Kapitel 1

Grundstrukturen

1.1 Motivation

Ziel: Lösung folgender Problemtypen: ↓2.4.07

1) Einfache (kombinatorische) Modelle: Mit welcher W'keit (Wahrscheinlichkeit) ist die Augenzahl beim Wurf von drei Würfeln größer als 10? Was ist die W'keit für 6 Richtige im Lotto?

   Bedingte W'keiten: Von zwei Münzen ist eine manipuliert. Man nimmt eine und wirft sie N mal, wobei k mal „Kopf" auftritt. Mit welcher W'keit handelt es sich um die manipulierte Münze?

2) Zufallsvariablen, Erwartungswerte etc: Ein Gesamtsystem besteht aus N unabhängigen Teilsystemen, die pro Arbeitszyklus mit den W'keiten $p_1, \ldots, p_N$ ausfallen. Das Gesamtsystem ist funktionsfähig, wenn mindestens 90% seiner Teile arbeiten. Wieviele Zyklen wird das System „im Durchschnitt" funktionsfähig bleiben?

3) Grenzwertsätze: Am Wahltag haben die Parteien bislang 42%, 40%, 7%, 6% der Stimmen erhalten (Hochrechnung vor Ende der Auszählung), der Nachrichtensprecher behauptet, die Stimmverteilung würden mit dem Endergebnis bis auf 1% übereinstimmen. Macht so eine Aussage Sinn (kritisches Bewusstsein)? Können die Grünen noch die absolute Mehrheit erlagen? Wenn ja, mit welcher W'keit? (Das Problem hierbei ist, dass das Wahlverhalten des Einzelwählers unbekannt ist. Durch Grenzwertsätze kann man dennoch Aussagen machen.)

Motivation zur Definition eines „stochastischen Modells":

Beispiel 1.1: Bei einem „fairen“ Würfel sind die Ergebnisse $1, \ldots, 6$ eines Wurfs „gleichwahrscheinlich“. Betrachte die Ereignisse:

1) Ich würfle eine gerade Zahl.
2) Ich würfle mindestens eine 3.
3) Ich würfle 2 mal, die Augensumme ist 10.

Wie wahrscheinlich sind diese Ereignisse? Intutitiv (bei gleichwahrscheinlichen „Elementarereignissen“):

$$P(\text{Ereignis}) \;=\; \frac{\text{Anzahl aller positiven Fälle}}{\text{Anzahl aller möglichen Fälle}}$$

Das Symbol P (engl: probability) steht in dieser Veranstaltung immer für „W'keit“.
Mathematisierung:

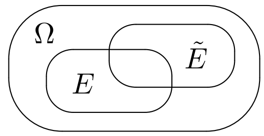
1) Menge aller möglichen Fälle $\Omega_1 = \{1, 2, \ldots, 6\}$.
   Ereignis E_1 = Menge aller positiven Fälle = $\{2, 4, 6\}$:
   $$\Rightarrow\; P(E_1) = \frac{|E_1|}{|\Omega_1|} = \frac{3}{6} = \frac{1}{2}\ .$$
2) Menge aller möglichen Fälle $\Omega_2 = \{1, 2, \ldots, 6\}$.
   Ereignis E_2 = Menge aller positiven Fälle = $\{3, 4, 5, 6\}$:
   $$\Rightarrow\; P(E_2) = \frac{|E_2|}{|\Omega_2|} = \frac{4}{6} = \frac{2}{3}\ .$$
3) Menge aller möglichen Fälle $\Omega_3 = \{(1, 1), (1, 2), \ldots, (6, 6)\}$.
   Ereignis E_3 = Menge aller positiven Fälle = $\{(5, 5), (6, 4), (4, 6)\}$:
   $$\Rightarrow\; P(E_3) = \frac{|E_3|}{|\Omega_3|} = \frac{3}{36} = \frac{1}{12}\ .$$

Strukturen, die diesen Beispielen gemeinsam sind:

i) Eine Menge Ω, die „alle möglichen Fälle“ beschreibt.
ii) Ein Ereignis E ist eine Teilmenge der Menge Ω aller möglichen Fälle.
iii) Es ist ein „W'keitsmaß“ P gegeben, welches jedem Ereignis E eine W'keit $P(E) \in [0, 1]$ zuordnet.

