

# STOCHASTIK

## H Wahrscheinlichkeitstheorie und Grundfragen ihrer Anwendung

Jeder Mensch, der aufmerksam seine Lebensumwelt beobachtet, bemerkt Vorgänge, deren Ereignisse sich nicht mit Sicherheit vorhersagen lassen, da ihr Eintreten vom Zufall abhängt. Derartige Vorgänge zu beschreiben und zu beherrschen ist das Anliegen der Wahrscheinlichkeitstheorie und der Statistik, für die der zusammenfassende Begriff „Stochastik“ üblich geworden ist.

Die historischen Wurzeln der **Wahrscheinlichkeitstheorie** liegen im Interesse an Glücksspielen und im Streben nach quantitativen Aussagen über Gewinnchancen. So wandte sich 1654 Chevalier Antonie Gombaud DE MÉRÉ (1610–1685), ein am Hof Ludwig des XIV. lebender Philosoph und Literat, an den berühmten Mathematiker Blaise PASCAL (1623–1662) mit der Frage: *Was ist wahrscheinlicher: bei vier Würfeln mit einem Würfel mindestens eine Sechs zu werfen, oder bei 24 Würfeln mit zwei Würfeln mindestens eine Doppelsechs zu werfen?* (s. Beispiel H 58) Durch diese und ähnliche Fragen zu Glücksspielen wurde zwischen den Mathematikern PASCAL und Pierre DE FERMAT (1601–1665) ein interessanter Briefwechsel angeregt, der als Geburtsurkunde der Wahrscheinlichkeitstheorie gilt.

Lotterien als eine Form von Glücksspielen ziehen die Menschen von jeher in ihren Bann. Ihre Wurzeln reichen bis in die Republik Genua des 17. Jahrhunderts zurück. Nach Deutschland kam das so genannte Genuesische Lotto „5 aus 90“ im 18. Jahrhundert. Frühzeitig verband man mit dem Lotto auch soziale Anliegen. So ordnete man z. B. jeder der 90 Zahlen den Namen eines unverheirateten Mädchens zu. Die fünf Mädchen, deren Losnummern gezogen wurden, erhielten eine Aussteuer von 50 Talern .... Wenngleich der Anwendungsbereich der Wahrscheinlichkeitsrechnung schon längst weit über Glücksspielprobleme hinausreicht, finden solche Fragen natürlich auch noch heute gerade im Alltag immer wieder reges Interesse – sicher mehr noch als viele andere mathematische Probleme. Nehmen wir ein Beispiel:

Anton, Katrin und Jonas – wieder einmal in „Kleingeldnot“ – diskutieren über einen erfolgversprechenden Tipp bei „6 aus 49“. Vor ihnen liegt eine Tabelle mit den Ziehungshäufigkeiten der einzelnen Zahlen in den 2320 Ausspielungen seit 9. 10. 1955.

Anton meint: „Vielleicht sollten wir diejenigen Zahlen nehmen, die in den vergangenen Jahren am wenigsten gezogen wurden, also die 13, die 28 oder die 34“ – worauf Katrin entgegnet: „Dann kannst du auch gleich 1, 2, 3, 4, 5, 6, die ersten 6 Primzahlen oder so etwa Ähnliches nehmen, denn die Ziehungsmaschine ‚merkt‘ sich doch nicht, welche Zahlen bereits früher gezogen wurden – bei jeder neuen Ziehung hat jede Zahl die gleiche Chance...., nämlich  $\frac{1}{49}$ !“

Anton ist damit nicht einverstanden: „Aber wenn alle Zahlen die gleiche Chance haben, muss sich die Ziehungshäufigkeit der einzelnen Zahlen doch zumindest in der Tendenz  $\frac{1}{49}$  nähern – und wie soll das geschehen, wie soll im Lauf der Zeit ein ‚Ausgleich‘ erfolgen, wenn die bislang ‚benachteiligten‘ Zahlen nicht doch eine größere Chance haben, gezogen zu werden?“

Hier mischt sich nun auch noch Jonas in die Diskussion ein: „Ob hier der Wert  $\frac{1}{49}$  überhaupt richtig ist? Vielleicht muss man doch bei jeder Ziehung im Auge behalten, was zuvor geschehen ist, ob also die Wahrscheinlichkeit, dass bei der nächsten Ziehung die ‚13‘ drankommt, davon beeinflusst wird, dass sie bei den vorausgegangenen 2320 Ziehungen nur 230-mal gezogen wurde?“

Welchen Tipp sollte man den Dreien für ihren Tipp geben? (s. Beispiel H 44)

13	28	34	8
230	251	256	261
45	16	24	47
265	268	272	272
14	4	23	30
273	274	274	274
11	12	15	29
279	279	279	280
20	35	37	5
284	284	284	286
17	41	39	9
288	288	289	290
33	42	26	36
295	295	297	300
16	25	31	46
265	268	272	272
4	23	30	10
273	274	274	274
12	15	29	40
279	279	279	280
35	37	5	46
284	284	286	287
41	39	9	27
288	288	289	290
42	26	36	3
295	295	297	300
28	34	8	
251	256	261	
16	24	47	7
268	272	272	273
4	23	30	10
274	274	274	277
12	15	29	40
279	279	279	280
35	37	5	46
284	284	286	287
41	39	9	27
288	288	289	290
42	26	36	3
295	295	297	300
31	48	38	49
303	303	310	323
27	31	2	6
291	291	292	293
9	27	31	4
303	303	310	323
21	48	38	49
303	303	310	323

## H 1 Zufallsexperimente

### H 1.1 Ein- und mehrstufige Zufallsexperimente; Ergebnismengen; Baumdiagramme

Die Wahrscheinlichkeitstheorie untersucht Vorgänge, deren Ausgang jeweils vom Zufall abhängt und demzufolge nicht mit Sicherheit vorhersagbar ist. Derartige Vorgänge werden (im Gegensatz zu deterministischen Vorgängen) als *Zufallsexperimente*, *zufällige Vorgänge* oder auch als *Vorgänge mit zufälligem Ergebnis* bezeichnet.

H 1

Definition H 1:

Einen Vorgang nennt man **Zufallsexperiment**, wenn dabei mindestens zwei Ergebnisse möglich sind und es vor Ablauf des Vorganges nicht vorhersagbar ist, welches der möglichen Ergebnisse eintreten wird. Außerdem kann ein Zufallsexperiment (wenigstens prinzipiell) beliebig oft und in gleicher Weise (d.h. unter einem bestimmten Komplex von Bedingungen) ablaufen.

Bei solchen vom Zufall abhängigen Vorgängen können die möglichen Ereignisse also nur mehr oder weniger „wahrscheinlich“, mehr oder weniger wahrscheinlich sein.

Damit die Wahrscheinlichkeitstheorie als *mathematische* Theorie des sinnvollen Vermutens Zufallsexperimente untersuchen kann, muss man den Vorgang mathematisch beschreiben. Zu diesem Zweck werden in einem mathematischen **Modell** die für die Untersuchung wesentlichen Eigenschaften des realen Vorganges erfasst und die unwesentlichen unberücksichtigt gelassen. Bei dieser Modellbildung fasst man in einem ersten Schritt alle möglichen interessierenden Ergebnisse des Zufallsexperiments zu einer Menge (im Sinne der Mathematik) zusammen. Es ist üblich, diese Menge als Ergebnismenge zu bezeichnen und mit  $\Omega$  zu symbolisieren.

H 2

Definition H 2:

Eine Menge  $\Omega$  heißt **Ergebnismenge** eines Zufallsexperiments, wenn jedem für die Beobachtung möglichen Ergebnis genau ein Element aus  $\Omega$  zugeordnet wird.

H 1

Beispiel H 1:

Zufallsexperiment: Einmaliges Werfen eines Tetraeders mit den Augenzahlen 1 bis 4

Beobachtungsziel

*geeignete Ergebnismengen:*

- Welche Augenzahl wird geworfen?  $\Omega_1 = \{1; 2; 3; 4\}$ ;  $\Omega_2 = \{1; 2; 3; 4; 5\}$   
( $\Omega$  kann ein Ergebnis enthalten, das niemals eintritt.)
- Wird eine 4 geworfen?  $\Omega_3 = \{4; \text{keine } 4\}$
- Wird eine 1 oder eine 3 geworfen?  $\Omega_4 = \{1; 3; \text{weder eine 1 noch eine 3}\}$

*ungeeignete Ergebnismengen:*

$\Omega_5 = \{1; \text{gerade Zahl}\}$   
(Ergebnis „3“ wäre nicht erfasst.)  
 $\Omega_6 = \{1; 3; \text{keine } 3\}$   
(„1“ wäre durch „1“ und „keine 3“ erfasst.)

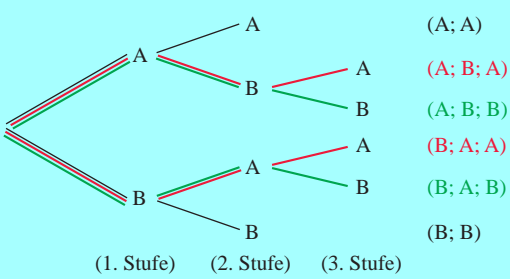
Bei der Angabe der Ergebnismenge  $\Omega$  gehen wir von idealisierten (modellhaften) Bedingungen aus. In obigem Beispiel darf so das Tetraeder nicht „auf Kippe“ stehen bleiben; es darf nicht entzwei oder verloren gehen usw. Auch darf kein Ergebnis eintreten können, das nicht in  $\Omega$  erfasst ist oder das durch *zwei oder mehrere* Elemente von  $\Omega$  beschrieben wird. Es ist allerdings gestattet, dass  $\Omega$  Ergebnisse enthält, die niemals eintreten können. Man wird bei der Konstruktion einer geeigneten Ergebnismenge  $\Omega$  ferner darauf achten, dass sie einerseits möglichst klein ist, also keine unnötigen Elemente enthält, und dass sie andererseits hinreichend groß, hinreichend fein ist, d.h. alle dem Beobachtungsziel entsprechenden Ergebnisse umfasst.

**Definition H 3:**  
Besteht ein zufälliger Vorgang aus mehreren, nacheinander ablaufenden Teilvorgängen, so spricht man von einem **mehrstufigen Zufallsexperiment**, bei  $k$  Teilvorgängen ( $k \in \mathbb{N} \setminus \{0\}$ ) von einem **k-stufigen Zufallsexperiment**.

H 3

Die Ergebnisse eines  $k$ -stufigen Zufallsexperiments lassen sich in einem **Baumdiagramm der Ergebnisse** erfassen, woraus man ihre Darstellung als  $k$ -Tupel leicht ablesen kann.

**Beispiel H 2:**  
Axel und Bernd spielen gegeneinander. Sieger ist derjenige, der zuerst zwei Spiele gewonnen hat. Von jedem Spiel wird der Gewinner notiert; ein Remis gibt es nicht.



H 2

$$\Omega = \{(A; A), (A; B; A), (A; B; B), (B; A; A), (B; A; B), (B; B)\}$$

Interpretation	Baumdiagramm der Ergebnisse	k-Tupel
Gewinner des ersten Spiels	erste Stufe	erste Koordinate
Gewinner des zweiten Spiels	zweite Stufe	zweite Koordinate
Gewinner des dritten Spiels, wenn dies ausgetragen wird	dritte Stufe (wenn vorhanden)	dritte Koordinate (wenn vorhanden)

Für  $k = 1$  entarten  $k$ -Tupel zu „Eintupeln“, die i. Allg. ohne Klammern geschrieben werden.

**Beispiel H 3:**  
Zufallsexperiment: Ein normaler Spielwürfel wird so lange geworfen, bis zum ersten Mal die Augenzahl 3 erscheint, höchstens aber viermal.  
Beobachtungsziel: Beobachtet wird bei jedem Wurf, ob die Augenzahl 3 fällt oder eine von 3 verschiedene Augenzahl  $\bar{3}$ .  
Ergebnismenge:  $\Omega = \{3, (\bar{3}; 3), (\bar{3}; \bar{3}; 3), (\bar{3}; \bar{3}; \bar{3}; 3), (\bar{3}; \bar{3}; \bar{3}; \bar{3})\}$

H 3

Würde man im Beispiel H 3 den Spielwürfel so lange werfen, bis zum ersten Mal die Augenzahl 3 erscheint, so ergäbe sich bei sonst gleichem Beobachtungsziel die Ergebnismenge  $\Omega = \{3, (\bar{3}; 3); (\bar{3}; \bar{3}; 3), \dots\}$ . Dieses  $\Omega$  ist keine endliche Ergebnismenge, sondern besitzt (abzähl-

bar) unendlich viele Elemente. Eine Ergebnismenge mit endlich vielen oder höchstens abzählbar unendlich vielen Ergebnissen heißt **diskrete Ergebnismenge**. Bei den weiteren Betrachtungen beschränken wir uns zunächst auf *endliche Ergebnismengen*.

Umfasst die Ergebnismenge sehr viele Elemente, so verwendet man häufig nicht die aufzählende Schreibweise.

H 4

Beispiel H 4:

Zufallsexperiment: Ein normaler Spielwürfel wird fünfmal hintereinander geworfen.

Beobachtungsziel: Beobachtet wird bei jedem Wurf die gefallene Augenzahl.

Ergebnismenge:  $\Omega = \{(w_1; w_2; w_3; w_4; w_5) \mid w_1, w_2, \dots, w_5 \in \{1; 2; 3; 4; 5; 6\}\}$

Das heißt: Die Ergebnismenge  $\Omega$  ist die Menge aller 5-Tupel  $(w_1; w_2; w_3; w_4; w_5)$ , wobei jede Koordinate  $w_i$  die Zahlen 1 bis 6 annehmen kann. Die Anzahl der Elemente von  $\Omega$  beträgt daher  $|\Omega| = 6^5 = 7776$ .

H 5

Beispiel H 5:

Zufallsexperiment: Aus den fünf Schülerinnen **Martha, Anita, Tina, Heike** und **Esther** sollen genau drei für die Kommission „ABI-Zeitung-1“ ausgewählt werden.

Beobachtungsziel: Beobachtet wird, welche drei Schülerinnen in die Kommission gewählt werden, und zwar unabhängig von der Reihenfolge ihrer Wahl.

Das (in Gedanken erstellte) Baumdiagramm der Ergebnisse enthält in der *ersten Stufe* fünf Verzweigungen für die Ergebnisse M, A, T, H und E, in der *zweiten Stufe* bei jedem Ergebnis der ersten Stufe jeweils vier Verzweigungen für jeweils alle *die* vier Schülerinnen, die in der ersten Stufe nicht gewählt worden sind, sowie in der *dritten Stufe* jeweils drei Verzweigungen für jeweils alle *die* drei Schülerinnen, die weder in der ersten noch in der zweiten Stufe gewählt worden sind.

Die dabei erhaltene Ergebnismenge  $\Omega_T$  umfasst also  $5 \cdot 4 \cdot 3 = 60$  *Tripel*. Dem Beobachtungsziel besser angepasst wäre es aber, wenn man z.B. die sechs *Tripel*-Ergebnisse (A; E; H), (A; H; E), (H; A; E), (H; E; A), (E; A; H), (E; H; A) aus  $\Omega_T$  aus denen man jeweils ablesen kann, in welchem Wahlgang A, E bzw. H gewählt worden ist, zu dem *einen Dreiermengen*-Ergebnis {A; E; H} zusammenfassen würde. Aus diesem ist dann nur noch ablesbar, dass A, E und H gewählt worden sind – unabhängig von der Reihenfolge, in der dies geschah. Damit ergibt sich eine zehnelementige Ergebnismenge:

$\Omega_M = \{\{A; E; H\}, \{A; E; M\}, \{A; E; T\}, \{A; H; M\}, \{A; H; T\}, \{A; M; T\}, \{E; H; M\}, \{E; H; T\}, \{E; M; T\}, \{H; M; T\}\}$

## H 1.2 Zufällige Ereignisse; Verknüpfen von Ereignissen

Beim Beobachten eines Zufallsexperiments interessiert man sich meist nicht nur dafür, welches Ergebnis eintritt, sondern auch, ob ein Ergebnis mit einer bestimmten Eigenschaft festzustellen ist. Besitzt ein Ergebnis die Eigenschaft A, so sagt man, das **Ereignis** A ist eingetreten. Für die Kennzeichnung eines Ereignisses A verwendet man im entsprechenden mathematischen Modell neben der verbalen Darstellung auch *die Menge aller der Ergebnisse aus  $\Omega$ , welche die Eigenschaft A besitzen*. Damit muss einerseits jedes Ereignis als Teilmenge von  $\Omega$  darstellbar und andererseits jede Teilmenge von  $\Omega$  als Ereignis interpretierbar sein.

## Definition H 4:

- (1) Jede Teilmenge  $A$  der endlichen Ergebnismenge  $\Omega$  heißt **Ereignis**  $A$ .
- (2) Stellt sich das Ergebnis  $e$  ein und gilt  $e \in A$ , so sagt man, das **Ereignis  $A$  tritt ein**.
- (3) Die Menge aller Teilmengen von  $\Omega$  nennt man **Ereignisraum**<sup>1)</sup> und bezeichnet sie (siehe auch Satz H 1) mit  $2^\Omega$  (oder in Anlehnung an *Potenzmenge* mit  $\wp(\Omega)$ ).

H 4

## Satz H 1:

Besitzt  $\Omega$  genau  $n$  Elemente ( $|\Omega| = n, n \in \mathbb{N} \setminus \{0; 1\}$ ), so gibt es  $2^n$  verschiedene Teilmengen von  $\Omega$ , d.h.  $2^n$  unterschiedliche Ereignisse in  $2^\Omega$ . Damit gilt also:  $|2^\Omega| = 2^{|\Omega|}$ .

H 1

Der Beweis könnte mittels der Methode der vollständigen Induktion erbracht werden.

## Beispiel H 6:

Zufallsexperiment: Einmaliges Werfen eines Tetraeders mit den Augenzahlen 1; 2; 3; 4

Ergebnismenge:  $\Omega = \{1; 2; 3; 4\}$  mit  $|\Omega| = 4$

einige Ereignisse:  $A = \{\text{Augenzahl ist weder prim noch ungerade}\} = \{4\}$ <sup>2)</sup>

$B = \{\text{Augenzahl ist gerade oder prim}\} = \{2; 3; 4\}$

$C = \{\text{Augenzahl ist kein Teiler von 12}\} = \emptyset$

$D = \{\text{Augenzahl ist nicht größer als 4}\} = \Omega$

Ereignisraum:  $2^\Omega = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1; 2\}, \{1; 3\}, \{1; 4\}, \{2; 3\}, \{2; 4\}, \{3; 4\}, \{1; 2; 3\}, \{1; 2; 4\}, \{1; 3; 4\}, \{2; 3; 4\}, \Omega\}$   
 $|2^\Omega| = 2^{|\Omega|} = 2^4 = 16$

Wird eine 3 geworfen, so treten die folgenden acht Ereignisse ein, die alle dadurch gekennzeichnet sind, dass sie das Fallen der Augenzahl 3 – ggf. kombiniert mit anderen Augenzahlen – umfassen:  $\{3\}, \{1; 3\}, \{2; 3\}, \{3; 4\}, \{1; 2; 3\}, \{1; 3; 4\}, \{2; 3; 4\}, \{1; 2; 3; 4\} = \Omega$

H 6

## Definition H 5:

Die  $n$  einelementigen Teilmengen der  $n$ -elementigen Ergebnismenge heißen **Elementarereignisse** oder auch **atomare Ereignisse**.  $A$  nennt man **unmögliches Ereignis**, wenn  $A = \emptyset$  gilt, und **sicheres Ereignis**, wenn  $A = \Omega$  gilt. Es seien  $A, B \in 2^\Omega$

H 5

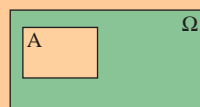
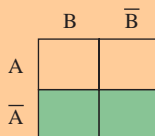
Symbol:

Sprechweise:

Mengenbild:

 $\bar{A}$ 

Das **Gegenereignis (komplementäres Ereignis)**  $\bar{A}$  (lies:  $A$  quer) tritt genau dann ein, wenn  $A$  nicht eintritt.



- 1) Der Begriff **Ereignisraum** wird statt des näherliegenden Begriffs Ergebnismenge verwendet, weil im Ereignisraum noch Operationen (z.B.  $\cap$  und  $\cup$  nach Definition H 5) zwischen seinen Ereignissen erklärt sind. In Analogie dazu sind die Begriffe **Vektorraum** und **Zahlenbereich** (mit den Operationen Addition, Multiplikation usw.) statt der Begriffe **Vektormenge** und **Zahlenmenge** gebräuchlich.
- 2) Wir *definieren* hier Ereignisse, was man häufig durch die Bezeichnung „ $A := \{\text{Augenzahl nicht 3}\}$ “ (also durch die Verwendung von „ $:=$ “ im Sinne von „*sei definiert als*“) kennzeichnet. In diesem Buch wird auf eine solche Schreibweise verzichtet.

Symbol:	Sprechweise:	Mengenbild:
$B \subseteq A$	Das Ereignis B <b>zieht</b> das Ereignis A <b>nach sich</b> . Das heißt: Immer wenn B eintritt, tritt auch A ein.	
$A \cap B$	Das Ereignis <b>A und B</b> (A geschnitten B) tritt genau dann ein, wenn <i>sowohl A als auch B</i> eintritt.	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$A \cup B$	Das Ereignis <b>A oder B</b> (A vereinigt B) tritt genau dann ein, wenn <i>mindestens eines</i> der Ereignisse A, B eintritt.	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$A \setminus B$	Das Ereignis <b>A und nicht B</b> (A minus B) tritt genau dann ein, wenn <i>A eintritt und gleichzeitig B nicht eintritt</i> . $A \setminus B = A \cap \bar{B}$	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$\overline{A \cap B}$	<b>Höchstens eines</b> der Ereignisse A, B tritt ein, wenn <i>entweder A oder B oder keines von beiden</i> eintritt. $\overline{A \cap B} = \bar{A} \cup \bar{B}$ <sup>1)</sup>	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$\overline{A \cup B}$	Das Ereignis <b>Weder A noch B</b> tritt genau dann ein, wenn <i>keines der beiden Ereignisse A, B eintritt</i> . $\overline{A \cup B} = \bar{A} \cap \bar{B}$ <sup>1)</sup>	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$(A \cap \bar{B}) \cup (\bar{A} \cap B)$	Das Ereignis <b>Entweder A oder B</b> tritt genau dann ein, wenn <i>genau eines</i> der Ereignisse A, B eintritt.	<div> <div> <div>B</div> <div><math>\bar{B}</math></div> </div> <div> <div>A</div> <div><math>\bar{A}</math></div> </div> </div>
$A \cap B = \emptyset$	Die Ereignisse A und B sind <b>unvereinbar</b> . Das heißt: A und B können nicht gleichzeitig eintreten.	

Die in der obigen Definition jeweils links stehenden Mengenbilder werden als **Vierfeldertafeln** und die rechts stehenden als **VENN-Diagramme** bezeichnet.

<sup>1)</sup>  $\overline{A \cap B} = \bar{A} \cup \bar{B}$  und  $\overline{A \cup B} = \bar{A} \cap \bar{B}$  werden als DE MORGANSche Regel bezeichnet (Augustus DE MORGAN, britischer Mathematiker und Logiker (1806–1871))





H 6

**Definition H 6:**  
Die Zahl  $H_n(A)$ , die angibt, wie oft bei  $n$ -maligem Realisieren eines Zufallsexperiments das Ereignis  $A$  eingetreten ist, heißt **absolute Häufigkeit** von  $A$ .

Absolute Häufigkeiten erhalten erst einen hinreichenden Informationswert, wenn man sie im Vergleich zur Gesamtzahl  $n$  der Realisierungen des Zufallsexperiments betrachtet:

H 7

**Definition H 7:**  
Ist  $H_n(A)$  die absolute Häufigkeit eines Ereignisses  $A$  bei  $n$ -maligem Realisieren eines Zufallsexperiments, so heißt  $h_n(A) = \frac{H_n(A)}{n}$  die **relative Häufigkeit** des Ereignisses  $A$ .

Durch das Bezeichnen von absoluten bzw. relativen Häufigkeiten eines Ereignisses  $A$  bei  $n$  Realisierungen des Zufallsexperiments mit den Symbolen  $H_n(A)$  bzw.  $h_n(A)$  kann leicht der Eindruck entstehen, dass diese Zahlen nur von  $n$  und  $A$  abhängen. In Wirklichkeit erhält man trotz jeweils gleichen  $n$  und  $A$  bei verschiedenen Realisierungsreihen des Zufallsexperiments in der Regel verschiedene  $H_n(A)$  und demzufolge ebenfalls verschiedene  $h_n(A)$ . Diese Erfahrung lässt auch die relative Häufigkeit als ein in der Mathematik unbrauchbares Maß für die Zufälligkeit des Eintretens eines Ereignisses erscheinen.

H 8

**Beispiel H 8:**  
a) Zufallsexperiment: Einmaliges Werfen einer Münze mit den Seiten *Wappen* und *Zahl*  
Ergebnismenge:  $\Omega = \{W; Z\}$   
Urlisten zu 13 Realisierungen des Zufallsexperiments  $h_{13}(\{\text{es fällt Zahl}\})$   

WWZZZWZZZZWW	$\frac{7}{13} \approx 0,54$
ZWWZWWZZWZZZ	$\frac{8}{13} \approx 0,62$
WZZZWZZZWZWZ	$\frac{9}{13} \approx 0,69$
ZZWZWWWZWZZWW	$\frac{6}{13} \approx 0,46$

  
b) Zufallsexperiment: Einmaliges Bestimmen des Funktionswertes der *RANDOM*-Funktion  $rand(2) - 1$  mithilfe des Taschenrechners  
Ergebnismenge:  $\Omega = \{0; 1\}$   
Urlisten zu 100 Realisierungen durch Erstellen von Listen (Folgen) aus jeweils 100  $(rand(2) - 1)$ -Funktionswerten mithilfe des **Sequence**-Befehls:  $seq(rand(2) - 1, i, 1, 100)$  (Fig. H 4; „i, 1, 100“ bedeutet hier „für alle natürlichen Zahlen  $i$  von 1 bis 100“.)  
 $h_{100}(\{\text{Funktionswert 1 wird ausgewählt}\})$  erhält man, indem man die Anzahl (absolute Häufigkeit) der Einsen als die Summe aller Funktionswerte 0 und 1 (wegen  $1 + 0 = 1$ ) aus der Liste berechnet und diese durch 100 dividiert:  
 $sum(seq(rand(2) - 1, i, 1, 100)) : 100$  (Fig. H 5)

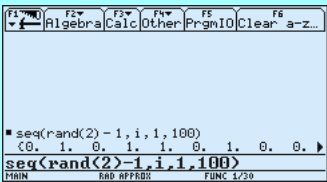


Fig. H 4

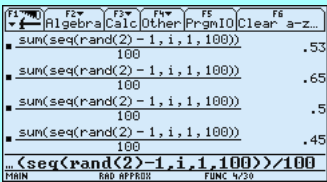


Fig. H 5



Dass die relative Häufigkeit  $h_n(A)$  nicht nur von  $n$  und  $A$  abhängt, sei am Beispiel eines einfachen stochastischen Vorgangs etwas genauer untersucht:

Ein Kronenverschluss wurde mittels Würfelbecher 1000-mal geworfen und in einem Versuchsprotokoll eine „1“ registriert, wenn der Kronenverschluss nach oben geöffnet, bzw. eine „0“, wenn er nach unten geöffnet auffiel. Für das Ereignis  $A = \{\text{der Kronenverschluss fällt nach oben geöffnet auf}\}$

berechnet man die relativen Häufigkeiten

- $h_{20}(A)$  für die ersten 20 Würfe, für die zweiten 20 Würfe, ..., für die 50-ten 20 Würfe,
- $h_{100}(A)$  für die ersten 100 Würfe, für die zweiten 100 Würfe, ..., für die zehnten 100 Würfe,
- $h_{20i}(A)$  nach insgesamt 20-mal- $i$  Würfeln mit  $i = \{1; 2; \dots; 50\}$ , also nach 20, 40, ..., 980, 1000 Würfeln.

i		$h_{20}(A)$	$h_{100}(A)$	$h_{20i}(A)$	i		$h_{20}(A)$	$h_{100}(A)$	$h_{20i}(A)$
1	11011 11111 11111 11111	0,95	0,80	0,9500	26	01111 11111 11101 01011	0,80	0,79	0,7115
2	01010 01111 10111 11011	0,70		0,8250	27	11011 11111 11111 10111	0,90		0,7185
3	11011 10101 00111 10111	0,70		0,7833	28	00010 10101 00011 10111	0,50		0,7107
4	11110 10100 11101 11110	0,70		0,7625	29	01111 01101 11111 11111	0,85		0,7155
5	11111 11111 11111 11011	0,95		0,8000	30	11111 11110 11101 11111	0,90		0,7217
6	10011 11101 11110 00110	0,65	0,69	0,7750	31	01111 10111 00111 11111	0,80	0,72	0,7242
7	10110 11100 11011 11111	0,75		0,7714	32	10111 11111 11111 10011	0,85		0,7281
8	10111 11100 01111 11101	0,75		0,7688	33	11010 11100 11001 11111	0,70		0,7273
9	01011 11010 11011 11111	0,75		0,7667	34	11001 11110 01001 00111	0,60		0,7235
10	10110 11111 11000 00001	0,55		0,7450	35	10110 11011 11101 01010	0,65		0,7214
11	01111 11011 00101 01111	0,70	0,73	0,7409	36	01111 01011 01111 11111	0,80	0,76	0,7236
12	11111 00111 10111 11011	0,80		0,7458	37	01101 01101 01011 01111	0,65		0,7216
13	01100 01011 11001 10001	0,50		0,7270	38	11011 11111 11111 00111	0,85		0,7250
14	11110 11000 11111 11111	0,80		0,7321	39	00011 01101 11010 11111	0,85		0,7231
15	11101 11100 11111 11111	0,85		0,7400	40	11111 11010 11011 11111	0,85		0,7263
16	11010 10011 01011 01111	0,65	0,67	0,7344	41	11011 11011 10111 10000	0,65	0,74	0,7244
17	11110 01101 11001 10100	0,60		0,7265	42	11011 11011 11101 11110	0,80		0,7262
18	11111 00111 11110 01011	0,75		0,7278	43	11111 11111 11111 00111	0,90		0,7302
19	00111 11001 11011 11110	0,70		0,7263	44	01000 11101 11111 01111	0,70		0,7295
20	11011 11110 11000 01101	0,65		0,7225	45	01010 10110 01111 11011	0,65		0,7278
21	10111 11110 11101 10101	0,75	0,65	0,7238	46	10011 11110 11111 01011	0,75	0,69	0,7283
22	11011 11101 01110 10110	0,70		0,7227	47	01101 00000 11001 11110	0,50		0,7234
23	11111 10101 01011 11011	0,75		0,7239	48	01101 11010 11011 00000	0,50		0,7188
24	10110 11000 00001 01111	0,50		0,7145	49	11111 11101 11001 11111	0,85		0,7214
25	00101 10000 11011 11011	0,55		0,7080	50	10111 11111 10101 11111	0,85		0,7240

Tab. H 1

Für das Erfassen und Auswerten der Ergebnisse des obigen Experiments kann der GTA wirkungsvoll eingesetzt werden. Man gibt dazu die Ergebnisse der einzelnen Wurfserien in Form von Listen ( $s_1, s_2, \dots, s_{50}$ ) ein und ermittelt daraus die interessierenden relativen Häufigkeiten.

Würde man zusätzlich die relativen Häufigkeiten  $h_{200}(A)$  errechnen, so ergäbe sich:

für die	ersten	zweiten	dritten	vierten	fünften	200 Würfe
$h_{200}(A)$	0,745	0,700	0,720	0,740	0,715	

Der Auswertungstabelle H 1 und obiger Ergänzung ist zu entnehmen, dass die relativen Häufigkeiten  $h_{20}(A)$  recht stark (nämlich zwischen 0,50 und 0,95) schwanken, die relativen Häufigkeiten  $h_{100}(A)$  hingegen nur noch zwischen 0,65 und 0,80 und die relativen Häufigkeiten  $h_{200}(A)$  sogar nur noch

zwischen 0,700 und 0,745 variieren, was die nachfolgende Graphik veranschaulicht. Die Stabdiagramme zeigen, wie sich die Verteilungen mit zunehmender Länge der Serie enger um eine (fiktive) Achse gruppieren.

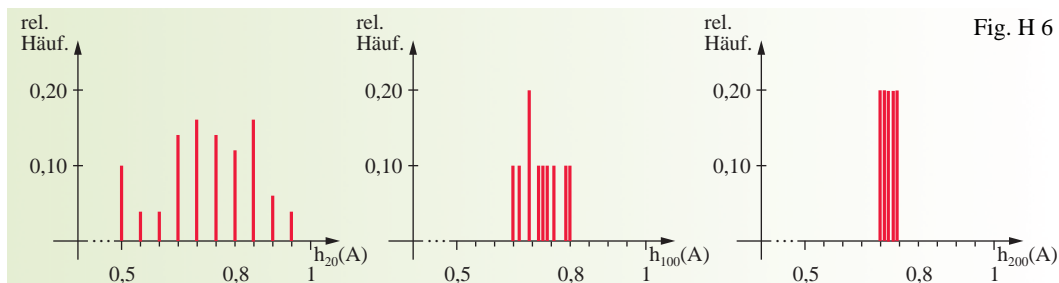


Fig. H 6

Besonders deutlich ist in der Spalte  $h_{20i}(A)$  der Tab. H 1 sowie in dem nachfolgenden Diagramm (Fig. H 7) abzulesen, dass die relativen Häufigkeiten mit zunehmender Anzahl der Realisierungen des zufälligen Vorgangs in der Tendenz immer weniger schwanken, dass sie sich stabilisieren, wenn gleich (im Unterschied zu Grenzwerten) auch immer wieder einmal etwas größere Abweichungen auftreten können.

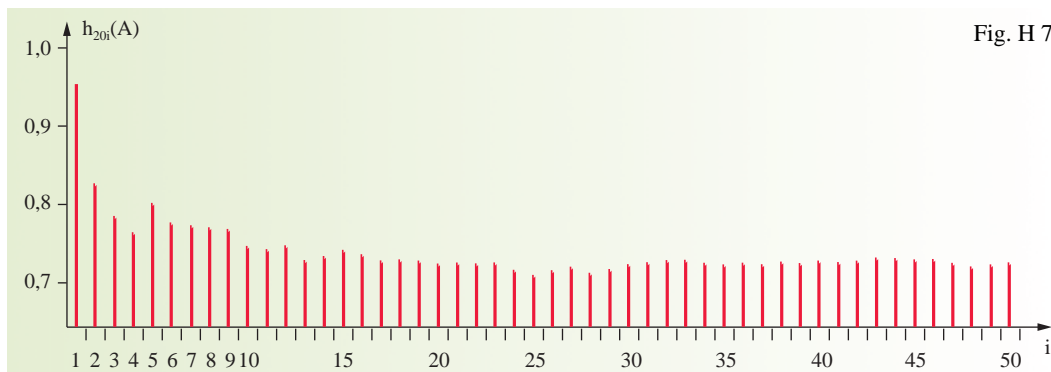


Fig. H 7

Dieses Stabilisieren von relativen Häufigkeiten und damit die *abnehmenden* Raten des Informationsgewinns bei *größer werdenden* Realisierungsanzahlen des Zufallsexperiments sind im Laufe der systematischen Untersuchung zufälliger Vorgänge immer wieder beobachtet worden. Man fasst diese Erkenntnis in folgenden Satz:

## H 2

**Satz H 2: Empirisches Gesetz der großen Zahlen**

Ist  $A$  ein Ereignis, das bei einem zufälligen Vorgang beobachtet werden kann, dann stabilisieren sich die relativen Häufigkeiten  $h_n(A)$  mit wachsender Anzahl  $n$  von Vorgangsrealisierungen jeweils gegen einen bestimmten Wert.

## H 9

Beispiel H 9:

Zufallsexperiment: Bestimmen des *RANDOM*-Funktionswertes  $\text{rand}(2) - 1$ , interpretiert als „einmaliges Werfen einer Taschenrechner-Münze“

Ergebnismenge:  $\Omega = \{0; 1\}$

Gegen welchen Wert stabilisieren sich die relativen Häufigkeiten  $h_n(A)$  des Ereignisses  $A = \{\text{der Funktionswert ist } 1\}$ ?



Es stellt sich aber nun die Frage, wie dieser stabile Wert (der zwar existiert, aber unbekannt ist) bestimmt werden kann, wenn es nicht möglich ist, bei hinreichend vielen Realisierungen des Zufallsexperiments das Eintreten bzw. Nichteintreten von A zu beobachten. Einen Ausweg fanden Mathematiker dadurch, dass sie aufhörten, ein Maß für die Zufälligkeit „*explizit*“ durch eine umgangssprachliche Bedeutungsangabe exakt definieren zu wollen. Vielmehr versuchten sie, eine Definition „*implizit*“ zu geben, indem sie – ähnlich wie bei der Charakterisierung der Schachfiguren – die Gültigkeit gewisser Regeln fordern für die Beziehungen zwischen dem Maß der Zufälligkeit und den Ereignissen, deren Zufälligkeit gemessen werden soll. Solche „*normative*“ Festlegungen, die in diesem mathematischen System nicht zu beweisen sind, werden als **Axiome** (von dem griechischen Wort  $\alpha\chi\iota\omega\mu\alpha$  – Verlangen) bezeichnet. Ein (möglichst kleines und in sich widerspruchsfreies) System von Axiomen sollte einerseits einen bestimmten Teil der Realität modellhaft widerspiegeln und andererseits jene mathematiktypische begriffliche Exaktheit aufweisen, die für logisch zwingende Schlussfolgerungen, für mathematische Beweisführungen notwendig ist.

Um einen **axiomatischen Wahrscheinlichkeitsbegriff** zu finden, um das Maß der Zufälligkeit implizit durch Axiome sinnvoll zu charakterisieren, betrachten wir – unter Berücksichtigung des empirischen Gesetzes der großen Zahlen (Satz H 2) – einige Eigenschaften der relativen Häufigkeiten  $h_n(A)$ . Diese müssten nämlich beim Übergang zu ihren stabilen Werten, d.h. beim Übergang zu den Wahrscheinlichkeiten  $P(A)$ <sup>1)</sup>, erhalten bleiben.

Für  $h_n$  gilt:

Wenn  $A, B \in 2^\Omega$ , so

- (1)  $0 \leq h_n(A)$ ,
- (2)  $h_n(A) \leq 1$ ,
- (3)  $h_n(\Omega) = 1$ ,
- (4)  $h_n(\emptyset) = 0$ ,
- (5)  $h_n(A \cup B) = h_n(A) + h_n(B)$  für  $A \cap B = \emptyset$ ,
- (6)  $h_n(A \cup B) = h_n(A) + h_n(B) - h_n(A \cap B)$ ,
- (7)  $h_n(\bar{A}) = 1 - h_n(A)$ ,
- (8)  $A \subseteq B \Rightarrow h_n(A) \leq h_n(B)$ .

Für  $P$  müsste dann gelten:

Wenn  $A, B \in 2^\Omega$ , so

- (1)  $0 \leq P(A)$ ,
- (2)  $P(A) \leq 1$ ,
- (3)  $P(\Omega) = 1$ ,
- (4)  $P(\emptyset) = 0$ ,
- (5)  $P(A \cup B) = P(A) + P(B)$  für  $A \cap B = \emptyset$ ,
- (6)  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ,
- (7)  $P(\bar{A}) = 1 - P(A)$ ,
- (8)  $A \subseteq B \Rightarrow P(A) \leq P(B)$ .

Der russische Mathematiker Andrej Nikolajewitsch KOLMOGOROW (1903–1987) fand im Jahre 1933, dass bereits drei dieser acht Regeln für ein entsprechendes Axiomensystem genügen.

Eingeschränkt auf endliche Ergebnismengen, lautet das **Axiomensystem der Wahrscheinlichkeitstheorie von KOLMOGOROW**:

## H 8

Definition H 8:

Eine Funktion  $P$ , die jeder Teilmenge  $A$  einer endlichen (Ergebnis-)Menge  $\Omega$  eine reelle Zahl  $P(A)$  zuordnet, heißt **Wahrscheinlichkeitsverteilung** (Wahrscheinlichkeitsfunktion oder auch Wahrscheinlichkeitsmaß), wenn sie folgenden drei Bedingungen genügt:

- Axiom 1* (Nichtnegativität):  $P(A) \geq 0$ ,  
*Axiom 2* (Normiertheit):  $P(\Omega) = 1$ ,  
*Axiom 3* (Additivität):  $P(A \cup B) = P(A) + P(B)$ , falls  $A \cap B = \emptyset$

<sup>1)</sup> Das Symbol  $P$  wird gewählt in Anlehnung an das lateinische Wort *probabilitas*, das *Wahrscheinlichkeit* bedeutet.

Aus den drei KOLMOGOROWSchen Axiomen lassen sich für die Wahrscheinlichkeitsverteilung  $P$  eine Reihe von Folgerungen gewinnen, die alle den Eigenschaften der relativen Häufigkeit entsprechen.

**Satz H 3: Eigenschaften einer Wahrscheinlichkeitsverteilung**

**H 3**

- |  |  |                             |
|--|--|-----------------------------|
| (1) $P(A) \leq 1$  | (2) $P(\emptyset) = 0$                   | (3) $P(\bar{A}) = 1 - P(A)$ |
| (4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ( <b>Additionssatz</b> <sup>1)</sup> ) |  |                             |
| (5) $A \subseteq B \Rightarrow P(A) \leq P(B)$                                       | (6) $P(A) = \sum_{e_i \in A} P(\{e_i\})$ |                             |

Die grundlegende Beweisidee zu den einzelnen Aussagen des Satzes H 3 liegt wegen des Axioms 3 in der Zerlegung eines Ereignisses in zwei geeignete (unvereinbare) Ereignisse. Zum Finden einer geeigneten Zerlegung können entsprechende VENN-Diagramme oder Vierfelder-Tafeln hilfreich sein. Am Beispiel des Beweises für die Aussagen (2), (4) und (5) soll dies illustriert werden.

*Beweis zu (2):*

$1 = P(\Omega)$	nach Axiom 2
$1 = P(\Omega \cup \emptyset)$ mit $\Omega \cap \emptyset = \emptyset$	eine mögliche Zerlegung von $\Omega$ in die zwei unvereinbaren Ereignisse $\Omega$ und $\emptyset$
$1 = P(\Omega) + P(\emptyset)$	nach Axiom 3
$1 = 1 + P(\emptyset)$	nach Axiom 2
$0 = P(\emptyset)$ w.z.b.w.	

*Oder:*

$P(\Omega) = P(\Omega \cup \emptyset)$ mit $\Omega \cap \emptyset = \emptyset$ ,	also
$P(\Omega) = P(\Omega) + P(\emptyset)$	nach Axiom 3 und damit
$0 = P(\emptyset)$ w.z.b.w.	

*Beweis zu (4):*

$P(A \cup B) = P(A \cup (\bar{A} \cap B))$	eine mögliche Zerlegung von $A \cup B$ in die zwei unvereinbaren Ereignisse $A$ und $\bar{A} \cap B$
$P(A \cup B) = P(A) + P(\bar{A} \cap B)$	nach Axiom 3
$P(A \cup B) = P(A) + P(\bar{A} \cap B)$	
$\quad + P(A \cap B) - P(A \cap B)$	0 addiert
$P(A \cup B) = P(A) + P((\bar{A} \cap B) \cup (A \cap B))$	
$\quad - P(A \cap B)$	nach Axiom 3
$P(A \cup B) = P(A) + P((\bar{A} \cup A) \cap B)$	nach <b>Distributivgesetzen</b> für Mengen oder der entsprechenden Vierfeldertafel
$\quad - P(A \cap B)$	
$P(A \cup B) = P(A) + P(\Omega \cap B) - P(A \cap B)$	nach Ereignis-Gegenereignis-Gesetzen für Ereignisse
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ w.z.b.w.	

*Beweis zu (5):*

$A \subseteq B \Rightarrow B = A \cup (B \cap \bar{A})$	mit $A \cap (B \cap \bar{A}) = \emptyset$
$\Rightarrow P(B) = P(A) + P(B \cap \bar{A})$	mit $P(B \cap \bar{A}) \geq 0$ nach Axiomen 3 und 1
$\Rightarrow P(B) \geq P(A)$ w.z.b.w.	

Liegen die Wahrscheinlichkeiten  $P(\{e\})$  aller  $n$  atomaren Ereignisse  $\{e\} \in 2^\Omega$  vor, so kann man insbesondere unter Verwendung von Satz H 3 (6) die Wahrscheinlichkeiten aller zugehörigen  $2^n$  Ereignisse ermitteln.

<sup>1)</sup> Die Aussage des Axioms 3 der Definition H 8 bezeichnet man als *spezieller Additionssatz*.

## H 10

Beispiel H 10:

Zufallsexperiment: Dreistufiger Materialtest auf Reißfestigkeit (stufenweise Erhöhung der Belastung bis zum Zerreißen)

Ergebnis  $i$  ( $i \in \{1; 2; 3\}$ ): Material reißt bei Belastungsstufe (Bst.)  $i$ 

Ergebnis 4: Material hält den Belastungsstufen 1 bis 3 stand.

