Kapitel PTS:II

II. Wahrscheinlichkeitsbegriff

- Zufallsexperimente
- Ergebnisräume
- Ereignisräume
- Relative Häufigkeit
- Klassischer Wahrscheinlichkeitsbegriff
- Weiterentwicklung des Wahrscheinlichkeitsbegriffs

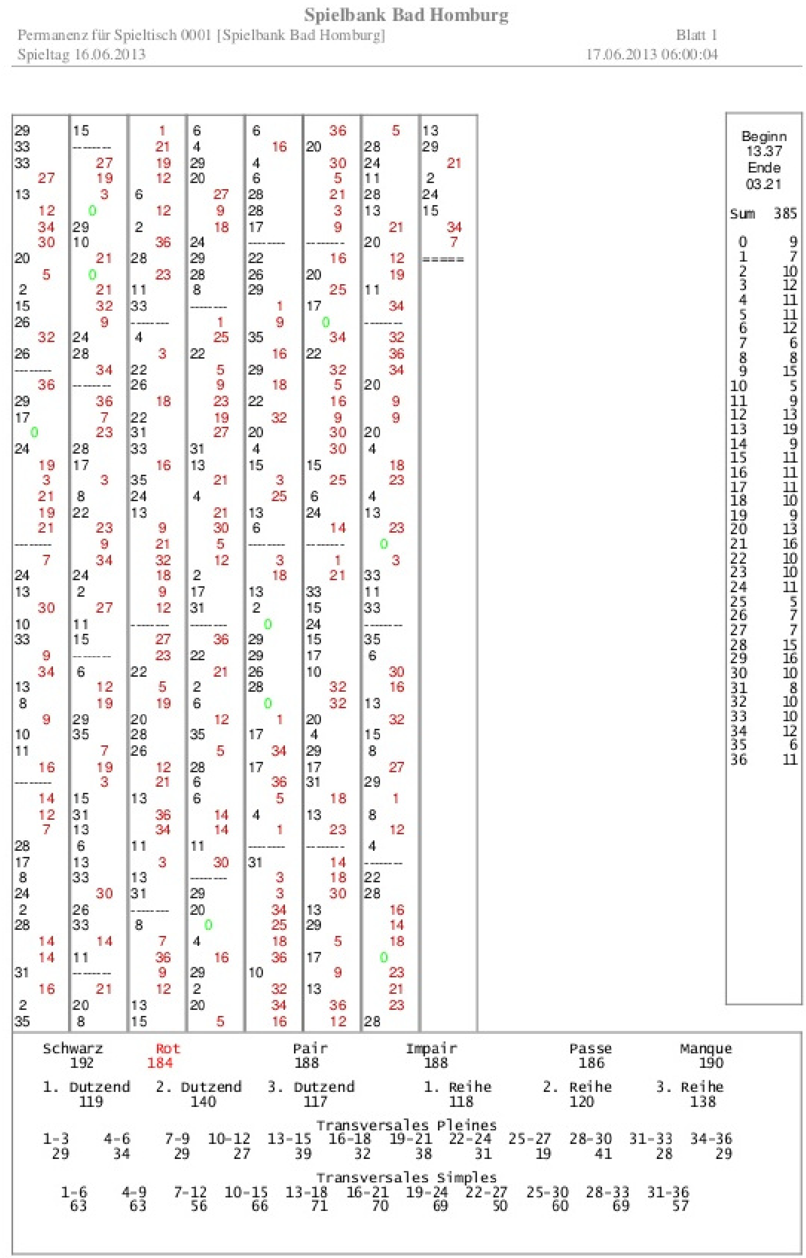
Relative Häufigkeit

Ziel:	Mathematische Modellierung des Zufalls
Schritt 1:	Beschreibung zufälliger Vorgänge als Zufallsexperiment
Schritt 2:	Zusammenfassung interessierender Ausgänge zum Ergebnisraum Ω
Schritt 3:	Identifikation interessierender Ereignisse im Ergebnisraum
Schritt 4:	Bestimmung der Häufigkeit des Ereigniseintritts
Schritt 5:	Statistische Wahrscheinlichkeit $\rightarrow$ Wahrscheinlichkeitsbegriff
Schritt 6:	Axiomatisierung des Wahrscheinlichkeitsbegriffs

Relative Häufigkeit

Beispiel: Französisches Roulette

- „Permanenzen“ zählen Ergebnisse und Ereignisse beim Roulette.
- Ereignisse an Tisch 1 in der Spielbank Bad Homburg am 16.06.2013:
  - 9-mal die Null
  - 192-mal „Schwarz“
  - 184-mal „Rot“
  - 119-mal das „1. Dutzend“
- Die Zählung von Ergebnissen und Ereignissen eines Zufallsexperiments ist der erste Schritt der beschreibenden Statistik.