In den Beispielen 1.1 war das W'keitsmaß jeweils $P(E) \;=\; |E|/|\Omega|$ (das „kombinatorische W'keitsmaß“), es hat die Eigenschaft

$$P(E \cup \tilde{E}) = P(E) + P(\tilde{E}) - P(E \cap \tilde{E})\ .$$



1.2 Grundlegende Definitionen (Modelle)

↓4.4.07

Zunächst eine ganz formale Definition der mathematischen Situation, mit der W'keitsaufgaben modelliert werden. Diese abstrakte Definition wird dann in den folgenden Teilabschnitten („kombinatorische Modelle“ etc.) mit konkreten Inhalten und Beispielen gefüllt.

Definition 1.2: (Stochastisches Modell)

Ein **stochastisches Modell** *(oder auch „***W'keitsraum***“ oder auch „***Experiment***“) ist ein Tripel $(\Omega, \mathcal{E}, P)$ bestehend aus*

i) einer beliebigen Menge $\Omega \neq \emptyset$ (der **Stichprobenraum***),*

ii) einer Menge $\mathcal{E} \subset \mathcal{P}(\Omega)$ von Teilmengen von Ω (die **Menge aller möglichen Ereignisse***). Die Elemente von $\mathcal{E}$ (Teilmengen von Ω) heißen* **Ereignisse***. $\mathcal{E}$ muss die folgenden Eigenschaften haben:*

1) Zu jedem Ereignis in $\mathcal{E}$ liegt auch das **Komplementereignis** *in $\mathcal{E}$:*

$$E \in \mathcal{E} \quad \Rightarrow \quad \Omega \setminus E \in \mathcal{E}.$$

2) Die leere Menge $\emptyset$ (und wegen 1. auch Ω) sind Ereignisse:
$\emptyset$ = „das unmögliche Ereignis“, Ω = „das sichere Ereignis“.

3) Die Vereinigung endlich vieler oder auch abzählbar ∞ vieler Ereignisse ist wieder ein Ereignis:

$$E_1, E_2, \ldots \in \mathcal{E} \quad \Rightarrow \quad \cup_i E_i \in \mathcal{E}.$$

iii) einer Abbildung $P : \mathcal{E} \mapsto [0, 1]$ (das **W'keitsmaß***) mit den Eigenschaften*

1) $P(\emptyset) = 0$, $P(\Omega) = 1$

2) Seien $E_1, E_2, \ldots \in \mathcal{E}$ disjunkt, d.h. $E_i \cap E_j = \emptyset$ für $i \neq j$. Dann gilt (auch für abzählbar ∞ viele Ereignisse):

$$P(\cup_i E_i) = \sum_i P(E_i) \qquad (\text{„}\sigma\text{-\textbf{Additivität}“}).$$

Bemerkung 1.3: *In der mathematischen Literatur und in weiterführenden Vorlesungen zur Stochastik wird eine Menge $\mathcal{E} \subset \mathcal{P}(\Omega)$ mit den Eigenschaften 1.2.ii.1 - 1.2.ii.3 eine* **Sigma-Algebra** *in der Menge Ω genannt.*

Bemerkung 1.4: *In dieser Vorlesung reicht es meist aus, sich unter $\mathcal{E}$ die Potenzmenge $\mathcal{P}(\Omega)$ (also die Menge aller Teilmengen) von Ω vorzustellen, d.h.,* jede *Teilmenge von Ω ist ein zulässiges Ereignis. Trivialerweise erfüllt $\mathcal{E} = \mathcal{P}(\Omega)$ die in Definition 1.2.ii geforderten Eigenschaften.*

