

# Der Begriff der Entropie

## in der Informationstheorie

(entwickelt an diversen Beispielen)

*Autor:* Jochen Ziegenbalg

*Email:* ziegenbalg.edu@gmail.com

*Internet:* <https://jochen-ziegenbalg.github.io/materialien/>

### ■ Lotto: Die Problemstellung

In der einfachsten Form sind beim Lotto 6 aus 49 nummerierten Kugeln (ohne Zurücklegen) zu ziehen. Dabei geht man davon aus, dass jede der 49 Kugeln mit derselben Wahrscheinlichkeit gezogen wird. Bei der Bewertung des End-Ergebnisses kommt es nicht auf die Reihenfolge der gezogenen Kugeln an, sondern nur darauf, ob 0 oder 1 oder 2 oder 3 oder 4 oder 5 oder 6 "Richtige" getippt worden waren.

Daraus ergeben sich in natürlicher Weise die folgenden *Problemstellungen*:

- (a) Wie viele Möglichkeiten gibt es (bzw. wie groß sind die Chancen, wie groß ist die Wahrscheinlichkeit) für *genau* 0, 1, 2, 3, 4, 5, 6 "Richtige".
- (b) Wie viele Möglichkeiten gibt es (bzw. wie groß sind die Chancen, wie groß ist die Wahrscheinlichkeit) für *mindestens* 0, 1, 2, 3, 4, 5, 6 "Richtige".

### ■ Kombinatorische Grundaufgabe: Ziehungen ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge

Beim Lotto werden 6 Kugeln aus einer Menge von 49 nummerierten Kugeln gezogen und (nach dem Ziehen) beiseite gelegt. Obwohl die Kugeln jeweils einzeln gezogen werden, spielt die Reihenfolge der Ziehung für das Ergebnis keine Rolle. Vom kombinatorischen Typus her entspricht das Lotto-Spiel offensichtlich dem Grund-Typ "Ziehen ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge" im *Urnenmodell*.

Wir gehen im folgenden zunächst der Frage nach, wie viele mögliche Lotto-Ziehungen es gibt.

In mengentheoretischer Interpretation lautet die entsprechende Frage: Wie viele  $k$ -elementige Teilmengen besitzt eine  $n$ -elementige Menge? Die Lösung lautet bekanntlich:  $\binom{n}{k}$ . Im folgenden wird eine Begründung dafür am Beispiel des Lotto-Spiels gegeben.

*Behauptung:* Es gibt  $\binom{49}{6}$  Möglichkeiten, 6 aus 49 nummerierten Kugeln ohne Berücksichtigung der Reihenfolge und ohne Zurücklegen zu ziehen.

*Begründung:* Wenn man zunächst die Reihenfolge der Ziehung beachtet, dann gibt es 49 Möglichkeiten, *eine* Kugel, 49·48 Möglichkeiten, *zwei* Kugeln, 49·48·47 Möglichkeiten *drei* Kugeln u.s.w. und schliesslich 49·48·47·46·45·44 Möglichkeiten *sechs* Kugeln auszuwählen. Da es beim Lotto nicht auf die Reihenfolge der Ziehung ankommt, gehen (bei 6 Kugeln) jeweils  $6! (= 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6)$  Fälle durch Permutation auseinander hervor; sie bilden eine *Permutationsklasse*. Jede dieser Permutationsklassen liefert genau ein Lotto-Ergebnis. Wenn man dementsprechend jede dieser Permutationsklassen nur einmal zählt, dann ergibt dies

$$\frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6!} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 54 \cdot 44 \cdot (43!)}{6! \cdot 43!} = \frac{49!}{6! \cdot (49-6)!} = \binom{49}{6}$$

verschiedene Möglichkeiten.

Numerische Auswertung mit *Mathematica*:

```
Binomial[49, 6]
```

```
13983816
```

*Satz:* Es gibt  $\binom{n}{k}$  Möglichkeiten,  $k$  Kugeln "ohne Zurücklegen" aus einer Urne mit  $n$  (unterschiedlichen) Kugeln zu ziehen, wenn die Reihenfolge der Ziehungen keine Rolle spielt.

*Beweis:* Paradigmatisch - im Sinne der obigen Argumentation.

## ■ Lotto: Anzahl der Möglichkeiten für *genau* $k$ "Richtige"

Wir betrachten wieder ein konkretes Beispiel: Wie groß sind die Chancen (wie groß ist die Wahrscheinlichkeit) für *genau* 4 "Richtige" im Lotto? Durch eine Lottoziehung wird die Menge der 49 Kugeln in zwei disjunkte Teilmengen zerlegt; die 6 *gezogenen* Kugeln und sie 43 *nicht gezogenen* Kugeln. Ein Lotto-Tipp besteht aus der Angabe von 6 Zahlen. "Vier Richtige" hat man genau dann, wenn 4 dieser Zahlen in der 6-er Teilmenge und 2 dieser Zahlen in der 43-er-Teilmenge liegen. Dies sind offenbar  $\binom{6}{4} \cdot \binom{43}{2}$  Möglichkeiten.