Spielbank Bad Homburg

Permanenz für Spieltisch 0001 [Spielbank Bad Homburg] Blatt 1

Spieltag 16.06.2013 17.06.2013 06:00:04

Beginn 13.37 Ende 03.21

Sum 385

Schwarz	Rot	Pair	Impair	Passe	Manque
192	184	188	188	186	190

1. Dutzend	2. Dutzend	3. Dutzend	1. Reihe	2. Reihe	3. Reihe
119	140	117	118	120	138

Transversales Pleines

1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-30	31-33	34-36
29	34	29	27	39	32	38	31	19	41	28	29

Transversales Simples

1-6	4-9	7-12	10-15	13-18	16-21	19-24	22-27	25-30	28-33	31-36
63	63	56	66	71	70	69	50	60	69	57

Bemerkungen:

- Richard Jarecki hat in den 1960er und 1970er Jahren in europäischen Spielkasinos mehr als 1 Million Dollar im Roulette gewonnen. Jarecki und sein Frau Carol haben Permanenzen zehntausender Roulette-Ereignisse erhoben und ausgewertet. Sie erkannten bei verschiedenen Tischen Tendenzen zu bestimmten Ergebnissen bzw. Ereignissen, die auf Fertigungsfehler und Abnutzung zurückzuführen waren. Mit diesem Wissen ausgestattet konnten die Jareckis ihre Gewinnchancen signifikant erhöhen. In der Öffentlichkeit behauptete Jarecki hingegen, das Roulette-Spiel mit Hilfe einer Datenauswertung auf einem Großrechner „geknackt“ zu haben.

Relative Häufigkeit

Definition 7 (Absolute und relative Häufigkeit)

Tritt bei n-maliger Wiederholung eines Zufallsexperiments ein Ereignis A genau $H_n(A)$-mal auf, so heißt diese Größe die absolute Häufigkeit des Ereignisses A und

$$h_n(A) = \frac{H_n(A)}{n}$$

die relative Häufigkeit des Ereignisses A.

Relative Häufigkeit

Definition 7 (Absolute und relative Häufigkeit)

Tritt bei n-maliger Wiederholung eines Zufallsexperiments ein Ereignis A genau $H_n(A)$-mal auf, so heißt diese Größe die absolute Häufigkeit des Ereignisses A und

$$h_n(A) = \frac{H_n(A)}{n}$$

die relative Häufigkeit des Ereignisses A.

Beispiel: Französisches Roulette (Fortsetzung)

- Ereignisse an Tisch 1 in der Spielbank Bad Homburg am 16.06.2013:
  - $H_{385}(\{0\}) = 9 \quad \rightarrow \quad h_{385}(\{0\}) = \frac{9}{385}$
  - $H_{385}(\text{„Schwarz“}) = 192 \quad \rightarrow \quad h_{385}(\text{„Schwarz“}) = \frac{192}{385}$
  - $H_{385}(\text{„Rot“}) = 184 \quad \rightarrow \quad h_{385}(\text{„Rot“}) = \frac{184}{385}$
  - $H_{385}(\text{„1. Dutzend“}) = 119 \quad \rightarrow \quad h_{385}(\text{„1. Dutzend“}) = \frac{119}{385}$

Relative Häufigkeit

Eigenschaften

Beispiele: Roulette-Permanenzen Bad Homburg ($n = 385$)

- $h_n(\text{„1. Dutzend“}) \;=\; (H_n(\{1\}) + \ldots + H_n(\{12\}))/n \;=\; \frac{119}{385}$
  Das Ereignis „1. Dutzend“ entspricht der Summe der absoluten Häufigkeiten der enthaltenen Elementarereignisse, geteilt durch die Zahl der Spiele.

Relative Häufigkeit

Eigenschaften

Beispiele: Roulette-Permanenzen Bad Homburg ($n = 385$)

- h_n(„1. Dutzend“) $= (H_n(\{1\}) + \ldots + H_n(\{12\}))/n = \frac{119}{385}$

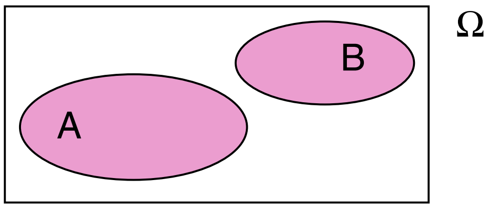
  Das Ereignis „1. Dutzend“ entspricht der Summe der absoluten Häufigkeiten der enthaltenen Elementarereignisse, geteilt durch die Zahl der Spiele.