In einigen Modellen ist die Einschränkung auf gewisse Teilmengen aus rein technischen Gründen *nötig: wenn W'keiten durch Integrale gegeben sind, dürfen nur Teilmengen betrachtet werden, über denen Integrale definiert werden können.*
Mit den in Definition 1.2.ii geforderten Eigenschaften von $\mathcal{E}$ folgen automatisch zahlreiche weitere Eigenschaften, z.B.:

$$E_1, E_2 \in \mathcal{E} \quad \Rightarrow \quad \Omega \setminus E_1 \text{ und } \Omega \setminus E_2 \in \mathcal{E} \quad \Rightarrow \quad (\Omega \setminus E_1) \cup (\Omega \setminus E_2) \in \mathcal{E}$$
$$\Rightarrow \quad \Omega \setminus \Big((\Omega \setminus E_1) \cup (\Omega \setminus E_2) \Big) = \Omega \setminus \Big(\Omega \setminus (E_1 \cap E_2) \Big) = E_1 \cap E_2 \in \mathcal{E} \ .$$

Damit sind nicht nur Vereinigungen, sondern auch Durchschnitte von (endlich vielen oder abzählbar unendlich vielen) Ereignissen wieder Ereignisse.

Bemerkung 1.5: *Das W'keitsmaß P werden wir (im ersten Teil der Vorlesung) stets explizit vorgeben (z.B. durch kombinatorisches Abzählen oder als Modellvorgabe). Später werden wir über Grenzwertsätze auch Aussagen bekommen, ohne das Maß genau zu kennen.*

Folgerung (einige allgemeingültige Rechenregeln) 1.6:

Aus den in Definition 1.2.iii gegebenen Eigenschaften von P folgen unmittelbar einige allgemeingültige Regeln für das Rechnen mit W'keiten:

a) *Für* disjunkte *Ereignisse ($E_1 \cap E_2 = \emptyset$) gilt in jedem Modell:*
$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \ .$$

b) *Allgemein gilt in jedem Modell:*
$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \ .$$

c) *Für komplementäre Ereignisse E und $\Omega \setminus E$ gilt in jedem Modell:*
$$P(\Omega \setminus E) = 1 - P(E) \ .$$

d) *Für $E_1 \subset E_2$ gilt[1] in jedem Modell $P(E_1) \leq P(E_2)$.*

[1] Achtung: es gilt nicht $P(E_1) < P(E_2)$ für $E_1 \subset E_2$, $E_1 \neq E_2$. Gegenbeispiel: Ich entscheide

Beweis: a) Folgt unmittelbar aus der Definition 1.2.
b) Benutze die disjunkten Zerlegungen

$$\begin{array}{rcl} E_1 \cup E_2 & = & (E_1 \setminus E_2) \cup (E_1 \cap E_2) \cup (E_2 \setminus E_1), \\ E_1 & = & (E_1 \cap E_2) \cup (E_1 \setminus E_2), \\ E_2 & = & (E_1 \cap E_2) \cup (E_2 \setminus E_1). \end{array}$$

Mit a) folgt

$$\begin{array}{rcl} P(E_1 \cup E_2) & = & P(E_1 \setminus E_2) + P(E_1 \cap E_2) + P(E_2 \setminus E_1), \\ P(E_1) & = & P(E_1 \cap E_2) + P(E_1 \setminus E_2), \\ P(E_2) & = & P(E_1 \cap E_2) + P(E_2 \setminus E_1). \end{array}$$

Subtrahiert man die zweite und dritte Gleichung von der ersten, erhält man

$$P(E_1 \cup E_2) - P(E_1) - P(E_2) = -P(E_1 \cap E_2).$$

c) folgt mittels

$$\Omega = E \cup (\Omega \setminus E) \quad \stackrel{a)}{\Rightarrow} \quad 1 = P(\Omega) = P(E) + P(\Omega \setminus E)\ .$$

d) folgt mittels

$$E_1 \subset E_2 \ \Rightarrow\ E_2 = E_1 \ \cup\ (E_2 \setminus E_1) \stackrel{a)}{\Rightarrow} P(E_2) = P(E_1) + P(E_2 \setminus E_1) \geq P(E_1)\ .$$

Q.E.D.