In *Mathematica*:

```
Binomial[6, 4]*Binomial[43, 2]
```

```
13545
```

Im Sinne des Laplace'schen Wahrscheinlichkeitbegriffs sind dies die "günstigen" Fälle. Ihr Anteil an "allen" Fällen ist somit  $\frac{\binom{6}{4} \cdot \binom{43}{2}}{\binom{49}{6}}$

In *Mathematica*:

```
Binomial[6,4]*Binomial[43,2] / Binomial[49,6]
```

$$\frac{645}{665896}$$

In Dezimalzahl-Darstellung ("numerischer" Darstellung):

```
% // N
0.00096862
```

Entsprechend ergibt sich: Alle 7 Fälle (genau 0, 1, 2, 3, 4, 5, 6 "Richtige"):

```
Table[Binomial[6, i] * Binomial[43, (6 - i)], {i, 0, 6}]
{6096454, 5775588, 1851150, 246820, 13545, 258, 1}
```

Für alle oben genannten 7 Fälle (genau 0, 1, 2, 3, 4, 5, 6 "Richtige") ergeben sich dementsprechend die folgenden "Chancen":

```
Table[Binomial[6, i] * Binomial[43, (6 - i)] / Binomial[49, 6], {i, 0, 6}]
{
  435461 / 998844, 68757 / 166474, 44075 / 332948, 8815 / 499422, 645 / 665896, 43 / 2330636, 1 / 13983816
}
% // N
{0.435965, 0.413019, 0.132378, 0.0176504, 0.00096862, 0.0000184499, 7.15112 × 10-8}
```

Hier eine Gesamttabelle:

```
Table[{i, N[Binomial[6, i] * Binomial[43, (6 - i)] / Binomial[49, 6]]}, {i, 0, 6}] //
TableForm
0      0.435965
1      0.413019
2      0.132378
3      0.0176504
4      0.00096862
5      0.0000184499
6      7.15112 × 10-8
```

Für das weitere sei die folgende Wahrscheinlichkeitsfunktion eingeführt:

```
p[k_] := Binomial[6,k]*Binomial[43,(6-k)] / Binomial[49,6]
```

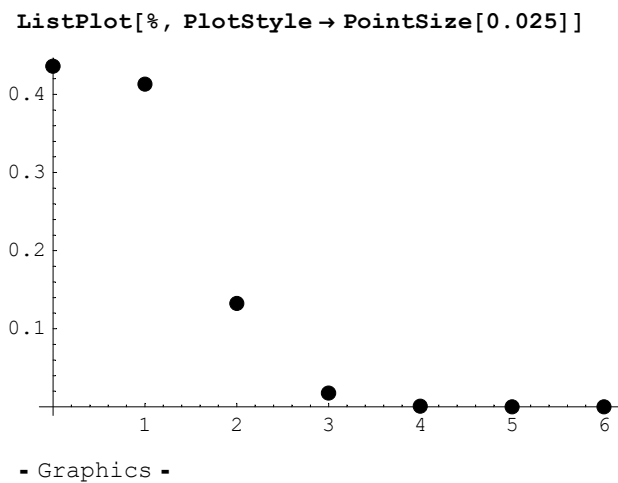
```
Table[{k, p[k]}, {k,0,6}]
```

```
{
  {0, 435461 / 998844}, {1, 68757 / 166474}, {2, 44075 / 332948},
  {3, 8815 / 499422}, {4, 645 / 665896}, {5, 43 / 2330636}, {6, 1 / 13983816}
}
```

Numerisch in Dezimalzahl-Darstellung:

```
Table[{k, p[k]}, {k,0,6}] // N
{
  {0., 0.435965}, {1., 0.413019}, {2., 0.132378}, {3., 0.0176504},
  {4., 0.00096862}, {5., 0.0000184499}, {6., 7.15112 × 10-8}
}
```

Das Ganze in graphischer Darstellung:



## ■ Lotto: Anzahl der Möglichkeiten für *mindestens* $k$ "Richtige"

Wir betrachten wieder ein konkretes Beispiel: Wie groß sind die Chancen (wie groß ist die Wahrscheinlichkeit) für mindestens 4 "Richtige"? Dies sind offenbar genau die Fälle "genau 4 Richtige" oder "genau 5 Richtige" oder "genau 6 Richtige".