Wahrscheinlichkeitsverteilung:

e	1	3	4
$P(\{e\})$	0,07	0,62	0,09

$$P(\{2\}) = 1 - P(\{1\}) - P(\{3\}) - P(\{4\}) = 1 - 0,07 - 0,62 - 0,09 = 0,22$$

$$P(\{\text{hält der Bst. 2 stand}\}) = P(\{3; 4\}) = 0,62 + 0,09 = 0,71$$

$$P(\{\text{ist auf Bst. 2 zu testen}\}) = P(\{2; 3; 4\}) = 0,22 + 0,62 + 0,09 = 0,93$$

oder

$$= P(\{\bar{1}\}) = 1 - 0,07 = 0,93$$

## H 1.5 Vier- und Mehrfeldertafeln; Zerlegungen der Ergebnismenge

Beim Berechnen der Wahrscheinlichkeiten von Ereignissen ist es oft zweckmäßig, sich die entsprechenden Wahrscheinlichkeiten mittels einer Vier- oder Mehrfeldertafel zu veranschaulichen. In diesem Zusammenhang geht es immer um eine *Zerlegung* der Ergebnismenge  $\Omega$  in Ereignisse, von denen bei jeder Realisierung des entsprechenden zufälligen Vorganges stets *genau eines* eintritt.

## H 9

Definition H 9:

Ereignisse  $A_1, A_2, \dots, A_n$  aus  $2^\Omega$  mit den folgenden drei Eigenschaften bilden eine **Zerlegung** der Ergebnismenge  $\Omega$ :

- (1) Jedes der Ereignisse besitzt eine positive Wahrscheinlichkeit, d.h.  $P(A_i) > 0$  für alle  $i \in \{1; 2; \dots; n\}$ .
- (2) Die Ereignisse sind paarweise unvereinbar, d.h.  $A_i \cap A_j = \emptyset$  für  $i \neq j$ .
- (3) Die Vereinigung aller Ereignisse ist das sichere Ereignis, d.h.  $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$ .

Sind  $E$  und  $F$  zwei Ereignisse aus  $2^\Omega$ , so lässt sich  $\Omega$  grob zerlegen in

$$\Omega = E \cup \bar{E}$$

bzw.

$$\Omega = F \cup \bar{F}$$

E

 $\bar{E}$ 


F

 $\bar{F}$ 


oder feiner in  $\Omega = (E \cap F) \cup (\bar{E} \cap F) \cup (E \cap \bar{F}) \cup (\bar{E} \cap \bar{F})$ :

	E	$\bar{E}$
F	$E \cap F$	$\bar{E} \cap F$
$\bar{F}$	$E \cap \bar{F}$	$\bar{E} \cap \bar{F}$

Trägt man im Innern und an den Rändern der Felder die entsprechenden Wahrscheinlichkeiten ein, so ergibt sich die Wahrscheinlichkeitsverteilung für die verfeinerte Zerlegung von  $\Omega$ .

	E	$\bar{E}$	
F	$P(E \cap F) = p_1$	$P(\bar{E} \cap F) = p_2$	$P(F) = p_1 + p_2$
$\bar{F}$	$P(E \cap \bar{F}) = p_3$	$P(\bar{E} \cap \bar{F}) = p_4$	$P(\bar{F}) = p_3 + p_4$
	$P(E) = p_1 + p_3$	$P(\bar{E}) = p_2 + p_4$	

**Beispiel H 11:**

Aufgrund von Beobachtungen weiß man, dass beim Einschalten einer Anlage das Bauteil A mit der Wahrscheinlichkeit 0,070 und das Bauteil B mit der Wahrscheinlichkeit 0,010 ausfällt. Dass A und B ausfallen, tritt mit der Wahrscheinlichkeit 0,022 ein. Mit welcher Wahrscheinlichkeit fällt wenigstens eines dieser beiden Bauteile beim Einschalten der Anlage aus?

Um diese Aufgabe zu lösen, trägt man zuerst die gegebenen Wahrscheinlichkeiten

$P(A) = 0,070$ ;  $P(B) = 0,10$  und  $P(A \cap B) = 0,022$

in eine entsprechende **Vierfeldertafel** (I) ein und vervollständigt dann diese Tafel (II):

(I)			(II)		
	B	$\bar{B}$		B	$\bar{B}$
A	0,022		0,07	0,022	$0,07 - 0,022 = 0,048$
$\bar{A}$				$0,10 - 0,022 = 0,078$	$0,93 - 0,078 = 0,852$
	0,10			0,10	$1 - 0,10 = 0,90$

Berechnen der gesuchten Wahrscheinlichkeit:  $P(A \cup B) = 0,078 + 0,022 + 0,048 = 0,148$

oder  $P(A \cup B) = 1 - 0,852 = 0,148$

oder  $P(A \cup B) = 0,07 + 0,10 - 0,022 = 0,148$

**H 11****Beispiel H 12: Vierfeldertafel mit Parametern**

Gegeben:  $P(E) = 0,3$

gesucht:  $P(E \cup F)$

$P(F) = 0,8$

$P((E \cap F) \cup (\bar{E} \cap \bar{F}))$

Da man in die Vierfeldertafel aufgrund der gegebenen Werte (durch Fettdruck markiert) nur die Wahrscheinlichkeiten an den Rändern eintragen kann, belegt man eine der Wahrscheinlichkeiten im Innern mit einem Parameter, um die weiteren Wahrscheinlichkeiten dann in Abhängigkeit von diesem Parameter zu bestimmen.

	E	$\bar{E}$	
F	$\alpha$	$0,8 - \alpha$	<b>0,8</b>
$\bar{F}$	$0,3 - \alpha$	$\alpha - 0,1$	0,2
	<b>0,3</b>	0,7	

Bedingungen an  $\alpha$ :

$0 \leq \alpha \leq 1$

$0 \leq 0,8 - \alpha \leq 1 \Rightarrow -0,2 \leq \alpha \leq 0,8$

$0 \leq 0,3 - \alpha \leq 1 \Rightarrow -0,7 \leq \alpha \leq 0,3$

$0 \leq \alpha - 0,1 \leq 1 \Rightarrow 0,1 \leq \alpha \leq 1,1$

$\Rightarrow 0,1 \leq \alpha \leq 0,3$

$P(E \cup F) = P(E) + P(F) - P(E \cap F) = 0,3 + 0,8 - \alpha = 1,1 - \alpha$

mit  $1,1 - 0,3 \leq 1,1 - \alpha \leq 1,1 - 0,1$  und damit  $0,8 \leq 1,1 - \alpha \leq 1$ .

$P((E \cap F) \cup (\bar{E} \cap \bar{F})) = P(E \cap F) + P(\bar{E} \cap \bar{F}) = \alpha + \alpha - 0,1 = 2\alpha - 0,1$

mit  $2 \cdot 1 - 0,1 \leq 2\alpha - 0,1 \leq 0,6 - 0,1$  und damit  $0,1 \leq 2\alpha - 0,1 \leq 0,5$ .

**H 12**



## H 13

Beispiel H 13: **Mehrfeldertafel** (mit zwei Zerlegungen von  $\Omega$ )

Die Ereignisse  $E_1$ ,  $E_2$  und  $E_3$  mögen eine Zerlegung der Ergebnismenge  $\Omega$  bilden. Mittels der vervollständigten Sechsfeldertafel sind die Wahrscheinlichkeiten  $P(\{E_1 \text{ oder } E_2 \text{ tritt ein}\})$ ,  $P(\{E_2 \text{ und } E_3 \text{ treten ein}\})$  und  $P(\{\text{es tritt } E_3 \text{ oder nicht } A \text{ ein}\})$  zu berechnen.

	$E_1$	$E_2$	$E_3$			
$A$	0,10	$p_1$	0,30	0,80	$p_1 = 0,80 - 0,10 - 0,30 = 0,40$	$p_4 = 1 - 0,80 = 0,20$
					$p_3 = 0,50 - p_1 = 0,10$	$p_2 = p_4 - 0,01 - p_3 = 0,09$
$\bar{A}$	$p_2$	$p_3$	0,01	$p_4$	$p_5 = 0,10 + p_2 = 0,19$	$p_6 = 1 - p_5 - 0,50 = 0,31$
	$p_5$	0,50	$p_6$			

$$P(\{E_1 \text{ oder } E_2 \text{ tritt ein}\}) = P(E_1 \cup E_2) = p_5 + 0,50 = 0,69$$

$$P(\{E_2 \text{ und } E_3 \text{ treten ein}\}) = P(E_2 \cap E_3) = P(\emptyset) = 0$$

$$P(E_3 \text{ oder nicht } A \text{ tritt ein}) = P(E_3 \cup \bar{A}) = 0,30 + 0,01 + p_3 + p_2 = 0,30 + p_4 = 0,50$$

Das im Beispiel H 13 durch die Mehrfeldertafel gegebene Modell eines Zufallsexperiments ließe sich beispielsweise in folgender Weise interpretieren: In einer Schule ist es Tradition, dass am Ende der 10. Klasse jeder Schüler (genau) eine vierstündige Klausur zu schreiben hat, wobei er einmal zwischen den Fächern Physik, Mathematik und Biologie und zudem jeweils zwischen zwei Leistungsanforderungsstufen wählen kann. Erfahrungsgemäß entscheiden sich 10 % für Physik auf der höheren Anforderungsstufe 1 und 30 % für Biologie auf der Anforderungsstufe 1. Nur 1 % wollen in Biologie in der niederen Anforderungsstufe 2 schreiben. 50 % der Schüler wählen Mathematik. 80 % der Schüler entscheiden sich für eine Klausur auf der Anforderungsstufe 1.

## H 14

Beispiel H 14: **Mehrfeldertafel** (mit drei Zerlegungen von  $\Omega$ )

Schüler eines Gymnasiums in Wahrstadt müssen mindestens einen Leistungskurs aus den Fächern Biologie (B), Chemie (C), Physik (Ph) besuchen. Erfahrungsgemäß wählen 70 % der Schüler Biologie und 60 % Chemie. Die Wahrscheinlichkeit, dass ein Schüler nur Physik auswählt, beträgt 0,10, dass er nur Biologie wählt, 0,25. 15 % der Schüler entscheiden sich nur für Physik und Chemie, während sich 30 % der Schüler für alle drei Fächer entscheiden. Es sind die folgenden Wahrscheinlichkeiten  $p_1$  bis  $p_6$  zu berechnen, wobei der betrachtete Schüler jeweils auf gut Glück ausgewählt worden sei:

$$p_1 = P(\{\text{der Schüler entscheidet sich für Chemie und Biologie}\})$$

$$p_2 = P(\{\text{der Schüler entscheidet sich für mindestens eines der Fächer Chemie und Biologie}\})$$

$$p_3 = P(\{\text{der Schüler entscheidet sich für höchstens eines der Fächer Chemie und Biologie}\})$$

$$p_4 = P(\{\text{der Schüler entscheidet sich für genau eines der Fächer Chemie und Biologie}\})$$

$$p_5 = P(\{\text{der Schüler entscheidet sich für genau eines der drei Fächer}\})$$

$$p_6 = P(\{\text{der Schüler entscheidet sich für mindestens zwei der drei Fächer}\})$$

		C	$\bar{C}$	
0,70	B	a	0,25	$\bar{Ph}$
		0,30	b	$Ph$
		0,15	0,10	$f$
d	$\bar{B}$	c	0	$\bar{Ph}$
		0,60	e	

Aus der Vierfeldertafel lassen sich folgende Gleichungen ablesen und sukzessive lösen:

$$d = 1 - 0,70 = 0,30$$

$$e = 1 - 0,60 = 0,40$$

$$c = d - 0,15 - 0,10 - 0 = 0,05$$

$$b = e - 0 - 0,10 - 0,25 = 0,05$$

$$a = 0,60 - 0,30 - 0,15 - c = 0,10 \quad f = 0,30 + b + 0,15 + 0,10 = 0,60$$

Aus der ergänzten Mehrfeldertafel erhält man für die gesuchten Wahrscheinlichkeiten:

$$p_1 = P(C \cap B) = a + 0,30 = 0,40$$

$$p_2 = 1 - P(\bar{C} \cap \bar{B}) = 1 - (0,10 + 0) = 0,90$$

$$p_3 = P(\bar{C} \cap \bar{B}) + P(\bar{C} \cap B) + P(C \cap \bar{B}) = (0,10 + 0) + (0,25 + b) + (0,15 + c) = 0,60$$

$$p_4 = P(\bar{C} \cap B) + P(C \cap \bar{B}) = 0,50$$

$$p_5 = P(C \cap \bar{B} \cap \bar{Ph}) + P(\bar{C} \cap B \cap \bar{Ph}) + P(\bar{C} \cap \bar{B} \cap Ph) = c + 0,25 + 0,10 = 0,40$$

$$p_6 = P(\bar{C} \cap B \cap Ph) + P(C \cap \bar{B} \cap Ph) + P(C \cap B \cap \bar{Ph}) + P(C \cap B \cap Ph) \\ = b + 0,15 + a + 0,30 = 0,60$$

## H 2 Gleichverteilung

### H 2.1 Der Begriff *Gleichverteilung* (LAPLACE-Experimente)

Wir betrachten das Zufallsexperiment „*einmaliges Werfen eines Würfels*“ mit der Ergebnismenge  $\Omega = \{1; 2; 3; 4; 5; 6\}$ . Die Erfahrungen beim Werfen eines normalen (nicht gezinkten) Spielwürfels sowie das Wissen über die symmetrische Form und die homogene Masseverteilung eines solchen Würfels besagen, dass jede seiner sechs Augenzahlen mit derselben Wahrscheinlichkeit gewürfelt wird. Die sechs atomaren Ereignisse  $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$  erscheinen uns als **gleichwahrscheinlich**.

Definition H 10:

Die Wahrscheinlichkeitsverteilung eines Zufallsexperiments (mit endlicher Ergebnismenge) heißt **Gleichverteilung**, wenn alle zugehörigen atomaren Ereignisse die gleiche Wahrscheinlichkeit besitzen. Diese Bedingung nennt man **LAPLACE-Annahme**.<sup>1)</sup> Kann man bei einem Zufallsexperiment von der Gültigkeit der LAPLACE -Annahme ausgehen, so spricht man von einem **LAPLACE -Experiment**.

H 10

Es sei  $\Omega = \{e_1; e_2; \dots; e_n\}$  die Ergebnismenge eines zufälligen Vorganges: Die Wahrscheinlichkeitsverteilung  $P$  heißt also dann Gleichverteilung, wenn  $P(\{e_1\}) = \dots = P(\{e_n\}) = p$  gilt. Diese Bedingung wird erfüllt, wenn die Wahrscheinlichkeit  $p$  den Wert  $\frac{1}{n}$  hat (in Analogie zum Würfel mit  $p = \frac{1}{6}$ ). Dieser Wert für  $p$  ist aber auch die einzige Möglichkeit, um der genannten Bedingung zu genügen. Deshalb stellt  $p = \frac{1}{n}$  nicht nur eine hinreichende, sondern auch eine notwendige Bedingung für das Vorliegen einer Gleichverteilung dar.

#### Satz H 4: **Wahrscheinlichkeit bei LAPLACE-Experimenten**

Für jedes LAPLACE-Experiment gilt: Besteht  $\Omega = \{e_1; e_2; \dots; e_n\}$  aus  $n$  Ergebnissen, so tritt jedes Ergebnis  $e_i$  ( $i \in \{1; 2; \dots; n\}$ ) mit der Wahrscheinlichkeit  $P(\{e_i\}) = \frac{1}{n} = \frac{1}{|\Omega|}$  ein.

H 4

<sup>1)</sup> Zufallsexperimente, bei denen eine Gleichverteilung sinnvollerweise angenommen werden kann, wurden insbesondere von dem französischen Mathematiker Marquis Pierre Simon de LAPLACE (1749–1827) untersucht. Ihm zu Ehren werden sie daher als LAPLACE -Experimente bezeichnet.

*Beweis:*

Da  $\Omega = \{e_1; e_2; \dots; e_n\}$  die Ergebnismenge eines LAPLACE-Experiments ist, so müssen nach der Definition H 10 alle atomaren Ereignisse die gleiche Wahrscheinlichkeit  $P(\{e_1\}) = \dots = P(\{e_n\}) = p$  besitzen. Folglich gilt  $P(\{e_1\}) + \dots + P(\{e_n\}) = n \cdot p$ . Die linke Seite der Gleichung ist nach Axiom 3 (Definition H 8) äquivalent zu  $P(\{e_1\} \cup \dots \cup \{e_n\})$ , da für alle  $i \in \{1; \dots; n-1\}$  jeweils die Unvereinbarkeit  $\{e_1; \dots; e_i\} \cap \{e_{i+1}\} = \emptyset$  erfüllt ist. Wegen  $\{e_1; e_2; \dots; e_n\} = \Omega$  und  $P(\Omega) = 1$  (nach Axiom 2) ergibt sich damit  $1 = n \cdot p$  bzw.  $p = \frac{1}{n}$  w.z.b.w.

Die Grundidee dieses Beweises besteht wiederum (wie bei der Herleitung weiterer Eigenschaften der Wahrscheinlichkeitsfunktion  $P$  aus den KOLMOGOROWschen Axiomen) darin, ein geeignetes Ereignis (hier  $\Omega$ ) in geeignete (unvereinbare) Ereignisse (hier in atomare Ereignisse) zu zerlegen.

Diese Modellannahme „Gleichverteilung“ bei einem Zufallsexperiment (mit endlicher Ergebnismenge) kann mit sehr unterschiedlichen sprachlichen Wendungen „signalisiert“ werden – z.B.:

- „Man kann auf jedes Ergebnis des Zufallsexperiments *mit der gleichen Chance* setzen.“
- „*Keines* der möglichen Ergebnisse des Zufallsexperiments ist hinsichtlich seines Eintretens *bevorzugt*.“
- „Es wird mit einem *idealen (symmetrischen, fairen, einwandfreien, ungezinkten, homogenen, nicht manipulierten, LAPLACE-, L-) Würfel* geworfen.“
- Vier äußerlich gleiche Kugeln werden „*auf gut Glück*“ („*blind*“, „*rein zufällig*“, „*wahllos*“) einer Urne entnommen usw.

Wenn man keinen Grund hat, das Eintreten irgendeines der Ergebnisse des Zufallsexperiments für wahrscheinlicher als das der jeweiligen anderen zu halten, dann geht man von der Gültigkeit der LAPLACE-Annahme aus. Dieses von LAPLACE angegebene *Prinzip des unzureichenden Grundes* wird vor allem durch geometrische oder physikalische Symmetrien objektiviert.

H 15

Beispiel H 15:

a) Einmaliges Werfen eines „fairen“ Würfels:

$$\Omega = \{1; 2; 3; 4; 5; 6\}; \quad |\Omega| = 6; \quad P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}$$

b) Werfen einer „idealen“ Münze:

$$\Omega = \{W; Z\}; \quad |\Omega| = 2; \quad P(\{W\}) = P(\{Z\}) = \frac{1}{2}$$

c) Inge entnimmt einer Urne mit vier gleichartigen von 1 bis 4 nummerierten Kugeln „auf gut Glück“ zwei Kugeln gleichzeitig.

$$\Omega = \{\{1; 2\}, \{1; 3\}, \{1; 4\}, \{2; 3\}, \{2; 4\}, \{3; 4\}\}; \quad |\Omega| = 6;$$

$$P(\{\{1; 2\}\}) = P(\{\{1; 3\}\}) = \dots = P(\{\{3; 4\}\}) = \frac{1}{6}$$

## H 2.2 Rechenregel für die Gleichverteilung (LAPLACE-Regel)

In Kenntnis der gleich großen Wahrscheinlichkeiten  $P(\{e_1\}) = P(\{e_2\}) = \dots = P(\{e_n\}) = \frac{1}{n}$  für jedes der atomaren Ereignisse sollen nun die Wahrscheinlichkeiten für entsprechende zusammengesetzte Ereignisse berechnet werden.

H 5

Satz H 5: **LAPLACE-Regel**

Für jedes Ereignis  $A \in 2^\Omega$  gilt die Rechenregel:

$$P(A) = \frac{|A|}{|\Omega|} \quad \text{bzw.} \quad P(A) = \frac{\text{Anzahl der für A günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}}.$$

*Beweis:*

Aus der Definition der Gleichverteilung und der Forderung des KOLMOGOROWSchen Axioms 3 (Definition H 8)  $P(A \cup B) = P(A) + P(B)$  für  $A \cap B = \emptyset$  folgt für ein beliebiges  $k$ -elementiges Ereignis  $A = \{e_1; e_2; \dots; e_k\}$  aus dem Ereignisraum  $2^\Omega$

$P(A) = P(\{e_1; e_2; \dots; e_k\}) = P(\{e_1\} \cup \{e_2\} \cup \dots \cup \{e_k\}) = P(\{e_1\}) + P(\{e_2\}) + \dots + P(\{e_k\})$ , also

$$P(A) = \frac{1}{|\Omega|} + \frac{1}{|\Omega|} + \dots + \frac{1}{|\Omega|} = \frac{k}{|\Omega|}. \text{ Danach gilt } P(A) = \frac{|A|}{|\Omega|} \text{ w.z.b.w.}$$

Den mit obiger Rechenregel aufgrund der LAPLACE-Annahme ermittelten Wert  $P(A) = \frac{|A|}{|\Omega|}$  nennt man häufig **LAPLACE-Wahrscheinlichkeit** von A oder auch **klassische Wahrscheinlichkeit** von A.<sup>1)</sup>

Beispiel H 16:

Für das einmalige Werfen eines idealen Würfels gilt  $\Omega = \{1; 2; 3; 4; 5; 6\}$ ;  $|\Omega| = 6$  sowie  $P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}$ . Daraus folgt beispielsweise:

$$P(\{\text{Augenzahl ist 2 oder 3}\}) = P(\{2; 3\}) = \frac{2}{6} = 0, \bar{3}$$

$$P(\{\text{Augenzahl ist ungerade}\}) = P(\{1; 3; 5\}) = \frac{3}{6} = 0,5$$

$$P(\{\text{Augenzahl ist eine Primzahl}\}) = P(\{2; 3; 5\}) = \frac{3}{6} = 0,5$$

$$P(\{\text{Augenzahl ist gerade und prim}\}) = P(\{2\}) = \frac{1}{6} = 0,1 \bar{6}$$

$$P(\{\text{Augenzahl ist 7}\}) = P(\emptyset) = \frac{0}{6} = 0$$

$$P(\{\text{Augenzahl ist kleiner als 10}\}) = P(\Omega) = \frac{6}{6} = 1$$

H 16

Beispiel H 17:

Gegeben sei ein vierreihiges GALTON-Brett. Dies besteht aus vier Reihen von Hindernissen, wobei sich in der  $k$ -ten Reihe genau  $k$  ( $k \in \{1; 2; 3; 4\}$ ) Hindernisse befinden (Fig. H 11). Unter der vierten Reihe sind fünf Fächer angebracht. Wir lassen nun eine Kugel geeigneter Größe von oben (den Gesetzen der Schwerkraft folgend) durch alle Reihen herunterrollen. An jedem Hindernis hat die Kugel dieselbe Chance, nach rechts (r) bzw. nach links (l) abgelenkt zu werden. Wie groß ist die Wahrscheinlichkeit dafür, dass die Kugel in das Fach II fällt?

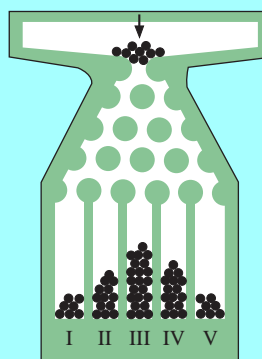


Fig. H 11

H 17

*Lösung:*

I	II	III	IV	V
		(r; r; l; l)		
		(r; l; r; l)		
	(r; r; r; l)	(l; r; r; l)	(r; l; l; l)	
	(r; r; l; r)	(r; l; l; r)	(l; r; l; l)	
	(r; l; r; r)	(l; r; l; r)	(l; l; r; l)	
(r; r; r; r)	(l; r; r; r)	(l; l; r; r)	(l; l; l; r)	(l; l; l; l)

<sup>1)</sup> 1812 hielt LAPLACE den Begriff *Wahrscheinlichkeit* in dieser ersten formalen, aber eingeschränkten Form fest.

Wie man sieht, enthält die zugehörige Ergebnismenge  $\Omega$  genau 16 Elemente ( $|\Omega| = 16$ ). Für den Weg der Kugel gibt es also insgesamt 16 Möglichkeiten, wobei jeder dieser Wege dieselbe Chance hat, von der Kugel eingeschlagen zu werden. Damit gilt die LAPLACE-Annahme. Für das Ereignis, dass die Kugel in das Fach II fällt, sind alle *die* Ergebnisse günstig, die *genau eine* Linksablenkung (l) beschreiben:

$$P(\{\text{die Kugel fällt ins Fach II}\}) = P(\{(r; r; l), (r; r; l; r), (r; l; r; r), (l; r; r; r)\}) = \frac{4}{16} = 0,25$$

Die mathematische Untersuchung von LAPLACE-Experimenten erweist sich nicht nur für Glücksspiele als nützlich, sondern sehr viele reale zufällige Vorgänge lassen sich – zumindest angenähert – als LAPLACE-Experimente auffassen: Fällt beim einmaligen Münzwurf „Zahl“ oder „Wappen“? Wird das zu erwartende Kind ein Mädchen oder ein Junge? Beide zufälligen Vorgänge können – zumindest angenähert (es gibt einen Überschuss an Jungengeburten im Verhältnis von etwa 51 : 49) – durch dasselbe stochastische Modell, nämlich eine zweielementige Ergebnismenge  $\Omega = \{1; 0\}$  und die Gleichverteilung mit  $P(\{1\}) = P(\{0\}) = \frac{1}{2}$ , beschrieben werden. Doch nicht immer ist die Annahme der Gleichverteilung so offensichtlich oder so direkt möglich.

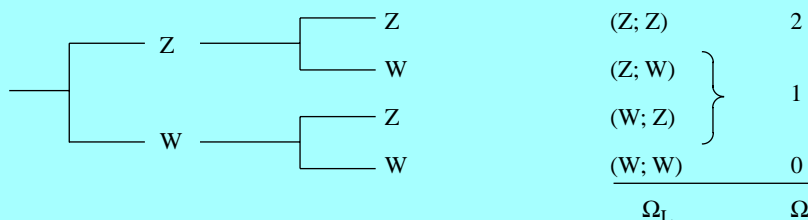
## H 18

Beispiel H 18: (D’ALEMBERTsches Paradoxon)

Zwei Münzen mit den Seiten „Zahl“ und „Wappen“ werden gleichzeitig geworfen. Es interessiert nur, welche Seiten oben liegen (und nicht, welche Seite bei jeweils welcher Münze oben liegt). Als Ergebnismenge bietet sich an

$$\begin{aligned}\Omega &= \{\text{zweimal Zahl, einmal Zahl und einmal Wappen, zweimal Wappen}\} \\ &= \{\text{zweimal Zahl, genau einmal Zahl, keinmal Zahl}\} = \{2; 1; 0\}.\end{aligned}$$

Die Anzahl aller möglichen Ergebnisse beträgt also 3. Hieraus zog der französische Mathematiker D’ALEMBERT (1717–1783) den Schluss, dass jedes der drei atomaren Ereignisse mit der Wahrscheinlichkeit  $\frac{1}{3}$  eintritt. In Wirklichkeit sind aber die drei atomaren Ereignisse nicht gleichwahrscheinlich. Dies wird deutlich, wenn man sich z.B. vorstellt, dass die beiden Münzen *nacheinander* geworfen werden. Dann werden nämlich aus dem *einen* bezüglich  $\Omega$  atomaren Ereignis  $\{1\} = \{\text{einmal Zahl und einmal Wappen}\}$  die *zwei* bezüglich  $\Omega_L$  atomaren Ereignisse  $\{(Z; W)\} = \{\text{zuerst Zahl und dann Wappen}\}$  und  $\{(W; Z)\} = \{\text{zuerst Wappen und dann Zahl}\}$ . Man erhält auf diese Weise also zwei atomare Ereignisse bezüglich einer „aufgeblasenen“ vierelementigen Ergebnismenge  $\Omega_L = \{(Z; Z), (Z; W), (W; Z), (W; W)\}$ .



Bei der Ergebnismenge  $\Omega_L$  haben die vier atomaren Ereignisse alle die gleiche Chance einzutreten, besitzen also dieselbe Wahrscheinlichkeit  $\frac{1}{|\Omega_L|} = \frac{1}{4}$ . Damit ergeben sich folgende Wahrscheinlichkeiten für die bezüglich  $\Omega$  atomaren Ereignisse:

$$P(\{2\}) = P(\{(Z; Z)\}) = \frac{1}{4}; \quad P(\{1\}) = P(\{(Z; W), (W; Z)\}) = \frac{2}{4}; \quad P(\{0\}) = P(\{(W; W)\}) = \frac{1}{4}.$$

Beispiel H 18 zeigt, wie man manchmal eine Wahrscheinlichkeitsverteilung, die keine Gleichverteilung bez. einer Ergebnismenge  $\Omega$  ist, über den Umweg einer Gleichverteilung einer anderen Ergebnismenge  $\Omega_L$  ermitteln kann. Bei diesem Umweg „verfeinert“, d.h. vergrößert man die ursprüngliche Ergebnismenge  $\Omega$  so, dass alle *neuen* atomaren Ereignisse von  $\Omega_L$  tatsächlich gleichverteilt sind. Sie treten dann mit der gleichen Wahrscheinlichkeit ein – für  $\Omega_L$  gilt nun also die LAPLACE-Annahme.

### H 2.3 Verschiedene Modelle für ein und dasselbe Zufallsexperiment

Die Wahl der Ergebnismenge und damit die Wahl des mathematischen Modells für ein Zufallsexperiment ist durch die Beschreibung des zufälligen Vorgangs häufig nicht eindeutig festgelegt. Drei solchen Wahlmöglichkeiten sowie deren Einfluss auf den Rechenweg beim Bestimmen von Wahrscheinlichkeiten auf der Basis der LAPLACE-Annahme wird im folgenden Beispiel nachgegangen.

Beispiel H 19:

Im Halbfinale einer im K.o.-System ausgetragenen Schulmeisterschaft im Schachspiel stehen zwei Schülerinnen und zwei Schüler. Die Aufteilung dieser vier Personen auf die beiden zu bildenden Besetzungen am Schachbrett soll „auf gut Glück“ erfolgen. Wie groß ist die Wahrscheinlichkeit dafür, dass das Halbfinale in gemischten Besetzungen ausgetragen wird?

*Modell 1:* Man mischt vier Zettel mit jeweils dem Namen von genau einem der vier Teilnehmer und zieht nacheinander „blind“ genau zwei Zettel, ohne dass der zuerst gezogene vor der zweiten Ziehung wieder eingemischt wird. Die Personen, deren Namen die beiden gezogenen Zettel tragen, bilden die eine Besetzung, die beiden übrigen die andere. Entsprechend der Fragestellung interessiert nur die Unterscheidung *Mädchen* (M), *Junge* (J). Sie führt bei Berücksichtigung der Reihenfolge der Ziehung zu der vierelementigen Ergebnismenge  $\Omega = \{(M; M), (M; J), (J; M), (J; J)\}$  sowie zu dem uns interessierenden Ereignis  $A = \{(M; J), (J; M)\}$ .

Die atomaren Ereignisse von  $2^\Omega$  sind aber nicht gleichwahrscheinlich. Während es für das Eintreten des Ergebnisses (M; M) die zwei Realisierungsmöglichkeiten „zuerst Mädchen  $M_1$  und dann Mädchen  $M_2$  ausgelost“ und „zuerst Mädchen  $M_2$  und dann Mädchen  $M_1$  ausgelost“, d.h. die beiden geordneten Paare  $(M_1; M_2)$  und  $(M_2; M_1)$  gibt, existieren für das Ergebnis (M; J) die vier Realisierungsmöglichkeiten  $(M_1; J_1)$ ,  $(M_1; J_2)$ ,  $(M_2; J_1)$ ,  $(M_2; J_2)$ . Analoges gilt auch für die Ergebnisse (J; J) und (J; M) aus  $\Omega$ . „Blasen“ wir also die Ergebnismenge  $\Omega$  künstlich auf, verfeinern wir die vierelementige Ergebnismenge  $\Omega$  zu der zwölfelementigen Ergebnismenge  $\Omega_L = \{(M_1; M_2), (M_2; M_1), (M_1; J_1), (M_1; J_2), (M_2; J_1), (M_2; J_2), (J_1; M_1), (J_1; M_2), (J_2; M_1), (J_2; M_2), (J_1; J_2), (J_2; J_1)\}$ ,

so ist jetzt die LAPLACE-Annahme gerechtfertigt. Für das Ereignis A gilt demzufolge:

$$\begin{aligned} P(A) &= P(\{\text{eine gemischte Besetzung wird ausgelost}\}) = P(\{(M; J); (J; M)\}) \\ &= P(\{(M_1; J_1), (M_1; J_2), (M_2; J_1), (M_2; J_2), (J_1; M_1), (J_1; M_2), (J_2; M_1), (J_2; M_2)\}) \\ &= \frac{8}{|\Omega_L|} = \frac{8}{12} = \frac{2}{3} \end{aligned}$$

*Modell 2:* Zeichnet man einen der am Auslosungsverfahren beteiligten Halbfinalisten (z.B.  $J_1$ ) willkürlich aus, dann lassen sich die möglichen Ergebnisse des Zufallsexperiments „Ziehen ‚auf gut Glück‘ genau eines aus den drei Namenszetteln von  $M_1$ ,  $M_2$  und  $J_2$ “ durch die Angabe seines möglichen Spielgegners interpretieren. Dies führt zu der Ergebnismenge  $\Omega = \{M_1; M_2; J_2\}$ ,

für welche die LAPLACE-Annahme gerechtfertigt ist. Somit ergibt sich als Wahrscheinlichkeit für das uns interessierende Ereignis, dass  $J_1$  gegen  $M_1$  oder gegen  $M_2$  zu spielen hat:

$$P(A) = P(\{M_1; M_2\}) = \frac{2}{3}$$

*Modell 3:* Registriert man bei der im *Modell 1* beschriebenen Auslosung nur, ob die zwei ausgelosten Personen entweder beide Mädchen sind oder nicht, so führt dies zu einer zweielementigen Ergebnismenge  $\Omega = \{m; nm\}$  (mit  $m$ : *nur Mädchen ausgelost* und  $nm$ : *nicht nur Mädchen ausgelost*). Diese Ergebnismenge entspricht aber nicht der gestellten Frage, da das zu betrachtende Ereignis  $A = \{\text{eine gemischte Besetzung wird ausgelost}\}$  nicht als Teilmenge von  $\Omega$  darstellbar ist.

Beim Vergleich dieser drei Modelle erweist sich einerseits eine dem Zufallsexperiment und der gestellten Frage gut angepasste, nahe liegende Ergebnismenge (z.B.  $\Omega$  vom *Modell 1*) nicht unbedingt auch bezüglich des rechnerischen Aufwands als so gut geeignet.

Andererseits muss eine „grobe“ Ergebnismenge ( $\Omega$  vom *Modell 2*) mit minimalem Rechenaufwand zur Bestimmung der gesuchten Wahrscheinlichkeit nicht auch so nahe liegend, so leicht zu finden sein, weil erst eine gewisse Umstrukturierung des eigentlichen Ablaufs des Zufallsexperiments erforderlich ist.

Aus diesen Erfahrungen ergeben sich drei **Tipps für die Wahl einer geeigneten Ergebnismenge**:

- Die Ergebnismenge  $\Omega$  muss sowohl dem Zufallsexperiment als auch dem diesbezüglich betrachteten Ereignis  $A$  entsprechen.  $A$  muss also stets eine Teilmenge des gewählten  $\Omega$  sein.
- Man sollte versuchen, die Ergebnismenge  $\Omega$  so zu wählen, dass alle atomaren Ereignisse gleichwahrscheinlich sind, dass also die LAPLACE -Annahme gerechtfertigt ist und damit die LAPLACE-Regel  $P(A) = \frac{|A|}{|\Omega|}$  angewandt werden kann.
- Unter Wahrung der LAPLACE -Annahme sollte die Ergebnismenge möglichst „grob“ sein, d.h. möglichst wenige Elemente enthalten.

Das nicht selten mühsame Suchen nach einer geeigneten Ergebnismenge  $\Omega$  unter Wahrung der LAPLACE-Annahme kann man unter Zuhilfenahme eines Baumdiagramms zu umgehen versuchen.

## H 2.4 Baumdiagramme; Pfadregeln

Kann man bei einem Zufallsexperiment davon ausgehen, dass alle seine atomaren Ereignisse gleichverteilt sind, so lassen sich die gesuchten Wahrscheinlichkeiten  $P(A) = \frac{|A|}{|\Omega|}$  durch Abzählen der „für  $A$  günstigen“ und der „möglichen“ Ergebnisse bestimmen. Dieses Abzählen war bei den bisher betrachteten Beispielen sehr einfach. Damit man sich auch bei etwas komplizierteren, mehrstufigen LAPLACE -Experimenten nicht verzählt, d.h. einerseits *wirklich alle* Möglichkeiten zählt und andererseits auch *keine* Möglichkeit *mehrfach* zählt, lässt sich wiederum ein **Baumdiagramm** nutzen.

Um daraus aber nicht nur wie im *Baumdiagramm der Ergebnisse* alle zugehörigen Ergebnisse als  $n$ -Tupel ablesen zu können, sondern auch die entsprechende Wahrscheinlichkeitsverteilung, schreibt man an jede Verzweigung, an jeden Ast die Wahrscheinlichkeit, mit der das entsprechende Ergebnis bzw. Ereignis dieser Stufe eintritt.

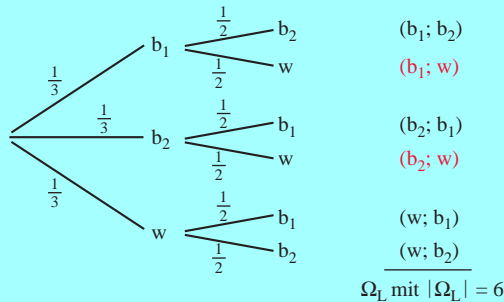


Beispiel H 20:

Zufallsexperiment: Aus einer Urne mit genau drei Kugeln (zwei blauen und einer weißen) werden nacheinander, ohne Zurücklegen und „auf gut Glück“ zwei Kugeln entnommen.

Gesucht:  $P(A) = P(\{\text{die weiße Kugel wird als zweite Kugel entnommen}\})$

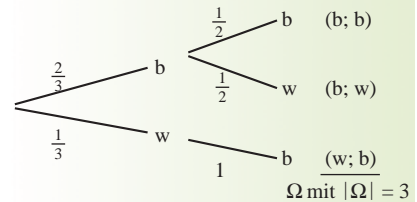
Baumdiagramm:



Da die LAPLACE-Annahme für  $\Omega_L$  gerechtfertigt ist, tritt jedes seiner Ergebnisse mit der Wahrscheinlichkeit  $\frac{1}{|\Omega_L|} = \frac{1}{6} = 0,1\bar{6}$  ein. Nach der LAPLACE-Regel erhält man somit für die gesuchte Wahrscheinlichkeit

$$P(A) = P(\{(b_1; w), (b_2; w)\}) = \frac{|A|}{|\Omega_L|} = \frac{2}{6} = \frac{1}{3}.$$

Würde man das Baumdiagramm aus obigem Beispiel stärker dem gestellten Problem anpassen, so vereinfachte es sich, wie nebenstehende Darstellung zeigt. Außerdem erhält man eine kleinere Ergebnismenge  $\Omega$ , für die aber die LAPLACE-Annahme nicht gerechtfertigt ist. Somit stellt sich die Frage, ob man für die Berechnung von  $P(A)$  nicht doch auf das im Beispiel H 20 dargestellte kompliziertere



Baumdiagramm und damit auf die größere Ergebnismenge  $\Omega_L$  zurückgreifen muss.

Um darauf eine Antwort zu finden, betrachten wir das Beispiel H 20 noch einmal etwas genauer.

Augenfällig ist dabei, dass für jedes atomare Ereignis  $\{\omega\}$  mit  $\omega \in \Omega_L$  einerseits  $P(\{\omega\}) = \frac{1}{|\Omega_L|} = \frac{1}{6}$  gilt und dass andererseits das Produkt der beiden Wahrscheinlichkeiten  $\frac{1}{3}$  und  $\frac{1}{2}$ , die längs des zu  $\omega$  führenden Weges stehen, ebenfalls  $\frac{1}{6}$  ist.

Im vereinfachten Baumdiagramm stehen längs des Pfades zu dem atomaren Ereignis  $\{(b; w)\} \subseteq \Omega$  die Wahrscheinlichkeiten  $\frac{2}{3}$  und  $\frac{1}{2}$  mit dem Produkt  $\frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$ , d.h. der Wahrscheinlichkeit von  $P(\{(b; w)\}) = P(A)$ . Multipliziert man jeweils die Wahrscheinlichkeiten längs der Pfade, die zu den atomaren Ereignissen  $\{(b; b)\}$  bzw.  $\{(w; b)\}$  führen, so erhält man analog  $\frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$  bzw.  $\frac{1}{3} \cdot 1 = \frac{1}{3}$ , d.h. dieselben Werte wie  $P(\{(b_1; b_2), (b_2; b_1)\})$  bzw.  $P(\{(w; b_1), (w; b_2)\})$ . Diese Auffälligkeiten legen die Vermutung nahe, dass bei einem mehrstufigen Zufallsexperiment die Wahrscheinlichkeit eines atomaren Ereignisses gleich ist dem Produkt der Wahrscheinlichkeiten entlang des Pfades, der dem entsprechenden Ergebnis im Baumdiagramm entspricht. Das wäre eine sehr einfache, praktische Rechenregel. Versuchen wir daher diese Vermutung zu untermauern, indem wir uns in einem weiteren Beispiel diese „Produktregel“ mithilfe der relativen Häufigkeiten plausibel machen.

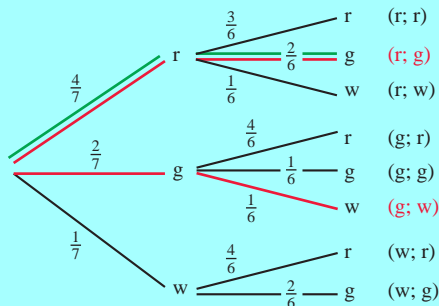
## H 21

## Beispiel H 21:

Einer Urne mit (genau) sieben Kugeln (vier roten, zwei grünen und einer weißen) werden „auf gut Glück“ nacheinander zwei Kugeln entnommen, und zwar

- a) Ziehen *ohne* Zurücklegen                      oder                      b) Ziehen *mit* Zurücklegen.

Baumdiagramm:



$$\Omega = \{(r; r), (r; g), \dots, (w; g)\} \quad |\Omega| = 8$$

Wahrscheinlichkeiten der atomaren Ereignisse

Es gilt:

$$\frac{4}{7} = P(\{r\}) \approx h_n(\{r\}) = \frac{H_n(\{r\})}{n} \text{ für eine}$$

hinreichend große Anzahl  $n$  von Realisierungen der ersten Ziehung

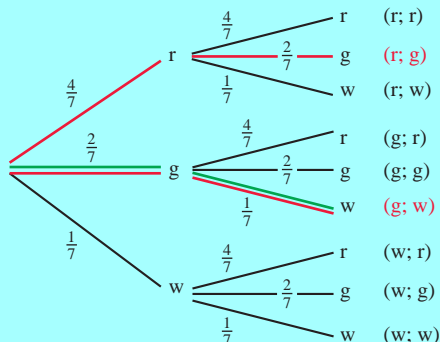
$$\Rightarrow H_n(\{r\}) \approx \frac{4}{7} \cdot n = \frac{4}{7} \text{ von } n$$

$$\Rightarrow H_n(\{(r; g)\}) \approx \frac{2}{6} \text{ von } H_n(\{r\}) = \frac{2}{6} \cdot \left(\frac{4}{7} \cdot n\right) = \left(\frac{2}{6} \cdot \frac{4}{7}\right) \cdot n$$

$$\Rightarrow h_n(\{(r; g)\}) = \frac{H_n(\{(r; g)\})}{n} \approx \frac{2}{6} \cdot \frac{4}{7} = \frac{4}{7} \cdot \frac{2}{6}$$

Es ist also sinnvoll, das Produkt  $\frac{4}{7} \cdot \frac{2}{6}$  der Wahrscheinlichkeiten längs des grün markierten Pfades, der zum Ergebnis  $(r; g)$  führt, als Wahrscheinlichkeit des atomaren Ereignisses  $\{(r; g)\}$  zu interpretieren. Führen wir diese Überlegung in analoger Weise für die weiteren sieben Ergebnisse von  $\Omega$  durch, so ergibt sich die gesuchte Wahrscheinlichkeitsverteilung.