1.2.1 Kombinatorische Modelle

Eine kombinatorische Betrachtungsweise bietet sich für Probleme an, die aus endlich vielen *gleichwahrscheinlichen* Elementarereignissen bestehen (Beispiele 1.1):

Definition 1.7: (Kombinatorisches Modell)

Ein **kombinatorisches Modell** $(\Omega, \mathcal{E}, P)$, *auch* **„Laplace-Experiment“** *genannt, besteht aus*

i) einer endlichen Menge $\Omega = \{\omega_1, \ldots, \omega_n\}$,
ii) den Ereignissen $\mathcal{E} = \mathcal{P}(\Omega)$ = *Potenzmenge von* Ω,
iii) dem W'keitsmaß $P(E) = |E|/|\Omega|$ *(das „***Zählmaß***“).*

Die Elemente ω_i *des Stichprobenraums* Ω *(genauer: die einelementigen Teilmengen* $\{\omega_i\}$*) heißen „***Elementarereignisse***“.*

---

mich, einen Würfel durch $\Omega = \{1, \ldots, 7\}$ zu modellieren (das ist meine legitime Entscheidung: ist zwar unnötig kompliziert, aber zulässig). Damit dieses Modell realistisch ist, muss ich $P(\{7\}) = 0$ setzen, es gilt $P(\{1, \ldots, 6\}) = P(\{1, \ldots, 7\}) = 1$.

Man kann sich unschwer überlegen, dass dies in der Tat ein stochastisches Modell im Sinne der Definition 1.2 ist: die Ereignismenge $\mathcal{E} = \mathcal{P}(\Omega)$ und das Maß P erfüllen die in 1.2.ii und 1.2.iii geforderten Eigenschaften.

Die Definition 1.7.iii) liefert den sauberen mathematischen Kontext für das intuitive Konzept

$$P(\text{Ereignis}) \;=\; \frac{\text{Anzahl aller positiven Fälle}}{\text{Anzahl aller möglichen Fälle}}\ .$$

In konkreten Aufgabenstellungen gilt es, in der „Modellierungsphase" ein geeignetes Ω und ein Ereignis $E \subset \Omega$ zu konstruieren, das der Aufgabe entspricht. Die „mathematische" Herausforderung ist dann, die Anzahl der Elemente $|E|$ und $|\Omega|$ zu bestimmen, um die W'keit des Ereignisses zu berechnen.

Beispiel 1.8: Lottoziehung: Ich mache einen Tip $T = \{t_1, \ldots, t_6\}$ von 6 unterschiedlichen Zahlen t_i aus $\{1, \ldots, 49\}$. Eine Ziehung ist die Auswahl einer 6-elementigen Teilmenge der Menge $\{1, \ldots, 49\}$, also:

$$\Omega = \text{Menge aller Ziehungen „6 aus 49"} = \{Z \subset \{1, \ldots, 49\};\ |Z| = 6\}\ .$$

Alle Ziehungen sind intuitiv gleichwahrscheinlich. Das Ereignis E = „6 Richtige" bedeutet, dass die Ziehung mit meinem Tip T übereinstimmt, also: $E = \{T\}$. Mit $|\Omega| = \binom{49}{6} = \frac{49!}{6!\cdot 43!} = 13\,983\,816$ folgt $P(E) = 1/13\,983\,816$.