Die entsprechenden Anzahlen ergeben sich offensichtlich durch Summierung.

$$\text{MindestensRichtig}[k\_]:= \sum_{i=k}^6 \text{Binomial}[6, i] * \text{Binomial}[43, (6 - i)]$$

```
MindestensRichtig[4]
```

```
13804
```

```
Table[MindestensRichtig[t], {t, 0, 6}]
```

```
{13983816, 7887362, 2111774, 260624, 13804, 259, 1}
```

Die folgende Tabelle gibt die entsprechenden Anteile an allen Möglichkeiten wieder:

```
Table[MindestensRichtig[t] / Binomial[49, 6], {t, 0, 6}]
```

```
{1, 563383/998844, 150841/998844, 4654/249711, 493/499422, 37/1997688, 1/13983816}
```

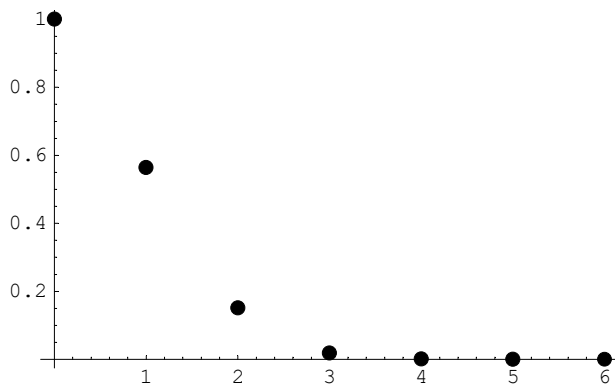
```
% // N
```

```
{1., 0.564035, 0.151016, 0.0186375, 0.000987141, 0.0000185214, 7.15112 × 10-8}
```

```
Table[{t, N[MindestensRichtig[t] / Binomial[49, 6]]}, {t, 0, 6}]
```

```
{ {0, 1.}, {1, 0.564035}, {2, 0.151016}, {3, 0.0186375},  
{4, 0.000987141}, {5, 0.0000185214}, {6, 7.15112 × 10-8}}
```

```
ListPlot[%, PlotStyle -> PointSize[0.025]]
```



- Graphics -

## ■ Histogramm / BarChart

```
Remove[BarChart]
```

```
Get["Graphics`Graphics`"]
```

```
(* alternativ:
```

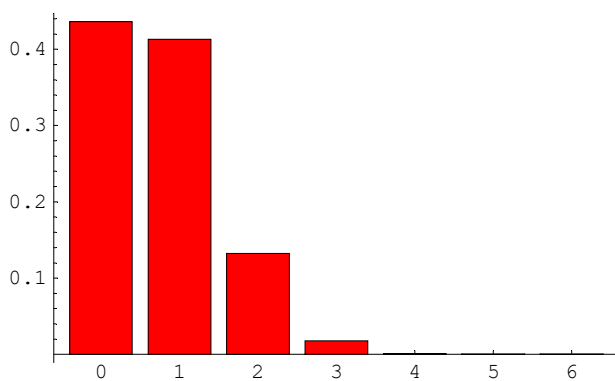
```
Remove[BarChart]
```

```
<<Graphics`Graphics` *)
```

```
Table[{N[Binomial[6, i] * Binomial[43, (6 - i)] / Binomial[49, 6]], i}, {i, 0, 6}]
```

```
{ {0.435965, 0}, {0.413019, 1}, {0.132378, 2}, {0.0176504, 3},  
  {0.00096862, 4}, {0.0000184499, 5}, {7.15112 × 10-8, 6} }
```

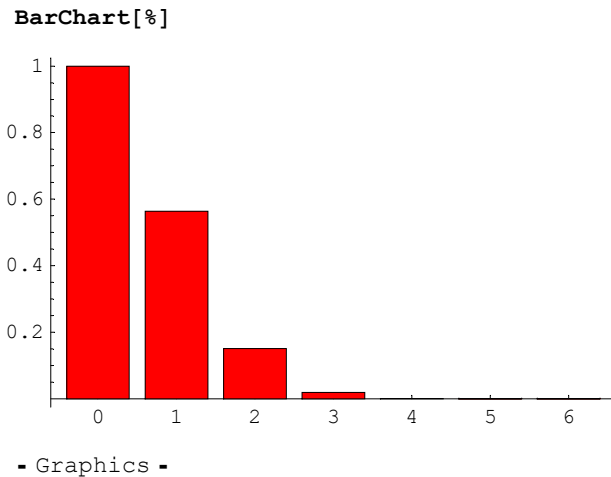
```
BarChart[%]
```



- Graphics -

```
Table[{N[MindestensRichtig[t] / Binomial[49, 6]], t}, {t, 0, 6}]
```

```
{ {1., 0}, {0.564035, 1}, {0.151016, 2}, {0.0186375, 3},  
  {0.000987141, 4}, {0.0000185214, 5}, {7.15112 × 10-8, 6} }
```