- Seien $A =$ „1. Dutzend“ und $B =$ „3. Dutzend“ unvereinbare Ereignisse:

  – $h_n(A) = \frac{119}{385}$, $h_n(B) = \frac{117}{385}$

  – $h_n(A \cup B) = \frac{236}{385}$, $h_n(A \cap B) = 0$

  ➜ $h_n(A \cup B) = h_n(A) + h_n(B)$



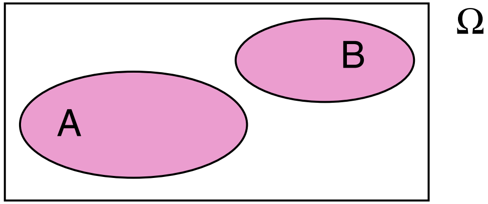
Relative Häufigkeit

Eigenschaften

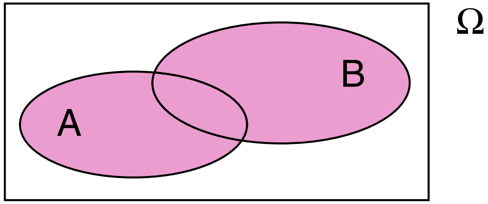
Beispiele: Roulette-Permanenzen Bad Homburg ($n = 385$)

- h_n(„1. Dutzend“) $= (H_n(\{1\}) + \ldots + H_n(\{12\}))/n = \frac{119}{385}$
  Das Ereignis „1. Dutzend“ entspricht der Summe der absoluten Häufigkeiten der enthaltenen Elementarereignisse, geteilt durch die Zahl der Spiele.

- Seien $A =$ „1. Dutzend“ und $B =$ „3. Dutzend“ unvereinbare Ereignisse:
  - $h_n(A) = \frac{119}{385}, \quad h_n(B) = \frac{117}{385}$
  - $h_n(A \cup B) = \frac{236}{385}, \quad h_n(A \cap B) = 0$
  - ➜ $h_n(A \cup B) = h_n(A) + h_n(B)$



- Seien $A =$ „1. Dutzend“ und $B =$ „ungerade Zahl“ vereinbare Ereignisse.
  - $h_n(A) = \frac{119}{385}, \quad h_n(C) = \frac{188}{385}$
  - $h_n(A \cup C) = \frac{247}{385}, \quad h_n(A \cap C) = \frac{60}{385}$
  - ➜ $h_n(A \cup C) < h_n(A) + h_n(C)$
  - ➜ $h_n(A \cup C) = h_n(A) + h_n(C) - h_n(A \cap C)$



Relative Häufigkeit
Eigenschaften

Summenregel:

- Die relative Häufigkeit von Ereignis A ist die Summe der relativen Häufigkeiten der Elementarereignisse, aus denen A zusammengesetzt ist:

$$h_n(A) = \sum_{\omega \in A} h_n(\{\omega\})$$

- Daraus folgt:

$$h_n(\emptyset) = 0, \quad h_n(\Omega) = 1, \text{ und } 0 \leq h_n(A) \leq 1$$

- Bemerkung:
  - Aus $h_n(A) = 0$ oder $h_n(A) = 1$ darf nicht geschlossen werden, dass A das unmögliche oder das sichere Ereignis ist.
  - $h_n(A)$ ist ja nur auf Basis von n Versuchen bestimmt worden und beim $(n+1)$-ten Versuch könnte A ja zum ersten Mal (nicht) eintreten.

Relative Häufigkeit

Eigenschaften

Additionsregel für unvereinbare Ereignisse (nur eines kann eintreten)

- $h_n(A \cup B) = h_n(A) + h_n(B), \quad$ falls $A \cap B = \emptyset$

Relative Häufigkeit

Eigenschaften

Additionsregel für unvereinbare Ereignisse (nur eines kann eintreten)

- $h_n(A \cup B) = h_n(A) + h_n(B), \quad$ falls $A \cap B = \emptyset$
- Beispiel $B = \bar{A}$: $\quad h_n(A \cup \bar{A}) = h_n(A) + h_n(\bar{A}) = 1 = h_n(\Omega)$
- Daraus folgt: $\quad h_n(A) = 1 - h_n(\bar{A})$

Relative Häufigkeit

Eigenschaften

Additionsregel für unvereinbare Ereignisse (nur eines kann eintreten)