Hierbei ist die Aussage $|\Omega| = \binom{49}{6}$ direkt dem folgenden Hilfssatz zu entnehmen:

Hilfssatz 1.9: (Hilfsmittel für kombinatorische Aufgaben)

a) *Es gibt $n!$ unterschiedliche Anordnungen („Permutationen") von n unterscheidbaren Objekten.*

b) *Es gibt $\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$ verschiedene m-elementige Teilmengen einer Menge mit n Elementen.*

c) *(Verallgemeinerung von b)) Es gibt*

$$\frac{n!}{n_1!\ \cdot\ n_2!\ \cdot\ \ldots\ \cdot\ n_m!}$$

verschiedene Möglichkeiten, eine Menge mit $n = n_1 + \cdots + n_m$ Elementen in m durchnummerierte Teilmengen zu jeweils $n_1, \ldots, n_m$ Elementen zu zerlegen (die n_i dürfen dabei auch 0 sein, wobei $0! = 1$ gesetzt wird).

d) *Es gibt $\binom{n+m-1}{n}$ Möglichkeiten, n nicht unterscheidbare Objekte auf m durchnummerierte Boxen aufzuteilen (wobei Boxen leer bleiben können).*

Beweisskizze: a) Es gibt n Möglichkeiten, das „erste Element" der Anordnung zu wählen. Dann gibt es $n-1$ Möglichkeiten, das „zweite Element" der Anordnung aus den verbleibenden Objekten zu wählen. Danach $n-2$ Möglichkeiten für das nächste Element usw. Macht zusammen $n \cdot (n-1) \cdot (n-2) \cdot \ \cdots = n!$ Möglichkeiten.
c) Die n Elemente werden in eine Reihe gelegt, die an den ersten n_1 Stellen liegenden Elemente kommen in die erste Teilmenge, die nächsten n_2 Elemente in die zweite Teilmenge usw. Es gibt nach a) insgesamt $n!$ Möglichkeiten, eine solche Reihe zu erzeugen. Werden die ersten n_1 Elemente als Menge interpretiert, so ist ihre Reihenfolge egal, d.h., nach a) sind $n_1!$ der Anordnungen zu identifizieren. Werden auch die nächsten n_2 Elemente als Menge interpretiert, so sind jeweils $n_1!\, n_2!$ der Anordnung zu identifizieren. Usw.
b) Spezialfall von c): zerlege die Menge in zwei Teilmengen mit jeweils $n_1 = m$ und $n_2 = n - m$ Elementen. Nach c) gibt es $n!/m!/(n-m)!$ Möglichkeiten.
d) Ich lege die Objekte in eine Reihe, lege ganz links und ganz rechts einen „Trennbalken" und sortiere $m-1$ weitere Trennbalken zwischen den Objekten ein. Der Raum zwischen zwei Trennbalken wird als Box angesehen (es liegen nun m Boxen vor). Da ich die Objekte nicht unterscheiden kann, ist die Aufteilung auf die Boxen eindeutig durch die Lage der Trennbalken gegeben. Es gibt nach a) $\binom{n+m-1}{m-1}$ Möglichkeiten, die Lage der „inneren" $m-1$ Trennbalken aus den insgesamt $n+m-1$ möglichen Positionen in der Reihe von n Objekten und $m-1$ Trennbalken zu wählen. Damit ist die Anzahl der Boxenaufteilungen

$$\binom{n+m-1}{m-1} \equiv \binom{n+m-1}{n} .$$

Q.E.D.

Beispiel 1.10: ↓11.4.07
Eine Urne enthält N Kugeln, davon S schwarz und $W = N - S$ weiß. Es werden n ($\leq N$) Kugeln *ohne Zurücklegen* gezogen. Mit welcher W'keit sind s der gezogenen Kugeln schwarz und $w = n - s$ weiß?
Lösung: Seien

$$\mathcal{W} = \{w_1, \ldots, w_W\} , \quad \mathcal{S} = \{s_1, \ldots, s_S\}$$

die weißen bzw. schwarzen Kugeln. Der Stichprobenraum sei die Menge aller Ziehungen einer n-elementigen Teilmenge aus der Urne $\mathcal{U} = \mathcal{S} \cup \mathcal{W}$:

$$\Omega = \{Z \subset \mathcal{S} \cup \mathcal{W};\ |Z| = n\} .$$

Das Ereignis $E =$ „s Kugeln sind schwarz, $w = n - s$ Kugeln sind weiß" besteht aus den folgenden Ziehungen:

$$E = \{Z \subset \mathcal{S} \cup \mathcal{W};\ Z = \tilde{\mathcal{S}} \cup \tilde{\mathcal{W}};\ \tilde{\mathcal{S}} \subset \mathcal{S};\ |\tilde{\mathcal{S}}| = s;\ \tilde{\mathcal{W}} \subset \mathcal{W};\ |\tilde{\mathcal{W}}| = w\} .$$

Nach 1.9.b) gilt $|\Omega| = \binom{N}{n}$. Es gibt es $\binom{S}{s}$ Möglichkeiten, die s-elementige Teilmenge $\tilde{\mathcal{S}}$ aus der S-elementigen Menge $\mathcal{S}$ zu wählen. Es gibt es $\binom{W}{w}$ Möglichkeiten, die w-elementige Teilmenge $\tilde{\mathcal{W}}$ aus der W-elementigen Menge $\mathcal{W}$ zu wählen. Damit folgt $|E| = \binom{S}{s}\binom{W}{w}$, also

$$P(E) \;=\; \frac{\binom{S}{s}\,\binom{W}{w}}{\binom{N}{n}}\;.$$

Statt von einer „Urne“ redet man auch von einer „**Gesamtpopulation**“ von N Elementen, die sich in eine „**Erfolgs**(teil)**population**“ von S Elementen und eine komplementäre „Misserfolgs(teil)population“ von $W = N - S$ Elementen zerlegt. Wählt man n Elemente ohne Zurücklegen aus, so ist die W'keit, dabei genau s Erfolge und $w = n - s$ Misserfolge zu ziehen, gegeben durch

$$P(\text{„genau } s \text{ Erfolge bei } n \text{ Ziehungen“}) \;=\; \frac{\binom{S}{s}\,\binom{N-S}{n-s}}{\binom{N}{n}}.$$

Man nennt die Abbildung $s \to P(\text{„genau } s \text{ Erfolge bei } n \text{ Ziehungen“})$ die „**hypergeometrische Verteilung**“ mit den vorgegebenen Parametern N (Populationsgröße), S (Erfolgspopulationsgröße), n (Anzahl der Ziehungen).

Z.B.: Mit welcher W'keit erhält man beim Skat s Asse? Es werden $n = 10$ Karten aus $N = 32$ Karten an mich verteilt. Unter den 32 Karten befinden sich $S = 4$ Asse. Die W'keit, s Asse zu erhalten, ist durch die hypergeometrische Verteilung gegeben:

$$P(\text{„}s \text{ Asse“}) = \frac{\binom{4}{s}\,\binom{32-4}{10-s}}{\binom{32}{10}}.$$

Intuitiv: es gibt $\binom{4}{s}$ Möglichkeiten, von den 4 Assen s zu erhalten, es gibt $\binom{28}{10-s}$ Möglichkeiten, die restlichen $10 - s$ Karten meiner Hand aus den verbleibenden 28 Karten zu erhalten, es gibt $\binom{32}{10}$ Möglichkeiten, insgesamt 10 Karten aus 32 zu erhalten.