## ■ Informationsgehalt und Entropie

Die folgende Theorie des Informationsgehalts wurde für diskrete Quellen "ohne Gedächtnis" entwickelt. Eine diskrete Quelle ist eine Quelle, die Signale aus einem endlichen Alphabet verschickt, wobei jedes Signal mit einer bestimmten (konstanten, positiven) Wahrscheinlichkeit verschickt wird. Mit dem Zusatz „ohne Gedächtnis“ ist gemeint, dass das Auftreten je zweier aufeinanderfolgender Signale  $a_i$  und  $a_j$  unabhängig voneinander ist (im Sinne der Wahrscheinlichkeitsrechnung) - dass m.a.W. also gilt: Die Wahrscheinlichkeit für das Auftreten des Wortes  $a_i a_j$  ist gleich dem Produkt der Wahrscheinlichkeiten für das Signal  $a_i$  und das Signal  $a_j$ :  $P(a_i a_j) = P(a_i) \cdot P(a_j)$ . Die Ereignisse "Signal  $a_i$  wird verschickt" und "Signal  $a_j$  wird verschickt" sind unabhängige Ereignisse im Sinne der Wahrscheinlichkeitsrechnung; die Quelle „weiß“ beim Senden eines Signals nicht, welche Signale sie vorher verschickt hat (daher: „Quelle ohne Gedächtnis“).

*Hinführung zum Begriff des Informationsgehalts:* Information muss gesucht werden. Je spezifischer, eingegrenzter, rarer, seltener, genauer eine Mitteilung ist, desto größer ist ihr Informationsgehalt. Wenn die Polizei z.B. einen Verbrecher sucht, dann ist der Hinweis „Er wurde am Kaiserstuhl gesehen“ aus informationstheoretischer (und sicher auch aus kriminalistischer) Sicht wertvoller als der Hinweis „Er wurde in Süddeutschland gesehen“.

*Ein weiteres Beispiel:* Nehmen wir an, wir haben Lotto gespielt und eine Quelle mit dem Alphabet  $A = \{0, 1, 2, 3, 4, 5, 6\}$  versendet nach der Lotto-Auslosung an uns eines der Signale:  $x$  (aus  $A$ ) mit der Bedeutung „Du hast (genau)  $x$  Richtige“. Dann ist offensichtlich, dass die Signale  $0, 1, \dots, 6$  mit unterschiedlicher Wahrscheinlichkeit auftreten und es erscheint plausibel, dass die Nachricht „6“ einen höheren Informationsgehalt hat als z.B. die Nachricht „3“.

Im folgenden wird also eine diskrete Quelle ohne Gedächtnis betrachtet. Die Quelle versendet Signale  $a_k$  aus einem *endlichen* Quellenalphabet jeweils mit (positiver) Wahrscheinlichkeit  $p_k$  ( $k = 1, \dots, n$ ).

Der *Informationsgehalt* des Signals  $a_k$  ist dann definiert als  $I(a_k) := \log_2\left(\frac{1}{p_k}\right)$ .

Als Maßeinheit wird festgelegt: 1 bit oder auch 1 Shannon - zu Ehren von Claude Shannon, 1916-2001, amerikanischer Ingenieur und Mathematiker; Begründer und Pionier der Informationstheorie.

Der durch die obige Definition gegebene Informationsgehalt eines Signals bzw. einer Nachricht ist insbesondere unabhängig von der Bedeutung der Nachricht.

Nach den Logarithmengesetzen gilt (für jede Basis  $b$ )  $\log_b\left(\frac{1}{x}\right) = -\log_b(x)$ . Deshalb gilt für den Informationsgehalt

$$I(a_k) = \log_2\left(\frac{1}{p_k}\right) = -\log_2(p_k).$$

*Plausibilitätsbetrachtung:* Ist die Wahrscheinlichkeit  $p_a$  für ein Signal  $a$  gleich 1, dann ist das Eintreffen des Signals  $a$  das "sichere" Ereignis. Entsprechend der obigen Definition ist dann  $I(a) = \log_2\left(\frac{1}{p_a}\right) = \log_2(1) = 0$ . Dies ist hochgradig plausibel, denn die Nachricht, dass ein sicheres Ereignis eingetreten ist, hat ersichtlich keinen Informationswert.

Ein weiteres *Beispiel:* Ist die Wahrscheinlichkeit  $p_a$  für ein Signal  $a$  gleich  $\frac{1}{2}$ , dann gilt:

$$I(a) = \log_2\left(\frac{1}{\left(\frac{1}{2}\right)}\right) = \log_2(2) = 1.$$

Im Sinne einer optimalen Codierung könnte man z.B. die Ereignisse "Das gesendete Signal war  $a$ " bzw. "Das gesendete Signal war nicht  $a$ " folgendermaßen codieren:

- 1: Das Signal  $a$  wurde gesendet
- 0: Das Signal  $a$  wurde nicht gesendet

D.h.: Eine Nachricht mit dem (informationstheoretischen) Informationsgehalt "1 bit" kann genau in einem Computer-Bit verschlüsselt werden. Die informationstheoretische Fassung der Einheit "bit" passt also genau zusammen mit dem Begriff des Computer-"Bits".