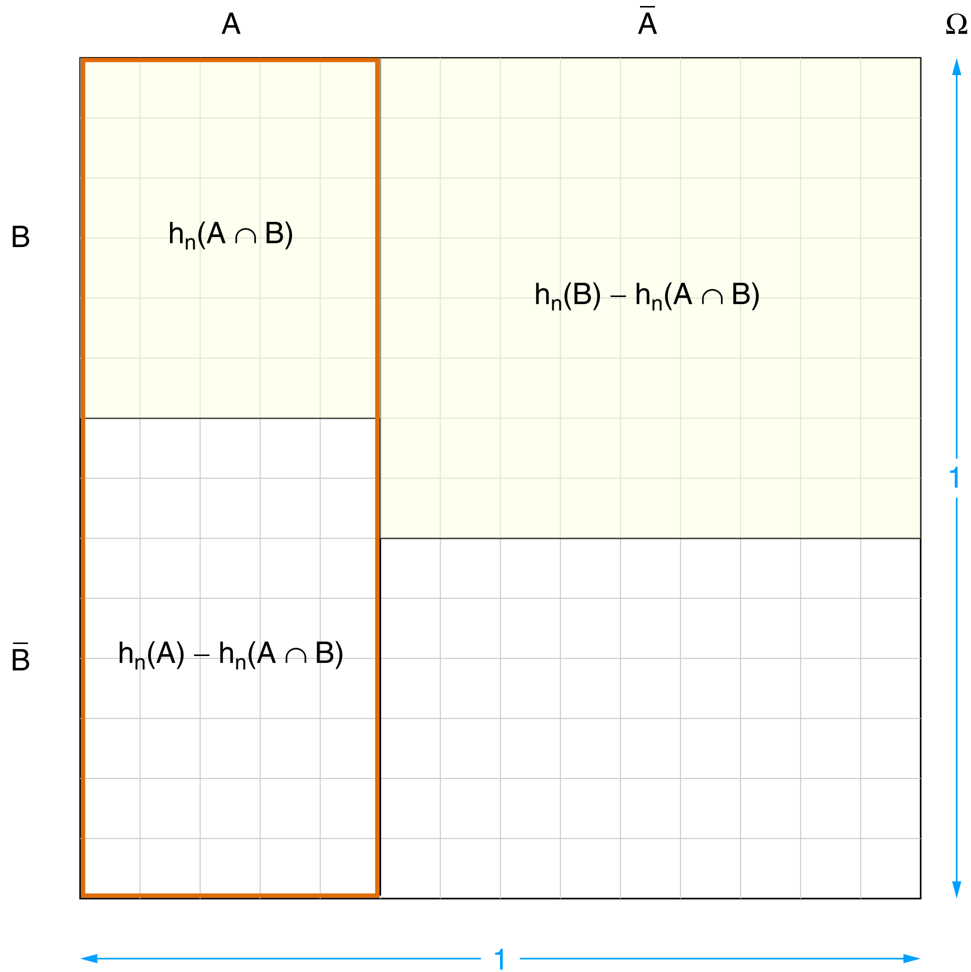
- $h_n(A \cup B) = h_n(A) + h_n(B), \quad$ falls $A \cap B = \emptyset$
- Beispiel $B = \bar{A}$: $\quad h_n(A \cup \bar{A}) = h_n(A) + h_n(\bar{A}) = 1 = h_n(\Omega)$
- Daraus folgt: $\quad h_n(A) = 1 - h_n(\bar{A})$

„Additionsregel“ für vereinbare Ereignisse (beide können gleichzeitig eintreten)

- $h_n(A \cup B) = h_n(A) + h_n(B) - h_n(A \cap B), \quad$ falls $A \cap B \neq \emptyset$
- Bei $h_n(A) + h_n(B)$ werden die Elementarereignisse aus $A \cap B$ doppelt gezählt, und müssen daher abgezogen werden.
- Bemerkung: Wenn „doch“ $A \cap B = \emptyset$, dann wird der letzte Term Null und wir erhalten automatisch die Additionsregel für unvereinbare Ereignisse

Relative Häufigkeit

Eigenschaften: Additionsregel für zwei Ereignisse



[Feuerpfeil/Heigel 1999]

Die Summe der im Venn-Diagramm angegebenen relativen Häufigkeiten ergibt $h_n(A \cup B)$.

Relative Häufigkeit

Beispiel: Blutgruppen

- Die roten Blutkörperchen eines Menschen können das Antigen A, das Antigen B, beide („AB“) oder keines („0“) besitzen.

- Blutgruppen in einer Stichprobe von $n = 30.000$ Personen:
  A: 44%, B: 10%, AB: 4%, 0: 42%.

Relative Häufigkeit

Beispiel: Blutgruppen

- Die roten Blutkörperchen eines Menschen können das Antigen A, das Antigen B, beide („AB“) oder keines („0“) besitzen.

- Blutgruppen in einer Stichprobe von $n = 30.000$ Personen:
  A: 44%, B: 10%, AB: 4%, 0: 42%.

- Sei „Person trägt Antigen A“ das Ereignis A und das Ereignis B analog.
  A und B können wie folgt kombiniert werden: $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, und $\overline{A} \cap \overline{B}$.

- Tafel der relativen Häufigkeiten der möglichen Ereignisse (in Prozent):

	A	$\overline{A}$	Σ
B	4 (AB)	10 (B)	14
$\overline{B}$	44 (A)	42 (0)	86
Σ	48	52	100

Relative Häufigkeit
Beispiel: Blutgruppen

- Die roten Blutkörperchen eines Menschen können das Antigen A, das Antigen B, beide („AB“) oder keines („0“) besitzen.