Baumdiagramm:



$$\Omega = \{(r; r), (r; g), \dots, (w; w)\} \quad |\Omega| = 9$$

Wahrscheinlichkeiten der atomaren Ereignisse

Es gilt:

$$\frac{2}{7} = P(\{g\}) \approx h_n(\{g\}) = \frac{H_n(\{g\})}{n} \text{ für eine}$$

hinreichend große Anzahl  $n$  von Realisierungen der ersten Ziehung

$$\Rightarrow H_n(\{g\}) \approx \frac{2}{7} \cdot n = \frac{2}{7} \text{ von } n$$

$$\Rightarrow H_n(\{(g; w)\}) \approx \frac{1}{7} \text{ von } H_n(\{g\}) = \frac{1}{7} \cdot \left(\frac{2}{7} \cdot n\right) = \left(\frac{1}{7} \cdot \frac{2}{7}\right) \cdot n$$

$$\Rightarrow h_n(\{(g; w)\}) = \frac{H_n(\{(g; w)\})}{n} \approx \frac{1}{7} \cdot \frac{2}{7} = \frac{2}{7} \cdot \frac{1}{7}$$

Es ist also sinnvoll, das Produkt  $\frac{2}{7} \cdot \frac{1}{7}$  der Wahrscheinlichkeiten längs des grün markierten Pfades, der zum Ergebnis  $(g; w)$  führt, als Wahrscheinlichkeit des atomaren Ereignisses  $\{(g; w)\}$  zu interpretieren. Führen wir diese Überlegung in analoger Weise für die weiteren acht Ergebnisse von  $\Omega$  durch, so ergibt sich die gesuchte Wahrscheinlichkeitsverteilung.

	e	(r; r)	(r; g)	(r; w)	(g; r)	(g; g)	(g; w)	(w; r)	(w; g)	(w; w)
a)	$P(\{e\})$	$\frac{4}{7} \cdot \frac{3}{6}$	$\frac{4}{7} \cdot \frac{2}{6}$	$\frac{4}{7} \cdot \frac{1}{6}$	$\frac{2}{7} \cdot \frac{4}{6}$	$\frac{2}{7} \cdot \frac{1}{6}$	$\frac{2}{7} \cdot \frac{1}{6}$	$\frac{1}{7} \cdot \frac{4}{6}$	$\frac{1}{7} \cdot \frac{2}{6}$	entfällt
b)	$P(\{e\})$	$\frac{4}{7} \cdot \frac{4}{7}$	$\frac{4}{7} \cdot \frac{2}{7}$	$\frac{4}{7} \cdot \frac{1}{7}$	$\frac{2}{7} \cdot \frac{4}{7}$	$\frac{2}{7} \cdot \frac{2}{7}$	$\frac{2}{7} \cdot \frac{1}{7}$	$\frac{1}{7} \cdot \frac{4}{7}$	$\frac{1}{7} \cdot \frac{2}{7}$	$\frac{1}{7} \cdot \frac{1}{7}$

Wahrscheinlichkeiten zusammengesetzter Ereignisse:

$$P(\{(r; g); (g; w)\}) = P(\{(r; g)\}) + P(\{(g; w)\}) \\ = \frac{4}{7} \cdot \frac{2}{6} + \frac{2}{7} \cdot \frac{1}{6} = \frac{5}{21} \approx 0,24$$

Die Wahrscheinlichkeit von  $\{(r; g); (g; w)\}$  ist also gleich der Summe der „Pfadwahrscheinlichkeiten“ längs der rot markierten Pfade von  $\{(r; g)\}$  und  $\{(g; w)\}$ .

Wahrscheinlichkeiten zusammengesetzter Ereignisse:

$$P(\{(r; g); (g; w)\}) = P(\{(r; g)\}) + P(\{(g; w)\}) \\ = \frac{4}{7} \cdot \frac{2}{7} + \frac{2}{7} \cdot \frac{1}{7} = \frac{10}{49} \approx 0,20$$

Die Wahrscheinlichkeit von  $\{(r; g); (g; w)\}$  ist also gleich der Summe der „Pfadwahrscheinlichkeiten“ längs der rot markierten Pfade von  $\{(r; g)\}$  und  $\{(g; w)\}$ .

Aus den Überlegungen zu den Beispielen leiten sich die folgenden drei Regeln für ein Baumdiagramm eines mehrstufigen Zufallsexperiments ab.

Satz H 6:

**Erste Pfadregel (Produktregel):**

Die Wahrscheinlichkeit eines atomaren Ereignisses ist gleich seiner Pfadwahrscheinlichkeit (d.h. gleich dem Produkt der Wahrscheinlichkeiten entlang des Pfades, der dem zugehörigen Ergebnis entspricht).

**Zweite Pfadregel (Summenregel):**

Die Wahrscheinlichkeit eines Ereignisses ist gleich der Summe der Wahrscheinlichkeiten aller der Pfade, die zu seinen zugehörigen Ergebnissen führen.

**Verzweigungsregel:**

Die Summe aller Wahrscheinlichkeiten an den Ästen, die von ein und demselben Verzweigungspunkt ausgehen, ist stets 1.

H 6

Die exakten Beweise der beiden Pfadregeln erfolgen mit der Herleitung der Sätze H 12 und H 13.

## H 2.5 Zählprinzip bei k-Tupeln

Ein Zufallsexperiment bestehe darin, dass zuerst eine L-Münze, dann ein L-Würfel, anschließend ein L-Tetraeder und zum Schluss nochmals eine L-Münze geworfen wird. Wie groß ist die Wahrscheinlichkeit des Ereignisses

$A = \{\text{mit den beiden Münzen das Gleiche, mit dem Würfel eine gerade und mit dem Tetraeder eine Primzahl werfen}\}?$

Es handelt sich um ein vierstufiges Zufallsexperiment. Zu seiner näheren Untersuchung bietet sich ein Baumdiagramm an, wobei dessen Zeichnen hier allerdings recht aufwändig wäre. Aber das vierstufige Zufallsexperiment erfüllt auf jeder Stufe die Laplace-Annahme und die Wahrscheinlichkeit von A ist daher nach der LAPLACE-Regel  $P(A) = \frac{|A|}{|\Omega|}$  zu bestimmen. Damit reduziert sich das Problem des Berechnens von  $P(A)$  auf das Bestimmen der Anzahlen  $|\Omega|$  und  $|A|$ . Solche mehr oder weniger komplizierten Anzahlbestimmungen sind Gegenstand der **Kombinatorik**<sup>1)</sup>, der Kunst des geschickten und teilweise trickreichen Abzählens.

Stellt man das obige vierstufige Zufallsexperiment in Gedanken durch ein Baumdiagramm dar, so ist zu erkennen, dass der Baum auf der *ersten Stufe* zwei Verzweigungen, d.h. zwei mögliche Ergebnisse (nämlich *Zahl* und *Wappen*) aufweist. Auf der *zweiten Stufe* gibt es an jedem dieser beiden

<sup>1)</sup> combinare (lat.) - zusammenstellen, verbinden

Ergebnisse sechs Verzweigungen (für die Augenzahlen 1 bis 6 des Würfels), d.h., das bisher zweistufige Zufallsexperiment hat  $2 \cdot 6 = 12$  verschiedene Pfade, also 12 Ergebnisse. Auf der *dritten Stufe* bestehen nun an jedem dieser 12 Ergebnisse vier Verzweigungen (für die Augenzahlen 1 bis 4 des Tetraeders), d.h. insgesamt  $(2 \cdot 6) \cdot 4 = 48$  verschiedene Pfade bzw. 48 Ergebnisse. Auf der *vierten* und damit letzten *Stufe* gibt es schließlich an jedem der 48 Ergebnisse wieder 2 Verzweigungen (für *Zahl* und *Wappen*). Das heißt: Das vierstufige Zufallsexperiment weist insgesamt  $(2 \cdot 6 \cdot 4) \cdot 2 = 96$  verschiedene Pfade, also 96 Ergebnisse auf. Jedes dieser 96 Ergebnisse kann aufgrund der Vierstufigkeit des Zufallsexperiments als ein 4-Tupel dargestellt werden, wobei

- die *erste* Koordinate des 4-Tupels mit den möglichen Ergebnissen der *ersten Stufe*, d.h. mit den Ergebnissen W oder Z des ersten Münzwurfs,
- die *zweite* Koordinate mit den möglichen Ergebnissen der *zweiten Stufe*, d.h. mit den Augenzahlen 1, 2, 3, 4, 5, 6 des Würfelwurfs,
- die *dritte* Koordinate mit den möglichen Ergebnissen der *dritten Stufe*, d.h. mit den Augenzahlen 1, 2, 3, 4 des Tetraederwurfs,
- die *vierte* Koordinate mit den möglichen Ergebnissen der *vierten Stufe*, d.h. mit den Ergebnissen W oder Z des zweiten Münzwurfs

belegt werden kann. Als Ergebnismenge erhalten wir somit

$$\Omega = \{(m_1; w; t; m_2) \mid m_1, m_2 \in \{W; Z\}, w \in \{1; 2; 3; 4; 5; 6\}, t \in \{1; 2; 3; 4\}\} \text{ mit } |\Omega| = 2 \cdot 6 \cdot 4 \cdot 2.$$

Interessiert nun z.B., wie viele Ergebnisse zu dem eingangs genannten Ereignis  $A \in 2^\Omega$  mit

$$A = \{\text{mit den beiden Münzen das Gleiche, mit dem Würfel eine gerade und mit dem Tetraeder eine Primzahl werfen}\}$$

gehören, so zählt man analog ab:

$$|A| = |\{(m_1; w; t; m_2) \in \Omega \mid m_1 \in \{M; Z\}, w \in \{2; 4; 6\}, t \in \{2; 3\}, m_2 = m_1\}| = 2 \cdot 3 \cdot 2 \cdot 1$$

Als Wahrscheinlichkeit von A ergibt sich  $P(A) = \frac{|A|}{|\Omega|} = \frac{2 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 6 \cdot 4 \cdot 2} = \frac{1}{8} = 0,125$ .

## H 7

### Satz H 7: **Zählprinzip für k-Tupel**

Wenn ein Ereignis A aus den k-Tupeln  $(a_1; a_2; \dots; a_k)$  besteht, wobei für

- das 1. Tupelelement  $a_1$  genau  $n_1$  Auswahlmöglichkeiten,
- das 2. Tupelelement  $a_2$  genau  $n_2$  Auswahlmöglichkeiten,
- $\vdots$
- das k-te Tupelelement  $a_k$  genau  $n_k$  Auswahlmöglichkeiten

existieren, so gibt es  $n_1 \cdot n_2 \cdot \dots \cdot n_k$  verschiedene solcher k-Tupel, d.h.  $n_1 \cdot n_2 \cdot \dots \cdot n_k$  verschiedene Ergebnisse in A.

Man kann auch sagen:

Die Anzahl der

- k-Tupel, für deren Koordinaten es  $n_1, n_2, \dots$  bzw.  $n_k$  Auswahlmöglichkeiten gibt,
- Möglichkeiten, aus jeder der k Mengen mit  $n_1, n_2, \dots$  bzw.  $n_k$  Elementen genau ein Element auszuwählen,
- Ergebnisse eines k-stufigen Zufallsexperiments mit  $n_1, n_2, \dots$  bzw.  $n_k$  (unabhängig voneinander eintretenden) Ergebnissen auf den einzelnen Stufen

beträgt stets  $n_1 \cdot n_2 \cdot \dots \cdot n_k$ .

H 22

Beispiel H 22:

Wie viele Ergebnisse sind beim gleichzeitigen Werfen von fünf verschiedenfarbigen LAPLACE-Würfeln möglich?

*Lösung:*

Sinnvollerweise stellt man jedes mögliche Ergebnis dieses Zufallsexperiments als 5-Tupel  $(w_1; w_2; w_3; w_4; w_5)$  dar, bei dem jeweils die erste Koordinate  $w_1$  angibt, welche Augenzahl mit dem ersten Würfel geworfen wird, die zweite Koordinate  $w_2$ , welche Augenzahl mit dem zweiten Würfel geworfen wird usw. bis zur fünften Koordinate  $w_5$ . Als Ergebnismenge  $\Omega$  erhält man dann

$$\Omega = \{(w_1; w_2; w_3; w_4; w_5) \mid w_1, w_2, w_3, w_4, w_5 \in \{1; 2; 3; 4; 5; 6\}\}.$$

Nach dem Zählprinzip für n-Tupel erhalten wir  $|\Omega| = 6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 = 6^5$ , da alle Koordinaten  $w_1$  bis  $w_5$  jeweils die sechs Belegungsmöglichkeiten 1, 2, 3, 4, 5, 6 besitzen.

H 23

Beispiel H 23:

Aus einem Buchstabensatz wurde in einem Leserahmen

a) das Wort BLUME und b) das Wort ANANAS  
zusammengestellt. Diese Buchstaben werden von einem leseunkundigen Kind fallen gelassen.

Zu bestimmen ist die Wahrscheinlichkeit dafür, dass das Kind beim erneuten Legen eines Wortes aus denselben Buchstaben (rein zufällig) wieder das ursprüngliche Wort erhält.

*Lösung:*

zu a):

Den beschriebenen zufälligen Vorgang können wir als fünfstufiges Zufallsexperiment auffassen, wobei auf der ersten Stufe der erste Buchstabe  $b_1$  in den Leserahmen gelegt wird, auf der zweiten Stufe der zweite Buchstabe  $b_2$  usw., bis der fünfte Buchstabe  $b_5$  gelegt ist. Als Ergebnismenge  $\Omega$  bietet sich somit eine Menge von 5-Tupeln der Gestalt  $(b_1; b_2; b_3; b_4; b_5)$  an. Jede der fünf Koordinaten kann einer der Buchstaben B, L, U, M und E sein, d. h.

$b_1, b_2, b_3, b_4, b_5 \in \{B; L; U; M; E\}$ . Es muss dabei aber beachtet werden, dass keiner der Buchstaben mehrfach auftreten darf, d. h.,  $b_1, b_2, b_3, b_4$  und  $b_5$  müssen paarweise verschieden sein. Diese letztgenannte Bedingung kann man symbolisieren durch  $|\{b_1; b_2; b_3; b_4; b_5\}| = 5$ .

Zusammengefasst also:

$$\Omega = \{(b_1; b_2; b_3; b_4; b_5) \mid b_1, \dots, b_5 \in \{B; L; U; M; E\} \text{ und } |\{b_1; b_2; b_3; b_4; b_5\}| = 5\}.$$

Die Anzahl aller möglichen Ergebnisse in  $\Omega$  ergibt sich nach dem Zählprinzip für k-Tupel:

$$|\Omega| = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120.$$

Da für die Ergebnismenge  $\Omega$  im vorliegenden Fall die LAPLACE-Annahme sinnvoll ist, lässt sich die gesuchte Wahrscheinlichkeit nach der LAPLACE-Regel (Satz H 5) berechnen:

$$P(\{\text{BLUME}\}) = P(\{(B; L; U; M; E)\}) = \frac{1}{|\Omega|} = \frac{1}{120} \approx 0,0083.$$

Mit anderen Worten: Die Wahrscheinlichkeit, dass von allen jeweils aus den 5 Buchstaben B, E, L, M und U bestehenden 5-Tupeln gerade dasjenige mit der Anordnung B; L; U; M; E getroffen wird, ist rund 0,0083.

zu b):

Um für {ANANAS} eine geeignete Ergebnismenge  $\Omega$  aufzuschreiben, muss man beachten, dass hier – im Unterschied zu BLUME – nicht jeder der Buchstaben nur einmal im Wort erscheint. Mithilfe eines kleinen „Tricks“ kann man trotzdem analog zu Fall a) vorgehen: Dieser Trick besteht darin, dass man bei den im Wort ANANAS mehrfach auftretenden Buchstaben A und N durch Nummerierung künstlich eine Unterscheidbarkeit einführt. Statt des *einen* dreifach vorkommenden Buchstaben A definiert man die *drei* bezüglich der Auswahl verschiedenen Buchstaben  $A_1, A_2, A_3$  und statt des *einen* zweifach vorkommenden Buchstaben N die *zwei* bezüglich der Auswahl verschiedenen Buchstaben  $N_1, N_2$ . Damit ergibt sich:

$$\Omega = \{(b_1; b_2; b_3; b_4; b_5; b_6) \mid b_1, \dots, b_6 \in \{A_1; A_2; A_3; N_1; N_2; S\} \text{ und } |\{b_1; \dots; b_6\}| = 6\}.$$

$$|\Omega| = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$$

Aufgrund der wiederum gerechtfertigten LAPLACE-Annahme lässt sich nun die gesuchte Wahrscheinlichkeit folgendermaßen berechnen:

$$P(\{\text{ANANAS}\}) = P(\{(b_1; b_2; b_3; b_4; b_5; b_6) \in \Omega \mid b_1, b_3, b_5 \in \{A_1; A_2; A_3\} \text{ und } b_2, b_4 \in \{N_1; N_2\} \text{ und } b_6 = S\}) = \frac{3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 \cdot 1}{720} = \frac{12}{720}$$

$$P(\{\text{ANANAS}\}) \approx 0,017.$$

Im obigen Beispiel H 23 haben wir das Zählprinzip für k-Tupel in spezieller Weise angewendet: Die k-Tupel  $(a_1; a_2; \dots; a_k)$  mit  $|\{a_1; a_2; \dots; a_k\}| = k$  wurden hier aus den Elementen einer k-elementigen Menge gebildet. Das heißt: Für jedes der k-Tupel wurde die gesamte Menge „aufgebraucht“.

Die einzelnen k-Tupel unterscheiden sich also lediglich durch die Anordnung der zur Verfügung stehenden k Elemente. Dabei standen k Auswahlmöglichkeiten für die Koordinate  $a_1$ , für  $a_2$  noch  $(k - 1)$  Möglichkeiten, für  $a_3$  noch  $(k - 2)$  Möglichkeiten usw. und schließlich für  $a_k$  nur noch 1 „Auswahl“-Möglichkeit zur Verfügung. Die Gesamtzahl der Anordnungsmöglichkeiten oder **Permutationen** beträgt damit insgesamt  $k \cdot (k - 1) \cdot (k - 2) \cdot \dots \cdot 1$ . Man schreibt für ein solches Produkt  $k \cdot (k - 1) \cdot (k - 2) \cdot \dots \cdot 1$  auch **k!** (gesprochen „k Fakultät“). Fassen wir die verschiedenen Formulierungsvarianten für den obigen Sachverhalt zusammen, so lässt sich feststellen:

Die Anzahl der

- k-Tupel, deren Koordinaten verschiedene Elemente aus einer k-elementigen Menge sind;
- Möglichkeiten, die Elemente einer k-elementigen Menge anzuordnen;

oder (in der Sprechweise der Kombinatorik)

- die Anzahl der verschiedenen **Permutationen von k Elementen (ohne Wiederholung)** beträgt jeweils  **$k \cdot (k - 1) \cdot (k - 2) \cdot \dots \cdot 1 = k!$**  ( $k \in \mathbb{N} \setminus \{0\}$ ).

## H 2.6 Zählprinzip bei n-elementigen Mengen

In einer Urne mögen sich genau 16 Kugeln befinden, und zwar sechs weiße und zehn rote. Dieser Urne werden nacheinander „auf gut Glück“ genau sieben Kugeln entnommen, ohne dass man dabei eine Kugel wieder zurücklegt. Es ist die Wahrscheinlichkeit für das Ereignis  $A = \{\text{es werden genau vier weiße Kugeln entnommen}\}$  zu bestimmen.

Um  $P(A)$  als Quotienten aus der Anzahl  $|A|$  der „für A günstigen“ Ergebnisse und der Anzahl  $|\Omega|$  „aller möglichen“ Ergebnisse zu berechnen, ermitteln wir zuerst eine geeignete Ergebnismenge  $\Omega$ . Da die Kugeln nacheinander und ohne Zurücklegen entnommen werden, bietet es sich an, folgende Ergebnismenge zu verwenden:

$$\Omega_T = \{(k_1; \dots; k_7) \mid k_1, \dots, k_7 \in \{w_1; \dots; w_6; r_1; \dots; r_{10}\} \text{ und } |\{k_1; \dots; k_7\}| = 7\}$$

mit  $|\Omega_T| = 16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10$

Dabei ist  $w$  als weiße und  $r$  als rote Kugel zu interpretieren. Weil aber für das Ereignis  $A$  die *Reihenfolge* der herausgenommenen Kugeln uninteressant ist, wäre es sinnvoller, *statt der 7-Tupel* als Ergebnisse *siebenelementige Mengen* (also „ungeordnet“) zu wählen:

$$\Omega_M = \{\{k_1; \dots; k_7\} \mid k_1, \dots, k_7 \in \{w_1; \dots; w_6; r_1; \dots; r_{10}\} \text{ und } |\{k_1; \dots; k_7\}| = 7\}$$

Diese Ergebnismenge  $\Omega_M$  hat weniger Elemente als  $\Omega_T$ . Da *all den 7! Stück* von 7-Tupeln aus  $\Omega_T$ , die sich nur in der Reihenfolge ihrer Koordinatenbelegung unterscheiden, genau *ein Element* in  $\Omega_M$  entspricht, gilt

$$|\Omega_M| = \frac{|\Omega_T|}{7!} = \frac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10}{7!} = \frac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9!}{7! \cdot 9!} = \frac{16!}{7! \cdot 9!} = \binom{16}{7}$$

Hierbei bedeutet die als *Binomialkoeffizient* bezeichnete Kurzschreibweise  $\binom{n}{k}$  (gesprochen *n über k*):

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \text{ für } n, k \in \mathbb{N} \text{ und } k \leq n \text{ sowie } 0! = 1.$$

Um die gesuchte Wahrscheinlichkeit  $P(A)$  bestimmen zu können, benötigen wir noch die Anzahl  $|A|$  der für  $A$  „günstigen“ Ergebnisse, die sich analog zu  $|\Omega_M|$  berechnen lässt. Die Anzahl  $|A|$  ist gleich dem Produkt aus

- der Anzahl  $\binom{6}{4}$  der Möglichkeiten, aus 6 weißen Kugeln genau 4 auszuwählen, und
- der Anzahl  $\binom{10}{3}$  der Möglichkeiten, aus 10 roten Kugeln genau 3 auszuwählen.

$$\text{Somit ergibt sich } P(A) = \frac{\binom{6}{4} \cdot \binom{10}{3}}{\binom{16}{7}}$$

Beim Berechnen der Wahrscheinlichkeit  $P(A)$  wurde – ausgehend vom Zählprinzip für  $k$ -Tupel – ein Zählprinzip für Mengen verwendet:

#### Satz H 8: Zählprinzip für Mengen

Die Anzahl der  $k$ -elementigen Teilmengen einer  $n$ -elementigen Menge ( $k \leq n$ ) ist  $\binom{n}{k}$ .

H 8

Man kann auch sagen:

Die Anzahl

- aller Möglichkeiten,  $k$  Elemente aus einer  $n$ -elementigen Menge auszuwählen;
- der **ungeordneten Stichproben ohne Zurücklegen** vom Umfang  $k$  aus einer  $n$ -elementigen Menge;

- der Möglichkeiten, in einem  $n$ -Tupel genau  $k$  Plätze zu reservieren;
- oder (in der Sprechweise der Kombinatorik)

- die Anzahl der **Kombinationen ohne Wiederholung von  $n$  Elementen zur  $k$ -ten Klasse**

beträgt jeweils  $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$  (für  $n, k \in \mathbb{N}$ ,  $k \leq n$ ).<sup>1)</sup>

<sup>1)</sup> Wäre die Reihenfolge der  $k$  Elemente zu berücksichtigen, würde es sich also z. B. um eine *geordnete* Stichprobe (ohne Zurücklegen) handeln, so stünden aus jeder  $k$ -elementigen Teilmenge dann  $k!$  von  $k$ -Tupeln. Ihre Gesamtanzahl wäre  $\frac{n!}{k! \cdot (n-k)!} \cdot k! = \frac{n!}{(n-k)!} = n \cdot (n-1) \cdot \dots \cdot (n-k+1)$ . In der Kombinatorik spricht man dann von *Variationen ohne Wiederholung*.



## H 24

## Beispiel H 24:

Es ist die Wahrscheinlichkeit  $P(A)$  dafür zu bestimmen, dass beim Zahlenlotto „6 aus 49“ (ohne Zusatzzahl) die Zahl 40 als größte der sechs Gewinnzahlen gezogen wird.

## Lösung:

Ergebnismenge:  $\Omega = \{ \{z_1; \dots; z_6\} \mid z_1, \dots, z_6 \in \{1; \dots; 49\} \text{ und } z_1 < z_2 < \dots < z_6 \}$   
 $|\Omega| = \binom{49}{6} = 13983816$  ist die Anzahl der Möglichkeiten, aus den 49 Kugeln die sechs verschiedenen Gewinnkugeln auszuwählen.

Ereignis:  $A = \{ \{z_1; \dots; z_5; 40\} \mid z_1, \dots, z_5 \in \{1; \dots; 39\} \text{ und } z_1 < z_2 < \dots < z_5 \}$   
 $|A| = \binom{39}{5} = 575757$  ist die Anzahl der Möglichkeiten, aus den 39 Kugeln mit einer Nummer kleiner als 40 die fünf verschiedenen noch auszuwählenden Kugeln „auf gut Glück“ zu ziehen.

Wahrscheinlichk.:  $P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{39}{5}}{\binom{49}{6}} \approx 0,041$  (da LAPLACE-Annahme hier gerechtfertigt.)

Die für das Anwenden der Zählprinzipien für k-Tupel und für Mengen erforderlichen Rechnungen lassen sich durch den Einsatz eines GTA wesentlich vereinfachen, da in den entsprechenden Untermenüs (beim TI-92 erreichbar z.B. durch [2nd][5](MATH)[7](Probability)) sowohl die Anzahl  $n!$  der Permutation als auch die Anzahl  $\binom{n}{r}$  der Kombinationen ohne Wiederholung (mittels  $nCr$ ) sofort angegeben werden. In Fig. H 12 sind als Beispiele die Berechnung von  $17!$  und die Lösung der Aufgabe aus Beispiel H 24 wiedergegeben.<sup>1)</sup>

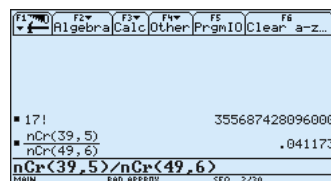


Fig. H 12

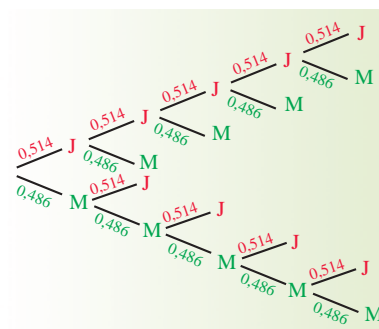
## H 2.7 Urnenmodelle; Ziehen mit und ohne Zurücklegen; hypergeometrische Verteilung

Henriette und August wünschen sich nichts sehnlicher, als eine Familie mit wenigstens einem Mädchen und wenigstens einem Jungen zu werden – freilich einerseits nur mit so vielen Kindern, wie für die Erfüllung ihres Wunsches notwendig sind, und andererseits mit höchstens fünf Kindern, auch wenn dabei ihr Hauptwunsch nach einem Pärchen nicht erfüllt sein sollte.

Um durch Umfragen auf empirischem Wege einen akzeptablen Schätzwert beispielsweise dafür zu erhalten, dass August und Henriette bei dieser Familienplanungsstrategie drei Kinder haben werden, ist im Rahmen des Stochastikunterrichts praktisch unmöglich.

Es wäre viel zu aufwändig (oder eventuell in Deutschland gar unmöglich), hinreichend viele Familien mit einer derartigen Strategie der Familienplanung ausfindig zu machen und zu befragen, dass eine fundierte Schätzung möglich würde. Daher suchen wir nach einer empirischen Ersatzvariante.

Das Umsetzen der Familienplanung von August und Henriette lässt sich aus stochastischer Sicht als das Realisieren eines fünfstufigen Zufallsexperiments auffassen, das durch das nebenstehende Baumdiagramm sinnvoll modelliert werden



<sup>1)</sup> Die Anzahl  $\frac{n!}{(n-k)!}$  der Variationen ohne Wiederholung erhält man mittels  $nPr$ .

kann. Dazu machen wir die Modellannahmen, dass keine Mehrlingsgeburten auftreten und dass – wie Geburtsstatistiken zu entnehmen ist – die Wahrscheinlichkeit einer Jungengeburt  $p_J = 0,514$  und die einer Mädchengeburt  $p_M = 0,486$  beträgt.

Dasselbe Baumdiagramm würde aber auch das folgende Zufallsexperiment charakterisieren:

In einer Urne befinden sich genau 1000 Kugeln (514 hellblaue und 486 rosafarbene)<sup>1)</sup>. Dieser Urne werden „auf gut Glück“ und mit Zurücklegen so viele Kugeln entnommen, bis erstmalig Kugeln bei der Farben oder bis fünf Kugeln entnommen worden sind.

Vergleicht man nun den zufälligen Vorgang „Kinder von Henriette und August“ entsprechend ihrer Strategie der Familienplanung und das beschriebene Zufallsexperiment mit einer Urne, so stellt man fest, dass zwischen beiden eine Strukturgleichheit besteht, d.h. beide zufälligen Vorgänge sind durch dasselbe mathematische Modell, d.h. durch

- dieselbe Ergebnismenge  $\Omega = \{0; 1\}$  und
- dieselben Wahrscheinlichkeiten ihrer atomaren Ereignisse  $P(\{1\}) = 0,514$  sowie  $P(\{0\}) = 0,486$  beschreibbar.

Eine derartige Strukturgleichheit zwischen einem praktischen zufälligen Vorgang und einem Urnenmodell ermöglicht es, reale Vorgänge mit zufälligen Ergebnis nachzuspielen, sie zu **simulieren**. Dadurch werden interessierende Wahrscheinlichkeiten von Ereignissen mithilfe eines entsprechenden **Urnenmodells** experimentell ermittelbar, d.h. sie lassen sich durch ihre relativen Häufigkeiten mit hinreichender Genauigkeit annähern. Das ist vor allem bei solchen ein- oder mehrstufigen zufälligen Vorgängen von großem Interesse, wie das Beispiel der Familienplanung von Henriette und August zeigt, die selbst gar nicht bzw. nur mit sehr hohem Kosten- oder Zeitaufwand hinreichend oft realisiert werden können.

Beim Übergang zum Urnenmodell ist zu beachten, dass ein solches nur dann nutzbar ist, wenn

- der zu simulierende zufällige Vorgang nur endlich viele mögliche Ergebnisse aufweist bzw. durch diese sinnvoll zu charakterisieren ist,
- die Wahrscheinlichkeitstabelle jedes einfachen (Teil-)Zufallsexperiments bekannt ist und deren Wahrscheinlichkeiten nur rationale Zahlen sind. Die letztgenannte Voraussetzung ist aber in der Praxis bedeutungslos, da einerseits irrationale Zahlen nur theoretischen Wert besitzen und durch rationale Zahlen stets mit hinreichender Genauigkeit zu approximieren sind und andererseits jedes Modell die Realität sowieso nur angenähert widerspiegeln kann.

Wie das folgende Beispiel zeigt, muss es zu einem Zufallsexperiment nicht nur *ein* Urnenmodell geben. Dabei soll von der Variante abgesehen werden, dass lediglich die Gesamtanzahl der Kugeln in der Urne bei Beibehaltung des Verhältnisses der Anzahl der verschiedenfarbigen Kugeln verändert wird, denn dies führt zu keinem qualitativ andersartigen Urnenmodell.

<sup>1)</sup> Zur Anzahl der hellblauen und rosafarbenen Kugeln kann man gelangen, in dem man die Wahrscheinlichkeiten der Jungen- und der Mädchengeburten als gleichnamige (gemeine) Brüche schreibt:  $p_J = \frac{514}{1000}$ ,  $p_M = \frac{486}{1000}$ . Die Zähler dieser Brüche geben an, wie viele Kugeln von den einzelnen Farben zu nehmen sind, der Hauptnenner entspricht der Gesamtzahl der Kugeln dieser Urne. Dieses Verfahren zur Bestückung der Urne mit verschiedenfarbigen Kugeln kann auch bei Sachverhalten angewandt werden, bei denen die zugehörige Wahrscheinlichkeitstabelle mehr als zwei Wahrscheinlichkeiten enthält.

## H 25

Beispiel H 25:

Zufallsexperiment: Einmaliges Werfen eines mit den Augenzahlen 1 bis 6 beschrifteten L-Würfels mit der Ergebnismenge  $\Omega = \{1; 2; 3; 4; 5; 6\}$  und den atomaren Ereignissen  $A_i = \{i\}$  für  $i \in \{1; 2; 3; 4; 5; 6\}$ .

Das Zufallsexperiment soll mithilfe eines Urnenmodells simuliert werden.

*Lösung:*

Simulationsvariante 1:

Einer Urne mit genau sechs Kugeln, die paarweise verschiedenfarbig sind, wird „auf gut Glück“ genau eine Kugel entnommen.

$$\Omega_1 = \{\text{Kugel mit der Farbe } i \text{ gezogen} \mid i \in \{1; 2; 3; 4; 5; 6\}\}, |\Omega_1| = 6$$

$$P(A_i) = P(\{\text{Kugel mit der Farbe } i \text{ gezogen}\}) = \frac{1}{|\Omega_1|} = \frac{1}{6}$$

Simulationsvariante 2:

Einer Urne mit genau fünf weißen und einer schwarzen Kugel wird „auf gut Glück“ nacheinander und ohne Zurücklegen so lange jeweils eine Kugel entnommen, bis man die schwarze Kugel herausgezogen hat.

$$\Omega_2 = \{(k_1; \dots; k_6) \mid k_1, \dots, k_6 \in \{w; s\} \text{ und genau eines der } k_1 \text{ bis } k_6 \text{ ist gleich } s\}; |\Omega_2| = 6$$

$$\begin{aligned} P(A_i) &= P(\{\text{die schwarze Kugel wird als } i\text{-te Kugel gezogen}\}) \\ &= P(\{\omega \in \Omega_2 \mid k_i = s\}) = \frac{1}{|\Omega_2|} = \frac{1}{6} \end{aligned}$$

Simulationsvariante 3:

Einer Urne mit genau drei Kugeln (einer weißen, einer schwarzen und einer roten) werden „auf gut Glück“ nacheinander und ohne Zurücklegen genau zwei Kugeln entnommen.

$$\Omega_3 = \{(k_1; k_2) \mid k_1, k_2 \in \{w; s; r\} \text{ und } k_1 \neq k_2\}; |\Omega_3| = 3 \cdot 2 = 6$$

Durch eine (eindeutige) Abbildung von  $\Omega_3$  auf  $\Omega$  wie z.B.

$(w; s) \rightarrow 1, (w; r) \rightarrow 2, (s; w) \rightarrow 3, (s; r) \rightarrow 4, (r; w) \rightarrow 5, (r; s) \rightarrow 6$  kann das Urnenmodell als Ersatzwürfel fungieren.

Simulationsvariante 4:

Einer Urne mit zwei weißen und zwei schwarzen Kugeln (und keiner weiteren außer diesen) werden „auf gut Glück“ nacheinander und ohne Zurücklegen drei Kugeln entnommen.

$$\begin{aligned} \Omega_4 &= \{(k_1; k_2; k_3; k_4) \mid k_1, k_2, k_3, k_4 \in \{w; s\} \text{ und genau zwei der } k_1, k_2, k_3, k_4 \text{ sind gleich } w\} \\ |\Omega_4| &= \binom{4}{2} \cdot 1 \cdot 1 \cdot 1 = 6 \end{aligned}$$

Durch eine (eindeutige) Abbildung von  $\Omega_4$  auf  $\Omega$  wie z.B.

$(w; w; s; s) \rightarrow 1, (w; s; w; s) \rightarrow 2, (w; s; s; w) \rightarrow 3, (s; w; s; w) \rightarrow 4, (s; s; w; w) \rightarrow 5, (s; w; w; s) \rightarrow 6$  kann das Urnenmodell als Ersatzwürfel fungieren.

Vergleicht man die verschiedenen Simulationsvarianten, so fällt auf, dass sie sich hinsichtlich der Gesamtanzahl der benötigten Kugeln, der Anzahl der benötigten verschiedenen Farben und hinsichtlich des Ziehungs Vorgangs unterscheiden. Bei der Simulationsvariante 1 ist der Ziehungs Vorgang sehr einfach, dafür kommen aber sechs Kugeln mit sechs verschiedenen Farben zum Einsatz. Bei der Simulationsvariante 4 werden nur vier Kugeln mit zwei verschiedenen Farben benötigt, dafür ist der Ziehungs Vorgang aber deutlich komplizierter. Will man ein Zufallsexperiment durch ein Urnenmodell praktisch, real simulieren, so hängt es von den jeweiligen Realisierungsmöglichkeiten ab, ob man sich für eine Variante mit möglichst wenigen Kugeln und möglichst wenigen verschiedenen Farben oder für eine Variante mit einem möglichst einfachen Ziehungs Vorgang entscheidet.

Nachdem wir ein einstufiges Zufallsexperiment analysiert haben, soll nun ein Beispiel mit einem mehrstufigen Zufallsexperiment untersucht werden. Dabei werden wir mit dem Sachverhalt konfrontiert, dass die Auswahl des Urnenmodells nicht nur vom Zufallsexperiment an sich abhängt, sondern auch – in Analogie zur Auswahl der Ergebnismenge – von der jeweiligen konkreten Fragestellung zum Zufallsexperiment.

Beispiel H 26:

Zufallsexperiment:

In einem Aufnahmetest sind zu jeder der genau vier Mathematikfragen jeweils genau fünf Antworten (zwei richtige und drei falsche) vorgegeben. Bei jeder Frage sind genau zwei dieser fünf vorgegebenen Antworten anzukreuzen, und zwar möglichst die beiden richtigen Antworten. Schüler „Enru von und zum Zufall“ verlässt sich auf den Zufall und setzt die jeweils zwei Kreuze „auf gut Glück“.

- a) Mit welcher Wahrscheinlichkeit beantwortet er Frage 3 so:

	X			X
--	---	--	--	---

Mögliches Urnenmodell

Aus einer Urne mit genau fünf Kugeln (zwei weißen für die als richtig anzukreuzenden Teilantworten und drei schwarzen für die als falsch einzustufenden Teilantworten) wird fünfmal nacheinander eine Kugel „auf gut Glück“ entnommen, ohne dass eine Kugel wieder in die Urne zurückgelegt wird;

**fünfmaliges Ziehen ohne Zurücklegen und mit Beachtung der Reihenfolge:**

$$P(\{(s; w; s; s; w)\}) = \frac{3 \cdot 2 \cdot 2 \cdot 1 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 0,1$$

- b) Mit welcher Wahrscheinlichkeit beantwortet er Frage 3 richtig?

Mögliches Urnenmodell

In einer Urne mit genau fünf Kugeln (zwei weißen für die richtigen Teilantworten und drei schwarzen für die falschen Teilantworten) werden gleichzeitig „auf gut Glück“ genau zwei Kugeln entnommen;

**zweimaliges Ziehen ohne Zurücklegen und ohne Beachtung der Reihenfolge:**

$$P(\{w\}) = \frac{\binom{2}{2} \cdot \binom{3}{0}}{\binom{5}{2}} = 0,1$$

- c) Mit welcher Wahrscheinlichkeit beantwortet er die vier Fragen so:

Frage 1	Frage 2	Frage 3	Frage 4
richtig	falsch	falsch	falsch

Mögliches Urnenmodell

Aus einer Urne mit genau zehn Kugeln (einer weißen für die richtig beantwortete Frage und neun schwarzen Kugeln für die falsch beantwortete Frage) werden nacheinander genau vier Kugeln „auf gut Glück“ entnommen, wobei die entnommene Kugel vor der nächsten Entnahme wieder in die Urne zurückgelegt wird;

**viermaliges Ziehen mit Zurücklegen und mit Beachtung der Reihenfolge:**

$$P(\{(w; s; s; s)\}) = \frac{1 \cdot 9 \cdot 9 \cdot 9}{10 \cdot 10 \cdot 10 \cdot 10} = 0,0729$$

Wie in den Beispielen H 25 und H 26 betrachtet, lassen sich für viele Zufallsexperimente adäquate Urnenmodelle konstruieren und anhand dieser die anstehenden Fragen klären.

### Hinweise für die Suche nach einem geeigneten Urnenmodell

- Bei einstufigen Zufallsexperimenten können die (endlich vielen) Ergebnisse durch die **Kugelfarben** (einschließlich Weiß und Schwarz) und die (rationalwertigen) Wahrscheinlichkeiten ihres Eintretens durch den jeweiligen **Farbenanteil** beschrieben werden.
- Bei mehrstufigen Zufallsexperimenten kann sich die Füllung der Urne von Stufe zu Stufe ändern.
- Das Ziehen der Kugeln aus der Urne kann entweder **mit Zurücklegen** oder **ohne Zurücklegen** der jeweils gezogenen Kugel vor der nächsten Kugelentnahme erfolgen.
- Das Ziehen einer Kugel aus der Urne geschieht jeweils „**auf gut Glück**“, d. h., jede Kugel besitzt die gleiche Wahrscheinlichkeit entnommen zu werden.
- Erfolgt das Ziehen der Kugeln *nacheinander* **unter Beachtung der Reihenfolge**, so nutzt man zum Berechnen der Wahrscheinlichkeiten das Zählprinzip für k-Tupel, erfolgt dagegen das Ziehen der Kugeln *mit einem Griff* **ohne Beachtung der Reihenfolge**, so nutzt man das Zählprinzip für Mengen.

Betrachten wir nochmals die im Abschnitt H 2.6 und im Beispiel H 26 Teil b untersuchten Zufallsexperimente, so erkennt man, dass ihnen allen ein Urnenmodell gleicher Struktur entspricht: Aus einer Urne, die nur Kugeln zweier Farben enthält, werden mit einem Griff (ohne Beachtung der Reihenfolge und ohne Zurücklegen) Kugeln entnommen. Die Wahrscheinlichkeit, dass sich unter den so entnommenen Kugeln eine gewisse Anzahl der einen der beiden Farben befindet, berechnet sich aufgrund des Zählprinzips für Mengen nach folgendem Satz:

H 9

#### Satz H 9: Hypergeometrische Verteilung

Werden einer Urne mit genau  $N$  Kugeln ( $M$  weiße,  $N - M$  schwarze) genau  $n$  Kugeln „auf gut Glück“ und ohne Zurücklegen entnommen, dann gilt

$$P(\{\text{genau } m \text{ weiße Kugeln entnommen}\}) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}}; \quad (N, M, n, m \in \mathbb{N}, 0 \leq m \leq n, m \leq M, n - m \leq N - M, M \leq N, n \leq N).$$

Diese Wahrscheinlichkeitsverteilung für die Anzahl der weißen gezogenen Kugeln nennt man **hypergeometrische Verteilung**.

Bisher sind wir davon ausgegangen, dass die betrachteten Urnenmodelle auch praktisch realisiert werden sollen. Man spricht in diesem Zusammenhang auch von *realen* Urnenmodellen.

Urnenmodelle können aber auch hilfreich sein, wenn für ein reales Zufallsexperiment ein geeignetes mathematisches Modell gesucht wird. Beim „Übersetzen“ des realen Zufallsexperiments in ein Urnenexperiment erkennt man vielfach leichter die meist einfachen Strukturen des realen Zufallsexperiments, das oft nur aufgrund seiner komplizierten (und nicht immer eindeutig interpretierbaren Beschreibung) in Worten recht kompliziert erscheint. Mitunter sind so auch leichter Analogien zu bereits früher modellierten Zufallsexperimenten festzustellen. Man spricht in diesem Zusammenhang von einem *gedanklichen* Urnenmodell. Wie nützlich ein solches gedankliches Urnenmodell sein kann, soll am folgenden Beispiel verdeutlicht werden:

H 27

#### Beispiel H 27:

Heinz und Ingo betreten am ersten Schultag nach den Sommerferien ihren Klassenraum, in dem die drei Sitzbänke mit je zwei Sitzplätzen der letzten Reihe noch frei sind. Beide Schüler setzen sich rein zufällig hin.

Mit welcher Wahrscheinlichkeit  $p$  setzen sie sich beide auf dieselbe Bank?

Lösung:

1	2
I	

3	4
II	

5	6
III	

**Variante 1:** Zuerst wählt Heinz eine *Bank* aus und dann entscheidet sich Ingo für eine *Bank*, und zwar jeweils „auf gut Glück“.

Ergebnismenge:  $\Omega_1 = \{(I; I), (I; II), (I; III), (II; I), (II; II), (II; III), (III; I), (III; II), (III; III)\}$   
 $|\Omega_1| = 3 \cdot 3 = 9$   
 $p = P(\{(I; I), (II; II), (III; III)\}) = \frac{3}{9} = \frac{1}{3}$

**Variante 2:** Zuerst wählt Heinz einen *Sitzplatz* aus und dann entscheidet sich Ingo für einen der noch freien *Sitzplätze* – jeweils „auf gut Glück“.

Ergebnismenge:  $\Omega_2 = \{(i; j) \mid i, j \in \{1; 2; 3; 4; 5; 6\} \text{ und } i \neq j\}$   
 $|\Omega_2| = 6 \cdot 5 = 30$   
 $p = P(\{(1; 2), (2; 1), (3; 4), (4; 3), (5; 6), (6; 5)\}) = \frac{6}{30} = \frac{1}{5}$

Beide Ergebnismengen scheinen sinnvoll gewählt zu sein, für beide ist auch die LAPLACE - Annahme gerechtfertigt, aber paradoxerweise widersprechen die errechneten Werte für die Wahrscheinlichkeit  $p$  einander. Wo steckt der Fehler? Welche Lösungsvariante ist die richtige?

Wir versuchen, diese Streitfrage durch Simulation mithilfe eines Urnenmodells zu klären.

Um ein geeignetes Urnenmodell zu konstruieren, ist zu entscheiden, wie der Vorgang „setzen sich rein zufällig“ modelliert werden soll. Dafür gibt es zwei Möglichkeiten

- zweimaliges Ziehen mit Zurücklegen aus einer Urne mit genau den drei Kugeln I, II und III oder
- zweimaliges Ziehen ohne Zurücklegen aus einer Urne mit genau den sechs Kugeln 1, 2, 3, 4, 5, 6.

Entscheidet man sich für das erste Urnenmodell, so bestätigt die Simulation den Wert  $p = \frac{1}{3}$ , beim zweiten Urnenmodell dagegen  $p = \frac{1}{5}$ . Somit wurde die Frage nach dem richtigen Wert für  $p$  noch immer nicht beantwortet. Aber beim Übergang zum Urnenmodell konnte man erkennen, dass das reale Zufallsexperiment ungenau beschrieben ist, dass die Formulierung „setzen sich rein zufällig“ auf zweierlei Weise interpretierbar ist und damit zwei verschiedene Urnenmodelle sowie zwei verschiedene Wahrscheinlichkeitswerte  $p$  zulässt.

## H 2.8 Simulation mithilfe von Zufallszahlen

Das Durchführen von umfangreichen Zufallsexperimenten „*Ziehen von Kugeln aus Urnen*“ zur Simulation praktischer zufälliger Vorgänge bleibt trotz der o.g. Vorzüge zeitaufwändig und mühsam. Daher liegt es nahe, diese Zufallsexperimente nochmals – teilweise oder vollständig – zu simulieren, und zwar durch „schnelle Rechner“, durch Computer. Für derartige Simulationen werden *Zufallszahlen* (genauer: Pseudozufallszahlen) genutzt.