In der Notation von *Mathematica* lautet der Informationsgehalt eines Signals  $a_k$  mit der Auftretenswahrscheinlichkeit  $p_k$ :

$$I(a_k) := \text{Log}\left[2, \frac{1}{p_k}\right] \quad (*)$$

Die *Entropie* der Quelle ist definiert durch  $H := \sum_{k=1}^n p_k * I(a_k)$ . Man kann sie deuten als das *gewichtete arithmetische Mittel* der Informationsgehalte der Quellensignale oder (im wahrscheinlichkeitstheoretischen Sinn) als den *Erwartungswert* für den Informationsgehalt eines Zeichens.

Mit Hilfe von (\*) folgt unmittelbar:  $H = \sum_{k=1}^n p_k * \text{Log}\left[2, \frac{1}{p_k}\right]$

Eine endliche, diskrete Quelle läßt sich (für  $n = 7$ ) wie folgt als Tabelle veranschaulichen:

```
LI = {"I(a0)", "I(a1)", "I(a2)", "I(a3)", "I(a4)", "I(a5)", "I(a6)"};
GridBox[Table[{ak, pk, LI[[k + 1]]}, {k, 0, 6}],
RowLines → True, ColumnLines → True] // DisplayForm
```

a <sub>0</sub>	p <sub>0</sub>	I (a <sub>0</sub> )
a <sub>1</sub>	p <sub>1</sub>	I (a <sub>1</sub> )
a <sub>2</sub>	p <sub>2</sub>	I (a <sub>2</sub> )
a <sub>3</sub>	p <sub>3</sub>	I (a <sub>3</sub> )
a <sub>4</sub>	p <sub>4</sub>	I (a <sub>4</sub> )
a <sub>5</sub>	p <sub>5</sub>	I (a <sub>5</sub> )
a <sub>6</sub>	p <sub>6</sub>	I (a <sub>6</sub> )

Ihre Entropie-Funktion läßt sich in *Mathematica* folgendermaßen realisieren (dabei sei L die Liste bestehend aus den Paaren  $\{p_k, I(a_k)\}$  bzw.  $\{p_k, \log_2\left(\frac{1}{p_k}\right)\}$ ):

```
Entropie[L_] := Apply[Plus, Map[Function[x, Apply[Times, x]], L] ]
```

## ■ 1. Beispiel: Lotto

Die Quellensignale beim Lotto sind die Signale aus der Quelle  $\{0, 1, 2, 3, 4, 5, 6\}$ . Die Wahrscheinlichkeiten der Quellensignale sind durch die weiter oben definierte Funktion  $p[k]$  gegeben.

Der Informationsgehalt des Signals  $a_k$  ( $= k$ ) ist dann gegeben durch

$$IL[k_] := \text{Log}[2, N[1/p[k]]]$$

Die folgende Tabelle gibt die Signale, sowie die Wahrscheinlichkeit und den Informationsgehalt für jedes Signal wieder.

```
Table[{k, N[p[k]], IL[k]}, {k, 0, 6}] // TableForm
```

0	0.435965	1.19772
1	0.413019	1.27572
2	0.132378	2.91726
3	0.0176504	5.82415
4	0.00096862	10.0118
5	0.0000184499	15.726
6	$7.15112 \times 10^{-8}$	23.7373

Der folgende Funktionsaufruf ermittelt die Entropie der Lotto-Quelle:

```
Entropie[Table[{N[p[k]], IL[k]}, {k, 0, 6}]]
```

1.54803

## ■ 2. Beispiel: Glücksrad (paradigmatisch für gleichverteilte Ereignisse)

Zum Vergleich betrachten wir ein Glücksrad mit den sieben gleichgroßen Feldern  $0,1,2,3,4,5,6$ . Alle Felder seien "gleichwahrscheinlich"; d.h., für alle  $k \in \{0, 1, 2, 3, 4, 5, 6\}$  ist die Signalwahrscheinlichkeit:  $pG[k] = \frac{1}{7}$ .

$$pG[k_] := 1/7$$

Somit ist der Informationsgehalt von allen Signalen aus der "Glücksrad"-Quelle gegeben durch die folgende Funktion:

```
IG[k_] := Log[2, N[1/pG[k]]]
```

```
Table[{k, IG[k]}, {k, 0, 6}]
```

```
{{0, 2.80735}, {1, 2.80735}, {2, 2.80735},
```

```
{3, 2.80735}, {4, 2.80735}, {5, 2.80735}, {6, 2.80735}}
```

Gesamt-Tabelle:



```

Table[{k, N[pG[k]], IG[k]},{k,0,6}] // TableForm
0      0.142857    2.80735
1      0.142857    2.80735
2      0.142857    2.80735
3      0.142857    2.80735
4      0.142857    2.80735
5      0.142857    2.80735
6      0.142857    2.80735