- Blutgruppen in einer Stichprobe von $n = 30.000$ Personen:
  A: 44%, B: 10%, AB: 4%, 0: 42%.

- Sei „Person trägt Antigen A“ das Ereignis A und das Ereignis B analog.
  A und B können wie folgt kombiniert werden: $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, und $\overline{A} \cap \overline{B}$.

- Tafel der relativen Häufigkeiten der möglichen Ereignisse (in Prozent):

	A	$\overline{A}$	Σ
B	4 (AB)	10 (B)	14
$\overline{B}$	44 (A)	42 (0)	86
Σ	48	52	100

- Mit dieser Tafel und der Additionsregel können wir bspw. berechnen:

$$h_n(A \cup B) = 48\% + 14\% - 4\% = 58\%$$

Relative Häufigkeit

Beispiel: Blutgruppen

- Die roten Blutkörperchen eines Menschen können das Antigen A, das Antigen B, beide („AB“) oder keines („0“) besitzen.

- Blutgruppen in einer Stichprobe von $n = 30.000$ Personen:
  A: 44%, B: 10%, AB: 4%, 0: 42%.

- Sei „Person trägt Antigen A“ das Ereignis A und das Ereignis B analog.
  A und B können wie folgt kombiniert werden: $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, und $\overline{A} \cap \overline{B}$.

- Tafel der relativen Häufigkeiten der möglichen Ereignisse (in Prozent):

	A	$\overline{A}$	Σ
B	4 (AB)	10 (B)	14
$\overline{B}$	44 (A)	42 (0)	86
Σ	48	52	100

- Mit dieser Tafel und der Additionsregel können wir bspw. berechnen:

$$h_n(A \cup B) = 48\% + 14\% - 4\% = 58\% \qquad h_n(\overline{A \cap B}) =$$

Relative Häufigkeit

Beispiel: Blutgruppen

- Die roten Blutkörperchen eines Menschen können das Antigen A, das Antigen B, beide („AB“) oder keines („0“) besitzen.

- Blutgruppen in einer Stichprobe von $n = 30.000$ Personen:
  A: 44%, B: 10%, AB: 4%, 0: 42%.

- Sei „Person trägt Antigen A“ das Ereignis A und das Ereignis B analog.
  A und B können wie folgt kombiniert werden: $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, und $\overline{A} \cap \overline{B}$.

- Tafel der relativen Häufigkeiten der möglichen Ereignisse (in Prozent):

	A	$\overline{A}$	Σ
B	4 (AB)	10 (B)	14
$\overline{B}$	44 (A)	42 (0)	86
Σ	48	52	100

- Mit dieser Tafel und der Additionsregel können wir bspw. berechnen:

$$h_n(A \cup B) = 48\% + 14\% - 4\% = 58\% \qquad h_n(\overline{A \cap B}) = 96\%$$

Relative Häufigkeit

Vierfeldertafel

- Zwei Ereignisse A und B zerlegen den Ergebnisraum in vier paarweise unvereinbare Ereignisse $A \cap B$, $A \cap \bar{B}$, $\bar{A} \cap B$ und $\bar{A} \cap \bar{B}$.

- Die absoluten / relativen Häufigkeiten der vier Ereignisse werden oft in einer sogenannten Vierfeldertafel notiert:

	A	$\overline{A}$	Σ
B	$H_n(A \cap B)$	$H_n(\bar{A} \cap B)$	$H_n(B)$
$\overline{B}$	$H_n(A \cap \bar{B})$	$H_n(\bar{A} \cap \bar{B})$	$H_n(\bar{B})$
Σ	$H_n(A)$	$H_n(\bar{A})$	n

	A	$\overline{A}$	Σ
B	$h_n(A \cap B)$	$h_n(\bar{A} \cap B)$	$h_n(B)$
$\overline{B}$	$h_n(A \cap \bar{B})$	$h_n(\bar{A} \cap \bar{B})$	$h_n(\bar{B})$
Σ	$h_n(A)$	$h_n(\bar{A})$	1

- Die absoluten / relativen Häufigkeiten der Einzelereignisse A, $\bar{A}$, B und $\bar{B}$ sind die Spalten- und Zeilensummen an den Rändern.

  Man nennt sie daher (absolute / relative) Randhäufigkeiten.