**Zufallszahlen** gibt man als Folge von „auf gut Glück“ aus der Menge  $\{0; 1; 2; 3; 4; 5; 6; 7; 8; 9\}$  ausgewählten Ziffern an, z.B.<sup>1)</sup>:

<sup>1)</sup> Die hier aufgeführten Zufallszahlen sind mit einem regulären Ikosaeder „erwürfelt“, das die Augenzahlen 0 bis 20 trägt, wobei jeweils nur die Ziffer der Einerstelle registriert wurde.

45566	18378	06086	79200	57766	72496	11419	81216	71452	73138
02620	63166	64032	54967	41605	93052	54432	89511	77634	50363
26772	81040	51060	18604	05609	15110	75197	42443	57202	83173
45917	60183	08511	63926	40226	66838	60523	82145	79492	66375
42130	89137	76338	44861	44227	27900	85568	34373	61283	68232 ...

Mithilfe einer solchen Ziffernfolge lässt sich der zu Beginn des Abschnitts H 2.7 betrachtete zufällige Vorgang „August und Henriette“ folgendermaßen simulieren:

Wir fassen die Zufallszahlen als Folge von Zifferntripeln auf, indem wir in den Fünftupeln jeweils die beiden rechten Ziffern (also die 4. und 5. Ziffer) ignorieren. Diese Folge der Zifferntripel ist dann eine Folge aus den 1000 Zahlen 000 bis 999, von denen wir die 514 Zahlen von 000 bis 513 als Jungegeburt und die 486 Zahlen von 514 bis 999 als Mädchengeburt interpretieren. Beginnt man mit dem aus 45566 entstandenen Tripel 455 in der ersten Zeile und setzt dann in der normalen Lese- richtung fort, so würde man in der angegebenen Interpretation folgende Gruppierungen erhalten (die immer enden, wenn sich ein „Pärchen“ oder ein Fünftupel gleicher Elemente M bzw. J ergeben hat):

JJJM MMJ MMMJ MMMM JM JJJJM JM MJ MJ MMMM  
JM MJ JJM ...

Zufallszahlen können auf sehr unterschiedliche Art und Weise erzeugt werden, z. B.

- durch Ziehen jeweils einer Kugel „auf gut Glück“ und mit Zurücklegen aus einer Urne mit genau zehn von 0 bis 9 durchnummerierten Kugeln;
- durch Ziehen jeweils einer Kugel „auf gut Glück“ und ohne Zurücklegen aus einer Urne mit genau zehn Kugeln (neun schwarzen und einer weißen) bis zur Ziehung der weißen Kugel;
- durch Drehen eines Glücksrades bzw. -kreisels, dessen zehn gleich große (Sektoren-) Felder mit den zehn Ziffern durchnummeriert sind;
- durch Werfen eines regulären Ikosaeders (Zwanzigflachs), bei dem jeweils genau zwei der 20 kongruenten Dreiecksflächen dieselbe Ziffer tragen;
- durch Werfen einer ungezinkten Münze – wobei man jeweils durch vier Münzwürfe die Dualdarstellung der Ziffern erzeugt und die sich dabei ergebenden Zahlen 10 bis 15 ignoriert;
- durch Beobachten geeigneter physikalischer Vorgänge, wie beispielsweise des radioaktiven Zerfalls (oder früher auch des Rauschens von Elektronenröhren).

Beim Erzeugen von Zufallszahlen mit einem solchen mechanischen oder „natürlichen“ Zufallsgenerator bleibt ein – wenn auch nur noch einmaliger – hoher Aufwand, der dem Aufwand einer Simulation mittels Urnenmodell gleicht. Die Suche nach einer Vereinfachungsmöglichkeit führte zu Pseudozufallsgeneratoren: Man fand heraus, dass es deterministische Algorithmen gibt, die Ziffernfolgen – so genannte **Pseudozufallszahlen** – mit weitgehend denselben Eigenschaften wie echte Zufallszahlen liefern.<sup>1)</sup> Bei „guten“ Zufallszahlen müssen die relativen Häufigkeiten des Eintretens von k-Tupeln (k-elementige Teilfolgen der Zufallszahl für alle möglichen positiven natürlichen Zahlen k) mit einer bestimmten Eigenschaft stabil werden gegen die (mittels Abzählverfahren oder Baumdiagramm zu errechnenden) Wahrscheinlichkeiten dieser Ereignisse.

<sup>1)</sup> Die Randomfunktion (*RND* bzw. *rand* oder *rand* (n)) als Zufallsgenerator von Taschenrechnern und Computern benutzt derartige deterministische Algorithmen und liefert Pseudozufallszahlen (zwischen 0 und 1 oder 1 und der natürlichen Zahl n).

Beispiel:  $z_1$  beliebig zu wählen mit  $0 < z_1 < 1$ ;  $z_{n+1}$  = gebrochener Teil von  $((z_n + \pi)^8)$



Zu den häufig angewandten Gütekriterien zählen:

- (1) Für die relativen Häufigkeiten des Auftretens der zehn Ziffern gilt
 
$$h_n(\{0\}) \approx h_n(\{1\}) \approx \dots \approx h_n(\{9\}) \approx \frac{1}{10}$$
- (2) Für die relativen Häufigkeiten des Auftretens der 100 Ziffernpaare gilt
 
$$h_n(\{(0; 0)\}) \approx h_n(\{(0; 1)\}) \approx \dots \approx h_n(\{(9; 9)\}) \approx \frac{1}{100}$$
- (3) Die  $10^3$  Zifferntripel, die  $10^4$  Ziffernquadrupel, ... besitzen die zu (1) und (2) analogen Eigenschaften.
- (4) Beim **Maximumtest**, für den die Zufallszahl als Folge von Zifferntripeln dargestellt worden ist, gilt:
 
$$h_n(\{\text{die mittlere Ziffer des Tripels ist größer als ihre beiden Nachbarn}\}) \approx \frac{1^2 + 2^2 + \dots + 9^2}{1000} = 0,285.$$
- (5) Beim **Pokertest**, für den die Zufallszahl als Folge von Fünftupeln dargestellt worden ist, gelten die nachfolgenden Näherungen:
 
$$h_n(\{\text{alle Ziffern des Fünftupels sind verschieden}\}) \approx \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{10^5} = 0,3024;$$

$$h_n(\{\text{genau vier verschiedene Ziffern sind im Fünftupel}\}) = h_n(\{\text{genau eine Ziffer tritt genau zweimal im Fünftupel auf}\}) \approx \binom{5}{2} \cdot \frac{10 \cdot 1 \cdot 9 \cdot 8 \cdot 7}{10^5} = 0,5040;$$

$$h_n(\{\text{genau drei verschiedene Ziffern sind im Fünftupel und davon genau zwei doppelt}\}) \approx \binom{5}{2} \cdot \binom{3}{2} \cdot \frac{10 \cdot 1 \cdot 9 \cdot 1 \cdot 8}{10^5} = 0,2160; \text{ usw.}$$
- (6) Beim **Run-Test** stimmen die relativen Häufigkeiten für das Auftreten langer Sequenzen von geraden Ziffern, von Primzahlziffern o.Ä. mit den entsprechenden zu berechnenden Wahrscheinlichkeiten gut überein.

Um die Güte einer Zufallszahl zu bestimmen, genügt es nicht, nur einen der obigen Tests (oder einen ähnlichen Test) durchzuführen. Die Zufallszahl 01234567890123456789012... beispielsweise genügt zwar hervorragend dem Test (1), ist aber nach dem Test (2) als sehr schlechte Zufallszahl einzustufen.

### Beispiel H 28:

a) Zufallsexperiment: Einmaliges Werfen eines L-Tetraeders

(1) Simulation einer Realisierung des Zufallsexperiments:

#### Variante 1

**rand(4)**-Funktion mit dem Wertebereich {1; 2; 3; 4} nutzen

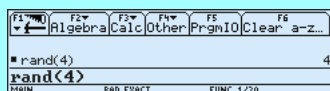


Fig. H 13

#### Variante 2

**rand()**-Funktion mit dem Wertebereich ]0; 1[; Ganzzahlbestimmung mit **int**-Funktion

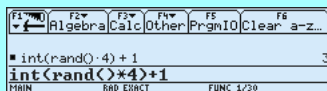


Fig. H 14

(2) Simulation von n Realisierungen des Zufallsexperiments:

#### Variante 1

n-mal einen Wert der **rand(4)**-Funktion im Home-Editor bestimmen

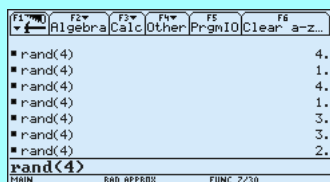


Fig. H 15

#### Variante 2

Funktion **tet** im Programm-Editor definieren

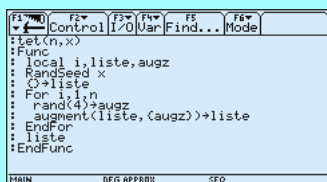


Fig. H 16



oder:

Liste im Home-Editor erstellen

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
seq(rand(4),1,1,5)→liste
└─ (3 4 3 3 2)
seq(rand(4),1,1,10)→liste
└─ (4 4 2 1 2 3 3 3 1)
seq(rand(4),1,1,100)→liste
└─ (2 2 1 1 4 2 4 3 3 1 3 1)
seq(rand(4),1,1,100)→liste
└─ (2 2 1 1 4 2 4 3 3 1 3 1)
MAIN RAD EXACT FUNC 3/20

```

Fig. H 17

und im Home-Editor als Liste ausgeben lassen

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
tet(10,1478888)
└─ (4 1 1 2 4 1 2 1 4 3)
tet(10,155555)
└─ (3 3 4 4 1 2 1 4 1 4)
tet(10,444444)
└─ (1 4 4 1 2 4 3 1 2 3)
tet(100,777777)
└─ (2 3 3 2 1 2 3 4 2 2 4 1)
tet(100,777777)
└─ (2 3 3 2 1 2 3 4 2 2 4 1)
MAIN RAD EXACT FUNC 4/20

```

Fig. H 18

oder als **Histogramm** veranschaulichen

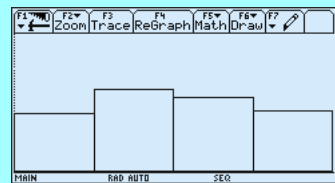


Fig. H 19

(3) Ermitteln von  $h_n(\{ \text{die Augenzahl beträgt höchstens } 3 \})$

Variante 1

Listen aus (2), linke Spalte, mit **int** und **sum** weiterverarbeiten

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
seq(rand(4),1,1,100)→liste
└─ (2 2 1 1 4 2 4 3 3 1 3 1)
int(liste)
└─ (0 0 0 0 1 0 1 0 0 0 0 0)
100-sum(int(liste)/4)
└─ 100
└─ 100-sum(int(liste/4))/100
└─ 72
MAIN RAD APPROR FUNC 2/20

```

Fig. H 20

Variante 2

Funktion **tetra** im Programm-Editor definieren

```

F1 F2 F3 F4 F5 F6
└─ Control I/O Var Find... Mode
tetra(n,a,x)
:Func
:Local i,liste,auz,relh
:RandSeed x
:O→liste
:For i,1,n
:rand(4)→auz
:If auz≤a Then
:augment(liste,(1))→liste
:Else
:augment(liste,(0))→liste
:EndIf
:approx(sum(liste)/n)→relh
:EndFor
:relh
:EndFunc
MAIN RAD AUTO FUNC

```

Fig. H 22

und im Home-Editor direkt ausgeben lassen

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
tetra(200,3,147897)
└─ .71
tetra(200,3,88458)
└─ .735
tetra(200,3,203040)
└─ .75
tetra(200,3,54321)
└─ .76
tetra(200,3,884427)
└─ .795
tetra(200,3,9753)
└─ .76
tetra(200,3,444)
└─ .72
MAIN RAD APPROR FUNC 2/20

```

Fig. H 23

Listen aus (2), mittl. Spalte, mit **int** und **sum** weiterverarbeiten

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
100-sum(int(100/4))
└─ 7
100-sum(int(tet(100,2247)/4))
└─ 78
100-sum(int(tet(100,3389)/4))
└─ 75
MAIN RAD APPROR FUNC 2/20

```

Fig. H 21

Vergleich mit der in diesem Fall leicht zu berechnenden Wahrscheinlichkeit

$$P(\{1; 2; 3\}) = \frac{3}{4} = 0,75$$

oder als Durchschnitt mehrerer Realisierungsfolgen

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
mean(seq(tetra(200,3,10^k*pi),k,1,10))
└─ .7435
tetra(200,3,10^k*pi,k,1,10)
└─ .78
MAIN RAD AUTO FUNC 1/20

```

Fig. H 24

b) Zufallsexperiment: Dreimaliges Werfen eines L-Tetraeders

- Ermitteln von  $h_{200}(\{ \text{die Augensumme beträgt } 7 \})$

Funktion **tetra3** im Programm-Editor definieren und

```

F1 F2 F3 F4 F5 F6
└─ Control I/O Var Find... Mode
tetra3(n,a,x)
:Func
:Local i,liste,auz,relh
:RandSeed x
:O→liste
:For i,1,n
:rand(4)+rand(4)+rand(4)→auz
:If auz=a Then
:augment(liste,(1))→liste
:Else
:augment(liste,(0))→liste
:EndIf
:approx(sum(liste)/n)→relh
:EndFor
:relh
:EndFunc
MAIN RAD AUTO FUNC

```

Fig. H 25

im Home-Editor ausgeben

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
tetra3(200,7,12)
└─ .175
tetra3(200,7,23)
└─ .165
tetra3(200,7,34)
└─ .205
tetra3(200,7,45)
└─ .195
tetra3(200,7,56)
└─ .17
tetra3(200,7,67)
└─ .22
tetra3(200,7,78)
└─ .17
tetra3(200,7,78)
└─ .17
MAIN RAD APPROR FUNC 7/20

```

Fig. H 26

