung für  $p = \frac{1}{4}$  und  $N = 12$ :

```
<< 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?

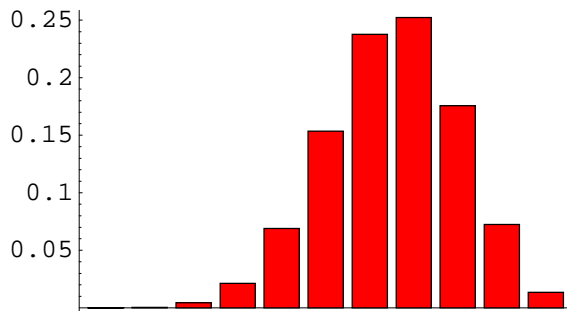
Lösung: Es ist  $p = 0.65$ ,  $N = 10$ . Die gesuchte Wahrscheinlichkeit ist nach der Summenregel zu berechnen:  
 $P(X \geq 8) = P(X = 8) + P(X = 9) + P(X = 10) = 0.26 \dots$

```
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**:  $0, 1, \dots, N$ .

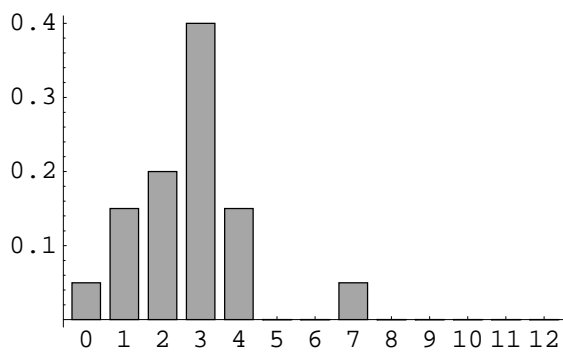
Als nächstes ist eine Stichprobe von Trefferzahlen herzustellen, z.B. vom Umfang 20:

```
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}
```

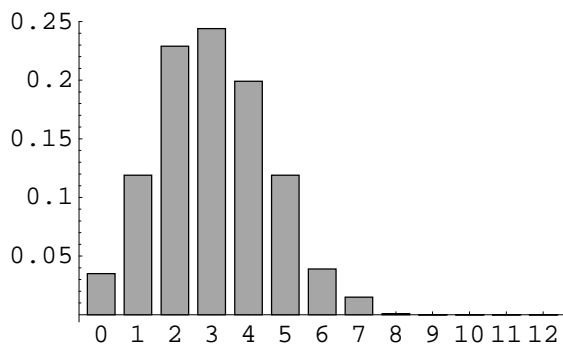
Die hierdurch gegebene Häufigkeitsverteilung über dem Merkmalraum  $\Omega = \{0, 1, \dots, 12\}$  ist die empirische Vorform der Binomialverteilung:

```
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:

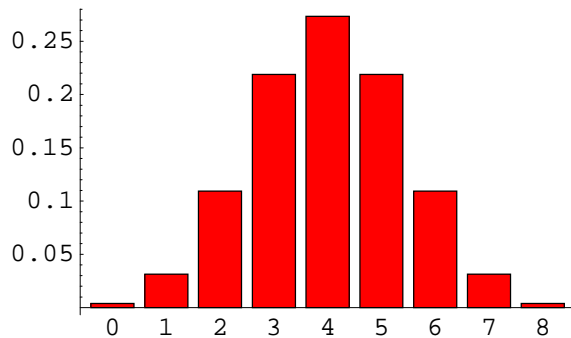
```
StabDiagramm[omega, Table[Trefferzahl[p, n], {1000}]]
```



## Normalverteilung

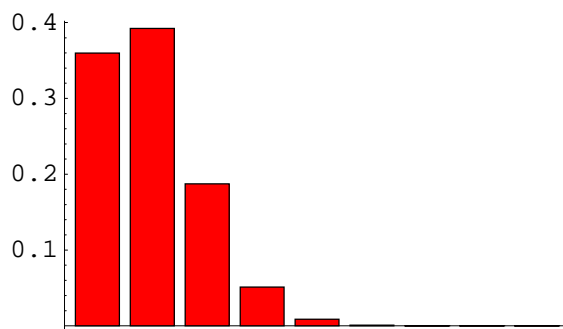
Binomialverteilungen (bzw. simulierte Trefferzahl-Verteilungen) sind eingipflig. Die beobachteten Merkmale verteilen sich mehr oder weniger gleichmäßig um den wahrscheinlichsten Wert herum. Bei  $p = \frac{1}{2}$  erweist sich eine Binomialverteilung als vollkommen symmetrisch:

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



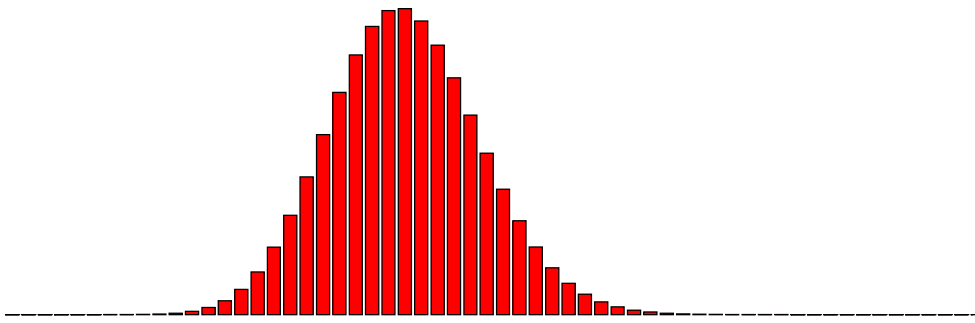
Je mehr  $p$  von diesem Wert abweicht, desto "schiefer" gerät das Verteilungsbild, hier z.B. für  $p = 0.12$ :

```
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]];
```