```

Die Entropie der Glücksrad-Quelle ist dann:

```

Entropie[Table[{N[pG[k]], IG[k]},{k,0,6}]]
2.80735

```

### ■ 3. Beispiel ("Extrem-Beispiel"): Sehr schiefe Verteilung

Wir betrachten nun das folgende Extrem-Beispiel: eine Quelle mit den Signalen aus dem Quellenalphabet  $\{0, 1, 2, 3, 4, 5, 6\}$  und den Signal-Wahrscheinlichkeiten:  $\{0.9994, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001\}$ .

```
pX[k_] := If[k==0, 0.9994, 0.0001]
```

Der Informationsgehalt der Signale aus der extrem schiefen Quelle ist gegeben durch:

```

IX[k_] := Log[2, N[1/pX[k]] ]
Table[{k, N[pX[k]], IX[k]},{k,0,6}] // TableForm
0      0.9994      0.000865877
1      0.0001      13.2877
2      0.0001      13.2877
3      0.0001      13.2877
4      0.0001      13.2877
5      0.0001      13.2877
6      0.0001      13.2877

```

Die Entropie der extrem schiefen Quelle ist gegeben durch

```

Entropie[Table[{N[pX[k]], IX[k]},{k,0,6}]]
0.00883798

```

Analogien: "Alles gleich"  $\iff$  hohe Entropie

"Sehr verschiedene Verhältnisse"  $\iff$  niedrige Entropie

Die Entropie lässt sich auch deuten als Maß für die "Gleichheit der Verhältniss".

## ■ Weitere Beispiele

### ■ Münzwurf mit zwei Münzen

Eine Quelle versende die Signale W, Z und G mit den Wahrscheinlichkeiten  $1/4$ ,  $1/4$  und  $1/2$ . (Deutung: Münzwurf mit zwei Münzen; W bedeute "beide Münzen zeigen Wappen"; Z bedeute "beide Münzen zeigen Zahl" und G bedeute "eine Münze zeigt Wappen und eine Münze zeigt Zahl".)

```
GridBox[{{"W", 1/4, Log[2, 1/(1/4)]},
         {"Z", 1/4, Log[2, 1/(1/4)]}, {"G", 1/2, Log[2, 1/(1/2)]}},
        RowLines -> True, ColumnLines -> True] // DisplayForm
```

W	$\frac{1}{4}$	2
Z	$\frac{1}{4}$	2
G	$\frac{1}{2}$	1

Entropie:

```
(1/4) * Log[2, 1/(1/4)] + (1/4) * Log[2, 1/(1/4)] + (1/2) * Log[2, 1/(1/2)] // N
1.5
```

Zum Vergleich: Die Entropie einer gleichverteilten Quelle mit den drei Signalen A, B und C ( $p(A) = p(B) = p(C) = 1/3$ ) ist

```
GridBox[{{"A", 1/3, N[Log[2, 1/(1/3)]]},
         {"B", 1/3, N[Log[2, 1/(1/3)]]}, {"C", 1/3, N[Log[2, 1/(1/3)]]}},
        RowLines -> True, ColumnLines -> True] // DisplayForm
```

A	$\frac{1}{3}$	1.5849625007211563
B	$\frac{1}{3}$	1.5849625007211563
C	$\frac{1}{3}$	1.5849625007211563

```
(1/3) * Log[2, 1/(1/3)] + (1/3) * Log[2, 1/(1/3)] + (1/3) * Log[2, 1/(1/3)] // N
1.58496
```

### ■ Binäre Quellen

Wir betrachten eine binäre Quelle, d.h. eine Quelle, deren Quellenalphabet aus nur zwei Elementen, etwa 0 und 1, besteht.

Die Signalwahrscheinlichkeiten seien (für  $x \in [0, 1]$ ) gegeben durch:

```
pb[0, x_] := x;
pb[1, x_] := (1 - x)
```

Der Informationsgehalt der Signale "0" bzw. "1" ist dann gegeben durch:

```
Ib[0, x_] := Log[2, 1 / pb[0, x] ] ;
Ib[1, x_] := Log[2, 1 / pb[1, x] ]
```

Und die Entropie der binären Quelle ist gegeben durch:

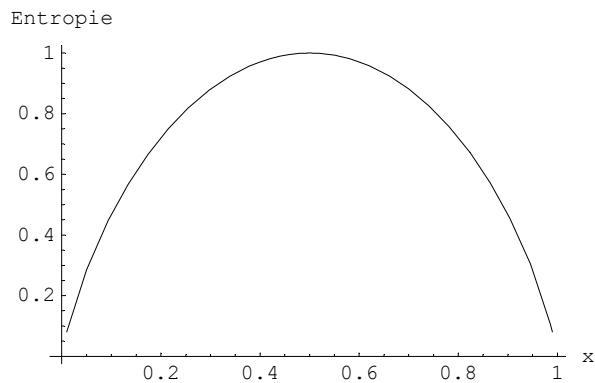
```
Hb[x_] := pb[0, x] * Ib[0, x] + pb[1, x] * Ib[1, x]
```

```
Hb[0.3]
```

```
0.881291
```