```

F1 F2 F3 F4 F5 F6
└─ Algebra Calc Other PrgmIO Clear a-z...
tetra3(200,7,12)
└─ .175
tetra3(200,7,23)
└─ .165
tetra3(200,7,34)
└─ .205
tetra3(200,7,45)
└─ .195
tetra3(200,7,56)
└─ .17
tetra3(200,7,67)
└─ .22
tetra3(200,7,78)
└─ .17
tetra3(200,7,78)
└─ .17
MAIN RAD APPROR FUNC 7/20

```

Fig. H 27

## H 3 Bedingte Wahrscheinlichkeiten

### H 3.1 Der Begriff *bedingte Wahrscheinlichkeit*; allgemeiner Multiplikationssatz

Nicht selten hören wir von Menschen, die in große Angst oder Lethargie verfallen, da sie bei Gleichsetzung von schweren Erkrankungen (etwa Krebs oder AIDS) mit dem Todesurteil folgendermaßen schlussfolgern: Bei mir fiel der entsprechende Test positiv aus, also bin ich erkrankt.

Betrachten wir diese Schlussfolgerung rational, so erkennen wir recht schnell, dass sie sehr voreilig ist, da man für einen derartigen Schluss unbedingt Zuverlässigkeitsaussagen über den durchgeführten Test, also „bedingte“ Wahrscheinlichkeiten der folgenden Art benötigt:

- Wahrscheinlichkeit, dass der Test positiv ausfällt unter der Bedingung, die Testperson ist (wirklich) erkrankt;
- Wahrscheinlichkeit, dass der Test negativ ausfällt unter der Bedingung, die Testperson ist (in Wirklichkeit) nicht erkrankt.

Die Zuverlässigkeit des Tests sei durch folgende Angaben gekennzeichnet:

- 96 % der erkrankten Personen werden entdeckt;
- 94 % der nicht erkrankten Personen werden als solche erkannt.

Wir ermitteln die Wahrscheinlichkeit  $P(\{\text{eine Person mit positivem Testergebnis ist erkrankt}\})$  mittels der uns bereits (für die Gleichverteilung) vertrauten Baumdiagramme. Der zu betrachtende zufällige Vorgang kann als zweistufig angesehen werden, wobei die erste Stufe die beiden möglichen atomaren Ereignisse  $K = \{\text{die Person ist erkrankt}\}$  und  $\bar{K} = \{\text{die Person ist nicht erkrankt}\}$  und die zweite Stufe jeweils die beiden möglichen atomaren Ereignisse

$[+] = \{\text{der Test fällt positiv aus}\}$  und

$[-] = \{\text{der Test fällt negativ aus}\}$

aufweist. Bereits beim Beschriften dieses Baumdiagramms

erkennt man, dass die Wahrscheinlichkeiten für die atomaren Ereignisse der ersten Stufe fehlen. Die Wahrscheinlichkeit

$P(K)$  dafür, dass eine rein zufällig ausgewählte Person (der ent-

sprechenden Bevölkerungsgruppe) erkrankt ist, wurde nicht angegeben. Auch wenn wir annehmen, diese Wahrscheinlichkeit sei uns zugänglich, können wir mittels des obigen Baumdiagramms und den beiden (uns bisher nur für die Gleichverteilung) zur Verfügung stehenden Pfadregeln die gesuchte Wahrscheinlichkeit nicht direkt berechnen, denn wir suchen die Wahrscheinlichkeit dafür, dass eine Person erkrankt ist unter der Bedingung, dass ihr Testergebnis positiv ausgefallen ist.

Damit ist die zeitliche Reihenfolge des zu beurteilenden Ereignisses  $K$  und der Bedingung  $[+]$  im Vergleich zu der im Baumdiagramm dargestellten kausalen Ursache-Wirkung-Beziehung, die der natürlichen Zeitachse folgt, vertauscht. Dieses Problem – das Bestimmen der **Wahrscheinlichkeit des Eintretens eines Ereignisses A unter der Bedingung B**, dass ein gewisses (jedoch nicht sicheres) Indiz für sein Eintreten bereits beobachtet worden ist – soll nun genauer untersucht werden.

Erinnern wir uns dazu, wie und aufgrund welcher Erfahrung die Wahrscheinlichkeit eines Ereignisses  $A$  definiert worden ist:

*Erfahrung*

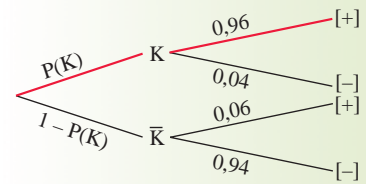
Stabilwerden relativer Häufigkeiten

$h_n(A) \rightsquigarrow P(A)$

*Definition H 8*

KOLMOGOROWSche Axiome

- (1)  $P(A) \geq 0$
- (2)  $P(\Omega) = 1$
- (3)  $P(A \cup B) = P(A) + P(B)$  für  $A \cap B = \emptyset$



Demzufolge ergäbe sich für die

$$\begin{aligned} \text{relative Häufigkeit des Ereignisses A} &= \frac{\text{Anzahl der Realisierungen, in denen A und B eintreten}}{\text{Anzahl der Realisierungen, in denen B eintritt}} \cdot \frac{1}{n} \\ \text{unter der Bedingung B} &= \frac{h_n(A \cap B)}{h_n(B)} \\ h_n(A|B) &= \frac{h_n(A \cap B)}{h_n(B)} \rightsquigarrow \frac{P(A \cap B)}{P(B)} = P(A|B)^1 = P_B(A) \end{aligned}$$

Wahrscheinlichkeit des Ereignisses A unter der Bedingung B; bedingte Wahrscheinlichkeit  $P_B$  von A (wenn  $P(B) \neq 0$ )

Gestützt auf diese Betrachtungen erscheint es sinnvoll, folgendermaßen zu definieren:

H 11

**Definition H 11:**

Sind  $P$  eine Wahrscheinlichkeitsverteilung auf dem endlichen Ereignisraum  $2^\Omega$  sowie  $A$  und  $B$  Ereignisse aus  $2^\Omega$  mit  $P(B) > 0$ , so heißt die durch  $P_B(A) = \frac{P(A \cap B)}{P(B)}$  definierte Funktion  $P_B$  die **bedingte Wahrscheinlichkeitsverteilung unter der Bedingung B**.

H 10

**Satz H 10:**

Ist  $P$  eine Wahrscheinlichkeitsverteilung auf dem endlichen Ereignisraum  $2^\Omega$  und gilt  $P(B) > 0$  für ein Ereignis  $B$  aus  $2^\Omega$ , so ist  $P_B$  eine Wahrscheinlichkeitsverteilung auf  $2^\Omega$ .

*Beweis:*

Es ist zu zeigen, dass  $P_B$  den drei KOLMOGOROWSchen Axiomen genügt. Dazu gelte  $A_1 \cap A_2 = \emptyset$ .

(1) Wegen  $P(A \cap B) \geq 0$  (nach Axiom (1) für  $P$ ) und  $P(B) > 0$  (nach Voraussetzung des Satzes) gilt

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \geq 0 \text{ für jedes } A \in 2^\Omega. \quad \text{w.z.b.w.}$$

(2)  $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$ , weil  $P(B) \neq 0$ . w.z.b.w.

$$\begin{aligned} (3) \quad P_B(A_1 \cup A_2) &= \frac{P((A_1 \cup A_2) \cap B)}{P(B)} = \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)} && \text{nach Definition H 11 und Distributivgesetz} \\ &= \frac{P(A_1 \cap B) + P(A_2 \cap B) - P(A_1 \cap B \cap A_2 \cap B)}{P(B)} && \text{nach Additionssatz H 3 (4)} \\ &= \frac{P(A_1 \cap B) + P(A_2 \cap B) - P(\emptyset)}{P(B)} && \text{nach Voraussetzung gilt } A_1 \cap A_2 = \emptyset \text{ und } \emptyset \cap A = \emptyset \text{ für alle } A \in 2^\Omega \\ &= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} && \text{nach Satz H 3 (2)} \\ &= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} = P_B(A_1) + P_B(A_2) && \text{w.z.b.w.} \end{aligned}$$

Jede bedingte Wahrscheinlichkeitsverteilung  $P_B$  ist also auch eine Wahrscheinlichkeitsverteilung.

Folglich gelten für  $P_B$  dieselben Rechenregeln wie für  $P$ :

$$P_B(\emptyset) = 0, \quad P_B(\bar{A}) = 1 - P_B(A), \quad P_B(A \cup B) = P_B(A) + P_B(B) - P_B(A \cap B) \quad \text{usw.}$$

Bringt man die definierende Gleichung bedingter Wahrscheinlichkeiten von der Quotientenform in die Produktform, so gibt man dieser einen gesonderten Namen:

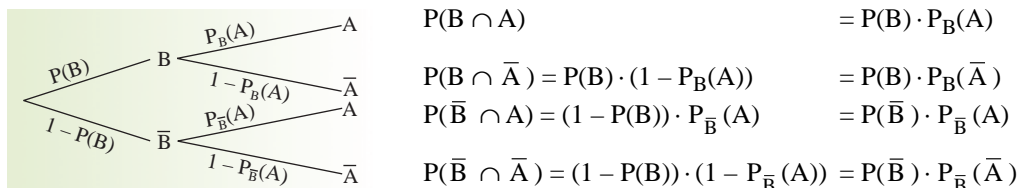
<sup>1)</sup> Diese ältere Schreibweise  $P(A|B)$  ist schreibtechnisch vorteilhafter als  $P_B(A)$ , jedoch hebt die Schreibweise  $P_B(A)$  stärker die veränderte Wahrscheinlichkeitsfunktion  $P_B$  hervor.

**Satz H 11: Allgemeiner Produkt- oder Multiplikationssatz**

Für zwei Ereignisse A und B mit  $P(B) > 0$  gilt stets  $P(A \cap B) = P_B(A) \cdot P(B)$ .

**H 11**

Satz H 11 gestattet, das **Baumdiagramm** auch in allgemeinerer Struktur zu verwenden.



Der allgemeine Produktsatz ist damit eine Verallgemeinerung der *Ersten Pfadregel* (Produktregel, s. Satz H 6) im zweistufigen Baumdiagramm.

Jede Datensammlung, die sich als Vierfeldertafel darstellen lässt, kann man in die Form eines **zweistufigen Baumdiagramms** bringen. Dabei wird auf der einen Stufe die eine und auf der anderen Stufe die andere Zerlegung von  $\Omega$  betrachtet. Mit welcher Zerlegungen man beginnt, spielt rechentech-nisch keine Rolle. Zu jeder Vierfeldertafel gehören also **zwei verschiedene Baumdiagramme**.

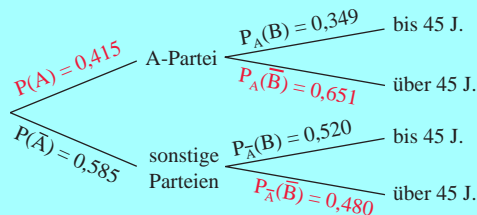
**Beispiel H 29:**

Nebenstehende Vierfeldertafel soll das Wahlverhalten zweier Altersgruppen gegenüber der A-Partei bei der Abgabe der Zweitstimme widerspiegeln.

		Wähler bis 45 J. B	Wähler über 45 J. $\bar{B}$	
A-Partei	A	14,5 %	27,0 %	41,5 %
sonstige Parteien	$\bar{A}$	30,4 %	28,1 %	58,5 %
		44,9 %	55,1 %	

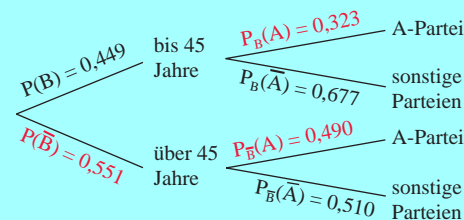
Presseartikel 1,  
aus dem sich Baumdiagramm und Vierfelder-tafel rekonstruieren lassen:

„Die A-Partei erreichte **41,5 %** der abgegebenen gültigen Zweitstimmen. Diese Stimmen kamen überwiegend (zu **65,1 %**) von Wählerinnen und Wählern über 45 Jahre. Bei den Wählern der übrigen Parteien macht diese Altersgruppe im Mittel nur **48,0 %** des Stimmenanteils aus.“



Presseartikel 2,  
aus dem sich Baumdiagramm und Vierfeldertafel rekonstruieren lassen:

„Besäßen nur Bürgerinnen und Bürger bis zum Alter von 45 Jahren das Wahlrecht, hätte die A-Partei noch nicht einmal ein Drittel der Stimmen erreicht (nämlich **32,3 %**). Unter den Wählerinnen und Wählern über 45, die **55,1 %** der Wählerschaft stellen, verpasste sie mit **49,0 %** knapp die absolute Mehrheit.“

**H 29**

Kehren wir nun – mit der Kenntnis der Definition bedingter Wahrscheinlichkeiten und des verallgemeinerten Baumdiagramms – zu dem eingangs beschriebenen Testproblem zurück (s. S. 389):

H 30

Beispiel H 30:

Die Zuverlässigkeit eines Testes für eine bestimmte Krankheit sei durch folgende Angaben gekennzeichnet:

- 96 % der erkrankten Personen werden durch den Test entdeckt,
- 94 % der nicht erkrankten Personen werden als solche erkannt.

Zu bestimmen ist die Wahrscheinlichkeit dafür, dass eine Person, bei welcher der Test positiv ausfiel, wirklich die betreffende Krankheit hat.

*Lösung:*

Wir definieren folgende Ereignisse:

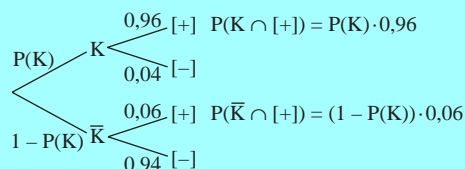
$K$  = {eine rein zufällig ausgewählte Person ist erkrankt}

$[+]$  = {der Test fällt positiv aus};  $[-]$  sei das Gegenereignis von  $[+]$ .

Gegeben:  $P_K([+]) = 0,96$ ;  $P_{\bar{K}}([-]) = 0,94$       gesucht:  $P_{[+]}(K)$

$$P_{[+]}(K) = \frac{P(K \cap [+])}{P([+] )}$$

- mit  $P(K \cap [+]) = 0,96 \cdot P(K)$  (wie dem nachfolgenden Baumdiagramm oder dem allgemeinen Multiplikationssatz zu entnehmen ist)
- mit  $P([+]) = P(K \cap [+]) + P(\bar{K} \cap [+])$  (nach Summenregel)  
 $= 0,96 \cdot P(K) + 0,06 \cdot (1 - P(K))$  (nach Baumdiagramm)



$$P_{[+]}(K) \Rightarrow \frac{0,96 \cdot P(K)}{0,96 \cdot P(K) + 0,06 \cdot (1 - P(K))} = \frac{0,96}{0,90 + \frac{0,06}{P(K)}}$$

Die gesuchte Wahrscheinlichkeit ist also bei den gegebenen Zuverlässigkeitsannahmen für den Test nur in Abhängigkeit von der Wahrscheinlichkeit  $P(K)$  der entsprechenden Erkrankung in der Bevölkerungsgruppe zu bestimmen. Ohne Kenntnis von  $P(K)$ , d. h. bei Vernachlässigen der Basisraten, ist ein Abschätzen der gesuchten bedingten Wahrscheinlichkeit  $P_{[+]}(K)$  nicht möglich. Das nachfolgende Zahlenbeispiel und der Graph (Fig. H 28) veranschaulichen diesen Sachverhalt noch einmal:

$$P(K) = 0,007 \quad P_{[+]}(K) \approx 0,10$$

$$P(K) = 0,07 \quad P_{[+]}(K) \approx 0,55$$

$$P(K) = 0,7 \quad P_{[+]}(K) \approx 0,97$$

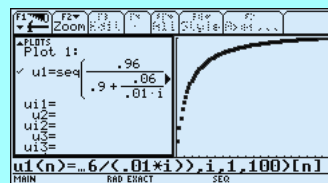


Fig. H 28

### H 3.2 Satz der totalen Wahrscheinlichkeit

Die in den Beispielen H 29 und H 30 angewendeten Rechenregeln für bedingte Wahrscheinlichkeiten sollen nun verallgemeinert werden.

Unter Verwendung des Baumdiagramms oder des allgemeinen Multiplikationssatzes erhielten wir im Beispiel H 30

$$P([+]) = P(K \cap [+]) + P(\bar{K} \cap [+]) = P(K) \cdot P_K([+]) + P(\bar{K}) \cdot P_{\bar{K}}([+]).$$

Um die Wahrscheinlichkeit des Ereignisses  $[+]$  zu berechnen, wird also

- zuerst  $\Omega$  in die zwei unvereinbaren Ereignisse  $K$  und  $\bar{K}$  zerlegt (Definition H 5) und dann
- $[+]$  in die zwei unvereinbaren Ereignisse  $K \cap [+]$  und  $\bar{K} \cap [+]$ .

Dadurch können die Additivität von  $P$  (Definition H 8) und der allgemeine Multiplikationssatz (Satz H 11) angewendet werden. Dieses Vorgehen muss nicht beschränkt bleiben auf eine Zerlegung von  $\Omega$  (Definition H 9) in zwei Ereignisse, sondern kann analog für eine Zerlegung in  $n$  Ereignisse ( $n \in \mathbb{N} \setminus \{0; 1\}$ ) durchgeführt werden.

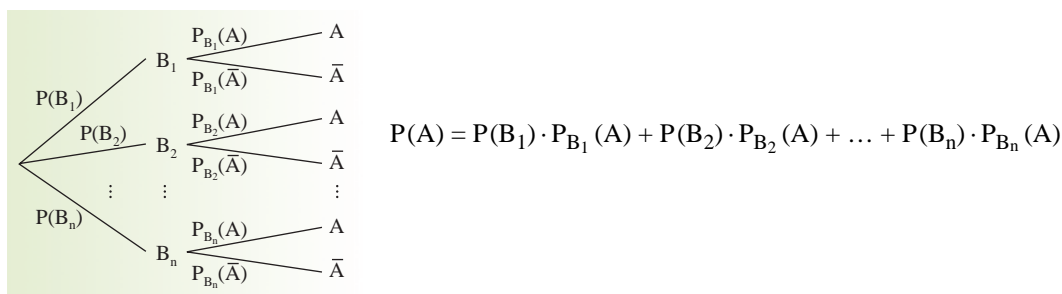
**Satz H 12: Satz der totalen (vollen) Wahrscheinlichkeit**

Bilden die Ereignisse  $B_1, B_2, \dots, B_n$  eine Zerlegung von  $\Omega$ , so gilt

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \dots + P(B_n) \cdot P_{B_n}(A).$$

H 12

Dieser Satz der totalen Wahrscheinlichkeit stellt eine Verallgemeinerung der *Zweiten Pfadregel* (Summenregel, vgl. Satz H 6) dar.



**Beispiel H 31:**

Ein Wanderer gehe vom Ort O aus und schlage an den Weggabelungen jeweils „auf gut Glück“ eine der möglichen Richtungen ein. Zu ermitteln ist die Wahrscheinlichkeit dafür, dass der Wanderer zum Punkt A gelangt, wenn folgendes Wegeschema (Fig. H 29) zugrunde liegt:

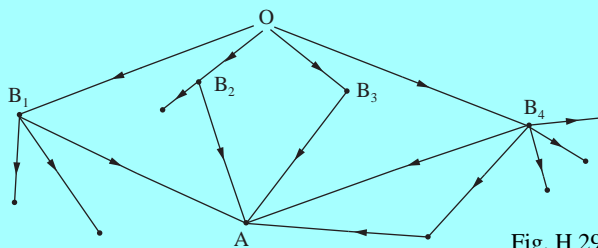


Fig. H 29

**Lösung:**

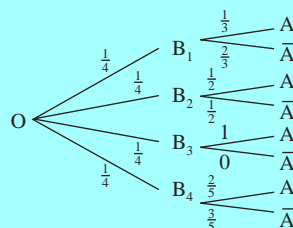
$A = \{\text{Wanderer kommt zum Punkt A}\};$

$B_i = \{\text{Wanderer kommt zum Punkt } B_i\}$  für  $i \in \{1; 2; 3; 4\}$

Die Ereignisse  $B_1, B_2, B_3$  und  $B_4$  bilden eine Zerlegung der zugehörigen Ergebnismenge  $\Omega$ . Somit gilt nach Satz H 12 für die gesuchte

Wahrscheinlichkeit  $P(A)$ :

$$\begin{aligned} P(A) &= \sum_{i=1}^4 P(B_i) \cdot P_{B_i}(A) \\ &= \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{2}{5} = \frac{67}{120} \approx 0,56 \end{aligned}$$



H 31

### H 3.3 BAYESSche Formel

Das Lösen stochastischer Probleme erfordert nicht selten, bei Kenntnis von bedingten Wahrscheinlichkeiten der Art  $P_{B_1}(A)$ ,  $P_{B_2}(A)$ , ... die „umgekehrten“ bedingten Wahrscheinlichkeiten der Art  $P_A(B_1)$ ,  $P_A(B_2)$ , ... zu berechnen (vgl. Beispiel H 30). Wir wenden hierfür zuerst die Definition H 11 der bedingten Wahrscheinlichkeit auf die zu ermittelnde Wahrscheinlichkeit an:

$$P_A(B_1) = \frac{P(A \cap B_1)}{P(A)}$$

Nutzt man nun zur Bestimmung des Zählers die erste Pfadregel, den allgemeinen Produktsatz (Satz H 11) und zur Bestimmung des Nenners die zweite Pfadregel, den Satz der totalen Wahrscheinlichkeit (Satz H 12), so ergibt sich die BAYESSche Formel:

H 13

Satz H 13: **BAYESSche Formel** (Satz von BAYES)<sup>1)</sup>

Bilden die Ereignisse  $B_1, B_2, \dots, B_n$  eine Zerlegung von  $\Omega$  und ist  $A$  ein beliebiges Ereignis aus  $2^\Omega$  mit  $P(A) > 0$ , so gilt für jedes  $i \in \{1; 2; \dots; n\}$

$$P_A(B_i) = \frac{P(B_i) \cdot P_{B_i}(A)}{P(A)} = \frac{P(B_i) \cdot P_{B_i}(A)}{P(B_1) \cdot P_{B_1}(A) + \dots + P(B_n) \cdot P_{B_n}(A)}.$$

H 32

Beispiel H 32:

In einem Gerätesystem seien die Geräte A, B, C in Reihe geschaltet, d.h., das Gerätesystem fällt genau dann aus, wenn eines der Geräte A, B, C ausfällt. Es sei nicht möglich, dass zwei Geräte gleichzeitig ausfallen. Langzeiterfahrungen besagen nun sowohl, dass das System wegen eines Defekts von A mit einer Wahrscheinlichkeit von 0,50 ausfällt, wegen B mit 0,30 und wegen C mit 0,20 – als auch, dass die Sicherung beim Ausfallen von A mit einer Wahrscheinlichkeit von 0,20 durchschlägt, beim Ausfallen von B im Prinzip immer und von C im Prinzip nie. (Ohne dass *eines* der Geräte ausfällt, schlage die Sicherung nicht durch.)

In welcher Reihenfolge sollte man zweckmäßigerweise die Geräte kontrollieren (um sie ggf. auszuwechseln), wenn das Gerätesystem ausgefallen und die Sicherung durchgeschlagen ist?

*Lösung:*

$S = \{\text{Gerätesystem fällt aus}\} = \Omega;$

$A = \{\text{Gerät A ist defekt}\};$

$B = \{\text{Gerät B ist defekt}\};$

$C = \{\text{Gerät C ist defekt}\}$  mit  $A \cup B \cup C = \Omega$  und  $A \cap B = \emptyset$ ,  
 $A \cap C = \emptyset, B \cap C = \emptyset$

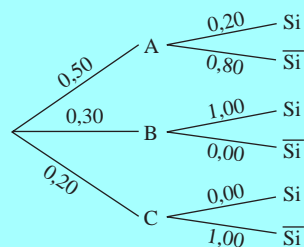
$Si = \{\text{Sicherung ist durchgeschlagen}\}$  mit  $Si \cap S = Si$

$P(\{\text{die Ursache für das Ausfallen des Systems und das Durchschlagen der Sicherung ist der Defekt von Gerät A}\})$

$$\begin{aligned} &= P_{S \cap Si}(A) = P_{Si}(A) = \frac{P(A \cap Si)}{P(Si)} = \frac{P(A) \cdot P_A(Si)}{P(A) \cdot P_A(Si) + P(B) \cdot P_B(Si) + P(C) \cdot P_C(Si)} \\ &= \frac{0,50 \cdot 0,20}{0,50 \cdot 0,20 + 0,30 \cdot 1,00 + 0,20 \cdot 0,00} = 0,25 \end{aligned}$$

Analog berechnen sich  $P_{S \cap Si}(B) = 0,75$  sowie  $P_{S \cap Si}(C) = 0,00$ .

Da die Ursachenwahrscheinlichkeit für B größer als die für A und die wiederum größer als die für C ist, würde es (statistisch gesehen) sinnvoll sein, die Geräte in der Reihenfolge B, A, C zu kontrollieren.



<sup>1)</sup> Thomas BAYES (1702–1761); engl. Pastor und Mathematiker

## Beispiel H 33:

In Sterbetafeln statistischer Jahrbücher werden zu jedem Alter  $x$  (in Jahren) zwischen 0 und 90 die Zahlen  $N(x)$  tabelliert, die angeben, wie viele von 100 000 Neugeborenen mindestens  $x$  Jahre alt werden. Wir betrachten die Ereignisse

$A_i = \{\text{ein weibliches Neugeborenes wird mindestens } i \text{ Jahre alt}\}$ ,

wobei die jeweiligen Personen stets „auf gut Glück“ ausgewählt seien.

Zu berechnen sind folgende Wahrscheinlichkeiten:

- $P(\{\text{ein weibliches Neugeborenes wird mindestens 70 Jahre alt}\})$
- $P(\{\text{eine gegenwärtig 40-jährige weibliche Person wird mindestens 60 Jahre alt}\})$
- $P(\{\text{eine weibliche Person wird mindestens 90 Jahre alt, wenn sie gegenwärtig } x \text{ Jahre alt ist}\})$
- $P(\{\text{eine gegenwärtig } x \text{ Jahre alte weibliche Person wird noch mindestens 10 Jahre leben}\})$

Vollendetes Altersjahr $x$	$N(x)$
0	100000
10	99393
20	99181
30	98835
40	98146
50	96362
60	92523
70	83332
80	60682
90	20623

Von 100000 weiblichen  
Lebendgeborenen erreichen  
 $N(x)$  das Alter  $x$   
(Sterbetafelauzug  
Deutschland 1995/97)

## Lösung:

Lösung mithilfe des GTA; die Daten der Spalte  $N(x)$  sind als Liste *liste* dargestellt:

$$\begin{aligned} \text{a) } P(A_{70}) &= \frac{N(70)}{N(0)} \\ &= \frac{83332}{100000} \\ &= 0,83332 \end{aligned}$$

Der Quotient  $\frac{N(70)}{N(0)}$  errechnet sich als Quotient des 8-ten Gliedes von *liste* durch das erste Glied von *liste*.

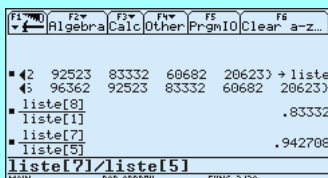


Fig. H 30

$$\begin{aligned} \text{b) } P_{A_{40}}(A_{60}) &= \frac{P(A_{40} \cap A_{60})}{P(A_{40})} \\ &= \frac{P(A_{60})}{P(A_{40})} = \frac{N(60) : N(0)}{N(40) : N(0)} \\ &= \frac{N(60)}{N(40)} = \frac{92523}{98146} = 0,942708 \end{aligned}$$

Der Quotient  $\frac{N(60)}{N(40)}$  errechnet sich als Quotient des 7-ten Gliedes von *liste* durch das 5-te Glied der Liste.

$$\begin{aligned} \text{c) } P_{A_x}(A_{90}) &= \frac{N(90)}{N(x)} \\ &\text{für } x \in \{0; 10; \dots; 90\} \end{aligned}$$

Die Quotienten  $\frac{N(90)}{N(x)}$  lassen sich mit dem **Sequenz**-Befehl als Liste *liste2* der Quotienten aus dem 10-ten und dem  $i$ -ten Glied von *liste* für  $i = 1; \dots; 10$  darstellen oder grafisch veranschaulichen.

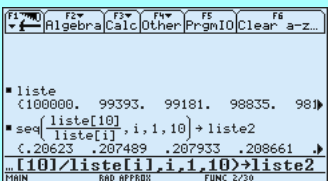


Fig. H 31

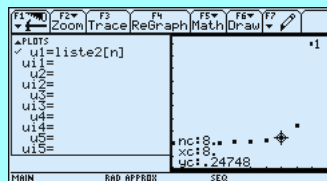


Fig. H 32



d)  $P_{A_x}(A_{x+10}) = \frac{N(x+10)}{N(x)}$   
für  $x \in \{0; 10; \dots; 80\}$

Die Quotienten  $\frac{N(x+10)}{N(x)}$  lassen sich mit dem *Sequenz*-Befehl als Liste *liste3* der Quotienten aus dem  $(i+1)$ -ten und dem  $i$ -ten Glied von *liste* für  $i = 1; \dots; 9$  darstellen oder grafisch veranschaulichen.

Fig. H 33

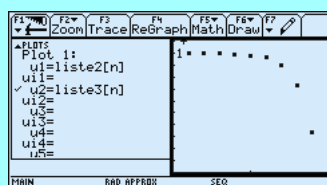


Fig. H 34

### H 3.4 Unabhängigkeit von Ereignissen; spezieller Multiplikationssatz

Hat die blaue Augenfarbe eines Kindes etwas mit seiner mathematischen Begabung zu tun? Vielleicht ja, denn Jana, das Matheass der Klasse, hat blaue Augen. Vielleicht nein, denn Eugen, das Matheass der Parallelklasse hat braune Augen. Um der Frage weiter nachzugehen, wählen wir 1000 Vierzehnjährige „auf gut Glück“ aus und überprüfen, ob bei ihnen die Merkmale „mathematisch begabt“ und „blauäugig“ ausgeprägt sind. Die (fiktiv) ermittelten absoluten Häufigkeiten werden in eine Vierfeldertafel eingetragen.

	B blauäugig	$\bar{B}$ nicht blauäugig	
M mathematisch begabt	7	43	50
$\bar{M}$ mathematisch nicht begabt	136	814	950
	143	857	1000

Dass die blaue Augenfarbe keinen Einfluss auf die Begabtheit für Mathematik besitzt, diese beiden Merkmale also *unabhängig voneinander* auftreten, wäre nachgewiesen, wenn der Anteil der 7 blauäugigen, mathematisch Begabten unter allen 50 mathematisch Begabten genauso groß ist wie der Anteil der 143 Blauäugigen unter allen 1000 getesteten Vierzehnjährigen:

$\frac{7}{50} = 0,14$  und  $\frac{143}{1000} = 0,143$ . Beide Anteile sind in guter Näherung (auf zwei Dezimalen gerundet) als gleich anzusehen. Mit absoluten bzw. relativen Häufigkeiten und Wahrscheinlichkeiten ließe sich diese Anteilsgleichheit folgendermaßen schreiben:

$$\left. \begin{aligned} \frac{7}{50} &= \frac{H_{1000}(B \cap M)}{H_{1000}(M)} = \frac{\frac{H_{1000}(B \cap M)}{1000}}{\frac{H_{1000}(M)}{1000}} = \frac{h_{1000}(B \cap M)}{h_{1000}(M)} \approx \frac{P(B \cap M)}{P(M)} = P_M(B) \\ \frac{143}{1000} &= \frac{H_{1000}(B)}{1000} = h_{1000}(B) \approx P(B) \end{aligned} \right\} \Rightarrow P_M(B) = P(B) \text{ in guter Näherung}$$

H 12

Definition H 12:

Zwei Ereignisse A und B des Ereignisraumes  $2^\Omega$  mit  $P(B) > 0$  heißen genau dann **voneinander (stochastisch)<sup>1)</sup> unabhängig**, wenn  $P_B(A) = P(A)$  gilt.

H 14

Satz H 14:

Wenn  $P_B(A) = P(A)$  mit  $P(A) > 0$ , so gilt auch  $P_A(B) = P(B)$ .

<sup>1)</sup> im Unterschied z.B. zur „linearen“ Unabhängigkeit von Vektoren (s. Definition G 14)

*Beweis:*

Nach Definition H 11 ist  $P_A(B) = \frac{P(A \cap B)}{P(A)}$ , woraus nach dem allgemeinen Multiplikationssatz (Satz H 11) und nach obiger Voraussetzung  $P_A(B) = \frac{P_B(A) \cdot P(B)}{P(A)} = \frac{P(A) \cdot P(B)}{P(A)} = P(B)$  folgt w.z.b.w.

Für unabhängige Ereignisse vereinfacht sich somit der allgemeine Multiplikationssatz:

**Satz H 15: Spezieller Multiplikationssatz**

Zwei Ereignisse A und B mit  $P(A) > 0$  und  $P(B) > 0$  sind genau dann voneinander unabhängig, wenn  $P(A \cap B) = P(A) \cdot P(B)$  gilt.

H 15

*Bemerkungen:*

Beim Gebrauch des Begriffes der (stochastischen) Unabhängigkeit von Ereignissen A, B, C mit  $P(A) > 0$ ,  $P(B) > 0$ ,  $P(C) > 0$  ist zu beachten:

(1) Sind A und B *unvereinbar*, so gilt  $P(A \cup B) = P(A) + P(B)$ .

Sind A und B *unabhängig*, so gilt  $P(A \cap B) = P(A) \cdot P(B)$ .

(2) Wenn A und B unabhängig sind, dann sind dies auch  $\bar{A}$  und  $\bar{B}$ ,  $\bar{A}$  und B sowie  $\bar{A}$  und  $\bar{B}$  (vgl. nebenstehende Vierfeldertafel).

	B	$\bar{B}$	
A	$P(A) \cdot P(B)$	$P(A) \cdot (1 - P(B))$	$P(A)$
$\bar{A}$	$(1 - P(A)) \cdot P(B)$	$(1 - P(A)) \cdot (1 - P(B))$	$1 - P(A)$
	$P(B)$	$1 - P(B)$	

(3) Drei Ereignisse A, B, C heißen **paarweise (stochastisch) unabhängig**, wenn

$P(A \cap B) = P(A) \cdot P(B)$  und  $P(A \cap C) = P(A) \cdot P(C)$  und  $P(B \cap C) = P(B) \cdot P(C)$ .

(4) Drei Ereignisse A, B, C heißen **(stochastisch) unabhängig**, wenn sie sowohl paarweise (stochastisch) unabhängig sind als auch der Gleichung  $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$  genügen.

Die *praktische* Bedeutung stochastischer Unabhängigkeit zeigen folgende Beispiele:

Nehmen wir an, ein „Wetterprophet“ sage für morgen schönes Wetter voraus und er irre sich in 20 % aller Fälle. Ein zweiter „Wetterprophet“ sage dasselbe voraus und die Wahrscheinlichkeit seines Irrtums sei ebenfalls 0,20. Schließt man nun daraus, dass die Wahrscheinlichkeit für einen Irrtum *beider*  $0,20 \cdot 0,20 = 0,040$  betrage, so ist genau das i. A. falsch. Man muss nämlich davon ausgehen, dass die beiden übereinstimmenden Wetterprognosen auf denselben Beobachtungsdaten basieren und deshalb *nicht voneinander (stochastisch) unabhängig*, sondern im Extremfall *identisch* sind. Gäbe nämlich der zweite Wetterprophet statt einer eigenen Wetterprognose nur das Zitat des ersten Wetterpropheten bekannt, so bliebe die Wahrscheinlichkeit für einen Irrtum gleich 0,20.

Man darf also nicht dem Trugschluss erliegen, dass die Übereinstimmung zweier Vorhersagen stets die Wahrscheinlichkeit eines Irrtums verkleinert.

**Beispiel H 34:**

In einer Urne befinden sich genau vier Kugeln – zwei weiße und zwei rote. Dieser Urne werden nacheinander a) mit Zurücklegen, b) ohne Zurücklegen genau zwei Kugeln entnommen. Sind die beiden Ereignisse

$A = \{\text{die erste gezogene Kugel ist weiß}\}$  und  $B = \{\text{die zweite gezogene Kugel ist weiß}\}$  voneinander (stochastisch) unabhängig?

*Lösung:*

Zu a): Nachstehendem Baumdiagramm sind die folgenden Wahrscheinlichkeiten zu entnehmen:

H 34

$$P(A) = \frac{2}{4} = \frac{1}{2}; \quad P(B) = \frac{2}{4} \cdot \frac{2}{4} + \frac{2}{4} \cdot \frac{2}{4} = \frac{1}{2}$$

$$P(A \cap B) = \frac{2}{4} \cdot \frac{2}{4} = \frac{1}{4}$$

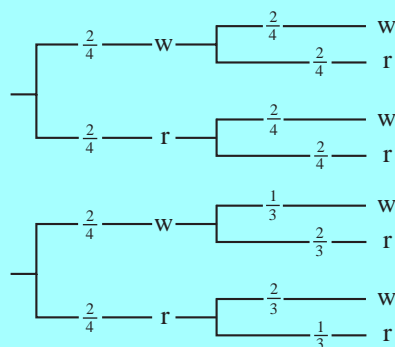
Folglich gilt  $P(A \cap B) = P(A) \cdot P(B)$ , d.h., A und B sind *voneinander (stochastisch) unabhängig*.

Zu b): Dem nebenstehenden Baumdiagramm sind die folgenden Wahrscheinlichkeiten zu entnehmen:

$$P(A) = \frac{2}{4} = \frac{1}{2}; \quad P(B) = \frac{2}{4} \cdot \frac{1}{3} + \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{2}$$

$$P(A \cap B) = \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{6}$$

Folglich gilt  $P(A \cap B) \neq P(A) \cdot P(B)$ , d.h., A und B sind *nicht voneinander unabhängig*.



Recht bemerkenswert ist das im Beispielteil b) erhaltene Ergebnis. Dabei verwundert sicher weniger, dass das Ereignis  $B = \{\text{die zweite gezogene Kugel ist weiß}\}$  von dem Ereignis  $A = \{\text{die erste gezogene Kugel ist weiß}\}$  stochastisch abhängig ist, da zwischen dem zuerst eingetretenen Ereignis A und dem danach eintretenden Ereignis B infolge des Nichtzurücklegens der zuerst gezogenen Kugel auch eine kausale Abhängigkeit besteht. Hingegen ist doch erstaunlich, dass auch A von B stochastisch abhängt, dass ein späteres Ereignis zur Beurteilung eines früheren herangezogen wird, obwohl es offensichtlich keine kausale Abhängigkeit geben kann. *Das Fehlen kausaler Abhängigkeit muss also nicht stochastische Unabhängigkeit nach sich ziehen.*

## H 35

Beispiel H 35:

Man vervollständige die nebenstehende Vierfeldertafel für den Fall, dass A und B voneinander unabhängig sind.

	B	$\bar{B}$
A	0,12	
$\bar{A}$		
	0,60	

Lösung:

$$P(B) = 1 - P(\bar{B}) = 1 - 0,60 = 0,40$$

$$P(A \cap B) = P(A) \cdot P(B) \Rightarrow P(A) = \frac{P(A \cap B)}{P(B)} = \frac{0,12}{0,40} = 0,30$$

$$P(\bar{A}) = 1 - P(A) = 1 - 0,30 = 0,70$$

$$P(\bar{A} \cap B) = P(\bar{A}) \cdot P(B) = 0,70 \cdot 0,40 = 0,28$$

$$P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = 0,70 \cdot 0,60 = 0,42$$

$$P(A \cap \bar{B}) = P(A) \cdot P(\bar{B}) = 0,30 \cdot 0,60 = 0,18$$

## H 36

Beispiel H 36:

Herr Theodor-Claus P. möchte heute Abend mit dem Auto von A(stadt) nach E(dorf) fahren. Der aktuellen Stauvorwarnkarte ist zu entnehmen, mit welcher Wahrscheinlichkeit auf den einzelnen Straßenabschnitten mit Stau zu rechnen ist.

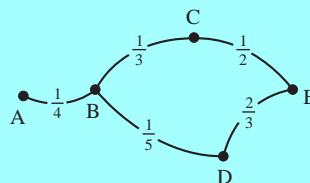
Mit welcher Wahrscheinlichkeit kann Herr P. über dieses Straßennetz von A nach E fahren, ohne in einen Stau zu geraten?

Zur Beantwortung dieser Frage werden folgende Ereignisse definiert:

$S_{AB} = \{\text{Stau auf dem Straßenabschnitt AB}\}; \quad S_{BC} = \{\text{Stau auf dem Straßenabschnitt BC}\}$

$S_{CE} = \{\text{Stau auf dem Straßenabschnitt CE}\}; \quad S_{BD} = \{\text{Stau auf dem Straßenabschnitt BD}\}$

$S_{DE} = \{\text{Stau auf dem Straßenabschnitt DE}\}$



Daraus ergibt sich:

$P(\{\text{von A nach E ohne Stau}\})$

=  $P(\{(AB \text{ ohne Stau und } BC \text{ ohne Stau und } CE \text{ ohne Stau}) \text{ oder } (AB \text{ ohne Stau und } BD \text{ ohne Stau und } DE \text{ ohne Stau})\})$

=  $P((\overline{S_{AB}} \cap \overline{S_{BC}} \cap \overline{S_{CE}}) \cup (\overline{S_{AB}} \cap \overline{S_{BD}} \cap \overline{S_{DE}}))$

=  $P(\overline{S_{AB}} \cap \overline{S_{BC}} \cap \overline{S_{CE}}) + P(\overline{S_{AB}} \cap \overline{S_{BD}} \cap \overline{S_{DE}}) - P(\overline{S_{AB}} \cap \overline{S_{BC}} \cap \overline{S_{CE}} \cap \overline{S_{BD}} \cap \overline{S_{DE}})$

(nach dem Additionssatz für Wahrscheinlichkeiten und bei Annahme der Unabhängigkeit der Ereignisse)

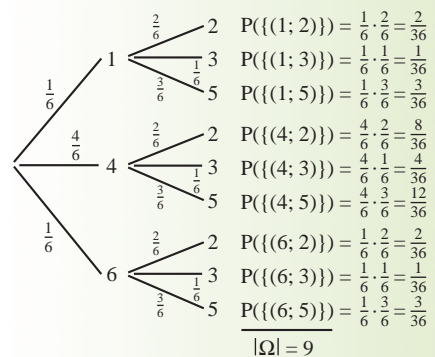
$$= \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{1}{3} - \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{4}{5} \cdot \frac{1}{3} \approx 0,383$$

## H 4 Zufallsgrößen

### H 4.1 Endliche Zufallsgrößen

Bei den bisherigen mathematischen Untersuchungen von zufälligen Vorgängen konzentrierten wir uns darauf, die Zufallsexperimente durch geeignete Ergebnismengen  $\Omega$  bzw. Ereignisräume  $2^\Omega$  zu beschreiben sowie die Wahrscheinlichkeiten von Ereignissen zu bestimmen. Es wurde also ermittelt, mit welcher Wahrscheinlichkeit  $P(A)$  ein Ereignis  $A$  zu *erwarten* ist. Was aber dieses Ereignis für uns *wert* ist, was sein Eintreten also an Gewinn bzw. Verlust bringt, wird erst jetzt mit dem so genannten *Erwartungswert* in die Betrachtungen einbezogen. Betrachtet sei folgendes Beispiel:

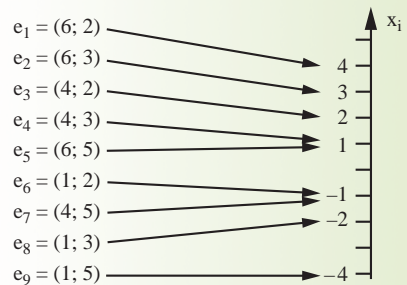
Andreas wird zu einem Würfelspiel eingeladen. Dazu liegen zwei ungezinkte Würfel bereit, von denen der eine Würfel (W1) die sechs Augenzahlen 1, 4, 4, 4, 4, 6 trägt und der andere (W2) die sechs Augenzahlen 2, 2, 3, 5, 5, 5. Als Gast darf er sich einen Spielwürfel auswählen. In jeder Spielrunde würfelt jeder der beiden Spieler genau einmal mit seinem Würfel. Nach jeder Würfelrunde zahlt der Spieler, der die niedrigere Augenzahl hat, an seinen Spielpartner die (positive) Differenz der Augenzahlen in Cent. Für welchen der Würfel sollte sich Andreas entscheiden, um zu gewinnen?



Das Werfen der beiden Würfel ist ein zweistufiges Zufallsexperiment, das durch das nebenstehende Baumdiagramm dargestellt werden kann. Daraus ist die neun-elementige Ergebnismenge

$\Omega = \{(w_1; w_2) \mid w_1 \in \{1; 4; 6\}, w_2 \in \{2; 3; 5\}\}$  ablesbar.

Die Wahrscheinlichkeiten ihrer atomaren Ereignisse ergeben sich nach der ersten Pfadregel. Um seine Entscheidung zu treffen, interessiert sich Andreas jedoch weniger für die neun möglichen Ergebnisse und die zugehörigen Wahrscheinlichkeitsverteilung als vielmehr für den Gewinn bzw. Verlust, den er durch die Wahl des einen bzw. des anderen Spielwürfels zu erwarten hat.



Fall 1: Andreas entscheidet sich für Würfel 1

Die Spielregeln ordnen jedem zufälligen Ergebnis einen positiven bzw. negativen Gewinn, seinen *Wert* in Cent zu. Mit anderen Worten: Jedem möglichen Ergebnis  $e \in \Omega$  wird (sieht man von der Maßeinheit ab) eine reelle Zahl zugeordnet.

## H 13

## Definition H 13:

Eine Funktion  $X: \Omega \mapsto \mathbb{R}$ , die jedem Ergebnis  $e \in \Omega$  eines Zufallsexperiments eine reelle Zahl  $x$  zuordnet, heißt **Zufallsgröße** (oder auch **Zufallsvariable**)  $X$ . Die Elemente des Wertebereichs von  $X$  nennt man **Werte** der Zufallsgröße  $X$ . Eine Zufallsgröße  $X$ , die nur endlich viele oder höchstens abzählbar unendlich viele verschiedene (Funktions-) Werte  $x_1, x_2, \dots, x_n, \dots$  besitzt, heißt **diskrete Zufallsgröße**. Diskrete Zufallsgrößen mit nur endlich vielen Werten bezeichnet man auch als **endliche Zufallsgrößen**.

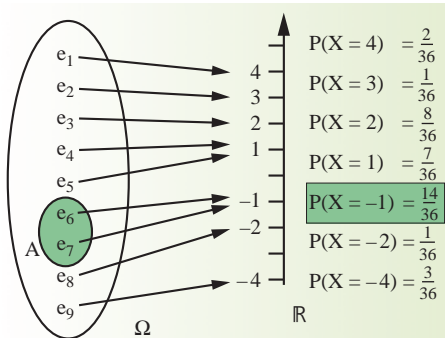
Zufallsgrößen werden i. Allg. mit den Großbuchstaben  $X, Y, Z$  bzw.  $X_i, Y_i, Z_i$  oder mit einem dem jeweiligen praktischen Problem angepassten Großbuchstaben (z.B.  $G$  für Gewinn) bezeichnet.

In dem Würfelspiel-Beispiel S. 399 haben wir die Bewertung der Ergebnisse  $e \in \Omega$  mithilfe der Zufallsgröße  $X$  beschrieben, die der nebenstehenden Funktionsgleichung genügt:

$$X(e) = \begin{cases} -4 & \text{für } e = (1; 5) \\ -2 & \text{für } e = (1; 3) \\ -1 & \text{für } e \in \{(4; 5), (1; 2)\} \\ +1 & \text{für } e \in \{(6; 5), (4; 3)\} \\ +2 & \text{für } e = (4; 2) \\ +3 & \text{für } e = (6; 3) \\ +4 & \text{für } e = (6; 2) \end{cases}$$

In der Stochastik verwendet man anstatt der in der Analysis üblichen Gleichungen für Funktionswerte wie z.B.  $X(e) = -1$  meist nur die Kurzschreibweise  $X = -1$ . In obigem Beispiel gilt  $X = -1$  genau dann, wenn das Ereignis  $A = \{(4; 5), (1; 2)\}$  eintritt. Dies erfolgt – nach der zweiten Pfadregel – mit der Wahrscheinlichkeit  $\frac{12}{36} + \frac{2}{36} = \frac{14}{36}$ . Dafür sagt man auch kürzer: Die Zufallsgröße  $X$  **nimmt den Wert  $-1$  mit der Wahrscheinlichkeit  $\frac{14}{36}$  an** und schreibt

$P(X = -1) = \frac{14}{36}$  oder  $P_X(\{-1\}) = \frac{14}{36}$ . Auf diese Weise ordnet man jedem Wert  $x_i$  der Zufallsgröße  $X$  die Wahrscheinlichkeit  $P(X = x_i)$  zu. Dabei ist  $P(X = x_i)$  gleich der Wahrscheinlichkeit  $P(A)$  des Ereignisses  $A \in 2^\Omega$ , das all die Ergebnisse umfasst, denen die Zufallsgröße  $X$  den Wert  $x_i$  zuordnet.



$e$	$e_1$	$e_2$	$e_3$	$e_4$	$e_5$	$e_6$	$e_7$	$e_8$	$e_9$
$P(\{e\})$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{8}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{12}{36}$	$\frac{1}{36}$	$\frac{3}{36}$

$x_i$	-4	-2	-1	1	2	3	4
$P(X = x_i)$	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{14}{36}$	$\frac{7}{36}$	$\frac{8}{36}$	$\frac{1}{36}$	$\frac{2}{36}$

## H 14

## Definition H 14:

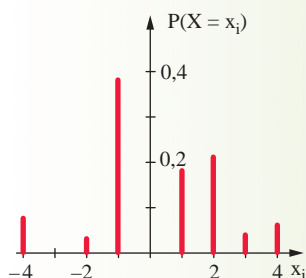
Kann eine diskrete Zufallsgröße  $X: \Omega \rightarrow \mathbb{R}$  nur die Werte  $x_1, x_2, \dots, x_n, \dots$  annehmen, so heißt die Funktion  $x_i \mapsto p_i = P(X = x_i) = P(\{e \in \Omega \text{ und } X(e) = x_i\})$  mit  $i \in \{1; 2; \dots; n; \dots\}$  **Wahrscheinlichkeitsverteilung** (Wahrscheinlichkeitsfunktion) der Zufallsgröße  $X$ . Die Funktion  $F$  mit  $F(x) = P(X \leq x)$  nennt man **Verteilungsfunktion** der Zufallsgröße  $X$ .

Als Darstellungsmöglichkeiten der Wahrscheinlichkeitsverteilung einer endlichen Zufallsgröße werden neben Wertetabellen auch Stab- oder Säulendiagramme (häufig als Histogramme<sup>1)</sup> bezeichnet) und insbesondere zweizeilige Matrizen genutzt. Die Daten aus dem obigen Würfelspiel ließen sich also auch folgendermaßen darstellen:

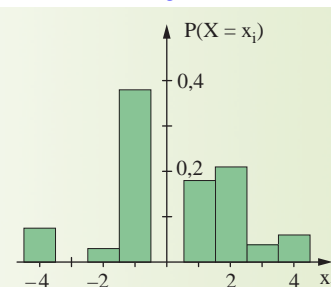
*zweizeilige Matrix:*

$$X \cong \begin{pmatrix} -4 & -2 & -1 & 1 & 2 & 3 & 4 \\ \frac{3}{36} & \frac{1}{36} & \frac{14}{36} & \frac{7}{36} & \frac{8}{36} & \frac{1}{36} & \frac{2}{36} \end{pmatrix}$$

*Stabdiagramm*



*Histogramm*

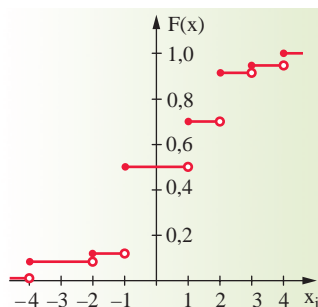


Mit dieser Wahrscheinlichkeitsverteilung bzw. mit der zugehörigen *Verteilungsfunktion*  $F$  lässt sich nun die Wahrscheinlichkeit bestimmen, mit der Andreas bei der Wahl des Würfels  $W_1$  das Spiel, an dem er sich beteiligt, auch gewinnt.

$$\begin{aligned} P(\{\text{Andreas gewinnt}\}) &= P(X > 0) = P(X = 1) + P(X = 2) + P(X = 3) \\ &= \frac{7}{36} + \frac{8}{36} + \frac{1}{36} + \frac{2}{36} = \frac{18}{36} = 0,5 \end{aligned}$$

$$\text{bzw.} \quad P(X > 0) = 1 - P(X \leq 0) = 1 - F(0) = 1 - 0,5 = 0,5$$

Eine Gewinnwahrscheinlichkeit von 0,5 zu haben ist eine recht schwache Information für Andreas, um sich für oder gegen eine Wahl des Spielwürfels 1 zu entscheiden. Fundierter könnte er seine Entscheidung treffen, wenn er berücksichtigt, welchen Gewinn bzw. Verlust er bei der Wahl des Würfels  $W_1$  zu erwarten hat.



## H 4.2 Erwartungswert

Beteiligte sich Andreas mit dem Würfel  $W_1$  beispielsweise an 3600 Runden des in im Abschnitt H 4.1 betrachteten Spieles, so wüssten wir auf Grund des empirischen Gesetzes der großen Zahlen und obiger Wahrscheinlichkeitsverteilung von  $X$ , dass er bei etwa  $\frac{3}{36}$  der Runden (d.h. im Mittel bei 300 Runden) jeweils 4 ct, bei etwa  $\frac{1}{36}$  der Runden (d.h. im Mittel bei 100 Runden) jeweils 2 ct und bei etwa  $\frac{14}{36}$  der Runden (d.h. im Mittel also bei 1400 Runden) jeweils 1 ct zu bezahlen hätte. Dagegen dürfte er im Mittel bei 700 Runden mit einem Gewinn von je 1 ct, bei 800 Runden von je 2 ct, bei 100 Runden von je 3 ct und bei 200 Runden mit einem Gewinn von je 4 ct rechnen. Der Gewinn (in ct) bei 3600 Spielen ergäbe sich damit näherungsweise aus der folgenden Rechnung:

$$(-4) \cdot 300 + (-2) \cdot 100 + (-1) \cdot 1400 + 1 \cdot 700 + 2 \cdot 800 + 3 \cdot 100 + 4 \cdot 200 = 600$$

Dividiert man diese Gleichung durch die Anzahl der Spielrunden, so ergibt sich ein *durchschnittlicher* Gewinn (oder auch *mittlerer* Gewinn) pro Runde von  $0,1\bar{6}$ . Das heißt: Die Wahl des Würfels  $W_1$  führt auf Dauer gesehen für Andreas zu einem Gewinn von jeweils  $0,1\bar{6}$  ct.

<sup>1)</sup> Bei dieser Veranschaulichung werden die Wahrscheinlichkeiten  $P(X = x_i)$  ( $i \in \{1; 2; \dots; n\}$ ) durch je eine Rechteckfläche dargestellt. In der Regel wählt man für diese Rechtecke einheitlich die Breite 1, so dass die Rechteckshöhe gleich  $P(X = x_i)$  ist. Entscheidend ist aber nur, dass die Rechtecke eines Histogramms alle dieselbe Breite und jeweils den Flächeninhalt  $P(X = x_i)$  aufweisen.

Dieser mittlere Gewinn pro Runde wurde also folgendermaßen berechnet:

$$0,1\bar{6} \approx (-4) \cdot \frac{H_{3600}(X=-4)}{3600} + (-2) \cdot \frac{H_{3600}(X=-2)}{3600} + (-1) \cdot \frac{H_{3600}(X=-1)}{3600} \\ + 1 \cdot \frac{H_{3600}(X=1)}{3600} + 2 \cdot \frac{H_{3600}(X=2)}{3600} + 3 \cdot \frac{H_{3600}(X=3)}{3600} + 4 \cdot \frac{H_{3600}(X=4)}{3600},$$

$$0,1\bar{6} \approx (-4) \cdot h_{3600}(X=-4) + (-2) \cdot h_{3600}(X=-2) + (-1) \cdot h_{3600}(X=-1) \\ + 1 \cdot h_{3600}(X=1) + 2 \cdot h_{3600}(X=2) + 3 \cdot h_{3600}(X=3) + 4 \cdot h_{3600}(X=4).$$

Auf Grund der Gültigkeit des empirischen Gesetzes der großen Zahlen werden nun bekanntlich relative Häufigkeiten  $h_n(A)$  bei hinreichend großen  $n$  stabil gegen die entsprechende Wahrscheinlichkeit  $P(A)$ . Der für Andreas zu erwartende Gewinn ergibt sich demzufolge aus der Gleichung

$$0,1\bar{6} = (-4) \cdot P(X=-4) + (-2) \cdot P(X=-2) + (-1) \cdot P(X=-1) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3) + 4 \cdot P(X=4).$$

Als Verallgemeinerung des oben dargestellten Berechnungswegs auf beliebige (endliche) Zufallsgrößen  $X$  lässt sich feststellen: Multipliziert man alle möglichen Werte  $x_i$  der Zufallsgröße  $X$  jeweils mit der Wahrscheinlichkeit  $P(X = x_i)$  ihres Eintretens (d.h. mit unserer *Erwartung* ihres Eintretens) und addiert diese Produkte, so erhält man den *Erwartungswert* der Zufallsgröße  $X$  (in obigem Beispiel den Erwartungswert des Gewinns pro Spielrunde). Ist dieser Erwartungswert für den Gewinn eines Spiels 0, so handelt es sich um ein **faies Spiel**.

## H 15

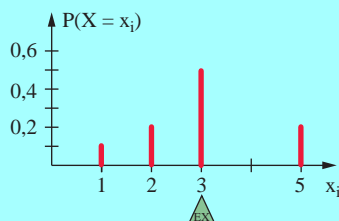
## Definition H 15:

$X$  sei eine endliche Zufallsgröße, die genau die Werte  $x_i (i \in \{1; 2; \dots; n\})$  annehmen kann, und zwar jeweils mit der Wahrscheinlichkeit  $P(X = x_i)$ . Dann nennt man die Kenngröße  $EX = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \dots + x_n \cdot P(X = x_n)$  den **Erwartungswert** der endlichen Zufallsgröße  $X$ . Für  $EX$  schreibt man auch  $E(X)$  bzw.  $\mu(X)$ ,  $\mu_X$  oder  $\mu^1$ .

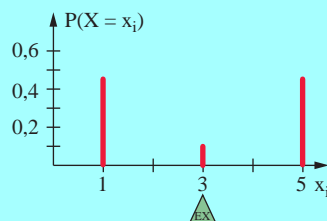
## H 37

## Beispiel H 37:

a)  $X \cong \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0,10 & 0,20 & 0,50 & 0,20 \end{pmatrix}$   
 $EX = 1 \cdot 0,10 + 2 \cdot 0,20 + 3 \cdot 0,50 + 5 \cdot 0,20$   
 $EX = 3,0$



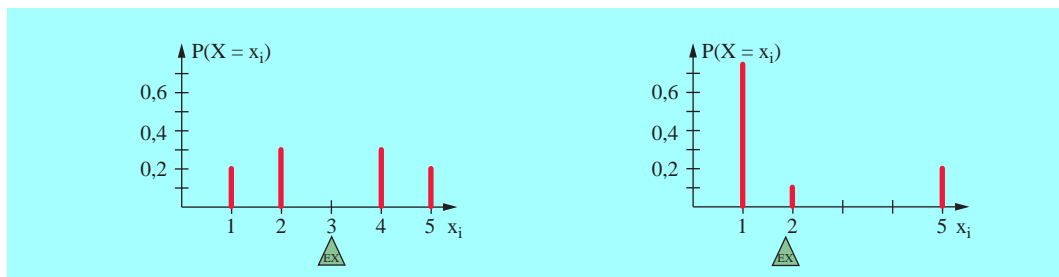
b)  $X \cong \begin{pmatrix} 1 & 3 & 5 \\ 0,45 & 0,10 & 0,45 \end{pmatrix}$   
 $EX = 1 \cdot 0,45 + 3 \cdot 0,10 + 5 \cdot 0,45$   
 $EX = 3,0$



c)  $X \cong \begin{pmatrix} 1 & 2 & 4 & 5 \\ 0,2 & 0,3 & 0,3 & 0,2 \end{pmatrix}$   
 $EX = 1 \cdot 0,2 + 2 \cdot 0,3 + 4 \cdot 0,3 + 5 \cdot 0,2$   
 $EX = 3,0$

d)  $X \cong \begin{pmatrix} 1 & 2 & 5 \\ 0,70 & 0,10 & 0,20 \end{pmatrix}$   
 $EX = 1 \cdot 0,70 + 2 \cdot 0,10 + 5 \cdot 0,20$   
 $EX = 1,9$

<sup>1)</sup> Für die Bezeichnung des Erwartungswerts wird der griechische Buchstabe  $\mu$  (gesprochen *mü*) in Anlehnung an den Begriff „Mittelwert“ verwendet.



Beispiel H 37 zeigt: Der Erwartungswert EX kann

- ein Wert der Zufallsgröße  $X$  sein (Beispiel H 37 a, b), muss es aber nicht (Beispiel H 37 c, d),
- in der Nähe des wahrscheinlichsten Wertes von  $X$  liegen (Beispiel H 37 a), muss es aber nicht (Beispiel H 37 b, d),
- in der Mitte aller möglichen Werte von  $X$  liegen (Beispiel. H 37 a, b, c), muss es aber nicht (Beispiel H 37 d),
- der Bedingung  $P(X \leq EX) = P(X \geq EX) = \frac{1}{2}$  genügen (Beispiel H 37 c), muss es aber nicht (Beispiel H 37 a, b, d).

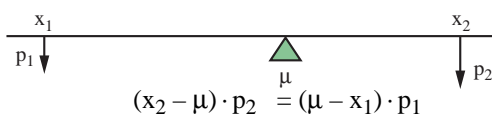
Der Erwartungswert EX kann auch als ein *Schwerpunkt* – im Sinne der Mechanik – interpretiert werden. Dazu fasst man die Wahrscheinlichkeitsverteilung von  $X \cong \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$  als Verteilung der Masse 1 auf die Punkte  $x_i$  eines masselosen Waagebalkens auf, wobei jedem Punkt  $x_i$  die Masse  $p_i = P(X = x_i)$  zugeordnet wird.

Es sei  $\mu$  der Schwerpunkt dieser Masseverteilung. Dann muss (in Verallgemeinerung des Hebelgesetzes) die Summe aller Drehmomente bezüglich  $\mu$  gleich null sein:

$$\begin{aligned} \sum_{i=1}^n (x_i - \mu) \cdot p_i &= 0 \\ \Rightarrow \sum_{i=1}^n (x_i \cdot p_i - \mu \cdot p_i) &= 0 \\ \Rightarrow \sum_{i=1}^n x_i \cdot p_i - \mu \cdot \sum_{i=1}^n p_i &= 0 \\ \Rightarrow EX - \mu \cdot 1 &= 0 \\ \Rightarrow \mu &= EX \end{aligned}$$

Spezialfall  $n = 2$

Kraft mal Kraftarm = Last mal Lastarm



$$(x_2 - \mu) \cdot p_2 = (\mu - x_1) \cdot p_1$$

$$x_2 p_2 - \mu p_2 = \mu p_1 - x_1 p_1$$

$$x_1 p_1 + x_2 p_2 = \mu p_1 + \mu p_2, \text{ also}$$

$$EX = \mu (p_1 + p_2) = \mu \cdot 1 = \mu$$

Der Waagebalken befindet sich genau dann im Gleichgewicht, wenn er im Punkt EX unterstützt wird. Der Erwartungswert EX ist also als Schwerpunkt einer Masseverteilung interpretierbar.

Beispiel H 38:

Beim einmaligen Werfen eines L-Würfels sei  $X$  die zufällige Augenzahl. Gesucht ist der Erwartungswert EX.

Lösung:

$$X \cong \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix};$$

$$EX = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3,5$$

Fig. H 35

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clear	a-z...
seq(i, i, 1, 6) → list1					
{1 2 3 4 5 6}					
seq(1/6, i, 1, 6) → list2					
{1/6 1/6 1/6 1/6 1/6 1/6}					
sum(seq(list1[i]·list2[i], i, 1, 6))					
					3.5
sum(list1·list2)					
					3.5
sum(list1*list2)					
3.5					
MAIN	RAD AUTO	FUNC 4/30			



## H 39

## Beispiel H 39:

Beim einmaligen Werfen eines fairen Würfels werde jeweils das um 1 vergrößerte Dreifache der Augenzahl notiert. Welcher Wert ist dann zu erwarten?

Lösung:

$$Y \cong \begin{pmatrix} 4 & 7 & 10 & 13 & 16 & 19 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

$$EY = 4 \cdot \frac{1}{6} + 7 \cdot \frac{1}{6} + 10 \cdot \frac{1}{6} + 13 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 19 \cdot \frac{1}{6}$$

$$EX = 11,5$$

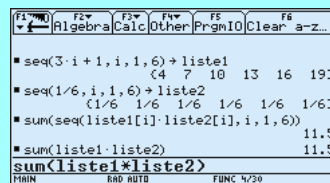


Fig. H 36

Vergleicht man bei den letzten beiden Beispielen die Zufallsgrößen  $X$  und  $Y$  sowie deren Erwartungswerte  $EX$  und  $EY$  jeweils miteinander, so fällt auf: Einerseits erhält man die möglichen Werte  $y_i$  von  $Y$  aus den ihnen entsprechenden Werten  $x_i$  von  $X$  nach der Vorschrift  $y_i = 3 \cdot x_i + 1$  und sie werden mit der gleichen Wahrscheinlichkeit  $\frac{1}{6}$  angenommen wie die zugehörigen  $x_i$ , weshalb man auch  $Y = 3 \cdot X + 1$  schreibt. Andererseits ergibt sich für den Erwartungswert von  $Y$  der Wert  $EY = 11,5 = 3 \cdot 3,5 + 1 = 3 \cdot EX + 1$ . Dieser Vergleich lässt die Gültigkeit folgenden Satzes vermuten.

## H 16

## Satz H 16:

Ist  $X \cong \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ P(X=x_1) & P(X=x_2) & \dots & P(X=x_n) \end{pmatrix}$  eine endliche Zufallsgröße, so gilt für den

Erwartungswert der Zufallsgröße  $aX + b \cong \begin{pmatrix} ax_1 + b & ax_2 + b & \dots & ax_n + b \\ P(X=x_1) & P(X=x_2) & \dots & P(X=x_n) \end{pmatrix}$ :

$$E(a \cdot X + b) = a \cdot EX + b \quad (a, b \in \mathbb{R}).$$

Beweis:

$$\begin{aligned} E(aX + b) &= \sum_{i=1}^n (ax_i + b) \cdot P(aX + b = ax_i + b) = \sum_{i=1}^n (ax_i + b) \cdot P(X = x_i) \\ &= a \cdot \sum_{i=1}^n x_i \cdot P(X = x_i) + b \cdot \sum_{i=1}^n P(X = x_i) = a \cdot EX + b \cdot 1 = a \cdot EX + b \quad \text{w.z. b. w.} \end{aligned}$$

Ist  $X$  eine endliche Zufallsgröße, so gilt also für lineare Funktionen  $f$ :  $E f(X) = f(EX)$ . Diese Aussage lässt sich aber nicht auf alle reellen Funktionen verallgemeinern, wie das nachfolgende Gegenbeispiel für die Betragsfunktion belegt:

$$X \cong \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

$$EX = (-1) \cdot \frac{1}{4} + 0 + 1 \cdot \frac{1}{4} = 0$$

$$|X| \cong \begin{pmatrix} 1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

$$E|X| = 1 \cdot \frac{1}{4} + 0 + 1 \cdot \frac{1}{4} = \frac{1}{2} \neq |EX|$$

## H 40

## Beispiel H 40:

Zufallsexperiment: Zweimaliges Werfen eines idealen Würfels mit den Augenzahlen 1, 2, 3, 4, 5, 6

Zufallsgröße  $Z$ : Augensumme (Summe der zwei gewürfelten Augenzahlen)

Gesucht: Erwartungswert  $EZ$

**Lösung:**

Z kann als Augensumme die elf Werte 2, 3, ..., 12 annehmen. Um deren Wahrscheinlichkeitsverteilung zu bestimmen, überlegen wir uns, welchen Ergebnissen welcher Wert von Z zugeordnet wird.

- 2 (1; 1)
- 3 (1; 2), (2; 1)
- 4 (1; 3), (2; 2), (3; 1)
- 5 (1; 4), (2; 3), (3; 2), (4; 1)
- 6 (1; 5), (2; 4), (3; 3), (4; 2), (5; 1)
- 7 (1; 6), (2; 5), (3; 4), (4; 3), (5; 2), (6; 1)
- 8 (2; 6), (3; 5), (4; 4), (5; 3), (6; 2)
- 9 (3; 6), (4; 5), (5; 4), (6; 3)
- 10 (4; 6), (5; 5), (6; 4)
- 11 (5; 6), (6; 5)
- 12 (6; 6)

Daraus ergibt sich:

$$\begin{aligned} EZ &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} \\ &\quad + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} \\ &\quad + 12 \cdot \frac{1}{36}, \text{ also } EZ = 7. \end{aligned}$$

Die zu erwartende mittlere Augensumme kann durch Simulation des Zufallsexperiments mithilfe des GTA ermittelt werden. Im Programm-Editor wird dazu die Funktion **erw** definiert. Mittels **rand(6) + rand(6)** und dem Listenverlängerungsbefehl **augment** ist eine Liste von n Augensummen zu erzeugen. Das Berechnen des Mittelwertes dieser n Listenwerte wird durch den **mean**-Befehl erreicht.