Relative Häufigkeit

Beispiel: Untersuchungen in der Sprachstatistik / Computerlinguistik / Websuche

Wörter: [Wikipedia]

- 30 Wörter machen etwa 32% der Wortanzahl eines deutschen Texts aus:

  die, der, und, in, zu, den, das, nicht, von, sie, ist, des, sich, mit, dem, dass, er, es, ein, ich, auf, so, eine, auch, als, an, nach, wie, im, für

Satzlänge: [dtv-Atlas zur deutschen Sprache]

- Kurze Sätze (4–12 Wörter) sind in Boulevard-Medien mit fast 50% aller Sätze deutlich häufiger relativ zu wissenschaftlichen Texten (etwa 10%).

„Neue“ Suchanfragen: [Google]

- Von den seit Jahren mehr als 3 Mrd. Suchanfragen pro Tag hat Google 15% noch nie vorher gesehen.

Relative Häufigkeit

Beispiel: Münzwurf

- Das Ergebnis eines Münzwurfs („Kopf“ oder „Zahl“) ist nicht vorhersagbar.
- Bei vielen Münzwürfen erwartet man aber, dass beides gleich oft auftritt.

Relative Häufigkeit

Beispiel: Münzwurf

- Das Ergebnis eines Münzwurfs („Kopf“ oder „Zahl“) ist nicht vorhersagbar.
- Bei vielen Münzwürfen erwartet man aber, dass beides gleich oft auftritt.

Experiment:

- 200-maliger Münzwurf (Ereignisse K.opf, Z.ahl, Zwischenstand alle 10 Würfe):

	#K	n	$h_n(\{\mathtt{K}\})$
ZKKKZKKZKK	7	10	0,700

Relative Häufigkeit

Beispiel: Münzwurf

- Das Ergebnis eines Münzwurfs („Kopf“ oder „Zahl“) ist nicht vorhersagbar.
- Bei vielen Münzwürfen erwartet man aber, dass beides gleich oft auftritt.

Experiment:

- 200-maliger Münzwurf (Ereignisse $\texttt{K}$.opf, $\texttt{Z}$.ahl, Zwischenstand alle 10 Würfe):

	#K	n	$h_n(\{\texttt{K}\})$
ZKKKZKKZKK	7	10	0,700
ZZKKZKKKZK	13	20	0,650

Relative Häufigkeit

Beispiel: Münzwurf

- Das Ergebnis eines Münzwurfs („Kopf“ oder „Zahl“) ist nicht vorhersagbar.
- Bei vielen Münzwürfen erwartet man aber, dass beides gleich oft auftritt.

Experiment:

- 200-maliger Münzwurf (Ereignisse K.opf, Z.ahl, Zwischenstand alle 10 Würfe):

	#K	n	$h_n(\{K\})$
ZKKKZKKZKK	7	10	0,700
ZZKKZKKKZK	13	20	0,650
ZZZZZKZKKZ	16	30	0,533
KKZZKKKKZK	23	40	0,575
KZKZZZZKZZ	26	50	0,520
KZKKZZKZKZ	31	60	0,517
ZZZKKZZZZZ	33	70	0,471
KKKZKZZZKK	39	80	0,488
ZZKZZKZKZK	43	90	0,478
ZZZZKKZZKZ	46	100	0,460

	#K	n	$h_n(\{K\})$
KKKKKZKZZK	53	110	0,482
KZKKZKKKKK	61	120	0,508
ZZKZZKKZKK	66	130	0,508
KZZZKZKKZZ	70	140	0,500
ZZKKZZZKZZ	73	150	0,486
KKZKKKKZKK	81	160	0,506
ZKZZKKZKKK	87	170	0,512
ZZKZZZZKZZ	89	180	0,494
KZKZZZKKZZ	93	190	0,489
KKZKZZZKKK	99	200	0,495

Relative Häufigkeit

Beispiel: Münzwurf

- Das Ergebnis eines Münzwurfs („Kopf“ oder „Zahl“) ist nicht vorhersagbar.
- Bei vielen Münzwürfen erwartet man aber, dass beides gleich oft auftritt.

Experiment:

- 200-maliger Münzwurf (Ereignisse `K`.opf, `Z`.ahl, Zwischenstand alle 10 Würfe):

	#K	n	$h_n(\{\mathrm{K}\})$
ZKKKZKKZKK	7	10	0,700
ZZKKZKKKZK	13	20	0,650
ZZZZZKZKKZ	16	30	0,533
KKZZKKKKZK	23	40	0,575
KZKZZZZKZZ	26	50	0,520
KZKKZZKZKZ	31	60	0,517
ZZZKKZZZZZ	33	70	0,471
KKKZKZZZKK	39	80	0,488
ZZKZZKZKZK	43	90	0,478
ZZZZKKZZKZ	46	100	0,460

	#K	n	$h_n(\{\mathrm{K}\})$
KKKKKZKZZK	53	110	0,482
KZKKZKKKKK	61	120	0,508
ZZKZZKKZKK	66	130	0,508
KZZZKZKKZZ	70	140	0,500
ZZKKZZZKZZ	73	150	0,486
KKZKKKKZKK	81	160	0,506
ZKZZKKZKKK	87	170	0,512
ZZKZZZZKZZ	89	180	0,494
KZKZZZKKZZ	93	190	0,489
KKZKZZZKKK	99	200	0,495

➜ $h_n(\{\mathrm{K}\})$ scheint sich mit wachsender Wurfanzahl um 0,5 zu stabilisieren.