Im Schaubild sieht das Ganze folgendermassen aus:

```
Plot[Hb[x], {x, 0.01, 0.99}, AxesLabel -> {x, Entropie}]
```



- Graphics -

Die Entropie nimmt also ihren Maximalwert dann an, wenn die Signalwahrscheinlichkeiten für "0" und "1" gleichgross sind.

```
Maximize[Hb[x], {x}]
```

```
{1, {x -> 1/2}}
```

## ■ Ternäre Quellen

Wir betrachten nun eine ternäre Quelle, d.h. eine Quelle, deren Quellenalphabet aus genau *drei* Elementen, etwa 0, 1 und 2 besteht.

Die Signalwahrscheinlichkeiten seien gegeben durch  $x$ ,  $y$  und  $z$  mit  $0 < x < 1$ ,  $0 < y < 1$ ,  $0 < z < 1$  mit  $x + y + z = 1$ :

```
pt[0, x_, y_] := x ;
pt[1, x_, y_] := y ;
pt[2, x_, y_] := 1 - (x + y) ;
```

Der Informationsgehalt der entsprechenden Signale lautet dann:

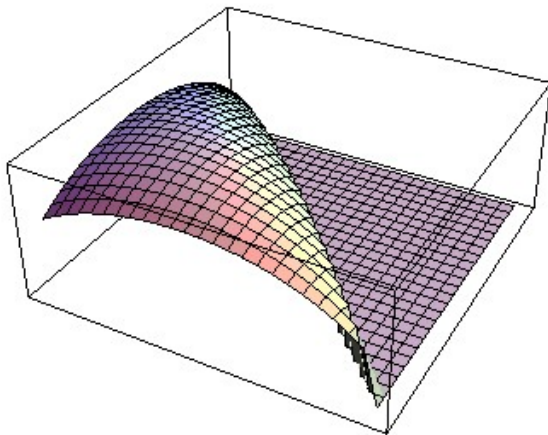
```
It[0, x_, y_] := Log[2, 1 / pt[0, x, y] ] ;
It[1, x_, y_] := Log[2, 1 / pt[1, x, y] ] ;
It[2, x_, y_] := Log[2, 1 / pt[2, x, y] ] ;
```

Und die Entropie der ternären Quelle ist gegeben durch:

```
Ht[x_, y_] :=  
  If[And[x > 0, y > 0, x + y < 1],  
    pt[0, x, y] * It[0, x, y] + pt[1, x, y] * It[1, x, y] + pt[2, x, y] * It[2, x, y], 0]  
  
Ht[0.3, 0.4]  
  
1.57095
```

Im Schaubild:

```
Plot3D[Ht[x, y], {x, 0.1, 0.9}, {y, 0.1, 0.9}]
```

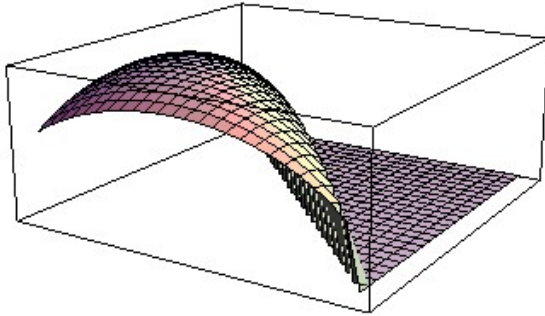


- SurfaceGraphics -

Nach Laden des folgenden Moduls kann die Darstellungsbox gedreht werden.

```
<< RealTime3D`
```

```
Plot3D[Ht[x, y], {x, 0.1, 0.9}, {y, 0.1, 0.9}]
```



```
- SurfaceGraphics -
```

```
Ht[1/3, 1/3] // N
```

```
1.58496
```

Zurück zur Standarddarstellung:

```
<< Default3D`
```

## ■ Mathematica Funktion Maximize

Im folgenden Beispiel müssen die Grenzen für  $x$  und  $y$  echt zwischen 0 und 1 liegen.

```
NMaximize[{Ht[x, y], x > 0.0000001 && x < 0.9999999 &&
  y > 0.0000001 && y < 0.9999999 && x + y < 0.999999999999}, {x, y}]
{1.58496, {x -> 0.333333, y -> 0.333333}}
```

## ■ Rückblick und Querverbindungen

Der Begriff der Entropie wird auch in anderen Wissenschaftszweigen (Physik, Chemie, Ästhetik, Systemtheorie) verwendet. Dabei gibt es Gemeinsamkeiten, die sich anhand der obigen Beispiele verdeutlichen lassen.