```

:erw(n,x)
:Func
:Local i,augs,liste
:RandSeed x
:O→liste
:For i,1,n
:rand(6)+rand(6)→augs
:augment(liste,(augs))→liste
:EndFor
:approx(mean(liste))
:EndFunc

```

Fig. H 37

Im Home-Editor wird die definierte Funktion **erw(n, x)** aufgerufen. Für n (Anzahl der Realisierungen des Zufallsexperiments) wählen wir hier 200. Die zur Initialisierung des Zufallsgenerators notwendige Zahl x wird beliebig gewählt, aber jeweils variiert.

erw(200, x)	Result
erw(200, 38512)	7.04
erw(200, 4791)	7.05
erw(200, 6844)	7.28
erw(200, 777)	7.345
erw(200, 9544)	7.075
erw(200, 14522)	6.755
erw(200, 32578)	7.005

Fig. H 38

Vergleicht man nun die Zufallsgrößen X (aus Beispiel H 38) und Z (aus dem letzten Beispiel H 40) sowie deren Erwartungswerte EX und EZ jeweils miteinander, so lässt sich erkennen:

Die möglichen Werte von Z erhält man einerseits als die verschiedenen Werte der Summen  $x_i + x_k$  für  $i, k \in \{1; 2; \dots; 6\}$ , wofür man auch  $Z = X + X$  schreibt, was aber nicht zu  $2X$  zusammengefasst werden darf.<sup>1)</sup>

Andererseits ergibt sich für den Erwartungswert von Z der Wert  $EZ = 7 = 3,5 + 3,5 = EX + EX$ . Dieser Vergleich lässt die Aussage des folgenden Satzes H 17 vermuten.

**Satz H 17:**

Für beliebige endliche Zufallsgrößen X und Y gilt  $E(X + Y) = EX + EY$ .

**H 17**

Der Beweis dieses Satzes kann analog zu dem für Satz H 16 erfolgen.

<sup>1)</sup> Für Zufallsgrößen X gilt also *nicht* das von den Zahlenbereichen und den Vektorräumen her bekannte Gesetz  $X + X = 2X$ .

## H 18

Satz H 18:

Für voneinander stochastisch *unabhängige* endliche Zufallsgrößen  $X$  und  $Y$  gilt die Gleichung  $E(X \cdot Y) = EX \cdot EY$ .

*Beweis:*

Die Zufallsgröße  $X$  nehme die Werte  $x_i$  mit den Wahrscheinlichkeiten  $P(X = x_i)$  für  $i \in \{1; 2; \dots; m\}$  und die Zufallsgröße  $Y$  die Werte  $y_k$  mit  $P(Y = y_k)$  für  $k \in \{1; 2; \dots; n\}$  an. Dann gilt:

$$\begin{aligned} E(X \cdot Y) &= \sum_{i=1}^m \sum_{k=1}^n x_i \cdot y_k \cdot P(X = x_i \text{ und } Y = y_k) \\ &= \sum_{i=1}^m \sum_{k=1}^n x_i \cdot y_k \cdot P(X = x_i) \cdot P(Y = y_k) \text{ wegen der Unabhängigkeit von } X \text{ und } Y \\ &= \sum_{i=1}^m x_i \cdot P(X = x_i) \cdot \sum_{k=1}^n y_k \cdot P(Y = y_k) = EX \cdot EY \quad \text{w. z. b. w.} \end{aligned}$$

**H 4.3 Streuung**

Zufallsgrößen können sich trotz desselben Erwartungswertes in ihren Wahrscheinlichkeitsverteilungen wesentlich voneinander unterscheiden:

$$X \cong \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$Y \cong \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

$$Z \cong \begin{pmatrix} 1 & 2 & 3 \\ \frac{9}{20} & \frac{1}{10} & \frac{9}{20} \end{pmatrix}$$

$$EX = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} = 2$$

$$EY = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} = 2$$

$$EZ = 1 \cdot \frac{9}{20} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{9}{20} = 2$$

Dieser beträchtliche Informationsverlust ist aber auch zu erwarten, da bei der Berechnung eines Erwartungswertes  $EX$  sowohl alle Werte  $x_i$  von  $X$  als auch alle zugehörigen Wahrscheinlichkeiten  $P(X = x_i)$  auf diese eine Kenngröße  $EX$  verdichtet werden. Zur näheren Kennzeichnung einer Wahrscheinlichkeitsverteilung wäre eine Zahl geeignet, die angibt, wie stark die einzelnen Werte  $x_i$  der Zufallsgröße  $X$  von ihrem Erwartungswert  $EX$  abweichen, wie weit sie also **streuen**, wie stark sie **variieren**.

Eine derartige Kenngröße für die Streuung einer Zufallsgröße  $X$  müsste die Abweichung der Werte  $X$  vom Erwartungswert  $EX$  charakterisieren. Es erscheint daher nahe liegend, dem gesuchten Streuungsmaß die Differenz  $X - EX$  zugrunde zu legen. Wir überlegen uns dazu:

$$X \cong \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

 $X - EX$ 

ist die zufällige Abweichung der Zufallsgröße  $X$  von ihrem Erwartungswert  $EX$  und damit eine Zufallsgröße, weshalb sie als Kenngröße ungeeignet ist.

 $E(X - EX)$ 

ist als zu erwartende Abweichung der Werte  $x_i$  von  $EX$  eine Zahl, die aber stets null ist.

Beispiel:

$$X \cong \begin{pmatrix} 1 & 2 & 3 \\ 0,60 & 0,10 & 0,30 \end{pmatrix}$$

$$EX = 1 \cdot 0,60 + 2 \cdot 0,10 + 3 \cdot 0,30 = 1,7$$

$$X - EX \cong \begin{pmatrix} -0,7 & 0,3 & 1,3 \\ 0,60 & 0,10 & 0,30 \end{pmatrix}$$

$$E(X - EX) = (-0,7) \cdot 0,60 + 0,3 \cdot 0,10 + 1,3 \cdot 0,30 = 0$$

$$\begin{aligned}
 E(X - EX) &= (x_1 - EX) p_1 + (x_2 - EX) p_2 + \dots + (x_n - EX) p_n \\
 &= x_1 p_1 - EX p_1 + x_2 p_2 - EX p_2 + \dots + x_n p_n - EX p_n \\
 &= x_1 p_1 + x_2 p_2 + \dots + x_n p_n - EX \cdot (p_1 + p_2 + \dots + p_n) \\
 &= EX - EX \cdot 1 = 0
 \end{aligned}$$

Dieses Zu-null-Ergänzen der positiven und der negativen Werte  $x_i - EX$  würde vermieden, wenn man jeweils statt der Abweichungen  $x_i - EX$  den Abstand  $|x_i - EX|$  wählt.

$E|X - EX| = |x_1 - EX| \cdot p_1 + |x_2 - EX| \cdot p_2 + \dots + |x_n - EX| \cdot p_n$  wäre ein sinnvolles Maß der Streuung, da durch das Setzen der Betragsstriche die Summe ungleich null wird. Diese Kenngröße wäre jedoch (erfahrungsgemäß bei allgemeinen Betrachtungen) wegen der Beträge nicht sehr praktikabel. Der gleiche gewünschte Effekt einer positiven Summe wird erreicht, wenn man an Stelle des Betragbildens jeweils quadriert.

$E(X - EX)^2 = (x_1 - EX)^2 \cdot p_1 + (x_2 - EX)^2 \cdot p_2 + \dots + (x_n - EX)^2 \cdot p_n$  scheint ein sinnvolles und praktikables Maß der Streuung zu sein, obwohl durch das Quadrieren die mehr als 1 von EX abweichenden Werte  $x_i$  stärker gewichtet werden als die Abweichungen kleiner als 1.

$$\begin{aligned}
 E|X - EX| &= 0,7 \cdot 0,60 + 0,3 \cdot 0,10 + 1,3 \cdot 0,30 \\
 &= 0,84
 \end{aligned}$$

$$\begin{aligned}
 E(X - EX)^2 &= (-0,7)^2 \cdot 0,60 + 0,3^2 \cdot 0,10 + 1,3^2 \cdot 0,30 \\
 &= 0,49 \cdot 0,60 + 0,09 \cdot 0,10 + 1,69 \cdot 0,30 \\
 &= 0,81
 \end{aligned}$$

Das gebräuchlichste Maß der Streuung bzw. der Varianz ist heute diese mittlere quadratische Abweichung, d.h. der Erwartungswert der quadratischen Abweichung der Zufallsgröße X von ihrem Erwartungswert EX. Die Maßeinheit dieser Varianz stimmt aber infolge des Quadrierens nicht mit der der Zufallsgröße überein. Deshalb führt man noch einen weiteren Begriff für die Quadratwurzel aus der Kenngröße Varianz ein.

#### Definition H 16:

Ist  $X \cong \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$  eine endliche Zufallsgröße mit dem Erwartungswert EX, so heißt  $E(X - EX)^2 = (x_1 - EX)^2 \cdot p_1 + (x_2 - EX)^2 \cdot p_2 + \dots + (x_n - EX)^2 \cdot p_n$  die **Streuung**<sup>1)</sup>  $D^2X$  oder auch **Varianz** Var X von X. Die Quadratwurzel aus der Streuung wird **Standardabweichung** genannt und mit DX bzw.  $\sqrt{\text{Var} X}$  oder auch mit  $\sigma(X)$ ,  $\sigma_X$ ,  $\sigma = \sqrt{\mu(X - \mu)^2}$  symbolisiert<sup>2)</sup>.

H 16

#### Satz H 19:

Für eine endliche Zufallsgröße X, die genau die Werte  $x_i$  mit  $i \in \{1; 2; \dots; n\}$  annehmen kann und die den Erwartungswert EX besitzt, gilt  $D^2X = E(X^2) - (EX)^2$ .

H 19

*Beweis:*

$$\begin{aligned}
 D^2X &= \sum_{i=1}^n (x_i - EX)^2 \cdot P(X = x_i) = \sum_{i=1}^n (x_i^2 - 2 \cdot x_i \cdot EX + (EX)^2) \cdot P(X = x_i) \\
 &= \sum_{i=1}^n x_i^2 \cdot P(X = x_i) - 2 \cdot EX \cdot \sum_{i=1}^n x_i \cdot P(X = x_i) + (EX)^2 \cdot \sum_{i=1}^n P(X = x_i) \\
 &= E(X^2) - 2 \cdot EX \cdot EX + (EX)^2 \cdot 1 = E(X^2) - (EX)^2 \quad \text{w. z. b. w.}
 \end{aligned}$$

<sup>1)</sup> Die für die Streuung gewählte Bezeichnung  $D^2X = E(X - EX)^2$  rührt von dem aus dem Lateinischen stammenden Wort *Dispersion* für *Streuung*.

<sup>2)</sup> Der griechische Buchstabe  $\sigma$  (gesprochen *sigma*) wird in Anlehnung an Standardabweichung gewählt.

Die zu Beginn dieses Abschnitts aufgeführten Zufallsgrößen  $X$ ,  $Y$  und  $Z$ , die zwar denselben Erwartungswert  $EX = EY = EZ = 2$ , aber verschiedene Wahrscheinlichkeitsverteilungen besitzen, weisen unterschiedliche Streuungen auf.

$$D^2X = E(X^2) - (EX)^2 = 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{3} + 3^2 \cdot \frac{1}{3} - (2)^2 = \frac{2}{3};$$

$$D^2Y = E(Y^2) - (EY)^2 = 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{2} + 3^2 \cdot \frac{1}{4} - (2)^2 = \frac{1}{2};$$

$$D^2Z = E(Z^2) - (EZ)^2 = 1^2 \cdot \frac{9}{20} + 2^2 \cdot \frac{1}{10} + 3^2 \cdot \frac{9}{20} - (2)^2 = \frac{9}{10}$$

## H 41

Beispiel H 41:

Lars Spielmann wird ein Glücksspiel in zwei Varianten angeboten, wobei folgende Spielregel gelten soll: Der Spieler wirft einmal mit zwei Würfeln, deren Seitenflächen jeweils mit den Augenzahlen 1 bis 6 durchnummeriert sind. Wirft der Spieler dabei einen Pasch ( $n; n$ ), so erhält er bei  $n \neq 6$  den  $n$ -fachen und für  $n = 6$  den fünfzehnfachen Einsatz zurück; wirft er die Augensumme 7, so bekommt er nur den Einsatz zurück. In allen übrigen Fällen verliert er den Einsatz an den Spielleiter.

Einsatz: Spielvariante 1:  $k \in (k > 0)$       Spielvariante 2:  $10 \cdot k \in (k > 0)$

Lars Spielmann will sich nur an einem fairen Spiel beteiligen. Um zu entscheiden, bei welchem  $k$ -Wert er welche Spielvariante wählen sollte, berechnet er die Erwartungswerte der jeweiligen Nettogewinne unter der Modellannahme, dass mit zwei LAPLACE-Würfeln geworfen wird.

*Lösung:*

Ergebnismenge:  $\Omega = \{(w_1, w_2) \mid w_1, w_2 \in \{1; 2; 3; 4; 5; 6\}\}$

$$P(\{(1; 1)\}) = \frac{1}{36}; \quad P(\{(2; 2)\}) = \frac{1}{36}; \quad \dots; \quad P(\{(6; 6)\}) = \frac{1}{36}$$

$$P(\{(1; 6), (2; 5), (3; 4), (4; 3), (5; 2), (6; 1)\}) = \frac{6}{36}$$

$$P(\{\text{weder Pasch noch Augensumme 7}\}) = \frac{24}{36}$$

Zufallsgröße  $N_e$ : zufälliger Nettogewinn des Spielers bei einem Einsatz von  $e \in (e \in \{k; 10 \cdot k\} \text{ für } k > 0)$

$$P(N_e = -e) = P(\{\text{weder Pasch noch Augensumme 7}\}) = \frac{24}{36}$$

$$P(N_e = 0) = P(\{(1; 6), (2; 5), (3; 4), (4; 3), (5; 2), (6; 1)\}) + P(\{(1; 1)\}) = \frac{7}{36}$$

$$P(N_e = e) = P(\{(2; 2)\}) = P(N_e = 2e) = P(\{(3; 3)\}) = P(N_e = 3e) = P(\{(4; 4)\}) = \frac{1}{36}$$

$$P(N_e = 4e) = P(\{(5; 5)\}) = P(N_e = 14e) = P(\{(6; 6)\}) = \frac{1}{36}$$

$$\text{Daraus ergibt sich: } N_e \cong \begin{pmatrix} -e & 0 & e & 2e & 3e & 4e & 14e \\ \frac{24}{36} & \frac{7}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} \end{pmatrix} \Rightarrow N_{10 \cdot k} = 10 \cdot N_k$$

$$\text{Erwartungswert: } EN_e = (-e) \cdot \frac{24}{36} + 0 \cdot \frac{7}{36} + e \cdot \frac{1}{36} + 2e \cdot \frac{1}{36} + 3e \cdot \frac{1}{36} + 4e \cdot \frac{1}{36} + 14e \cdot \frac{1}{36} = 0$$

Beide Spielvarianten sind *fair* – der Erwartungswert  $EN_e = 0$  lässt auf Dauer, auf lange Sicht sowohl keinen Gewinn als auch keinen Verlust erwarten. Die Varianten unterscheiden sich jedoch trotzdem wesentlich voneinander, was deutlich wird, wenn man von einer nur geringen Anzahl von Spielrunden ausgeht. Beteiligt sich Lars etwa genau einmal an dem Spiel, so gilt z.B.  $P(\{\text{Lars verliert genau seinen Einsatz}\}) = \frac{24}{36} = 0,6$ . Das heißt aber: Mit der Wahrscheinlichkeit von ca. 0,7 verliert er im Fall  $k = 1$  bei Spielvariante 1 nur 1 €, bei Spielvariante 2

jedoch 10 €. Die Spielvariante 2 ist daher als „risikoreicher“ anzusehen, was in der Standardabweichung  $DN_e$  zum Ausdruck kommen müsste:

$$\begin{aligned} DN_e &= \sqrt{E(N_e^2) - (EN_e)^2} \quad \text{mit } EN_e = 0 \\ &= \sqrt{(-e)^2 \cdot \frac{24}{36} + (0)^2 \cdot \frac{7}{36} + (2e)^2 \cdot \frac{1}{36} + (3e)^2 \cdot \frac{1}{36} + (4e)^2 \cdot \frac{1}{36} + (14e)^2 \cdot \frac{1}{36}} \\ &= \sqrt{\frac{249}{36}e^2} \approx 2,63e \quad \text{mit } e > 0 \quad \Rightarrow \quad D^2N_k \approx 6,92k \quad D^2N_{10k} = D^2(10 \cdot N_k) \approx 692k \\ &\quad DN_k \approx 2,6k \quad DN_{10k} = D(10 \cdot N_k) \approx 26k \end{aligned}$$

Die erheblich höheren Werte für die Streuung bzw. Standardabweichung von  $N_{10k}$  im Vergleich zu denen von  $N_k$  weisen auf das höhere Risiko bei Spielvariante 2 hin. Dies zeichnet die Spielvariante jedoch als die spannendere aus.

Um die Streuung einer Zufallsgröße mit einem möglichst geringen Rechenaufwand zu bestimmen, ist die Anwendung der nachfolgend genannten Eigenschaften dieser Kenngröße nützlich.

Satz H 20:

Ist  $X$  eine endliche Zufallsgröße, so gilt für alle  $a, b \in \mathbb{R}$ :  $D^2(aX + b) = a^2 \cdot D^2X$ .

H 20

*Beweis:*

$$\begin{aligned} D^2(aX + b) &= E(aX + b)^2 - (E(aX + b))^2 && \text{nach Satz H 19} \\ &= E(a^2X^2 + 2abX + b^2) - (a \cdot EX + b)^2 \\ &= a^2 \cdot EX^2 + 2ab \cdot EX + b^2 - a^2 \cdot (EX)^2 - 2ab \cdot EX - b^2 && \text{nach Satz H 16} \\ &= a^2 \cdot EX^2 - a^2 \cdot (EX)^2 = a^2 \cdot (EX^2 - (EX)^2) = a^2 \cdot D^2X && \text{nach Satz H 19} \quad \text{w. z. b. w.} \end{aligned}$$

Aufgrund dieses Satzes kann man die Spielvariante 2 aus dem Beispiel H 41 ohne weitere Rechnung als die mit der 100-fachen Streuung bzw. der 10-fachen Standardabweichung im Vergleich zur Spielvariante 1 charakterisieren. Währenddessen gilt nach Satz H 16 für die Erwartungswerte

$$EN_{10k} = E(10 \cdot N_k) = 10 \cdot EN_k.$$

Satz H 21:

Für beliebige voneinander unabhängige endliche Zufallsgrößen  $X$  und  $Y$  gilt  $D^2(X + Y) = D^2X + D^2Y$ .

H 21

Auf den Beweis dieses Satzes wird hier verzichtet.

Erwartungswert und Streuung ermöglichen es, den bisher nur umschriebenen Begriff des *Stabilwerdens relativer Häufigkeiten* (vgl. Satz H 2) mathematisch exakt zu definieren. Wir überlegen dazu:

Ein zufälliger Vorgang werde  $n$ -mal unabhängig voneinander realisiert. Man beobachtet dabei jeweils, ob das Ereignis  $A$  eintritt oder nicht. Dies wird für die  $i$ -te Realisierung ( $i \in \{1; 2; \dots; n\}$ ) beschrieben durch die Zufallsgröße  $X_i \in \begin{pmatrix} 1 & 0 \\ P(A) & 1 - P(A) \end{pmatrix}$ . Die Zufallsgröße  $\sum_{i=1}^n X_i$  gibt damit an, wie oft das Ereignis  $A$  bei den  $n$  Realisierungen eingetreten ist – sie ist also die als Zufallsgröße aufgefasste absolute Häufigkeit  $H_n(A)$ . Somit kann man die relative Häufigkeit  $h_n(A)$  als die Zufallsgröße  $h_n(A) = \frac{1}{n} \cdot \sum_{i=1}^n X_i$  auffassen und deren Erwartungswert berechnen:

$$E(h_n(A)) = E\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot \sum_{i=1}^n EX_i = \frac{1}{n} \cdot n \cdot EX_1 \quad (\text{nach Satz H 16 und H 17})$$

$$= EX_1 = 1 \cdot P(A) + 0 \cdot (1 - P(A)), \text{ also } E(h_n(A)) = P(A).$$

Des Weiteren gilt  $\lim_{n \rightarrow \infty} D^2(h_n(A)) = 0$ , denn

$$D^2(h_n(A)) = D^2\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot n \cdot D^2 X_1 \quad (\text{nach Satz H 20 und H 21, da } X_1, \dots, X_n \text{ unabhängig sind})$$

$$= \frac{1}{n} \cdot (EX_1^2 - (EX_1)^2) = \frac{1}{n} \cdot (P(A) - (P(A))^2) \quad \text{und} \quad \lim_{n \rightarrow \infty} \left[\frac{1}{n} \cdot (P(A) - (P(A))^2)\right] = 0.$$

Für große  $n$  streuen die relativen Häufigkeiten  $h_n(A)$  also nicht mehr um die Wahrscheinlichkeit  $P(A)$  – man kann sie als eine gute Näherung für  $P(A)$  verwenden. Damit ist zugleich der nachfolgende Satz bewiesen, durch den das *Empirische Gesetz der großen Zahlen* (Satz H 2) eine theoretische (auf dem KOLMOGOROWSchen Axiomensystem basierende) Interpretation und Rechtfertigung erhält:

H 22

**Satz H 22: Zusammenhang zwischen relativer Häufigkeit und Wahrscheinlichkeit**

Wird ein zufälliger Vorgang  $n$ -mal unabhängig voneinander wiederholt und das Eintreten des Ereignisses  $A$  beobachtet, so gilt für die als Zufallsgröße aufgefasste relative Häufigkeit

$$h_n(A) = \frac{1}{n} \cdot \sum_{i=1}^n X_i \quad \text{mit} \quad X_i \in \begin{pmatrix} 1 & 0 \\ P(A) & 1 - P(A) \end{pmatrix} \quad \text{die Beziehung}$$

$$E(h_n(A)) = P(A) \quad \text{sowie} \quad \lim_{n \rightarrow \infty} D^2(h_n(A)) = 0.$$

Der Erwartungswert  $EX$  ermöglicht somit eine Prognose für das arithmetische Mittel von vielen Beobachtungswerten der Zufallsgröße  $X$ .

Die Zweckmäßigkeit von  $D^2 X$  als ein „Streuungsmaß“ einer Zufallsgröße  $X$  zeigt auch der nachfolgende Satz.

H 23

**Satz H 23: TSCHEBYSCHESCHE<sup>1)</sup> Ungleichung**

Es sei  $X$  eine endliche Zufallsgröße mit dem Erwartungswert  $EX$  und der Streuung  $D^2 X$ . Dann beträgt die Wahrscheinlichkeit dafür, dass  $X$  einen Wert annimmt, der um mindestens  $\alpha$  ( $\alpha > 0$ ) von  $EX$  abweicht, höchstens  $\frac{D^2 X}{\alpha^2}$ . Für jeden positiven Wert von  $\alpha$  gilt also die Ungleichung

$$P(|X - EX| \geq \alpha) \leq \frac{1}{\alpha^2} \cdot D^2 X.$$

*Beweis:*

$$\begin{aligned} P(|X - EX| \geq \alpha) &= P((X - EX)^2 \geq \alpha^2) \\ &= \sum_{i: (x_i - EX)^2 \geq \alpha^2} P(X = x_i) = \sum_{i: \frac{(x_i - EX)^2}{\alpha^2} \geq 1} P(X = x_i) \\ &\leq \sum_{i: \frac{(x_i - EX)^2}{\alpha^2} \geq 1} \frac{(x_i - EX)^2}{\alpha^2} \cdot P(X = x_i) \\ &\leq \sum_{i=1}^n \frac{(x_i - EX)^2}{\alpha^2} \cdot P(X = x_i) = \frac{1}{\alpha^2} \cdot \sum_{i=1}^n (x_i - EX)^2 \cdot P(X = x_i) \\ &= \frac{1}{\alpha^2} \cdot D^2 X \quad \text{nach Definition H 16} \quad \text{w. z. b. w.} \end{aligned}$$

<sup>1)</sup> P. L. TSCHEBYSCHEW (1821–1894); russischer Mathematiker

Mithilfe dieser TSCHEBYSCHESchen Ungleichung ist man in der Lage, diejenige Wahrscheinlichkeit abzuschätzen, mit der eine Zufallsgröße  $X$  einen Wert annimmt, der um mehr als eine fest vorgegebene Zahl  $\alpha$  vom Erwartungswert  $EX$  abweicht; diese Wahrscheinlichkeit ist umso kleiner, je kleiner die Streuung  $D^2X$  ist.

#### Beispiel H 42:

Stephanie, die gerade 18 Jahre alt geworden ist, entnimmt einer kurzen Zeitungsnotiz, dass das Lebensalter, welches eine 18-Jährige erreicht, eine Zufallsgröße mit dem Erwartungswert von 75 und einer Standardabweichung von 5 Jahren ist.

Stephanie möchte daraufhin die Wahrscheinlichkeit abschätzen, dass sie ein Alter

- a) von mehr als 70 und weniger als 80,                      b) von mehr als 65 und weniger als 85,
- c) von mehr als 60 und weniger als 90 Jahren erreicht.

*Lösung* (mittels der TSCHEBYSCHESchen Ungleichung):

Modellannahme: Stephanie ist eine „auf gut Glück“ ausgewählte 18-Jährige.

Die Wahrscheinlichkeiten sollen abgeschätzt werden, obwohl von der Zufallsgröße  $X$  nur die beiden Kenngrößen  $EX = 75$  und  $DX = 5$  bekannt sind.

$$\begin{aligned} P(70 < X < 80) &= P(-5 < X - 75 < 5) \\ &= P(|X - EX| < 5) = 1 - P(|X - EX| \geq 5) \\ &\geq 1 - \frac{1}{5^2} \cdot D^2X \geq 1 - \frac{1}{25} \cdot 25 \end{aligned}$$

$$P(70 < X < 80) \geq 0$$

Die TSCHEBYSCHESche Ungleichung ist nur eine relativ grobe Abschätzung und liefert daher in diesem Falle keine spezifische Information zum Eintreten des untersuchten Ereignisses.

$$\begin{aligned} P(65 < X < 85) &= P(-10 < X - 75 < 10) \\ &= P(|X - EX| < 10) = 1 - P(|X - EX| \geq 10) \\ &\geq 1 - \frac{1}{10^2} \cdot D^2X = 1 - \frac{1}{100} \cdot 25 \end{aligned}$$

$$P(65 < X < 85) \geq 0,75$$

$$\begin{aligned} P(60 < X < 90) &= P(-15 < X - 75 < 15) \\ &= P(|X - EX| < 15) = 1 - P(|X - EX| \geq 15) \\ &\geq 1 - \frac{1}{15^2} \cdot D^2X = 1 - \frac{1}{225} \cdot 25 \end{aligned}$$

$$P(60 < X < 90) \geq 0,8\bar{8}$$

Wählt man in der TSCHEBYSCHESchen Ungleichung für den Parameter  $\alpha$  Vielfache der Standardabweichung  $\sigma$ , setzt man also  $\alpha = n \cdot \sigma$ , so erhält man  $P(|X - EX| \geq n \cdot \sigma) \leq \frac{1}{n^2}$ . Die Wahrscheinlichkeit, dass  $X$  einen Wert annimmt, der von  $EX$  um mindestens das  $n$ -fache der Standardabweichung  $\sigma$  abweicht, ist folglich höchstens  $\frac{1}{n^2}$ . Für die Spezialfälle  $n = 1; 2; 3$  ergibt sich demnach:

$$P(|X - EX| \geq \sigma) \leq 1, \quad P(|X - EX| \geq 2\sigma) \leq 0,25, \quad P(|X - EX| \geq 3\sigma) \leq 0,1\bar{1}$$

Diese aus der TSCHEBYSCHESchen Ungleichung gewonnenen Aussagen werden als  $\sigma$ -Regeln oder  $3\sigma$ -Regel bezeichnet.

#### Satz H 24: **3 $\sigma$ -Regel**

Die Wahrscheinlichkeit dafür, dass eine endliche Zufallsgröße  $X$  mit dem Erwartungswert  $EX = \mu$  und der Streuung  $D^2X = \sigma^2$  Werte

- im  $2\sigma$ -Intervall  $(\mu - 2\sigma; \mu + 2\sigma)$  annimmt, beträgt mindestens 0,75,
- im  $3\sigma$ -Intervall  $(\mu - 3\sigma; \mu + 3\sigma)$  annimmt, beträgt mindestens 0,88.

H 42

H 24



H 43

Beispiel H 43:

Es ist eine Aussage über den Wert der Wahrscheinlichkeit  $P(|X - EX| \geq 2 \cdot DX)$  für die Zufallsgröße  $X \cong \begin{pmatrix} -2 & 0 & 3 \\ 0,125 & 0,750 & 0,125 \end{pmatrix}$  zu treffen.

$$EX = (-2) \cdot 0,125 + 0 \cdot 0,750 + 3 \cdot 0,125 = 0,125$$

$$D^2X = EX^2 - (EX)^2 = 4 \cdot 0,125 + 0 \cdot 0,75 + 9 \cdot 0,125 - 0,125^2 = 1,609375$$

*Lösung 1* (exakte Berechnung):

$$\begin{aligned} P(|X - 0,125| \geq 2 \cdot \sqrt{1,609375}) &= 1 - P(|X - 0,125| < 2 \cdot \sqrt{1,609375}) \\ &= 1 - P(-2 \cdot \sqrt{1,609375} + 0,125 < X < 2 \cdot \sqrt{1,609375} + 0,125) \\ &= 1 - P(X = -2) - P(X = 0) = 1 - 0,125 - 0,750 = 0,125 \end{aligned}$$

*Lösung 2* (Abschätzung nach TSCHEBYSCHEW):

$$P(|X - EX| \geq 2 \cdot DX) \leq 0,25 \quad \text{nach Satz H 24}$$

H 44

Beispiel H 44:

Nunmehr sind alle Voraussetzungen geschaffen, auf die Probleme im einleitend (S. 351) wiedergegebenen Disput zwischen Anton, Katrin und Jonas über einen Erfolg versprechenden Tipp bei „6 aus 49“ einzugehen. Im Mittelpunkt des Gesprächs steht die Zahl 13, weil sie bei den bisherigen 2320 Ausspielungen des Samstags-Lotto am wenigsten, nämlich nur 230-mal gezogen wurde. Das ergibt die relative Häufigkeit  $h_{2320}(\{13\}) = \frac{230}{2320} \approx 0,099$ . Nach dem empirischen Gesetz der großen Zahlen stabilisieren sich die relativen Häufigkeiten  $h_n(\{i\})$  für das Auftreten einer bestimmten Zahl  $i$  ( $i \in \{1; \dots; 49\}$ ) mit wachsendem  $n$  gegen einen bestimmten Wert. Aber gegen welchen? Anton meint gegen  $\frac{1}{49} \approx 0,020$ . Jonas bezweifelt dies, denn dieser Wert muss die Wahrscheinlichkeit sein, dass diese Zahl  $i$  bei einer Ausspielung als Gewinnzahl auftritt.

(1) Wie groß ist die Wahrscheinlichkeit  $P(A)$ , dass eine bestimmte Zahl  $a$  bei *einer* Ausspielung als Gewinnzahl auftritt?

*Lösungsvariante 1:*

Urnenmodell: Sechs Ziehungen ohne Zurücklegen aus einer Urne mit einer goldenen und 48 schwarzen Kugeln; **hypergeometrische Verteilung** (Satz H 9)

$$P(A) = \frac{\binom{1}{1} \cdot \binom{48}{5}}{\binom{49}{6}} = \frac{\frac{48!}{5! \cdot 43!}}{\frac{49!}{6! \cdot 43!}} = \frac{48! \cdot 6! \cdot 43!}{5! \cdot 43! \cdot 49!} = \frac{6}{49} \approx 0,122$$

*Lösungsvariante 2:*

$A_i = \{\text{die Zahl } a \text{ wird bei einer Ausspielung als } i\text{-te Zahl gezogen}\} \quad (i \in \{1; \dots; 6\})$

$P(A) = P(A_1 \cup A_2 \cup \dots \cup A_6) = P(A_1) + P(A_2) + \dots + P(A_6)$ , da  $A_1, \dots, A_6$  unvereinbar sind

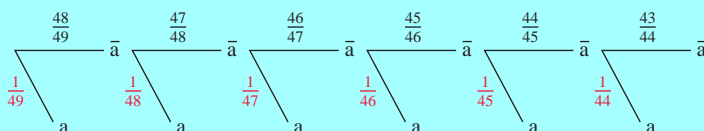


Fig. H 39

$$P(A_1) = \frac{1}{49} \quad (\text{jede Kugel hat die gleiche Chance gezogen zu werden})$$

$$P(A_2) = P(\bar{A}_1) \cdot P_{\bar{A}_1}(A_2) = \frac{48}{49} \cdot \frac{1}{48} = \frac{1}{49} \quad (\text{nach der ersten Pfadregel})$$

$$P(A_3) = P(\overline{A_1} \cap \overline{A_2}) \cdot P_{\overline{A_1} \cap \overline{A_2}}(A_3) = \frac{48}{49} \cdot \frac{47}{48} \cdot \frac{1}{47} = \frac{1}{49}$$

⋮

$$P(A_6) = P(\overline{A_1} \cap \dots \cap \overline{A_5}) \cdot P_{\overline{A_1} \cap \dots \cap \overline{A_5}}(A_6) = \frac{48}{49} \cdot \frac{47}{48} \cdot \frac{46}{47} \cdot \frac{45}{46} \cdot \frac{44}{45} \cdot \frac{1}{44} = \frac{1}{49}$$

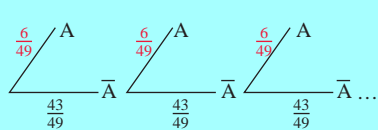
(Auch wenn die einzelnen Ziehungen einer Ausspielung *nicht* unabhängig voneinander sind, ist die Wahrscheinlichkeit, die Zahl  $a$  zu ziehen, bei jeder dieser sechs Ziehungen  $\frac{1}{49}$ .)

$$\Rightarrow P(A) = \frac{1}{49} \cdot 6 = \frac{6}{49}$$

Anton hat also übersehen, dass das zu betrachtende Zufallsexperiment die Ausspielung von *sechs* Zahlen ist und *nicht* die Ziehung einer *einzig*en Zahl aus den Zahlen 1 bis 49<sup>1)</sup>. Ihm ist aber zuzustimmen, wenn er meint, dass im Laufe der Zeit die bislang „benachteiligten“ Zahlen, insbesondere die 13, häufiger gezogen werden müssten. Aber dies gilt nur in der Tendenz, d.h. für eine sehr große Anzahl von Ausspielungen und es sagt überhaupt nichts aus über die kommende Ziehung. Eine Frage jedoch lässt sich noch etwas genauer beantworten:

(2) Wie viele Ausspielungen sind zu erwarten, bis eine bestimmte Zahl  $a$  ( $1 \leq a \leq 49$ ) gezogen wird?

Zufallsgröße  $X$ : zufällige Anzahl der Ausspielungen, die nötig sind, um die Zahl  $a$  zu ziehen (diskrete, nicht endliche Zufallsgröße)



$$P(X=1) = \frac{6}{49};$$

$$P(X=2) = \frac{43}{49} \cdot \frac{6}{49};$$

$$P(X=3) = \left(\frac{43}{49}\right)^2 \cdot \frac{6}{49};$$

$$\dots; P(X=k) = \left(\frac{43}{49}\right)^{k-1} \cdot \frac{6}{49}$$

Fig. H 41

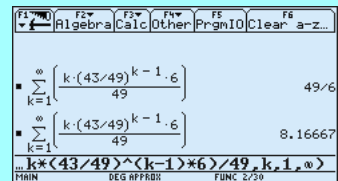


Fig. H 40

$$\text{Erwartungswert } EX = \sum_{k=1}^{\infty} k \cdot \left(\frac{43}{49}\right)^{k-1} \cdot \frac{6}{49} \approx 8,2$$

Der Abstand zwischen zwei Ausspielungen mit der Glückszahl 13 wird also auf *sehr lange Zeit* gesehen im Mittel etwa 8,2 betragen. Dieses Ergebnis darf aber zu keiner Interpretation der Art „nach ungefähr acht Ausspielungen tritt die 13 auf“ verleiten. Es erlaubt auch keine Aussage über die kommenden acht Ausspielungen.

Um zu einer handhabbaren Aussage für das Tippverhalten zu gelangen, schlägt Jonas vor, auf das Instrumentarium der bedingten Wahrscheinlichkeit zurückzugreifen.

(3) Wie groß ist die Wahrscheinlichkeit, dass bei der kommenden Ausspielung die Zahl 13 auftritt, unter der Bedingung, dass sie bei den vorangegangenen 2320 Ausspielungen 230-mal gezogen wurde?

$$P_{\{\text{unter den vorangegangenen 2320 Ausspielungen wurde die 13 gerade 230-mal gezogen}\}}(\{13\}) = P(\{13\}) = \frac{6}{49}$$

(wegen der Unabhängigkeit der Ausspielungen)

<sup>1)</sup> Die Lotto-Betreiber unternehmen große Anstrengungen, damit die Modellannahme, die einzelnen Ausspielungen werden immer unter gleichen Bedingungen durchgeführt, als gerechtfertigt anzusehen ist. So werden die Gewichte der Kugeln und der technische Zustand des Ziehungsgerätes vor jeder Ausspielung von einem Aufsichtsbeamten überprüft.

Die Kenntnis der vorangegangenen Ausspielung bringt *keinen* Erkenntnisgewinn für die kommende Ausspielung. Es ist in der Tat so, wie Katrin es zum Ausdruck bringt: Die Gewinnwahrscheinlichkeit ist bei jeder Ausspielung für jede Tippierei gleich – ob sie nun (1; 2; 3; 4; 5; 6), (1; 4; 9; 16; 25; 36), (7; 15; 16; 26; 38; 39) oder wie auch immer lauten mag.

Die Frage nach einem erfolgversprechenden Tippverhalten, wie sie Anton, Katrin und Jonas beschäftigt, zielt aber nicht nur auf die Gewinnwahrscheinlichkeit, sondern auch auf den Gewinnerwartungswert. Bekanntlich hängt die Höhe des Gewinns pro Tippierei für die einzelnen Gewinnklassen von der zufälligen Anzahl der Gewinne in dieser Gewinnklasse bei der jeweiligen Ausspielung ab. Diese Anzahl der Gewinne in einer Gewinnklasse wird aber auch davon beeinflusst, ob Gewinnzahlen ausgespielt werden, von denen gewisse Zahlen oder Zahlenkombinationen in der Regel öfter oder weniger oft getippt werden. Enthält ein Lotto-Tipp Zahlen, die erfahrungsgemäß weniger oft getippt werden, und gewinnt dieser Lotto-Tipp bei der kommenden Ausspielung, so dürfte eine höhere Gewinnausschüttung zu erwarten sein. Dies ist der einzige Tipp, den wir Anton, Katrin und Jonas hier geben können.

## H 5 Binomialverteilung

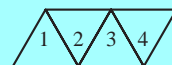
### H 5.1 BERNOULLI-Experimente

H 45

Beispiel H 45:

*Zufallsexperiment:*

Einmaliges Werfen eines L-Tetraeders mit dem nebenstehenden Netz



*Beobachtungsziel:*

*Zufallsgröße:*

*Interpretation:*

Welche Augenzahl wird gewürfelt?

$$Y \cong \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

Für  $i \in \{1; 2; 3; 4\}$

$Y = i$ : „Anzahl  $i$  gewürfelt“

Wird die 4 gewürfelt?

$$Z \cong \begin{pmatrix} 1 & 0 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$Z = 1$ : „Augenzahl 4 gewürfelt“

$Z = 0$ : „1, 2 oder 3 gewürfelt“

Wird eine Primzahl gewürfelt?

$$R \cong \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$R = 1$ : „2 oder 3 gewürfelt“

$R = 0$ : „1 oder 4 gewürfelt“

Wird ein Teiler von 4 gewürfelt?

$$T \cong \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

$T = 1$ : „1, 2 oder 4 gewürfelt“

$T = 0$ : „3 gewürfelt“

Aus obigem Beispiel H 45 sind neben *gleichverteilten* Zufallsgrößen wie  $Y$  und  $R$  solche Zufallsgrößen von besonderer Bedeutung, die wie  $Z$ ,  $R$  und  $T$  nur zwei Werte annehmen können, die also Zufallsexperimente mit nur zwei (interessierenden) Ergebnissen – genannt „**Erfolg**“ und „**Misserfolg**“ bzw. „**Treffer**“ und „**Niete**“ beschreiben. Sinnvollerweise ordnet man für eine solche Zufallsgröße  $X$  dem Ergebnis „Erfolg“ die Zahl 1 und dem Ergebnis „Misserfolg“ die Zahl 0 zu und bezeichnet  $P(X = 1)$  mit  $p$  als „**Erfolgswahrscheinlichkeit  $p$** “ oder „**Trefferwahrscheinlichkeit  $p$** “ sowie demzufolge  $P(X = 0)$  mit  $1 - p$  (oder häufig mit  $q$ ).

Definition H 17:

Eine Zufallsgröße  $X \cong \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$  heißt **BERNOULLI-Größe** und das zugehörige (einstufige) Zufallsexperiment **BERNOULLI-Experiment**<sup>1)</sup>.

H 17

Satz H 25: Erwartungswert einer BERNOULLI-Größe

Eine BERNOULLI-Größe  $X \cong \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$  besitzt den Erwartungswert  $EX = p$  und die Streuung  $D^2X = p \cdot (1-p)$ .

H 25

Beweis:

$$EX = 1 \cdot p + 0 \cdot (1-p) = p$$

nach Definition H 15

$$D^2X = E(X^2) - (EX)^2$$

nach Satz H 19

$$= 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 = p - p^2 = p \cdot (1-p)$$

w.z.b.w.

Beispiel H 46:

(1) Zufallsexperiment: Einmaliges Werfen eines Kronenverschlusses

Interpretation:

$$X \cong \begin{pmatrix} 1 & 0 \\ 0,72 & 0,28 \end{pmatrix}$$

1 – „fällt nach oben geöffnet“

$$EX = 0,72$$

0 – „fällt nach unten geöffnet“

$$D^2X = 0,72 \cdot 0,28 \approx 0,20$$

(2) Zufallsexperiment: Einmaliges Werfen eines (idealen) Tetraeders

Interpretation:

$$X \cong \begin{pmatrix} 1 & 0 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

1 – „Augenzahl 4“

$$EX = \frac{1}{4} = 0,25$$

0 – „Augenzahl ist nicht 4“

$$D^2X = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16} = 0,1875$$

H 46

## H 5.2 BERNOULLI-Ketten; binomialverteilte Zufallsgrößen

Im Zusammenhang mit BERNOULLI-Experimenten sind beispielsweise folgende Fragen (bezogen auf das Beispiel H 46 (2)) von besonderem Interesse:

Frage 1: Mit welcher Wahrscheinlichkeit fällt genau dreimal die Augenzahl 4, wenn das Tetraeder genau siebenmal geworfen wird?

Frage 2: Mit welcher Wahrscheinlichkeit fällt die Augenzahl 4 mindestens 25-mal, wenn das Tetraeder genau 100-mal geworfen wird?

Der Beantwortung dieser Fragen liegt als Zufallsexperiment ein sieben- bzw. 100-faches Realisieren ein und desselben BERNOULLI-Experiments zugrunde:

Definition H 18:

Eine n-fache und unabhängig voneinander ausgeführte Realisierung eines BERNOULLI-Experiments mit der Erfolgswahrscheinlichkeit p heißt **BERNOULLI-Kette der Länge n und mit der Erfolgswahrscheinlichkeit p** oder kurz **BERNOULLI-Kette mit den Parametern n und p**.

H 18

Eine solche BERNOULLI-Kette wird beschrieben durch eine Zufallsgröße X, welche die n + 1 Werte 0, 1, 2, ..., n annehmen kann.  $X = k$  wäre dann für  $k \in \{0; 1; 2; \dots; n\}$  zu interpretieren als das Ereignis

<sup>1)</sup> Diese Bezeichnung wurde zu Ehren des Mathematikers Jakob BERNOULLI (1654–1705) gewählt.

nis, dass in der BERNOULLI-Kette  $k$  „Erfolge“ und  $n - k$  „Misserfolge“ registriert werden.

Um eine Formel für die Wahrscheinlichkeit  $P(X = k)$  für  $k \in \{0; 1; 2; \dots; n\}$  herzuleiten, kehren wir zur obigen *Frage 1* zurück. Das dort zugrunde gelegte Zufallsexperiment ist eine BERNOULLI-Kette der Länge  $n = 7$  und mit der Erfolgswahrscheinlichkeit  $p = 0,25$ , d.h. ein siebenstufiges Zufallsexperiment mit den Verzweigungswahrscheinlichkeiten  $p = 0,25$  und  $1 - p = 0,75$ . Die Frage 1 soll nun mittels eines (wenn auch – wegen des hohen Aufwandes – nur in Gedanken gezeichneten) Baumdiagramms und der beiden Pfadregeln beantwortet werden.

Jedes Ergebnis, das zu dem hier interessierenden Ereignis  $A = \{\text{genau dreimal die Augenzahl 4 bei siebenmaligem Werfen}\}$  gehört, ist als 7-Tupel mit drei Einsen und vier Nullen darzustellen. So würde beispielsweise das 7-Tupel  $(0; 1; 0; 0; 0; 1; 1)$  bedeuten, dass beim zweiten, sechsten und siebenten Wurf „Erfolg“ eintrat, d.h. die Augenzahl 4 fiel, während bei den anderen vier Würfeln ein „Misserfolg“ zu verzeichnen war. Nach der *ersten Pfadregel* ergibt sich die Wahrscheinlichkeit für das Eintreten des Ergebnisses  $(0; 1; 0; 0; 0; 1; 1)$  durch Multiplikation der Wahrscheinlichkeiten längs des zu diesem Ergebnis führenden Pfades:

$$\frac{0,75}{0,75} \cdot \frac{0,25}{0,25} \cdot \frac{0,75}{0,75} \cdot \frac{0,75}{0,75} \cdot \frac{0,75}{0,75} \cdot \frac{0,25}{0,25} \cdot \frac{0,25}{0,25} = 0,25^3 \cdot (1 - 0,25)^4.$$

$$P(\{(0; 1; 0; 0; 0; 1; 1)\}) = 0,75 \cdot 0,25 \cdot 0,75 \cdot 0,75 \cdot 0,75 \cdot 0,25 \cdot 0,25 = 0,25^3 \cdot (1 - 0,25)^4.$$

Dieselbe Wahrscheinlichkeit ergibt sich für *alle möglichen* 7-Tupel mit drei Einsen und vier Nullen, d.h. für alle Ergebnisse aus  $A$ . Damit erhält man  $P(A)$  nach der *zweiten Pfadregel* als Summe dieser Wahrscheinlichkeiten, also durch Multiplikation von  $0,25^3 \cdot (1 - 0,25)^4$  mit der Anzahl  $|A|$  aller möglichen 7-Tupel mit drei Einsen und vier Nullen. Diese Anzahl ist nach dem Zählprinzip für Mengen  $\binom{7}{3} = \frac{7!}{3!4!}$ , d.h. gleich der Anzahl der Möglichkeiten, aus der 7-elementigen Menge  $\{1; 2; \dots; 7\}$  der Experimentnummern eine dreielementige Teilmenge zu bilden – und zwar jeweils bestehend gerade aus den Nummern derjenigen Experimente, bei denen „Erfolg“ erzielt worden ist.

Die Wahrscheinlichkeit, dass bei einer BERNOULLI-Kette der Länge  $n = 7$  und mit der Erfolgswahrscheinlichkeit  $p = 0,25$  genau dreimal „Erfolg“ eintritt, beträgt demzufolge:

$$P(X = 3) = P(A) = \binom{7}{3} \cdot 0,25^3 \cdot (1 - 0,25)^4 \approx 0,17.$$

Bezogen auf *Frage 1* bedeutet das: Beim *siebenmaligen „Würfeln“* fällt mit einer Wahrscheinlichkeit von ca. 0,17 *genau dreimal* die Augenzahl 4.

Als Verallgemeinerung erhalten wir den nachfolgenden Satz:

H 26

### Satz H 26: BERNOULLI-Formel

Bei einer BERNOULLI-Kette der Länge  $n$  und mit der Erfolgswahrscheinlichkeit  $p$  beträgt die Wahrscheinlichkeit

- für genau  $k$ -mal „Erfolg“  $B_{n;p}(\{k\}) = P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$  und
- für höchstens  $k$ -mal „Erfolg“

$$B_{n;p}(\{0; 1; \dots; k\}) = P(X \leq k) = P(X = 0) + P(X = 1) + \dots + P(X = k) \\ = B_{n;p}(\{0\}) + B_{n;p}(\{1\}) + \dots + B_{n;p}(\{k\})^{1)}$$

Mit dieser Formel kann man nun die zu einer BERNOULLI-Kette der Länge  $n$  und mit der Erfolgswahrscheinlichkeit  $p$  gehörende Zufallsgröße vollständig charakterisieren.

<sup>1)</sup> Anstelle von  $B_{n;p}(\{k\})$  bzw.  $B_{n;p}(\{0; 1; \dots; k\})$  werden auch z.B. (in Anlehnung an die Verteilungsfunktion als Stammfunktion ihrer Dichtefunktion bei stetigen Zufallsgrößen)  $f_B(k; n; p)$  bzw.  $F_B(k; n; p)$  oder die weniger genauen Schreibweisen  $b_{n;p}(k)$  bzw.  $B_{n;p}(k)$  verwendet.

## Beispiel H 47:

Einer Urne mit genau 11 Kugeln (5 weißen, 3 grünen und 3 roten) werden „auf gut Glück“ nacheinander und mit Zurücklegen vier Kugeln entnommen. Mit welcher Wahrscheinlichkeit werden dabei genau  $k$  ( $k \in \{0; 1; 2; 3; 4\}$ ) grüne Kugeln gezogen?

## Lösung:

Eine *einmalige* Ziehung aus der vorgegebenen Urne entspricht dem mathematischen Modell eines BERNOULLI-Experiments mit der Erfolgswahrscheinlichkeit  $p = \frac{3}{11}$ , das durch die Zufallsgröße

$$X \cong \begin{pmatrix} 1 & 0 \\ \frac{3}{11} & \frac{8}{11} \end{pmatrix} \text{ beschrieben wird.}$$

Eine *viermalige* Ziehung entspricht dem mathematischen Modell einer BERNOULLI-Kette der Länge  $n = 4$  und mit der Erfolgswahrscheinlichkeit  $p = \frac{3}{11}$ , das durch die Zufallsgröße  $Y$  beschrieben wird, die die zufällige Anzahl der „Erfolge“ angibt.

$$P(Y = 0) = B_4; \frac{3}{11}(\{0\}) = \binom{4}{0} \cdot \left(\frac{3}{11}\right)^0 \cdot \left(\frac{8}{11}\right)^4 \approx 0,280$$

$$P(Y = 1) = B_4; \frac{3}{11}(\{1\}) = \binom{4}{1} \cdot \left(\frac{3}{11}\right)^1 \cdot \left(\frac{8}{11}\right)^3 \approx 0,420$$

$$P(Y = 2) = B_4; \frac{3}{11}(\{2\}) = \binom{4}{2} \cdot \left(\frac{3}{11}\right)^2 \cdot \left(\frac{8}{11}\right)^2 \approx 0,236$$

$$P(Y = 3) = B_4; \frac{3}{11}(\{3\}) = \binom{4}{3} \cdot \left(\frac{3}{11}\right)^3 \cdot \left(\frac{8}{11}\right)^1 \approx 0,059$$

$$P(Y = 4) = B_4; \frac{3}{11}(\{4\}) = \binom{4}{4} \cdot \left(\frac{3}{11}\right)^4 \cdot \left(\frac{8}{11}\right)^0 \approx 0,006$$

## Punktdiagramm

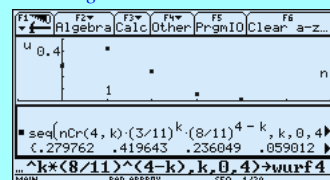


Fig. H 42

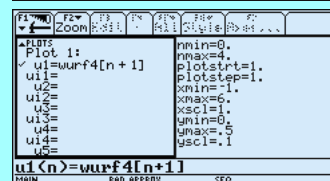


Fig. H 43

## Histogramm

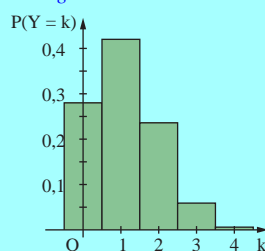


Fig. H 44

## Definition H 19:

Eine Zufallsgröße  $X$ , welche die Werte  $0; 1; 2; \dots; n$  mit den Wahrscheinlichkeiten

$$P(X = k) = B_{n;p}(\{k\}) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad \text{für } k \in \{0; 1; 2; \dots; n\}$$

annimmt, heißt **binomialverteilt mit den Parametern  $n$  und  $p$**  oder auch kurz  $B_{n;p}$ -verteilt (geschrieben:  $X \sim B_{n;p}$ ). Die zu  $X$  gehörende Wahrscheinlichkeitsverteilung nennt man **Binomialverteilung<sup>1)</sup> mit den Parametern  $n$  und  $p$** .

H 19

## Satz H 27:

Die Binomialverteilung genügt den drei KOLMOGOROWSchen Axiomen der Wahrscheinlichkeitstheorie.

H 27

<sup>1)</sup> Die diese Wahrscheinlichkeitsverteilung charakterisierenden Terme heißen *Binomialkoeffizienten*, denn sie treten beim Entwickeln der  $n$ -ten Potenz  $(a+b)^n$  eines Binoms nach dem binomischen Lehrsatz als Koeffizienten auf. Der binomische Lehrsatz ist eine Verallgemeinerung der binomischen Formel  $(a+b)^2 = a^2 + 2ab + b^2$  für das Quadrieren eines Binoms.

## H 48

## Beispiel H 48:

Beim mathematischen Känguru-Wettbewerb sind in 75 Minuten ohne Taschenrechner 30 Aufgaben zu lösen. Bei jeder Aufgabe sind fünf Antworten zur Auswahl angegeben, von denen genau eine richtig ist. In einer der Aufgaben für Klasse 11 (1997) war z.B. zu entscheiden:

„An der Einerstelle der Summe  $1! + 2! + \dots + 100!$  steht die Ziffer

- (a) 0    (b) 3    (c) 5    (d) 2    (e) eine andere Ziffer als die angegebenen.“

Der Schüler „Hans im Glück“ baut auf sein Glück und „würfelt“ seine Antworten mit einem Dodekaeder aus. (Dieser reguläre 12-Flächen wird durch genau 12 kongruente gleichseitige Fünfecke begrenzt.) Wirft Hans eine 11 oder eine 12, so ignoriert er diesen Wurf und würfelt nochmals. Lässt die geworfene Augenzahl (ungleich 11 oder 12) bei der Division durch 5 den Rest 0, so wählt er die Antwort (a), bei Rest 1 die Antwort (b) usw. bis zur Antwort (e), die er bei Rest 4 wählt.

Mit welcher Wahrscheinlichkeit treten dann die nachfolgenden Ereignisse A, B, C und D ein:

A = {genau 12 richtige Antworten};

B = {höchstens 2 richtige Antworten}

C = {mindestens 2 richtige Antworten}

D = {mindestens die Hälfte der Antw. richtig}

## Lösung:

Zufallsgröße X: zufällige Anzahl der vom Schüler „Hans im Glück“ richtig „gelösten“ Aufgaben

$$X \sim B_{30; 0,2}, \text{ d.h. } P(X = k) = \binom{30}{k} \cdot 0,2^k \cdot 0,8^{30-k} \text{ für } k \in \{0; 1; 2; \dots; 30\}$$

$$P(A) = P(X = 12) = B_{30; 0,2}(\{12\})$$

$$= \binom{30}{12} \cdot 0,2^{12} \cdot 0,8^{18} \approx 0,006 \text{ (Fig. H 45, 1. u. 3. Zeile)}$$

$$P(B) = P(X \leq 2) = B_{30; 0,2}(\{0; 1; 2\})$$

$$= \binom{30}{0} \cdot 0,2^0 \cdot 0,8^{30} + \binom{30}{1} \cdot 0,2^1 \cdot 0,8^{29} + \binom{30}{2} \cdot 0,2^2 \cdot 0,8^{28} \approx 0,044 \text{ (Fig. H 45, 2. u. 4. Zeile)}$$

i	bi(n,p,i)
0	0.000231
1	0.001479
2	0.006382
3	0.01479
4	0.02479
5	0.03479
6	0.04479
7	0.05479
8	0.06479
9	0.07479
10	0.08479
11	0.09479
12	0.10479
13	0.11479
14	0.12479
15	0.13479
16	0.14479
17	0.15479
18	0.16479
19	0.17479
20	0.18479
21	0.19479
22	0.20479
23	0.21479
24	0.22479
25	0.23479
26	0.24479
27	0.25479
28	0.26479
29	0.27479
30	0.28479

Fig. H 45

$$P(C) = P(X \geq 2) = 1 - P(X < 2) \quad (\text{oder: } P(2 \leq X \leq 30) - \text{Fig. H 45, 2. u. 5. Zeile})$$

$$= 1 - B_{30; 0,2}(\{0; 1\}) = 1 - \left( \binom{30}{0} \cdot 0,2^0 \cdot 0,8^{30} + \binom{30}{1} \cdot 0,2^1 \cdot 0,8^{29} \right) \approx 0,989$$

$$P(D) = P(X \geq 15) = 1 - P(X \leq 14) \quad (\text{oder: } P(15 \leq X \leq 30) - \text{Fig. H 45, 2. u. 6. Zeile})$$

$$= 1 - \left( \binom{30}{0} \cdot 0,2^0 \cdot 0,8^{30} + \binom{30}{1} \cdot 0,2^1 \cdot 0,8^{29} + \dots + \binom{30}{14} \cdot 0,2^{14} \cdot 0,8^{16} \right) \approx 0,00023.$$

## H 5.3 Tabellierungen zur Binomialverteilung

Für binomialverteilte Zufallsgrößen  $X \sim B_{n;p}$  erfordert das Berechnen von Wahrscheinlichkeiten der Art  $P(X = k)$ ,  $P(X \leq k)$ ,  $P(X \geq k)$  bei größeren  $n$  und  $k$  mittels der BERNOULLI-Formel einen erheblichen Rechenaufwand (so keine programmierbaren Taschenrechner oder Computer eingesetzt werden können). Aus diesem Grund sind für einige gebräuchliche  $n$  und  $p$  die Wahrscheinlichkeiten

$$B_{n;p}(\{k\}) = P(\{\text{genau } k \text{ Erfolge}\}) = P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

sowie die kumulierten, die **aufsummierten binomialen Wahrscheinlichkeiten**, d.h. die Werte der Verteilungsfunktion

$$B_{n;p}(\{0; 1; \dots; k\}) = P(\{\text{höchstens } k \text{ Erfolge}\}) = P(X \leq k) = B_{n;p}(\{0\}) + B_{n;p}(\{1\}) + \dots + B_{n;p}(\{k\})$$

$$= \sum_{i=0}^k \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \text{ tabelliert.}$$

**Beispiel H 49: Tabelle für  $B_{n;p}(\{k\})$** 

$$n = 5, \quad p = 0,4 \quad P(X = 3) = B_{10;0,4}(\{3\}) \approx 0,230$$

$$n = 50, \quad p = 0,5 \quad P(X = 22) = B_{50;0,5}(\{22\}) \approx 0,079$$

$$n = 50, \quad p = 0,8 \quad P(X = 38) = B_{50;0,8}(\{38\}) \approx 0,103$$

Ist  $B_{n;p}$  nur für  $p \leq 0,5$  tabelliert, so nutzt man folgende rechnerische Umformung

$$B_{50;0,8}(\{38\}) =$$

$$\binom{50}{38} \cdot 0,8^{38} \cdot 0,2^{12} = \frac{50!}{38!12!} \cdot 0,8^{38} \cdot 0,2^{12} = \frac{50!}{12!38!} \cdot 0,2^{12} \cdot 0,8^{38} = \binom{50}{12} \cdot 0,2^{12} \cdot 0,8^{38} = B_{50;0,2}(\{12\})$$

Zu derselben Erkenntnis gelangt man auch durch folgende inhaltliche Überlegung:

$$B_{50;0,8}(\{38\})$$

$$= P(\{\text{genau 38 Erfolge bei 50 Realisierungen mit einer Erfolgswahrscheinlichkeit von 0,8}\})$$

$$= P(\{\text{genau 12 Misserfolge bei 50 Realisierungen mit einer Misserfolgswahrscheinlichkeit von 0,2}\})$$

$$= B_{50;0,2}(\{12\})$$

$n = 5$

k	p = 0,4 $B_{n;p}(\{k\})$	$B_{n;p}(\{0; \dots; k\})$
0	0,07776	0,07776
1	0,25920	0,33696
2	0,34560	0,68256
3	0,23040	0,91296
4	0,07680	0,98976
5	0,01024	1,00000

H 49

Dieses Beispiel lässt vermuten, dass es genügt, die Wahrscheinlichkeiten  $B_{n;p}(\{k\})$  und damit auch  $B_{n;p}(\{0; 1; \dots; k\})$  für  $p \leq 0,5$  zu tabellieren bzw. Tabellen mit den zwei Eingängen  $p$  und  $1 - p$  anzugeben.

**Beispiel H 50: Tabelle für  $B_{n;p}(\{k\})$  mit zwei Eingängen**

$n = 4$

k	p = 0,05	p = 0,10	p = $\frac{1}{6}$	p = 0,20	
0	0,81451	0,65610	0,48225	0,40960	4
1	0,17148	0,29160	0,38580	0,40960	3
2	0,01354	0,04860	0,11574	0,15360	2
3	0,00047	0,00360	0,01543	0,02560	1
4	0,00001	0,00010	0,00077	0,00160	0
	p = 0,95	p = 0,90	p = $\frac{5}{6}$	p = 0,80	k

$$B_{4;0,10}(\{1\}) \approx 0,29160$$

bei  $k = 1$  über den hellgrünen Eingang

$$B_{4;0,10}(\{1\}) = B_{4;0,90}(\{3\}) \approx 0,29160$$

bei  $k = 4 - 1 = 3$  über den dunkelgrünen Eingang

$$B_{4;0,80}(\{1\}) \approx 0,02560$$

bei  $k = 1$  über den dunkelgrünen Eingang

$$B_{4;0,80}(\{1\}) = B_{4;0,20}(\{3\}) \approx 0,02560$$

bei  $k = 4 - 1 = 3$  über den hellgrünen Eingang

H 50

**Satz H 28: Rechenregeln für binomialverteilte Zufallsgrößen**

Bei binomialverteilten Zufallsgrößen  $X$  mit den Parametern  $n$  und  $p$  gilt für  $i, k \in \{0; 1; \dots; n\}$  und  $i \leq k$

$$(1) B_{n;p}(\{k\}) = B_{n;1-p}(\{n-k\})$$

$$(2) B_{n;p}(\{k\}) = B_{n;p}(\{0; 1; \dots; k\}) - B_{n;p}(\{0; 1; \dots; k-1\})$$

$$(3) B_{n;p}(\{0; 1; \dots; k\}) = 1 - B_{n;1-p}(\{0; 1; \dots; n-k-1\})$$

$$(4) B_{n;p}(\{k; k+1; \dots; n\}) = 1 - B_{n;p}(\{0; 1; \dots; k-1\})$$

$$(5) B_{n;p}(\{i; i+1; \dots; k\}) = B_{n;p}(\{0; 1; \dots; k\}) - B_{n;p}(\{0; 1; \dots; i-1\}).$$

H 28



*Beweis* der Aussage (1) vgl. Satz H 29

*Beweis* der Aussage (3):

$$B_{n;p}(\{0; 1; \dots; k\})$$

$$\begin{aligned}
 &= \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} && \text{nach Definition} \\
 &= 1 - \sum_{i=k+1}^n \binom{n}{i} p^i (1-p)^{n-i} && \text{nach } P(\bar{A}) = 1 - P(A) \\
 &= 1 - \left[ \binom{n}{k+1} p^{k+1} (1-p)^{n-(k+1)} + \binom{n}{k+2} p^{k+2} (1-p)^{n-(k+2)} + \dots + \binom{n}{n-1} p^{n-1} (1-p)^1 \right. \\
 &\quad \left. + \binom{n}{n} p^n (1-p)^0 \right] \\
 &= 1 - \left[ \binom{n}{n-(k+1)} (1-p)^{n-(k+1)} p^{k+1} + \binom{n}{n-(k+2)} (1-p)^{n-(k+2)} p^{k+2} + \dots \right. \\
 &\quad \left. + \binom{n}{1} (1-p)^1 p^{n-1} + \binom{n}{0} (1-p)^0 p^n \right] && \text{nach } \binom{n}{k} = \binom{n}{n-k} \text{ und Kommutativgesetz der Multiplikation} \\
 &= 1 - \sum_{i=0}^{n-k-1} \binom{n}{i} (1-p)^i p^{n-i} && \text{nach Assoziativ- und Distributivgesetz der Addition sowie Definition des Summensymbols} \\
 &= 1 - B_{n;1-p}(\{0; 1; \dots; n-k-1\})
 \end{aligned}$$

Die Beweise der Aussagen (2), (4) und (5) ergeben sich direkt aus der Anwendung der Rechenregel  $P(\bar{A}) = 1 - P(A)$  für Gegenereignisse.

H 51

Beispiel H 51: **Tabelle für  $B_{n;p}(\{0; 1; \dots; k\}) = F_{n;p}(k)$**

$$\begin{aligned}
 n = 5, \quad p = 0,4 \quad P(X \leq 3) &= B_{5;0,4}(\{0; 1; 2; 3\}) \approx 0,913 \\
 n = 200, \quad p = 0,4 \quad P(X \leq 74) &= B_{200;0,4}(\{0; 1; \dots; 74\}) \\
 &\approx 0,214 \\
 P(X > 74) &= 1 - P(X \leq 74) \\
 &= 1 - B_{200;0,4}(\{0; 1; \dots; 74\}) \\
 &\approx 0,786 \\
 P(X \geq 74) &= 1 - P(X \leq 73) \\
 &= 1 - B_{200;0,4}(\{0; 1; \dots; 73\}) \\
 &\approx 0,826
 \end{aligned}$$

n = 5

k	p = 0,4	
	$B_{n;p}(\{k\})$	$B_{n;p}(\{0; \dots; k\})$
0	0,07776	0,07776
1	0,25920	0,33696
2	0,34560	0,68256
3	0,23040	0,91296
4	0,07680	0,98976
5	0,01024	1,00000

H 52

Beispiel H 52: **Tabelle für  $B_{n;p}(\{0; 1; \dots; k\})$  mit zwei Eingängen**

n = 4

k	p = 0,05	p = 0,10	p = $\frac{1}{6}$	p = 0,20	
0	0,81451	0,65610	0,48225	0,40960	3
1	0,98598	0,94770	0,86806	0,81920	2
2	0,99952	0,99630	0,98380	0,97280	1
3	0,99999	0,99990	0,99923	0,99840	0
4	1,00000	1,00000	1,00000	1,00000	-1
	p = 0,95	p = 0,90	p = $\frac{5}{6}$	p = 0,80	k

$$B_{4;0,10}(\{0; 1; 2\}) \approx 0,99630$$

bei k = 2 über den hellgrünen Eingang

$$B_{4;0,10}(\{0; 1; 2\}) = 1 - B_{4;0,90}(\{0; 1\})$$

bei k = 4 - 2 - 1 über den dunkelgrünen

$$\approx 1 - (1 - 0,99630) = 0,99630$$

Eingang mit 1 - (abgelesener Wert)

$$B_{4; 0,80}(\{0; 1; 2\}) \approx 1 - 0,81920 \approx 0,1808$$

bei  $k = 2$  über den dunkelgrünen Eingang mit  $1 -$  (abgelesener Wert)

$$\begin{aligned} B_{4; 0,80}(\{0; 1; 2\}) &= 1 - B_{4; 0,20}(\{0; 1\}) \\ &\approx 1 - 0,81920 \approx 0,1808 \end{aligned}$$

bei  $k = 4 - 2 - 1 = 1$  über den hellgrünen Eingang

Beispiel H 53:

Wir greifen noch einmal die Frage (2) aus Abschnitt H 5.2 auf: Mit welcher Wahrscheinlichkeit fällt die Augenzahl 4 mindestens 25-mal, wenn ein ungezinktes, mit den Augenzahlen 1 bis 4 beschriftetes Tetraeder genau 100-mal geworfen wird?

*Lösung:*

Zufallsgröße  $X$ : zufällige Anzahl der Vieren bei 100 Würfeln;  $X \sim B_{100; 0,25}$

$$P(X \geq 25) = 1 - P(X < 25) = 1 - B_{100; 0,25}(\{0; 1; 2; \dots; 24\}) \approx 1 - 0,48167 \approx 0,54$$

H 53

Beispiel H 54:

In einer großen Gruppe von Skitouristen sind 90% der Personen „fortgeschrittene“ Skiläufer und 10% Anfänger.

- „Auf gut Glück“ werden nun zehn Personen der Gruppe ausgewählt. Mit welcher Wahrscheinlichkeit befinden sich unter diesen zehn Skitouristen genau acht Fortgeschrittene?
- Wie viele Skitouristen müssen „auf gut Glück“ *mindestens* ausgewählt werden, damit mit einer Wahrscheinlichkeit von *mindestens* 0,99 *mindestens* einer der Touristen ein Anfänger ist?

*Lösungsvariante 1:*

Zufallsgröße  $X_n$ : zufällige Anzahl der Fortgeschrittenen unter  $n$  „auf gut Glück“ Ausgewählten

$$X_n \sim B_{n; 0,90}$$

$$\begin{aligned} \text{a) } P(X_{10} = 8) &= B_{10; 0,90}(\{8\}) \\ &= B_{10; 0,10}(\{2\}) \approx 0,194 \end{aligned}$$

$$\begin{aligned} \text{b) } P(X_n \leq n - 1) &= 1 - P(X_n = n) \\ &= 1 - B_{n; 0,90}(\{n\}) \\ &= 1 - \binom{n}{n} \cdot 0,90^n \cdot 0,10^0 \\ &= 1 - 0,90^n \end{aligned}$$

*Lösungsvariante 2:*

Zufallsgröße  $Y_n$ : zufällige Anzahl der Anfänger unter  $n$  „auf gut Glück“ Ausgewählten

$$Y_n \sim B_{n; 0,10}$$

$$\text{a) } P(Y_{10} = 2) = B_{10; 0,10}(\{2\}) \approx 0,194$$

$$\begin{aligned} \text{b) } P(Y_n \geq 1) &= 1 - P(Y_n = 0) \\ &= 1 - B_{n; 0,10}(\{0\}) \\ &= 1 - \binom{n}{0} \cdot 0,10^0 \cdot 0,90^n \\ &= 1 - 0,90^n \end{aligned}$$

Es soll gelten:  $1 - 0,90^n \geq 0,99$  bzw.  $0,90^n \leq 0,01$ . Daraus folgt  $\ln(0,90^n) \leq \ln 0,01$ , also  $n \geq \frac{\ln 0,01}{\ln 0,90}$  und damit  $n > 43,7$ . Es müssen mindestens 44 Personen ausgewählt werden.

H 54

## H 5.4 Grafische Veranschaulichung der Binomialverteilung $B_{n; p}$

Jede Binomialverteilung  $B_{n; p}$  wird durch die beiden Parameter  $n$  und  $p$  charakterisiert. Dabei ist  $n$  als Länge der zugehörigen BERNOULLI-Kette, d.h. als Anzahl der Realisierungen des zugehörigen BERNOULLI-Experiments zu interpretieren und  $p$  als Wahrscheinlichkeit des „Erfolges“ beim BERNOULLI-Experiment. Die Abhängigkeit der  $B_{n; p}$ -Verteilung von  $n$  und  $p$  kann man sich anhand zweier Serien von Histogrammen veranschaulichen, in denen jeweils einer der beiden Parameter konstant gehalten wird.

### Histogramm für $X \sim B_{14; p}$ mit variablem $p$

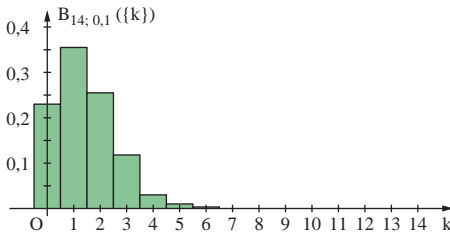


Fig. H 46

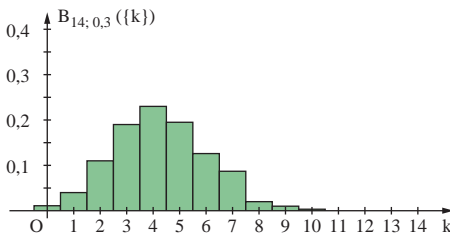


Fig. H 47

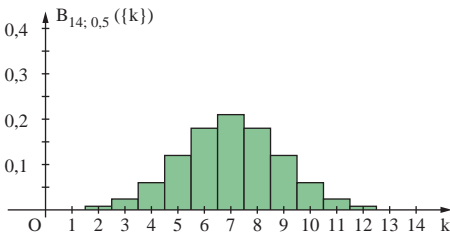


Fig. H 48

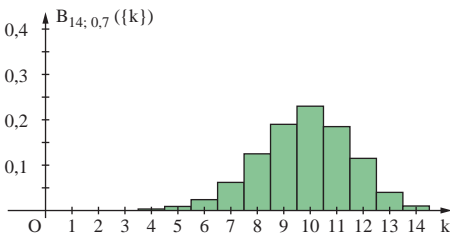


Fig. H 49

### Eigenschaften von $B_{14; p}$

- (1) Nur das Histogramm bei  $p = 0,5$  ist axialsymmetrisch (und zwar zur Geraden mit der Gleichung  $k = 14 \cdot 0,5$ ); je mehr  $p$  von 0,5 abweicht, desto asymmetrischer wird das Histogramm.
- (2) Das Histogramm von  $B_{14; p}$  ist das Spiegelbild von  $B_{14; 1-p}$  bezüglich der Spiegelgeraden mit der Gleichung  $k = 14 \cdot 0,5$ .
- (3) Die Lage des höchsten Rechteckes wandert mit wachsendem  $p$  nach rechts. Es befindet sich bei  $k \approx 14 \cdot p$ . Dabei nimmt die Höhe des höchsten Rechteckes mit wachsendem  $p$  für  $p \leq 0,5$  ab und für  $p \geq 0,5$  wieder zu. Das Histogramm „zerfließt“ am stärksten für  $p = 0,5$ .

- (4) Für  $p_1 < p_2$  gilt

$$B_{14; p_1}(\{0; 1; \dots; k\}) > B_{14; p_2}(\{0; 1; \dots; k\})$$

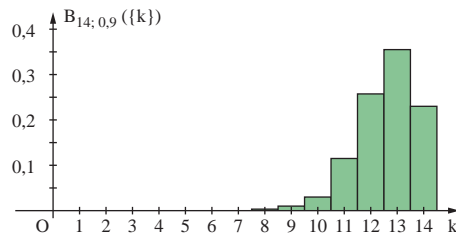


Fig. H 50

### Histogramm für $X \sim B_{n; 0,3}$ mit variablem $n$

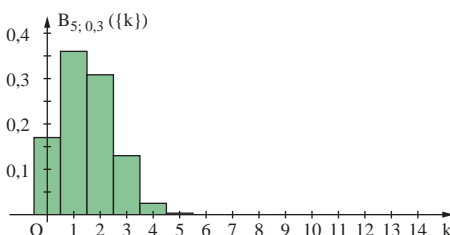


Fig. H 51

### Eigenschaften von $B_{n; 0,3}$

- (5) Die Histogramme werden mit wachsendem  $n$  zunehmend axialsymmetrischer.
- (6) Die Histogramme werden mit wachsendem  $n$  zunehmend breiter und zugleich flacher. (Die Wahrscheinlichkeitsmasse 1 „zerfließt“.)

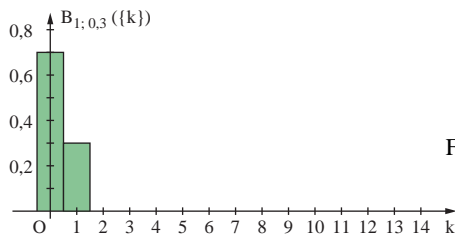


Fig. H 52

(7) Einige Histogramme (für die  $(n+1) \cdot 0,3$  ganzzahlig ist) besitzen zwei nebeneinander liegende (für  $k = (n+1) \cdot 0,3 - 1$  und für  $k = (n+1) \cdot 0,3$ ) maximal hohe Rechtecke. (Fig. H 53, H 55)

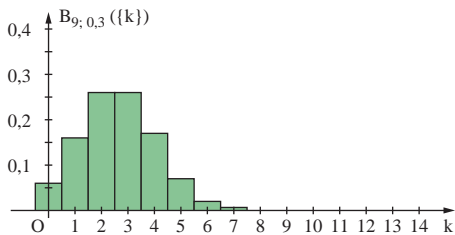


Fig. H 53

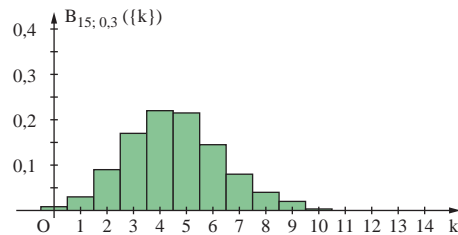


Fig. H 54

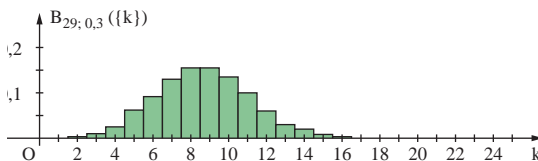


Fig. H 55

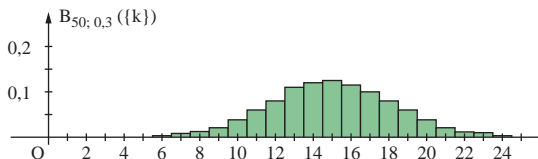


Fig. H 56

Die markante Symmetrieeigenschaft (2), die wir beim Vergleich der Histogramme von  $B_{14;p}$  und  $B_{14;1-p}$  beobachtet haben, lässt sich zu dem folgenden Satz verallgemeinern.

**Satz H 29: Gegenseitige Lage der Histogramme von  $B_{n;p}$  und  $B_{n;1-p}$**

Die Histogramme von  $B_{n;p}$  und  $B_{n;1-p}$  liegen bezüglich der Geraden mit der Gleichung  $k = 0,5n$  axialsymmetrisch zueinander.<sup>1)</sup>

Spezialfall: Das Histogramm von  $B_{n;0,5}$  ist axialsymmetrisch bezüglich der Geraden  $k = 0,5n$ .

H 29

*Beweis:*

Es ist nachzuweisen, dass für alle  $k$  ( $k \in \{0; 1; \dots; n\}$ )

gilt:  $B_{n;p}(\{k\}) = B_{n;1-p}(\{0,5n + (0,5n - k)\})$

Zu zeigen ist also die Gleichheit von  $B_{n;p}(\{k\})$  und  $B_{n;1-p}(\{n - k\})$ , die wir für  $B_{50;0,8}$  für  $k = 40$  im Beispiel H 49 bereits gezeigt haben. Allgemein gilt:

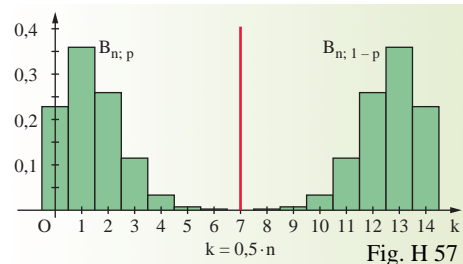


Fig. H 57

<sup>1)</sup> Dieser Satz stellt eine geometrische Interpretation der Aussage (1) von Satz H 28 dar.

$$\begin{aligned}
 B_{n;p}(\{k\}) &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\
 &= \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k} = \frac{n!}{(n-k)!k!} \cdot (1-p)^{n-k} \cdot p^k \\
 &= \binom{n}{n-k} \cdot (1-p)^{n-k} \cdot p^k = B_{n;1-p}(\{n-k\}) \quad \text{w.z.b.w.}
 \end{aligned}$$

### H 5.5 Erwartungswert und Streuung (Varianz) binomialverteilter Zufallsgrößen

Um den Erwartungswert einer binomialverteilten Zufallsgröße  $X \sim B_{n;p}$  auf Grund seiner Definition zu bestimmen, müsste die Summe

$$\begin{aligned}
 EX &= 0 \cdot B_{n;p}(\{0\}) + 1 \cdot B_{n;p}(\{1\}) + \dots + n \cdot B_{n;p}(\{n\}) \\
 &= 0 \cdot \binom{n}{0} \cdot p^0 \cdot (1-p)^{n-0} + 1 \cdot \binom{n}{1} \cdot p^1 \cdot (1-p)^{n-1} + \dots + n \cdot \binom{n}{n} \cdot p^n \cdot (1-p)^0
 \end{aligned}$$

vereinfacht werden. Dies erfordert aber einen hohen Aufwand an Umformungen. Um zu einer Hypothese für EX zu gelangen, könnte man versuchen, die Parameter n und p von  $B_{n;p}$  so zu wählen, dass EX sich mit einem relativ geringen Rechenaufwand ermitteln lässt.

H 55

Beispiel H 55:

(1)  $X \sim B_{5;0,2}$

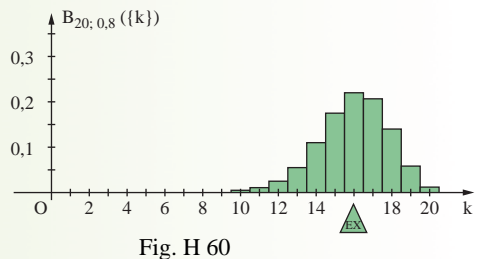
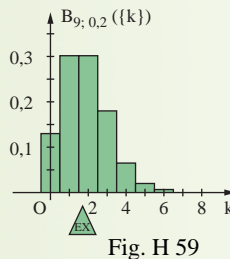
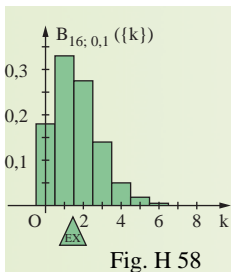
$$\begin{aligned}
 EX &= 0 \cdot \binom{5}{0} \cdot 0,2^0 \cdot 0,8^5 + 1 \cdot \binom{5}{1} \cdot 0,2^1 \cdot 0,8^4 + \dots + 5 \cdot \binom{5}{5} \cdot 0,2^5 \cdot 0,8^0 \\
 &= 0 + 0,4096 + 0,4096 + 0,1536 + 0,0256 + 0,0016 = 1 = 5 \cdot 0,2
 \end{aligned}$$

(2)  $X \sim B_{5;0,3}$

$$\begin{aligned}
 EX &= 0 \cdot \binom{5}{0} \cdot 0,3^0 \cdot 0,7^5 + 1 \cdot \binom{5}{1} \cdot 0,3^1 \cdot 0,7^4 + \dots + 5 \cdot \binom{5}{5} \cdot 0,3^5 \cdot 0,7^0 \\
 &= 0 + 0,36015 + 0,6174 + 0,3969 + 0,1134 + 0,01215 = 1,5 = 5 \cdot 0,3
 \end{aligned}$$

Beispiel H 55 legt die Vermutung nahe, dass  $EX = n \cdot p$  gilt.

Ein zweiter Zugang zu einer vernünftigen Vermutung für EX bei  $X \sim B_{n;p}$  ist die Interpretation des Erwartungswertes als Schwerpunkt der Verteilung der Wahrscheinlichkeitsmasse 1.



$$EX \approx 1,5$$

$$EX \approx 1,8$$

$$EX \approx 16$$

$$EX = n \cdot p = 16 \cdot 0,1 = 1,6$$

$$EX = n \cdot p = 9 \cdot 0,2 = 1,8$$

$$EX = n \cdot p = 20 \cdot 0,8 = 16$$

Den obigen drei Histogrammen ist zu entnehmen, dass der Erwartungswert in der Nähe derjenigen k mit einem maximal hohen Rechteck zu vermuten ist. Diese Rechtecke sind aber, wie auch die im Abschnitt H 5.4 diskutierten Histogramme bestätigen, jeweils in der Nähe von  $n \cdot p$ . Das bestärkt uns in der Hypothese:  $EX = n \cdot p$ .

Interpretieren wir den Erwartungswert  $EX$  als die zu erwartende mittlere Anzahl der Erfolge, so kann man einen Näherungswert für  $EX$  auch experimentell bestimmen. Wir erhalten so einen dritten Zugang zu einer sinnvollen Vermutung für  $EX$ . Dazu muss die entsprechende BERNOULLI-Kette hinreichend oft realisiert bzw. simuliert werden.

Beispiel H 56:

Zufallsgröße  $Y$  des BERNOULLI-Experiments:  $Y \cong \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$  mit der Erfolgswahrscheinlichkeit  $p = 0,38$

Zufallsgröße  $X$  der BERNOULLI-Kette: Zufällige Anzahl der Erfolge bei fünf Realisierungen von  $Y$ ;  $X \sim B_{n;p}$  mit  $n = 5$  und  $p = 0,38$

Simulation der BERNOULLI-Kette der Länge  $n = 5$  und  $p = 0,38$  durch Erzeugen von (Pseudo-) Zufallszahlen aus dem Intervall  $[1; 100]$  mit einem Taschenrechner: Die Zufallszahlen aus  $[1; 38]$  werden als „Erfolg“ interpretiert und mit „1“ protokolliert; die Zufallszahlen aus  $[39; 100]$  werden als „Misserfolg“ interpretiert und mit „0“ protokolliert.