Relative Häufigkeit

Beispiel: Geburtenstatistik

- Lebendgeborene: Deutschland, Monate, Geschlecht [Statistisches Bundesamt]

2020	Offizielle Statistik		
	$♀_i$	$♂_i$	n_i
Jan	31.049	32.664	63.713
Feb	28.512	30.188	58.700
Mär	30.584	31.646	62.230
Apr	29.835	31.404	61.239
Mai	31.247	33.457	64.704
Jun	32.194	33.916	66.110
Jul	34.436	36.626	71.062
Aug	33.800	35.892	69.692
Sep	33.533	35.924	69.457
Okt	32.164	33.855	66.019
Nov	29.041	30.461	59.502
Dez	29.364	31.352	60.716

Relative Häufigkeit

Beispiel: Geburtenstatistik

- Lebendgeborene: Deutschland, Monate, Geschlecht [Statistisches Bundesamt]

2020	Offizielle Statistik			$h_{n_i}(\{\male_i\})$
	$\female_i$	$\male_i$	n_i	
Jan	31.049	32.664	63.713	0.5127
Feb	28.512	30.188	58.700	0.5143
Mär	30.584	31.646	62.230	0.5085
Apr	29.835	31.404	61.239	0.5128
Mai	31.247	33.457	64.704	0.5171
Jun	32.194	33.916	66.110	0.5130
Jul	34.436	36.626	71.062	0.5154
Aug	33.800	35.892	69.692	0.5150
Sep	33.533	35.924	69.457	0.5172
Okt	32.164	33.855	66.019	0.5128
Nov	29.041	30.461	59.502	0.5119
Dez	29.364	31.352	60.716	0.5164

➜ Schwankungen von $h_{n_i}(\{\male\})$ im Intervall $[0{,}509; 0{,}517]$.

Relative Häufigkeit

Beispiel: Geburtenstatistik

- Lebendgeborene: Deutschland, Monate, Geschlecht [Statistisches Bundesamt]

2020	Offizielle Statistik			$h_{n_i}(\{♂_i\})$	kumuliert		$h_n(\{♂\})$
	$♀_i$	$♂_i$	n_i		$n = \sum n_i$	$z = \sum ♂_i$	
Jan	31.049	32.664	63.713	0.5127	63.713	32.664	0.5127
Feb	28.512	30.188	58.700	0.5143	122.413	62.852	0.5134
Mär	30.584	31.646	62.230	0.5085	184.643	94.498	0.5118
Apr	29.835	31.404	61.239	0.5128	245.882	125.902	0.5120
Mai	31.247	33.457	64.704	0.5171	310.586	159.359	0.5131
Jun	32.194	33.916	66.110	0.5130	376.696	193.275	0.5131
Jul	34.436	36.626	71.062	0.5154	447.758	229.901	0.5134
Aug	33.800	35.892	69.692	0.5150	517.450	265.793	0.5137
Sep	33.533	35.924	69.457	0.5172	586.907	301.717	0.5141
Okt	32.164	33.855	66.019	0.5128	652.926	335.572	0.5140
Nov	29.041	30.461	59.502	0.5119	712.428	366.033	0.5138
Dez	29.364	31.352	60.716	0.5164	773.144	397.385	0.5140

➜ Schwankungen von $h_{n_i}(\{♂\})$ im Intervall $[0{,}509; 0{,}517]$.

➜ Schwankungen von $h_n(\{♂\})$ im Intervall $[0{,}512; 0{,}514]$.

Relative Häufigkeit

Beispiel: Geburtenstatistik

- Lebendgeborene: Deutschland, Jahre, Geschlecht [Statistisches Bundesamt]

Jahr	Offizielle Statistik			$h_{n_i}(\{♂_i\})$	kumuliert		$h_n(\{♂\})$
	$♀_i$	$♂_i$	n_i		$n = \sum n_i$	$z = \sum ♂_i$	
2009	323.877	341.249	665.126	0.5131	665.126	341.249	0.5131
2010	330.710	347.237	677.947	0.5122	1.343.073	688.486	0.5126
2011	322.786	339.899	662.685	0.5129	2.005.758	1.028.385	0.5127
2012	327.915	345.629	673.544	0.5131	2.679.302	1.374.014	0.5128
2013	332.249	349.820	682.069	0.5129	3.361.371	1.723.834	0.5128
2014	348.092	366.835	714.927	0.5131	4.076.298	2.090.669	0.5129
2015	359.097	378.478	737.575	0.5131	4.813.873	2.469.147	0.5129
2016	386.546	405.585	792.141	0.5120	5.606.014	2.874.732	0.5128
2017	382.374	402.510	784.901	0.5128	6.390.915	3.277.242	0.5128
2018	383.471	404.052	787.523	0.5131	7.178.438	3.681.294	0.5128
2019	378.798	399.292	778.090	0.5132	7.956.528	4.080.586	0.5129
2020	375.759	397.385	773.144	0.5140	8.729.672	4.477.971	0.5130

➜ Deutliche Stabilisierung von $h_n(\{♂\})$ um einen Wert leicht geringer als $0{,}513$.