In den obigen Beispielen wurden unterschiedliche Wahrscheinlichkeitsverteilungen für die Signale der Quelle betrachtet. Die Beispiele zeigten, dass die Entropie dann am größten war, wenn alle Signale mit derselben Signalwahrscheinlichkeit auftraten.

Entsprechend ist allgemein in der Systemtheorie die Entropie eines (geschlossenen) Systems dann am größten, wenn alle Elemente eines Systems (z.B. die Teilchen eines abgeschlossenen physikalischen Systems) denselben Zustand haben. Dies bezeichnet man als den Zustand der größten Unordnung (maximales Chaos, maximale Komplexität).

Ein System, dessen Bestandteile "unterschiedliche Verhältnisse" aufweisen, wird als strukturiert angesehen. Strukturierte Systeme weisen einen höheren Grad an innerer Ordnung auf als unstrukturierte Systeme; sie sind in höherem Maße "geordnet".

Die Entropie wird in diesem Sinne als Maß für die Unordnung eines Systems verstanden. Je größer die Entropie desto größer der Grad an Unordnung.

### **Zum Entropie-Begriff in der Physik (Thermodynamik)**

*Energie* und *Entropie* sind zwei fundamentale Begriffe der Physik. Sie sind Gegenstand des ersten und des zweiten Hauptsatzes der Thermodynamik.

Der *erste* Hauptsatz der Thermodynamik (auch als *Energiesatz* bezeichnet) konstatiert, anschaulich gesprochen, die Äquivalenz von Wärme und mechanischer Arbeit. In physikalischen Systemen kann z.B. mechanische Energie in Wärmeenergie umgewandelt werden und umgekehrt; die Gesamtenergiebilanz verändert sich dabei nicht. So kann (z.B. durch die Joulesche Apparatur) mechanische Energie durch Reibung in Wärmeenergie umgewandelt werden. In Wärmekraftmaschinen wird Wärmeenergie in mechanische Energie gewandelt.

*Ein Beispiel:* Ein auf dem Dach liegender Ziegelstein hat eine bestimmte mechanische Energie (Lageenergie, potentielle Energie). Wenn er vom Dach herab fällt, verwandelt sich diese potentielle Energie in kinetische Energie (Bewegungsenergie) und zu einem geringen Teil aufgrund der "Luftreibung" auch in Wärmeenergie. Die Gesamtenergie bleibt gleich. Wenn der Ziegelstein am Boden zerschellt, zerbröseln er. Dabei verliert er seine potentielle und seine kinetische Energie. Durch den Vorgang des Aufpralls wird Verformungsenergie (Deformationsenergie) und nochmals Wärmeenergie erzeugt. Die Gesamtenergiebilanz bleibt wiederum gleich.

Über die "Richtung" der Energie-Änderung ist im "Energiesatz" zunächst nichts ausgesagt. So wäre es z.B. (wenn es eine solche "Gerichtetheit" physikalischer Vorgänge nicht gäbe) theoretisch denkbar, dass ein am Boden liegender Dachziegel auf das Dach springt, indem er aus seiner Wärmeenergie unter Temperaturenniedrigung die kinetische Energie gewinnt, die für die dafür benötigte Arbeit erforderlich ist. Er wäre dann kälter, läge dafür aber auf dem Dach.

Eine Vielzahl von Beobachtungen und Erfahrungstatsachen bildet jedoch die Grundlage für die Auffassung, dass die Veränderungen in physikalischen Systemen *gerichtet* sind, dass sie also einen *irreversiblen* Charakter haben. Genau dies bringt der *zweite Hauptsatz der Thermodynamik*, der auch als *Entropiesatz* bezeichnet wird, zum Ausdruck. Der Entropiesatz besagt, dass in einem abgeschlossenen physikalischen System Wärmeprozesse so ablaufen, dass die Entropie des Systems niemals abnimmt.

Anschaulich ausgedrückt: Ziegelsteine können vom Dach fallen, zerbröseln, und dabei Wärme- und Verformungsenergie erzeugen, aber zerbröselte Ziegelsteine können nicht, sich abkühlend, zu einem heilen Ziegelstein zusammenfügen und auf ein Dach springen. (Dabei entspricht der "heile" Ziegelstein einem System höherer Ordnung, d.h. niedriger Entropie, und der "zerbröselte" Ziegelstein einem System niedriger Ordnung, d.h. hoher Entropie.)

*Literaturhinweis:* Chr. Gerthsen; Physik, J. Springer Verlag, Berlin

### **Zum Entropie-Begriff in der Ästhetik**

Theorie der "numerischen" Ästhetik (Max Bense): Ein total einfarbiges Bild ist maximal ungeordnet, denn es gibt keinerlei Ordnungsmerkmale in dem Bild. Seine "Entropie" ist maximal.

## **■ Hilfsprogramme**