```
00100 01011 01101 10010 00000 11000 00100 10000 01001 00010 11001 01110 01000 01000 11101 00011
11001 10010 10110 01001 10110 10010 11100 11110 01000 01100 11100 01100 10000 10110 11001 01010
11100 00000 01000 10000 10000 10100 01100 00011 10000 00010 00100 10000 10001 01101 01010 10111
01000 11101 00001 01001 11111 00011 10000 10001 01010 11000 00001 01010 01111 10011 01000 01011
00101 00010 00110 11000 00001 10010 01101 01001 10001 01011 10000 10100 10001 11000 00001 10110
00100 00100 10010 00101 01001 11000 00100 00000 01101 01000 01001 01001 10000 00001 01101 10100
01111 01010 10110 00001
```

Anzahl der Erfolge bei fünf Realisierungen:

0		3	
1		4	
2		5	

Das arithmetische Mittel der Anzahl der Erfolge bei den 100 Simulationen ergibt

$$\frac{0 \cdot 3 + 1 \cdot 32 + 2 \cdot 37 + 3 \cdot 21 + 4 \cdot 6 + 5 \cdot 1}{100} = 1,98 \approx n \cdot p = 5 \cdot 0,38 = 1,9.$$

Auch dieses Beispiel bekräftigt die Hypothese  $EX = n \cdot p$ .

H 56

Satz H 30:

Eine binomialverteilte Zufallsgröße  $X \sim B_{n;p}$  besitzt den Erwartungswert  $EX = n \cdot p$ , die Varianz  $\text{Var}X = n \cdot p \cdot (1 - p)$  sowie die Standardabweichung  $\sqrt{\text{Var}X} = \sqrt{n \cdot p \cdot (1 - p)}$ .

H 30

Beweis/Variante 1:

$$\begin{aligned}
 EX &= \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n-1} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} + \frac{n!}{(n-1)!0!} p^n (1-p)^0 \\
 &= np \sum_{k=1}^{n-1} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} (1-p)^{(n-1)-(k-1)} + n \cdot p^n \\
 &= np \left( \sum_{k=0}^{n-2} \frac{(n-1)!}{k!(n-1-k)!} p^k (1-p)^{(n-1)-k} + p^{n-1} \right) \\
 &= np \left( \binom{n-1}{0} p^0 (1-p)^{n-1} + \binom{n-1}{1} p^1 (1-p)^{(n-1)-1} + \dots + \binom{n-1}{n-2} p^{n-2} (1-p)^1 + \binom{n-1}{n-1} p^{n-1} (1-p)^0 \right)
 \end{aligned}$$

$$= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} = n \cdot p \quad \text{w.z.b.w.}$$

*Beweis/Variante 2:*

$X \sim B_{n;p}$  ist aufzufassen als Summe der  $n$  unabhängigen Zufallsgrößen  $X_1, X_2, \dots, X_n$  mit

$$X_1 = X_2 = \dots = X_n \cong \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}, \text{ d.h. } X = X_1 + X_2 + \dots + X_n.$$

Demzufolge gilt

$$EX = E(X_1 + X_2 + \dots + X_n) = EX_1 + EX_2 + \dots + EX_n = n \cdot p \quad (\text{Satz H 17})$$

$$D^2X = D^2(X_1 + X_2 + \dots + X_n) = D^2X_1 + D^2X_2 + \dots + D^2X_n = n \cdot p \cdot (1-p) \quad (\text{Satz H 21})$$

w.z.b.w.

**H 57**

**Beispiel H 57:**

Ein ideales Tetraeder mit den Seitenbeschriftungen 1, 2, 3 und 4 werde 200-mal geworfen. Wie viele Vieren sind dabei zu erwarten? Welche Standardabweichung  $\sigma$  besitzt die (zufällige) Anzahl der Vieren? Mit jeweils welcher Wahrscheinlichkeit liegt die (zufällige) Anzahl  $X$  der Vieren im Intervall  $[EX - \sigma; EX + \sigma]$ ,  $[EX - 2\sigma; EX + 2\sigma]$  bzw.  $[EX - 3\sigma; EX + 3\sigma]$ ?

*Lösung:*

Zufallsgröße  $X$ :

zufällige Anzahl der Vieren beim 200-maligen Werfen des L-Tetraeders;  $X \sim B_{n;p}$  mit  $n = 200$  und  $p = 0,25$

Erwartungswert:

$$EX = n \cdot p = 200 \cdot 0,25 = 50 \text{ Es sind 50 Vieren zu erwarten.}$$

Standardabweichung:

$$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{200 \cdot 0,25 \cdot 0,75} \approx 6,12$$

$$P(X \in [EX - \sigma; EX + \sigma]) = P(X \in [43,87 \dots; 56,12 \dots])$$

$$= P(44 \leq x \leq 56)$$

$$= B_{200; 0,25}(\{0; 1; \dots; 56\}) - B_{200; 0,25}(\{0; 1; \dots; 43\})$$

$$\approx 0,85546 - 0,14376 \approx 0,712$$

$$P(X \in [EX - 2\sigma; EX + 2\sigma]) = P(38 \leq x \leq 62)$$

$$= B_{200; 0,25}(\{0; 1; \dots; 62\}) - B_{200; 0,25}(\{0; 1; \dots; 37\})$$

$$\approx 0,97745 - 0,01824 \approx 0,959$$

$$P(X \in [EX - 3\sigma; EX + 3\sigma]) = P(32 \leq x \leq 68)$$

$$= B_{200; 0,25}(\{0; 1; \dots; 68\}) - B_{200; 0,25}(\{0; 1; \dots; 31\}) \approx 0,99830 - 0,00079 \approx 0,998$$

**Punktdiagramm**

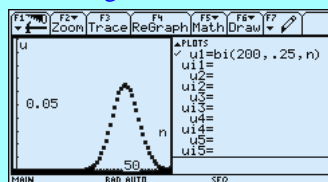


Fig. H 61

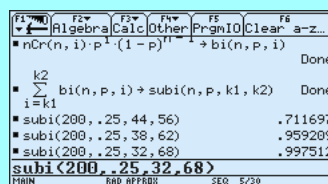


Fig. H 62

**H 58**

**Beispiel H 58:**

Wir kehren nun zu der einleitend (S. 351) aufgegriffenen Frage zurück, mit der sich der Chevalier DE MÉRÉ an den berühmten Mathematiker PASCAL gewandt hatte, um sie nun mit heutigen Mitteln zu beantworten: Was ist wahrscheinlicher: bei vier Würfeln mit einem Würfel mindestens eine Sechs zu werfen oder bei 24 Würfeln mit zwei Würfeln mindestens eine Doppelsechs zu werfen?

Als „Hobbymathematiker“<sup>(1)</sup> ging DE MÉRÉ davon aus, dass beide Wahrscheinlichkeiten gleich

<sup>1)</sup> PASCAL schrieb über DE MÉRÉ: „... er ist ein sehr tüchtiger Kopf, aber er ist kein Mathematiker (das ist ... ein großer Mangel), und er begreift nicht einmal, dass eine mathematische Linie bis ins Unendliche teilbar ist ...“

sein müssten, da sich 24 zu 36 wie 4 zu 6 verhält. Seine Beobachtungen als Spieler schienen dieser Annahme aber zu widersprechen.

**Lösung:**

**Modellannahme:** Die genannten Würfel sind L-Würfel, deren Seitenflächen jeweils mit den Augenzahlen 1 bis 6 durchnummeriert sind. Die Würfe erfolgen unabhängig voneinander.

**Zufallsgröße  $X_4$ :** zufällige Anzahl der geworfenen Sechsen bei vier Würfeln mit genau einem Würfel;  $X_4 \sim B_{4; 1/6}$

$P(\{\text{bei vier Würfeln mit einem Würfel mindestens eine Sechse}\})$

$$= P(X_4 \geq 1) = 1 - P(X_4 = 0) = 1 - \binom{4}{0} \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^4 \approx 0,5177$$

**Zufallsgröße  $Y_{24}$ :** zufällige Anzahl der geworfenen Doppelsechsen bei 24 Würfeln mit zwei Würfeln;  $Y_{24} \sim B_{24; 1/36}$

$P(\{\text{bei 24 Würfeln mit zwei Würfeln eine Doppelsechse zu werfen}\})$

$$= P(Y_{24} \geq 1) = 1 - P(Y_{24} = 0) = 1 - \binom{24}{0} \cdot \left(\frac{1}{36}\right)^0 \cdot \left(\frac{35}{36}\right)^{24} \approx 0,4914$$

Die Wahrscheinlichkeit, bei vier Würfeln mit einem Würfel mindestens eine Sechse zu werfen, ist also (wenn auch nur geringfügig) größer als die Wahrscheinlichkeit, bei 24 Würfeln mit zwei Würfeln mindestens eine Doppelsechse zu werfen.

Bliebe zu klären, ob es für die von DE MÉRÉ angenommene Verhältnisgleichheit  $\frac{4}{6} = \frac{24}{36}$  eine sinnvolle wahrscheinlichkeitstheoretische Deutung gibt.  $\frac{4}{6} = \frac{24}{36}$  kann auch als  $4 \cdot \frac{1}{6} = 24 \cdot \frac{1}{36}$  geschrieben werden und entspricht damit der Gleichheit der Erwartungswerte  $EX_4 = EY_{24} = 0, \bar{6}$ .  $X_4$  und  $Y_{24}$  haben dieselben Erwartungswerte, aber die  $Y_{24}$ -Werte streuen stärker als die  $X_4$ -Werte:

$$DX_4 = \sqrt{4 \cdot \frac{1}{6} \cdot \frac{5}{6}} \approx 0,7454 < 0,8051 \approx \sqrt{24 \cdot \frac{1}{36} \cdot \frac{35}{36}} = DY_{24}$$

## H 5.6 Simulation von BERNOULLI-Ketten mit dem Taschenrechner

Um eine BERNOULLI-Kette der Länge  $n$  und mit der Erfolgswahrscheinlichkeit  $p$  zu simulieren, definieren wir im Programm-Editor des GTA die Funktion **simbin**. Dazu wird nach Start dieses Editors mit **[APPS]** **[7]** (*Programmeditor*) **[3]** (*NEW*) in das Dialogfeld der Variablenname **simbin** eingetragen und dann diese Funktion entsprechend (der aus zwei Schirmbildern zusammengesetzten) Fig. H 63 definiert.

Gibt man nun im Home-Editor **simbin(4, 0.6, 390014)** ein, so wird auf dem Bildschirm eine Liste der Länge  $n = 4$  angezeigt, in der 1 für „Erfolg“ und 0 für „Misserfolg“ steht sowie die zufällige Anzahl  $X$  der Einsen binomialverteilt ist mit  $P(X = k) = B_{4; 0.6}(\{k\})$  für  $k \in \{0; 1; 2; 3; 4\}$ . Die zusätzliche Eingabe eines sich von Simulation zu Simulation ändernden (aber ansonsten beliebigen)  $x$ -Wertes (hier 390014) dient der Initialisierung des Zufallsgenerators. Verändert man den  $x$ -Wert nicht, so werden dieselben Listen erstellt.

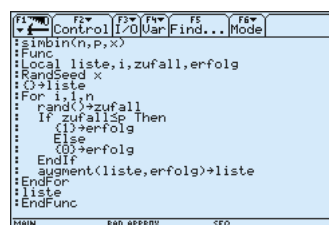


Fig. H 63

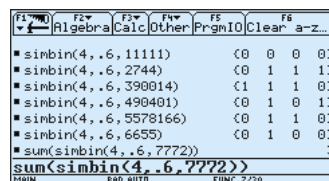


Fig. H 64



Um mithilfe *mehrerer* (m Stück) simulierter BERNOULLI-Ketten von jeweils derselben Länge  $n$  und mit derselben Erfolgswahrscheinlichkeit  $p$  Näherungswerte für die Wahrscheinlichkeiten  $B_{n,p}(\{k\})$  bzw.  $B_{n,p}(\{0; 1; \dots; k\})$  zu ermitteln, kann man im Programm-Editor Funktion **simbinm** definieren (Fig. H 65).

Im Home-Editor wird durch Eingabe von

**simbinm(200, 4, 0.6, 2888)** bzw.

**cumSum(simbinm(200,0.6,2888))**

eine Liste von Näherungswerten für

$B_{4,0.6}(\{k\}) = \binom{4}{k} \cdot 0.6^k \cdot 0.4^{4-k}$  bzw. für

die aufsummierten Wahrscheinlichkeiten  $B_{4,0.6}(\{0; 1; \dots; k\})$

ausgegeben, die durch 200 Simulationen einer BERNOULLI-Kette mit den Parametern  $n = 4$  und  $p = 0.6$  erzeugt worden sind.