Bemerkungen:

- Das statistische Bundesamtes erfasst das biologische Geschlecht bei Geburt. Hermaphroditismus („Doppelgeschlechtlichkeit“) wird nicht gesondert ausgewiesen.
- Die Stabilisierung ist nicht „fair“ bei $0{,}5$, es sind im Durchschnitt etwas mehr Jungen als Mädchen.
- Die Geschlechterverteilung direkt nach der Empfängnis ist bislang nicht final geklärt. Sekundär wird kann sie von einer geschlechtsspezifischen Sterberaten je nach Entwicklungsphase des Fötus beeinflusst, so dass das unbeeinflusste Verhältnis etwa $1{,}05 : 1$ ist. Tertiär wird das Verhältnis von kulturellen, gesellschaftlichen, und regionalen Faktoren beeinflusst, wie zum Beispiel geschlechtsspezifischer Geburtenverhinderung oder dem Aussetzen von Kindern unerwünschtem Geschlechts. [Wikipedia]

Relative Häufigkeit

Empirisches Gesetz der großen Zahlen / Statistische Wahrscheinlichkeit

Empirisches Gesetz der großen Zahlen:

> Es gibt Ereignisse, deren relative Häufigkeit nach einer hinreichend großen Anzahl von Versuchen ungefähr gleich einem festen Zahlenwert ist.

Relative Häufigkeit

Empirisches Gesetz der großen Zahlen / Statistische Wahrscheinlichkeit

Empirisches Gesetz der großen Zahlen:

> Es gibt Ereignisse, deren relative Häufigkeit nach einer hinreichend großen Anzahl von Versuchen ungefähr gleich einem festen Zahlenwert ist.

- Gesetzmäßigkeiten, die nur bei einer großen Anzahl von Versuchen feststellbar werden statistisches Gesetz genannt.
- Ereignisse, bei denen das empirische Gesetz der großen Zahlen feststellbar ist, werden statistische Ereignisse genannt.
- Je näher die relative Häufigkeit eines Ereignisses sich sich bei 1 stabilisiert, umso „wahrscheinlicher“ nennt man es umgangsprachlich.

→ Dieser Zahlenwert wird daher als die statistische Wahrscheinlichkeit des Ereignisses bezeichnet.

Relative Häufigkeit

Empirisches Gesetz der großen Zahlen / Statistische Wahrscheinlichkeit

Das empirische Gesetz der Großen Zahlen ist eine Erfahrungstatsache und damit letztlich analog zu einem physikalischen Gesetz.

Beispiel:

Die Wärme, Masse und Dichte einer Münze sind ebenso charakteristische, physikalische Eigenschaften, wie die statistische Wahrscheinlichkeit, bei einem Wurf Kopf zu erhalten.

Relative Häufigkeit

Empirisches Gesetz der großen Zahlen / Statistische Wahrscheinlichkeit

Das empirische Gesetz der Großen Zahlen ist eine Erfahrungstatsache und damit letztlich analog zu einem physikalischen Gesetz.

Analogie zur Festlegung eines objektiven Temperaturmaßes:

	Wärme	Wahrscheinlichkeit
Begriff	Temperatur als *Wärmegrad* eines Körpers	Wahrscheinlichkeit als *Grad der Sicherheit* eines Ereignisses
Subjektive Maß-Skala	kalt, warm, heiß, . . .	(sehr) unwahrscheinlich, wahrscheinlich, höchstwahrscheinlich, . . .
Erfahrungstatsache (physikal. Gesetz)	Die meisten Körper *dehnen* sich bei Erwärmung aus. (Wärmeausdehnungsgesetz)	Es gibt Ereignisse, deren relative Häufigkeit sich bei sehr vielen Versuchen um einen festen Zahlenwert *stabilisiert*. (empirisches Gesetz der großen Zahlen)
Objektives Maß	Volumen eines Körpers als *Maß für seine Temperatur* (Thermometer)	Der Wert der Stabilisierung der relativen Häufigkeit eines Ereignisses als *Maß für seine Wahrscheinlichkeit* (statistische Wahrscheinlichkeit)