```

: simbinm(n,p,x)
: Func
: Local list1,i,zaehler,j,zufall,anz,lis
: te2,k,x,y,lis
: RandSeed x
: ( )->list1
: For i,1,m
:   @zaehler
:   For j,1,n
:     rand()->zufall
:     If zufall<p
:       zaehler+1->zaehler
:   EndFor
:   augment(list1,(zaehler))>list1
: EndFor
: ( )->list2
: For k,0,n
:   @anz
:   For x,1,m
:     If list1(x)=k
:       anz+1->anz
:   EndFor
:   augment(list2,(anz))>list2
: EndFor
: seq(approx(list2(y)/m),y,1,n+1)>list
: list
: EndFunc
  
```

Fig. H 65

```

simbinm(200,4,.6,15799)
{.05 .135 .315 .345 .155}
simbinm(200,4,.6,2888)
{.005 .155 .295 .425 .12}
cumSum(simbinm(200,4,.6,2888))
{.005 .16 .455 .88 1.}
msum(simbinm(200,4,.6,2888))
  
```

Fig. H 66

### H 5.7 Grenzwertsatz von DE MOIVRE-LAPLACE zur Binomialverteilung

Bei vielen praktischen Problemen (z.B. im Zusammenhang mit dem Testen von Hypothesen – s. Abschnitt J 1) werden aufsummierte Binomialwahrscheinlichkeiten  $B_{n,p}(\{0; 1; \dots; k\})$  für große  $n$  (z.B.  $n > 200$ ) oder „krumme“ Werte von  $p$  oder  $n$  (z.B.  $p = 0.22$  oder  $n = 172$ ) benötigt, die üblicherweise nicht tabelliert vorliegen. Es war daher in einer Zeit, in der es noch keine elektronischen Hilfsmittel gab, naheliegend, nach entsprechenden **Näherungsformeln** zu suchen.

In der Analysis werden Flächeninhalte unter dem Graphen einer stetigen Funktion näherungsweise mittels der Rechteckmethode (Streifenmethode) bestimmt. Solcherart Rechtecke treten auch bei den zu betrachtenden Binomialwahrscheinlichkeiten auf, wenn man deren Histogramme betrachtet, jedoch ist die zugehörige stetige Funktion unbekannt. In einem **Histogramm** wird im Unterschied zum **Stabdiagramm**, wo man an der Stelle  $k$  einen „Stab“ der Länge  $P(X = k)$  einträgt, über dem Intervall  $[k - 0.5; k + 0.5]$  ein Rechteck mit der Breite 1 und der Höhe  $P(X = k)$  gezeichnet. Die Wahrscheinlichkeit  $B_{n,p}(\{k\})$  entspricht im Histogramm der Flächenmaßzahl des Rechtecks an der Stelle  $k$  sowie die aufsummierten Wahrscheinlichkeiten  $B_{n,p}(\{0; 1; \dots; k\})$  der Summe der Flächenmaßzahlen der Rechtecke (soweit sie zeichnerisch erfassbar sind) an den Stellen  $0, 1, \dots, k$  (Fig. H 67).

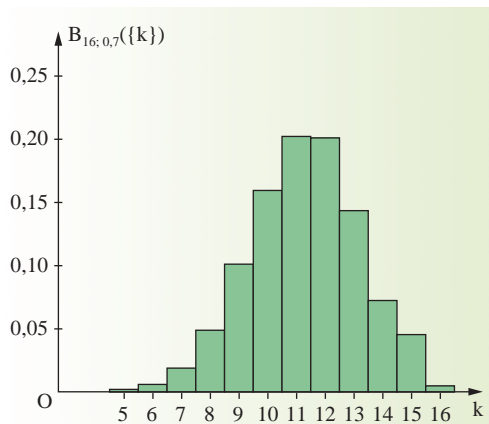


Fig. H 67

Betrachtet man Histogramme von Binomialwahrscheinlichkeit z.B. mittels Computer nochmals genauer, so erkennt man (in Ergänzung zu den Beobachtungen im Abschnitt H 5.4):

$$X \sim B_{5; 0,3}(\{k\})$$

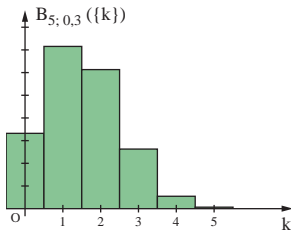


Fig. H 68a

$$X - 5 \cdot 0,3 = X - 1,5$$

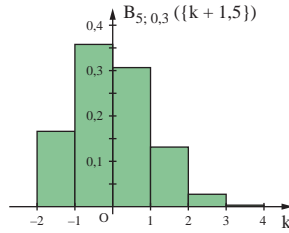


Fig. H 68b

$$\frac{X - 1,5}{\sqrt{1,05}}$$

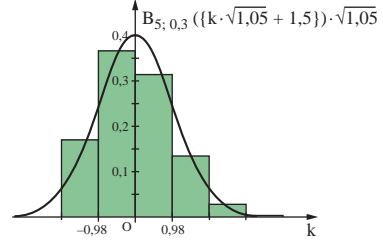


Fig. H 68c

$$X \sim B_{15; 0,3}(\{k\})$$

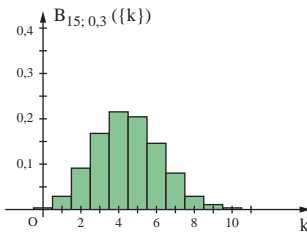


Fig. H 69a

$$X - 15 \cdot 0,3 = X - 4,5$$

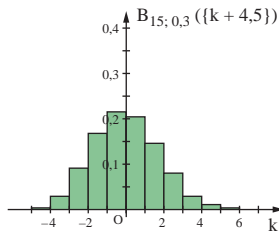


Fig. H 69b

$$\frac{X - 4,5}{\sqrt{3,15}}$$

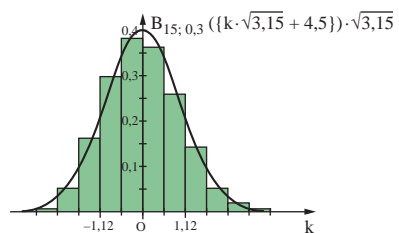


Fig. H 69c

$$X \sim B_{50; 0,3}(\{k\})$$

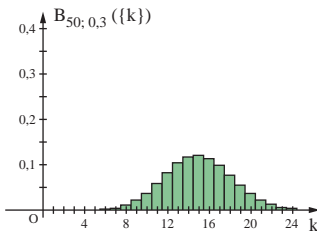


Fig. H 70a

$$X - 50 \cdot 0,3 = X - 15$$

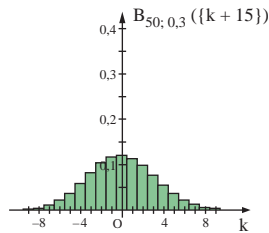


Fig. H 70b

$$\frac{X - 1,5}{\sqrt{10,5}}$$

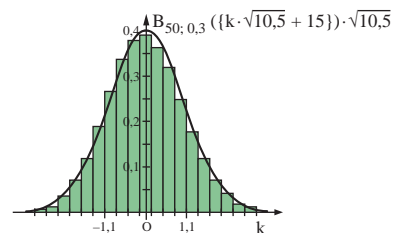


Fig. H 70c

- (1) Je größer  $\mu = EX = n \cdot p$  wird, desto weiter wandert das Histogramm „nach rechts“. Betrachtet man statt der Zufallsgröße  $X$  die (zentrierte) Zufallsgröße  $X - \mu$ , so beträgt deren Erwartungswert  $E(X - \mu) = EX - \mu = 0$  und das Histogramm hat sein höchstes Rechteck bei  $k = \mu - \mu = 0$ , da  $P(X - \mu = k - \mu) = P(X = k) = B_{n; p}(\{k\})$ ,  $P(X - \mu = i) = B_{n; p}(\{i + \mu\})$ .
- (2) Je größer  $\sigma^2 = n \cdot p \cdot (1 - p)$  wird, desto mehr „zerfließt“ das Histogramm, desto flacher und zugleich breiter wird das Histogramm. Geht man von der bereits zentrierten Zufallsgröße  $X - \mu$  über zu der standardisierten Zufallsgröße  $\frac{X - \mu}{\sigma}$  mit  $P(\frac{X - \mu}{\sigma} = \frac{k - \mu}{\sigma}) = B_{n; p}(\{k\})$ , so besitzt diese die Streuung  $D^2(\frac{X - \mu}{\sigma}) = D^2(\frac{1}{\sigma}(X - \mu)) = (\frac{1}{\sigma})^2 \cdot D^2X = (\frac{1}{\sigma})^2 \cdot \sigma^2 = 1$  und zerfließt daher nicht mehr. (Die Rechteckbreiten werden von der Breite 1 der Intervalle  $[k - 0,5; k + 0,5]$  bei  $X$  auf die Breite  $\frac{1}{\sigma}$  der Intervalle  $[\frac{k - \mu - 0,5}{\sigma}; \frac{k - \mu + 0,5}{\sigma}]$  bei  $\frac{X - \mu}{\sigma}$  gestaucht, demzufolge müssen die Rechteckhöhen von  $B_{n; p}(\{k\})$  bei  $X$  auf  $\sigma \cdot B_{n; p}(\{k\})$  bei  $\frac{X - \mu}{\sigma}$  gestreckt werden, so dass der Rechteckflächeninhalt der Maßzahl nach unverändert  $B_{n; p}(\{k\})$  bleibt.)

(3) Je größer  $n$  wird, desto mehr nähert sich die äußere Form der (für  $p \neq 0,5$  nicht symmetrischen) Histogramme der standardisierten Zufallsgröße  $\frac{X-\mu}{\sigma}$  einer zur y-Achse symmetrischen Glockenkurve. Diese **GAUSSsche Glockenkurve** ist der Graph der geraden Funktion  $\phi$  mit der Gleichung  $\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$  und ihrer Stammfunktion  $\Phi$ .

Betrachtet man die Streckung des Histogramms von  $\frac{X-\mu}{\sigma}$  mit dem Streckungsfaktor  $\sigma$ , so ergibt sich für die Einzelwahrscheinlichkeit  $B_{n;p}(\{k\}) = P(X = k) = P(\frac{X-\mu}{\sigma} = \frac{k-\mu}{\sigma}) \approx \frac{1}{\sigma} \cdot \phi(\frac{k-\mu}{\sigma})$  sowie für die aufsummierten Wahrscheinlichkeiten

$$\begin{aligned} B_{n;p}(\{0; 1; \dots; k\}) &= \sum_{i=0}^k B_{n;p}(\{i\}) \\ &= \sum_{i=0}^k (\text{Flächeninhalt des Rechtecks mit der Breite 1 und der Höhe } B_{n;p}(\{i\})) \\ &= \sum_{i=0}^k (\text{Flächeninhalt des Rechtecks mit der Breite } \frac{1}{\sigma} \text{ und der Höhe } \sigma \cdot B_{n;p}(\{i\})) \\ &\approx \int_{-\infty}^{\frac{k-\mu}{\sigma}} \phi(x) dx = \Phi(\frac{k-\mu}{\sigma}) \end{aligned}$$

H 31

### Satz H 31: Grenzwertsatz (Näherungsformel) von DE MOIVRE-LAPLACE

Es sei  $X$  eine binomialverteilte Zufallsgröße mit  $X \sim B_{n;p}$  und  $n \cdot p \cdot q > 9$ . Dann gelten die

*lokale Näherungsformel*  $P(X = k) = B_{n;p}(\{k\}) \approx \frac{1}{\sigma} \phi(\frac{k-\mu}{\sigma})$  und die

*globale Näherungsformel*  $P(X \leq k) = B_{n;p}(\{0; 1; \dots; k\}) \approx \Phi(\frac{k+0,5-\mu}{\sigma})$ ,

wobei  $\mu = EX = n \cdot p$  und  $\sigma = DX = \sqrt{n \cdot p \cdot (1-p)}$ .

Da die Funktion  $\phi: x \mapsto \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$  nicht geschlossen integrierbar ist, sind die Funktionswerte ihrer Stammfunktion  $\Phi$  tabelliert. Für die Nutzung dieser Tabellen ist die Eigenschaft  $\Phi(-x) = 1 - \Phi(x)$  wesentlich, da somit  $\Phi$  nur für nichtnegative Argumente tabelliert werden muss. Die Gültigkeit dieser Gleichung für alle  $x \in \mathbb{R}$

kann man aus dem zur y-Achse symmetrischen Graphen  $\phi$  entnehmen, wenn man beachtet, dass

$$\int_{-\infty}^{+\infty} \phi(x) dx = 1 \text{ gilt: } \Phi(-x) = \int_{-\infty}^{-x} \phi(t) dt = \int_{+\infty}^{-x} \phi(t) dt = \int_{-\infty}^{+\infty} \phi(t) dt - \int_{-\infty}^x \phi(t) dt = 1 - \Phi(x).$$

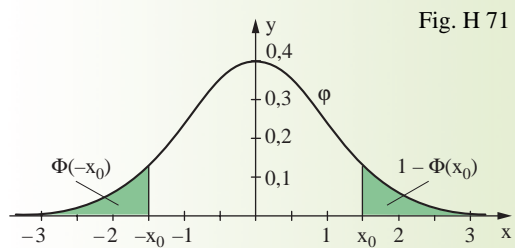


Fig. H 71

H 59

### Beispiel H 59:

Von einer Sorte Saatgetreide, für das eine Keimfähigkeit von 90% ausgewiesen ist, werden 1000 Körner ausgesät. Es soll ermittelt werden, mit welcher Wahrscheinlichkeit dann die folgenden Ereignisse A bis C eintreten:

- Der in der Formel enthaltene Summand 0,5 hat keinen mathematisch begründbaren Hintergrund. Sein Einfügen beruht auf Erfahrung. Die Formel wird auch ohne diesen Summanden 0,5 genutzt.

A = {es keimen höchstens 905 Körner};

B = {es keimen mindestens 890 Körner}

C = {es keimen mindestens 910 und höchstens 960 Körner}

**Lösung:**

Zufallsgröße X: zufällige Anzahl der keimenden Körner unter den 1000 ausgesäten;

$$X \sim B_{1000; 0,90}$$

**Erwartungswert:**  $\mu = 1000 \cdot 0,90 = 900$ ; Streuung  $\sigma^2 = 1000 \cdot 0,90 \cdot 0,10 = 90$

Die Faustregel  $n \cdot p \cdot q > 9$  des Satzes H 31 für die Anwendbarkeit der Tabellen der **Normalverteilung** zur näherungsweisen Berechnung obiger Binomialwahrscheinlichkeiten ist somit erfüllt.

- $P(A) = P(X \leq 905) \approx \Phi\left(\frac{905,5 - 900}{\sqrt{90}}\right) \approx \Phi(0,58) \approx 0,72$
- $P(B) = P(X \geq 890) = 1 - P(X < 890) = 1 - P(X \leq 889) \approx 1 - \Phi\left(\frac{889,5 - 900}{\sqrt{90}}\right) \approx 1 - \Phi(-1,11)$   
 $= 1 - (1 - \Phi(1,11)) = \Phi(1,11) \approx 0,87$
- $P(C) = P(910 \leq X \leq 960) \approx \Phi\left(\frac{960,5 - 900}{\sqrt{90}}\right) - \Phi\left(\frac{909,5 - 900}{\sqrt{90}}\right) \approx \Phi(6,38) - \Phi(1,00)$   
 $\approx 1,00000 - 0,8413 \approx 0,16$

Werden die Wahrscheinlichkeiten  $P(A)$ ,  $P(B)$  und  $P(C)$  mithilfe der bereits definierten **subi**-Funktion berechnet, so erhält man auf dem GTA-Bildschirm nebenstehende Angaben:

F1	F2	F3	F4	F5	F6
Algebra	Calc	Other	PrgmIO	Clear	a-z
nCr(n, i) · p <sup>i</sup> · (1 - p) <sup>n-i</sup> → bi(n, p, i)					
Done					
Σ <sub>i=k1</sub> bi(n, p, i) → subi(n, p, k1, k2)					
Done					
1 - subi(1000, .9, 906, 1000)					
.71578					
subi(1000, .9, 890, 1000)					
.865224					
subi(1000, .9, 910, 960)					
.158238					
subi(1000, .9, 910, 960)					
.158238					

Fig. H 72

Beispiel H 59 macht deutlich, dass für Berechnungen von  $B_{n; p}$  mit  $n \cdot p \cdot (1 - p) > 9$  eine Approximation der Binomialverteilung mithilfe des Satzes von DE MOIVRE-LAPLACE dann praktisch überflüssig wird, wenn Computer oder programmierbare Taschenrechner zugänglich sind.

## H 5.8 Normalverteilung

Der Grenzwertsatz von de Moivre-LAPLACE (Satz H 31) legt den Gedanken nahe, mittels der GAUSSschen Glockenkurve bzw. der auf der Menge der reellen Zahlen definierten Funktion  $\Phi$  nicht nur Näherungswerte von  $P(X \leq k) = B_{n; p}(\{0; 1; \dots; k\})$  für die *natürlichen* Zahlen  $k$  mit  $0 \leq k \leq n$  zu bestimmen, sondern für *beliebige reelle* Zahlen  $k$ . Da die Menge  $\mathbb{R}$  der reellen Zahlen nicht mehr „durchnummerierbar“ ist, sprengt eine solche Erweiterung freilich den Rahmen sowohl der endlichen als auch der diskreten Zufallsgrößen: Man gelangt zu **stetigen Zufallsgrößen**<sup>1)</sup> (oder auch *zufälligen reellen Zahlen*).

**Definition H 20:**

Eine Zufallsgröße  $X$  heißt **stetig**, wenn es eine nichtnegative Funktion  $f$  gibt, so dass

$$P(X \leq x) = F(x) = \int_{-\infty}^x f(t)dt \text{ für alle } x \in \mathbb{R} \text{ gilt. Die Funktion } f \text{ nennt man } \mathbf{Dichtefunktion} \text{ und } F$$

**Verteilungsfunktion** von  $X$ . Ihr Erwartungswert ist  $EX = \int_{-\infty}^{\infty} x \cdot f(x)dx$ .

H 20

<sup>1)</sup> Für die Betrachtung *stetiger* Zufallsgrößen wird das KOLMOGOROWSche Axiomensystem für endliche Zufallsgrößen verallgemeinert: Man erweitert den Ereignisraum  $2^\Omega$  (auch *Ereignisalgebra* oder kurz *Algebra* genannt) durch eine  $\sigma$ -Algebra, die Additivität durch eine  $\sigma$ -Additivität.

Für eine derartige stetige Zufallsgröße  $X$  gilt demzufolge auf Grund der Definition eines bestimmten Integrals als Grenzwert

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_a^b f(t) dt = F(b) - F(a) \text{ und damit } P(X = a) = 0.$$

H 21

Definition H 21:

Eine stetige Zufallsgröße  $X$  heißt **gleichverteilt über dem Intervall  $[a; b]$**  ( $a < b$ ), wenn für ihre Dichtefunktion gilt:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

$$\text{Dann ist } EX = \frac{a+b}{2}, D^2X = \frac{1}{12} (b-a)^2.$$

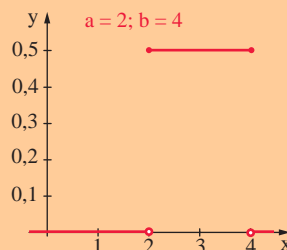


Fig. H 73

H 60

Beispiel H 60:

Anja verspricht zwischen 18.30 und 19.00 Uhr einzutreffen. Mit welcher Wahrscheinlichkeit  $p$  kommt sie erst innerhalb der letzten sieben Minuten an?

*Lösung:*

Können wir davon ausgehen, dass Anja kein Zeitintervall für ihr Eintreffen bevorzugen wird, so ergibt sich eine Gleichverteilung auf dem Intervall  $[0; 30]$ :

$$p = P(X \in [23; 30]) = \int_{23}^{30} \frac{1}{30-0} dx = \left[ \frac{1}{30} x \right]_{23}^{30} = 0,2\bar{3}.$$

Überträgt man die Gleichverteilung über einem Intervall auf den  $\mathbb{R}^2$ , so gelangt man zu einer weiteren Wahrscheinlichkeitsverteilung:

H 22

Definition H 22:

Ist die Ergebnismenge  $\Omega$  ein endliches Flächenstück mit dem Inhalt  $A_\Omega$ , so heißt die Funktion  $P$ , die jeder Teilfläche  $E$  (unabhängig von ihrer speziellen Form) von  $\Omega$  mit dem Inhalt  $A_E$  die

Wahrscheinlichkeit  $P(E) = \frac{A_E}{A_\Omega}$  zuordnet, **geometrische Wahrscheinlichkeitsverteilung**.

H 61

Beispiel H 61:

Es soll der Inhalt der in Fig. H 74 markierten Fläche bestimmt werden.

*Lösung:*

Die Simulation *stochastischer* Vorgänge ist die Grundlage der **MONTÉ-CARLO-Methode** zur numerischen Lösung verschiedener – auch *nicht-stochastischer* – mathematischer Probleme. So kann man beispielsweise

das bestimmte Integral  $I = \int_0^1 f(x) dx$  mit  $0 \leq f(x) \leq 1$  interpretieren als

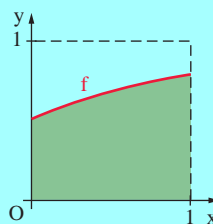


Fig. H 74

geometrische Wahrscheinlichkeit des Ereignisses  $E$ , dass ein „rein zufällig“ auf das Einheitsquadrat  $[0; 1] \times [0; 1]$  geworfener Punkt auf die unterhalb des Graphen von  $f$  liegende Fläche fällt. Die anhand von  $n$  Paaren von Pseudozufallszahlen aus  $[0; 1]$  auszählbare relative Häufigkeit  $h_n(E)$  dient als Schätzwert und damit Näherungswert des Integrals  $I$  bzw. des Inhalts der markierten Fläche.

Speziell mit einem Viertelkreisbogen als Graphen von  $f$  lässt sich  $\pi$  näherungsweise ermitteln: Die Wahrscheinlichkeit, dass ein zufällig ausgewählter Punkt obigen Einheitsquadrats auch im Viertelkreisbogen liegt, lässt sich einerseits auffassen als Quotient aus  $\frac{\pi \cdot 1^2}{4}$  und  $1^2$ , also  $\frac{\pi}{4}$ , und andererseits näherungsweise als Quotient aus der Anzahl  $m$  der Zufallszahlenpaare  $(x; y)$  mit  $x^2 + y^2 \leq 1$  und der Anzahl  $g$  aller aufgetretenen Zufallszahlenpaare  $(x; y)$  des Einheitsquadrats. Damit gilt:  $\frac{\pi}{4} \approx \frac{g}{m}$  bzw.  $\pi \approx \frac{4g}{m}$ . Das historisch erste Beispiel für das Anwenden der MONTE-CARLO-Methode ist das BUFFONSche<sup>1)</sup> Nadelwurfexperiment zur näherungsweisen Bestimmung von  $\pi$ , das der Pariser Akademie im Jahre 1733 vorgetragen worden ist.

Definition H 23:

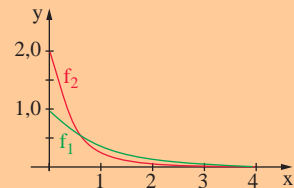
Eine stetige Zufallsgröße  $X$  heißt

**exponentiell verteilt mit dem Parameter  $\lambda$  ( $\lambda > 0$ ),**

wenn für ihre Dichtefunktion gilt:  $f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$ .

Dann ist  $EX = \frac{1}{\lambda}$ ,  $D^2X = \frac{1}{\lambda^2}$ .

Fig. H 75



H 23

Beispiel H 62:

Die zufällige Zerfallszeit  $X$  der Kerne eines radioaktiven Nuklids mit einer Halbwertszeit von 140 Tagen kann als exponentiell verteilt angesehen werden. Dabei versteht man unter Halbwertszeit diejenige Zeit, in deren Verlauf ein Kern dieses Nuklids mit der Wahrscheinlichkeit  $\frac{1}{2}$  zerfällt<sup>2)</sup>. In welcher Zeitdauer  $x$  (in Tagen) erfolgt ein Zerfall mit der Wahrscheinlichkeit 0,95?

*Lösung:*

Bestimmung von  $\lambda$ :

$$P(X \leq 140) = \int_0^{140} \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^{140} = 1 - e^{-140\lambda} = 0,5, \text{ also } e^{-140\lambda} = 0,5$$

$$\text{und damit } \lambda = \frac{\ln 0,5}{-140} \approx 0,00495.$$

Bestimmung von  $x$ :

$$P(X \leq x) \approx \int_0^x 0,00495 e^{-0,00495t} dt = 1 - e^{-0,00495x} = 0,95, \text{ also } e^{-0,00495x} = 0,05$$

$$\text{und damit } x = \frac{\ln 0,05}{-0,00495} \approx 605.$$

H 62

Definition H 24:

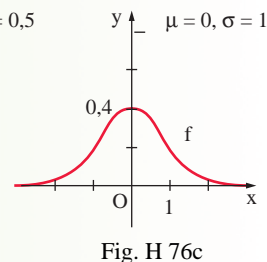
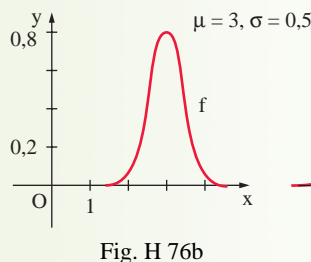
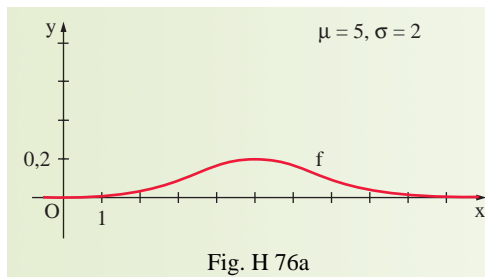
Eine (stetige) Zufallsgröße  $X$  heißt **normalverteilt mit den Parametern  $\mu$  und  $\sigma^2$**  oder  $(\mu; \sigma^2)$ -normalverteilt oder kurz  $N(\mu; \sigma^2)$ -verteilt, wenn

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad \text{mit } f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

H 24

1) Comte George-Louis Leclerc DE BUFFON (1707–1788); französischer Naturforscher und Mathematiker  
 2) Dies entspricht der physikalischen Sprechweise, wonach die Halbwertszeit angibt, in welcher Zeit sich jeweils die Hälfte der vorhandenen instabilen Atomkerne umwandelt.

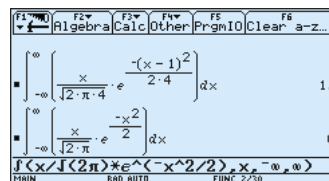
Die Fig. H 76a, H 76b und H 76c zeigen die Graphen von  $f$  für ausgewählte Werte von  $\mu$  und  $\sigma$ .



Die Dichtefunktion  $f$  besitzt die Symmetrieeigenschaft  $f(\mu - x) = f(\mu + x)$ . Der Graph von  $f$  ist demzufolge axialsymmetrisch zur Geraden mit der Gleichung  $x = \mu$ . Das heißt: Die  $x$ -Achse ist symmetrisch bezüglich  $x = \mu$  mit Wahrscheinlichkeiten bewichtet. Der Parameter  $\mu$  bildet den *Schwerpunkt* der Wahrscheinlichkeitsverteilung. Als Schwerpunkt einer Verteilung einer Zufallsgröße  $X$  wird er als **Erwartungswert EX** interpretiert. Demzufolge müsste

$$\int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu \quad \text{gelten, was man mit den Mitteln der}$$

Analysis (unter Verwendung von  $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ ) beweisen kann.



Nebenstehende Figur gibt die Resultate der Berechnung des

Integrals für die zwei Fälle  $\mu = 1, \sigma = 2$  (1. Zeile) und  $\mu = 0, \sigma = 1$  (2. Zeile) an.

Der Parameter  $\sigma^2$  charakterisiert die Wendestellen  $x_{W1} = \mu - \sigma$  und  $x_{W2} = \mu + \sigma$  des Graphen von  $f$ . Je größer  $\sigma^2$  ist, desto breiter ist der glockenförmige Graph von  $f$  und desto größer müsste demzufolge die **Streuung**  $D^2X$  der  $N(\mu; \sigma^2)$ -verteilten Zufallsgröße  $X$  sein.

H 32

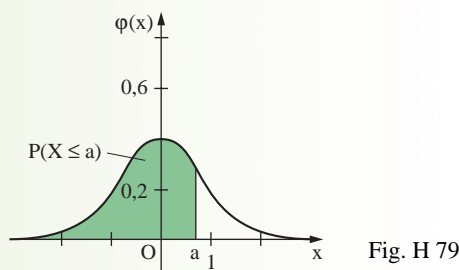
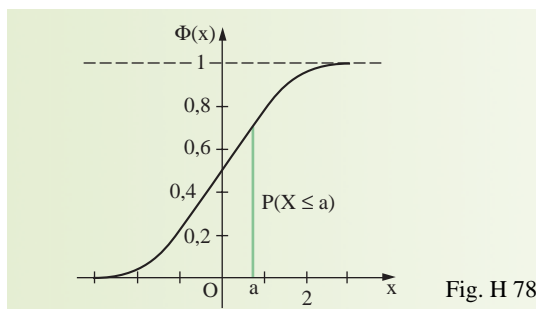
### Satz H 32: Erwartungswert und Streuung bei Normalverteilung

Für jede  $(\mu; \sigma^2)$ -normalverteilte Zufallsgröße  $X$  gilt  $EX = \mu$  und  $D^2X = \sigma^2$ .

Die Dichtefunktion  $f$  einer  $N(0; 1)$ -verteilten Zufallsgröße bezeichnet man mit  $\varphi$ .

Es gilt  $\varphi(-x) = \varphi(x)$ .

Die Verteilungsfunktion  $F$  einer  $N(0; 1)$ -verteilten Zufallsgröße wird mit  $\Phi$  bezeichnet und häufig **GAUSSsche Summenfunktion** genannt. Es gilt  $\Phi(-x) = 1 - \Phi(x)$ .



Die Wahrscheinlichkeit  $P(X \leq a) = \Phi(a) = \int_{-\infty}^a \varphi(x) dx$  kann als Inhalt der Fläche gedeutet werden, die durch den Graphen von  $\varphi$ , die  $x$ -Achse und die Gerade mit der Gleichung  $x = a$  begrenzt wird

(Fig. H 79). In Fig. H 78 ist die zur Abszisse  $x = a$  gehörende Ordinate  $\Phi(a)$  als die Wahrscheinlichkeit  $P(X \leq a)$  zu interpretieren.

Satz H 33:

Ist  $X$  eine  $N(\mu; \sigma^2)$ -verteilte Zufallsgröße, so ist die standardisierte Zufallsgröße  $Y = \frac{X - \mu}{\sigma}$   $N(0; 1)$ -normalverteilt. Es gilt  $P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$ .

H 33

Eine Zufallsgröße  $X \sim N(0; 1)$  bezeichnet man auch als **standardnormalverteilt**.

Da die Funktion  $\phi$  keine elementare Stammfunktion besitzt, sind die Funktionswerte von  $\Phi$  für Argumente  $x \geq 0$  tabelliert. Die dort enthaltenen Funktionswerte von  $\Phi$  wurden bestimmt auf Grund einer Reihendarstellung von  $\Phi$ :  $\Phi(x) = \frac{1}{2} + \frac{x}{\sqrt{2\pi}} \left(1 + \sum_{k=1}^n (-1)^k \frac{x^{2k}}{k! \cdot 2^k \cdot (2k+1)}\right)^{1)}$ .

Definiert man die  $n$ -te Partialsumme dieser Reihe als Funktions-term in Abhängigkeit von  $n$  und  $x$  auf dem Taschenrechner, so lässt sich  $\Phi(x)$  mit hinreichender Genauigkeit berechnen. Um die Funktionswerte  $\Phi(x)$  für negative Argumente zu bestimmen, nutzt man die Symmetrie der Dichtefunktion  $\phi$ , die durch die Gleichung  $\Phi(-x) = 1 - \Phi(x)$  beschrieben wird.

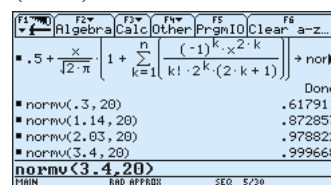


Fig. H 80

Beispiel H 63:

Zufallsgröße  $X \sim N(5; 2^2)$

H 63

- a) Die Wahrscheinlichkeit  $P(X \leq 6)$  ist sowohl mittels der zu  $X$  gehörenden Dichtefunktion  $f$  als auch mittels der Dichtefunktion  $\phi$  der standardisierten Normalverteilung grafisch zu veranschaulichen.

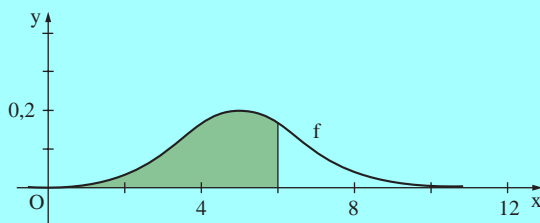


Fig. H 81

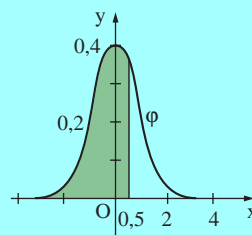


Fig. H 82

- b) Die Wahrscheinlichkeiten  $P(X \leq 6)$ ,  $P(X > 5.5)$ ,  $P(4 \leq X \leq 7)$ ,  $P(|X - 5| \leq 1.4)$  und  $P(|X - 5| > 1)$  sind zu berechnen und mittels  $\phi$  grafisch zu veranschaulichen.

allgemein  $N(\mu; \sigma^2)$

Beispiel  $N(5; 2^2)$

Veranschaulichung

$$P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$$\begin{aligned} P(X \leq 6) &= \Phi\left(\frac{6 - 5}{2}\right) \\ &= \Phi(0.5) \approx 0.691 \end{aligned}$$

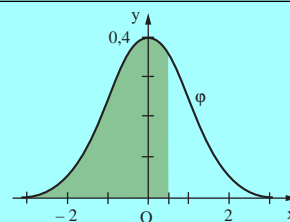


Fig. H 83

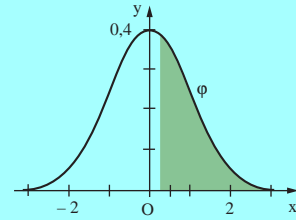
1) Diese Reihendarstellung von  $\Phi$  basiert auf der Integration der Potenzreihenentwicklung von  $\phi$ .



$$\begin{aligned}
 P(X > x) \\
 &= 1 - P(X \leq x) \\
 &= 1 - \Phi\left(\frac{x - \mu}{\sigma}\right)
 \end{aligned}$$

$$\begin{aligned}
 P(X > 5,5) \\
 &= 1 - P(X \leq 5,5) \\
 &= 1 - \Phi\left(\frac{5,5 - 5}{2}\right) \\
 &= 1 - \Phi(0,25) \approx 0,401
 \end{aligned}$$

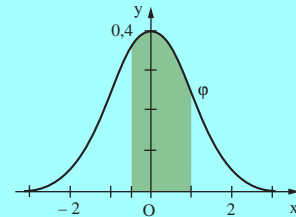
Fig. H 84



$$\begin{aligned}
 P(x_1 \leq X \leq x_2) \\
 &= P(X \leq x_2) - P(X < x_1) \\
 &= \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right)
 \end{aligned}$$

$$\begin{aligned}
 P(4 \leq X \leq 7) \\
 &= \Phi\left(\frac{7-5}{2}\right) - \Phi\left(\frac{4-5}{2}\right) \\
 &= \Phi(1) - \Phi(-0,5) \approx 0,533
 \end{aligned}$$

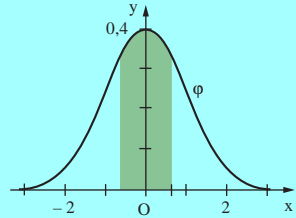
Fig. H 85



$$\begin{aligned}
 P(|X - \mu| \leq x) \\
 &= P(-x \leq X - \mu \leq x) \\
 &= P(-x + \mu \leq X \leq x + \mu) \\
 &= \Phi\left(\frac{x}{\sigma}\right) - \Phi\left(-\frac{x}{\sigma}\right) \\
 &= \Phi\left(\frac{x}{\sigma}\right) - (1 - \Phi\left(\frac{x}{\sigma}\right)) \\
 &= 2 \cdot \Phi\left(\frac{x}{\sigma}\right) - 1
 \end{aligned}$$

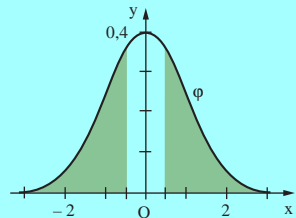
$$\begin{aligned}
 P(|X - 5| \leq 1,4) \\
 &= 2 \cdot \Phi\left(\frac{1,4}{2}\right) - 1 \\
 &= 2 \cdot \Phi(0,7) - 1 \\
 &\approx 0,516
 \end{aligned}$$

Fig. H 86



$$\begin{aligned}
 P(|X - \mu| > x) \\
 &= 1 - P(|X - \mu| \leq x) \\
 &= 1 - (2 \cdot \Phi\left(\frac{x}{\sigma}\right) - 1) \\
 &= 2 - 2 \cdot \Phi\left(\frac{x}{\sigma}\right) \\
 &= 2 \cdot (1 - \Phi\left(\frac{x}{\sigma}\right))
 \end{aligned}$$

$$\begin{aligned}
 P(|X - 5| > 1) \\
 &= 2 \cdot (1 - \Phi\left(\frac{1}{2}\right)) \\
 &= 2 \cdot (1 - \Phi(0,5)) \\
 &\approx 0,617
 \end{aligned}$$



## H 64

## Beispiel H 64:

Einem Messprotokoll von 125 neun-jährigen Kiefern einer Versuchsfläche lässt sich die nebenstehende Strichliste und das ihr entsprechende Streifendiagramm entnehmen.

Kann man die Verteilung der Höhen bei diesen Kiefern als normalverteilt ansehen? Um diese Frage zu beantworten, ist in das Streifendiagramm die Dichtefunktion  $f$  derjenigen Normalverteilung zu skizzieren, die gegebenenfalls als

Näherung der vorgegebenen Verteilung anzusehen ist.

Höhe in cm	Klassenmitten $x_i$	Strichliste
[50; 70[	60	I
[70; 90[	80	I
[90; 110[	100	II
[110; 130[	120	IIII IIII
[130; 150[	140	IIII IIII IIII
[150; 170[	160	IIII IIII IIII IIII II
[170; 190[	180	IIII IIII IIII IIII IIII IIII
[190; 210[	200	IIII IIII IIII IIII IIII II
[210; 230[	220	IIII IIII
[230; 250[	240	IIII I
[250; 270[	260	III

*Lösung:*

Die Strichliste bzw. das Streifendiagramm der Verteilung der Baumhöhen zeigt in groben Zügen den Verlauf einer Glockenkurve, die mit einer Normalverteilung  $N(\mu, \sigma^2)$  verglichen werden soll. Als ihren Erwartungswert  $\mu$  wählen wir das arithmetische Mittel

$$\begin{aligned}\bar{x}_{125} &= \frac{1}{125} \sum_{i=1}^{11} x_i \cdot H_{125}(\{x_i\}) \\ &= \frac{1}{125} (1 \cdot 60 + 1 \cdot 80 + 2 \cdot 100 + 9 \cdot 120 + 15 \cdot 140 + 22 \cdot 160 \\ &\quad + 30 \cdot 180 + 27 \cdot 200 + 9 \cdot 220 + 6 \cdot 240 + 3 \cdot 260) = 176,32\end{aligned}$$

und als ihre Streuung  $\sigma^2$  die empirische Streuung

$$\begin{aligned}\sigma_{125-1}^2 &= \frac{1}{125-1} \sum (x_i - \bar{x}_{125})^2 \cdot H_{125}(\{x_i\}) \\ &= \frac{1}{124} [(60 - 176,32)^2 \cdot 1 + (80 - 176,32)^2 \cdot 1 + \dots + (260 - 176,32)^2 \cdot 3] \approx 1344,4 \approx 36,7^2\end{aligned}$$

Um den Graphen der Dichtefunktion von  $N(176,32; 36,7^2)$  in das gegebene Streifendiagramm einzutragen, ist zu beachten, dass dieses Streifendiagramm kein Histogramm ist. Damit man die dargestellten absoluten Häufigkeiten  $H_{125}(\{x_i\})$  als relative Häufigkeiten

$h_{125}(\{x_i\}) = \frac{1}{125} \cdot H_{125}(\{x_i\})$  interpretieren kann, müssten wir nur den Maßstab der Ordinatenachse ändern, die Ordinatenwerte auf ihren 125-ten Teil verkleinern. Somit wäre aber noch nicht gewährleistet, dass – wie bei einem Histogramm erforderlich – die Summe der Rechteckflächeninhalte den Wert 1 ergibt, da die Rechteckbreiten alle gleich der Klassenbreite 20 sind. Infolgedessen wird aus dem ursprünglichen Streifendiagramm ein Histogramm, wenn die Werte auf der Ordinatenachse insgesamt auf ihren  $(125 \cdot 20)$ -ten Teil gebracht werden. In dieses Histogramm skizzieren wir den Graphen der Dichtefunktion  $f$  mit

$$f(x) = \frac{1}{\sqrt{2\pi \cdot 36,7^2}} \cdot e^{-\frac{(x - 176,32)^2}{2 \cdot 36,7^2}}.$$

Mit dem GTA ist die oben gestellte Aufgabe beispielsweise dadurch zu lösen, dass man das Histogramm im Graphmodus *Function* anstatt durch den Graphen der Treppenfunktion mit den Funktionswerten  $H_{125}(\{x_i\})$  durch den der Treppenfunktion mit den Funktionswerten  $\frac{1}{125 \cdot 20} \cdot H_{125}(\{x_i\})$  darstellt und den Graphen von  $f$  in dasselbe Koordinatensystem zeichnet.

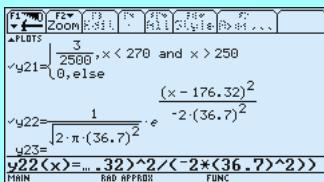


Fig. H 88

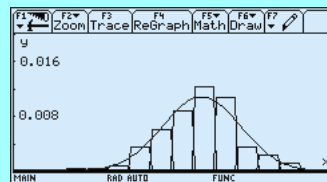


Fig. H 89

Beim Vergleich von Histogramm und Normalverteilung stellt man trotz des recht geringen Umfangs ( $m = 125$ ) der statistischen Ausgangsdaten keine zu große Abweichung fest. So weit die wenigen Messdaten erkennen lassen, entspricht die Verteilung der Höhen bei den neunjährigen Kiefern einer Normalverteilung.

Die Aussage der  $3\sigma$ -Regel (Satz H 24), die mittels der für alle Zufallsgrößen geltenden TSCHEBYSCHESCHEN Ungleichung gefunden worden war, kann für normalverteilte Zufallsgrößen verschärft werden: Es ist praktisch sicher, dass eine normalverteilte Zufallsgröße  $X \sim N(\mu; \sigma^2)$  nur Werte aus dem  $3\sigma$ -Intervall annimmt.

H 34

**Satz H 34: 3 $\sigma$ -Regel für normalverteilte Zufallsgrößen**

Eine  $N(\mu; \sigma^2)$ -verteilte Zufallsgröße  $X$  nimmt Werte

- im  $1\sigma$ -Intervall  $[\mu - 1\sigma; \mu + 1\sigma]$  mit einer Wahrscheinlichkeit von rund 0,68,
- im  $2\sigma$ -Intervall  $[\mu - 2\sigma; \mu + 2\sigma]$  mit einer Wahrscheinlichkeit von rund 0,95 und
- im  $3\sigma$ -Intervall  $[\mu - 3\sigma; \mu + 3\sigma]$  mit einer Wahrscheinlichkeit von mehr als 0,99 an.

**H 5.9 Zentraler Grenzwertsatz**

Während der anfänglichen stochastischen Untersuchungen zufälliger Vorgänge beschränkten wir uns auf das Modell einer *endlichen* Zufallsgröße. Angesichts der Endlichkeit jeder messbaren Realität erscheint eine Ausweitung auf nicht-endliche diskrete oder gar stetige Zufallsgrößen nicht recht sinnvoll. Die bisherigen Erfahrungen im Mathematikunterricht belegen aber sehr wohl den Nutzen des Arbeitens mit mathematischen Idealisierungen wie beispielsweise den geometrischen Körpern oder den stetigen Funktionen – und auch der Schritt zu den stetigen Zufallsgrößen erlaubt Zugänge zu sonst kaum lösaren Problemen.

Der Grenzwertsatz von DE MOIVRE-LAPLACE eröffnete – historisch gesehen – erst die Möglichkeit für ein praktikables Berechnen von langen, aber endlichen BERNOULLI-Ketten, und zwar dadurch, dass sie näherungsweise durch eine normalverteilte, also eine stetige Zufallsgröße beschrieben wurden. Dieser Grenzwertsatz für Summen von unabhängigen Zufallsgrößen beschränkt sich auf identisch verteilte Zufallsgrößen und zusätzlich auf BERNOULLI-Größen. Es lässt sich jedoch beweisen, dass die Summe von  $n$  unabhängigen, aber beliebig verteilten Zufallsgrößen (unter ansonsten nur noch schwachen zusätzlichen Bedingungen, die bei Anwendungsproblemen fast immer erfüllt sind) angenähert normalverteilt ist und die entsprechende standardisierte „Summengröße“ der  $N(0; 1)$ -Verteilung genügt.

H 35

**Satz H 35: Zentraler Grenzwertsatz**

Besitzen die voneinander unabhängigen Zufallsgrößen  $X_i$  für  $i \in \{1; 2; \dots; n\}$  jeweils sowohl einen endlichen Erwartungswert  $EX_i$  als auch eine endliche Streuung  $D^2X_i$ , so gilt

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n EX_i}{\sqrt{\sum_{i=1}^n D^2X_i}} < b\right) = \int_a^b \phi(x) dx.$$

Dieser Zentrale Grenzwertsatz gibt die theoretische Begründung dafür, dass so viele in der Praxis auftretende Häufigkeitsverteilungen mehr oder weniger glockenförmig sind. Diese Glockenform ist immer dann zu erwarten, wenn die betrachtete Zufallsgröße  $X$  ihre Werte unter dem Einfluss vieler unabhängiger und additiv wirkender Faktoren annimmt. Für derartige endliche Zufallsgrößen lassen sich also die entsprechenden Wahrscheinlichkeiten mit praktisch hinreichender Genauigkeit mittels der tabellierten stetigen  $N(0; 1)$ -Verteilung bestimmen.