Beispiel 1.11: Mit mir haben sich insgesamt 22 Leute für ein Turnier gemeldet. Es werden willkürlich 4 Gruppen $G_1, \ldots, G_4$ mit $|G_1| = |G_2| = 5$, $|G_3| = |G_4| = 6$ gebildet, die nacheinander die 1-te Runde austragen sollen. Es beginnt G_1 um 8^{oo}. Mit welcher W'keit muss ich früh aufstehen?
Lösung: a) Intuitiv und einfach: ich stelle mir vor, ich ziehe eine Kugel aus einer Urne mit 22 Kugeln, auf denen die Ziffern 1 bis 22 stehen. Die Ziffern 1 bis 5 bedeuten G_1, die Ziffern 6 bis 10 bedeuten G_2 usw.:

$$\Omega = \{1, 2, \ldots, 22\}, \quad E = \{1, 2, \ldots, 5\} \quad \Rightarrow \quad P(E) = \frac{|E|}{|\Omega|} = \frac{5}{22}.$$

b) Alternative (formalere) Lösung. Der Stichprobenraum sei die Menge aller möglichen Gruppeneinteilungen

$$\Omega = \{(G_1, G_2, G_3, G_4);\ G_1 \cup \cdots \cup G_4 = \{1, \ldots, 22\},\ |G_1| = |G_2| = 5, |G_3| = |G_4| = 6\}\ .$$

Nach 1.9.c) gilt $|\Omega| = \frac{22!}{5!\,5!\,6!\,6!} = 150\ 570\ 227\ 808$. In wievielen dieser Möglichkeiten ist eine bestimmte Person (ich, mit der Nummer 22, sagen wir) in der Gruppe G_1? Dies entspricht den Möglichkeiten, die anderen 21 Leute in Gruppen $\tilde{G}_1, G_2, G_3, G_4$ mit $|\tilde{G}_1| = 4$, $|G_2| = 5$, $|G_3| = |G_4| = 6$ aufzuteilen. Das Ereignis „ich werde in G_1 eingeteilt" entspricht damit

$$E = \{(\tilde{G}_1 \cup \{22\}, G_2, G_3, G_4);\ \tilde{G}_1 \cup \cdots \cup G_4 = \{1, \ldots, 21\},\\ |\tilde{G}_1| = 4, |G_2| = 5, |G_3| = |G_4| = 6\}\ .$$

Nach 1.9.c) gilt $|E| = \frac{21!}{4!\,5!\,6!\,6!} = 34\ 220\ 506\ 320$. Die W'keit, dass ich früh aufstehen muss, ist damit

$$P(E) = \frac{|E|}{|\Omega|} = \frac{21!}{4!\,5!\,6!\,6!}\ \frac{5!\,5!\,6!\,6!}{22!} = \frac{21!}{4!}\ \frac{5!}{22!} = \frac{5}{22} \approx 0.227...\ .$$

Beispiel 1.12: Eine Abzählübung: Wieviele Möglichkeiten gibt es, eine ganze Zahl n in genau m verschiedene Summanden $n = n_1 + \cdots + n_m$ mit $n_i \in \{0, 1, 2, \ldots\}$ zu zerlegen? Die Reihenfolge der Summanden soll dabei berücksichtigt werden.
Z.B. $n = 3, m = 2$: 4 Möglichkeiten $3 = 0 + 3 = 1 + 2 = 2 + 1 = 3 + 0$.
Z.B. $n = 4, m = 2$: 5 Möglichkeiten $4 = 0 + 4 = 1 + 3 = 2 + 2 = 3 + 4 = 4 + 0$.
Lösung: Dies lässt sich unmittelbar auf den Hilfssatz 1.9.d) zurückführen. Wir wiederholen das Argument:
Symbolisiere jeden Summanden n_i durch n_i aufgereihte Einsen. Diese „Einser-Reihen" werden dann mit Pluszeichen verbunden. Es bleibt damit das Problem, alle Möglichkeiten zu finden, eine Kette aus n Einsen und $m-1$ Pluszeichen anzuordnen. Da weder die Einsen noch die Pluszeichen unterschieden werden können, ist die Aufteilung durch die Position der $m-1$ Pluszeichen in der Gesamtkette von $n+m-1$ Objekten gegeben. Die Position der Pluszeichen entspricht einer Auswahl von $n-1$ Platznummern aus den möglichen Plätzen $1, 2, \ldots, n + m - 1$. Nach 1.9.b) ist die Anzahl der Möglichkeiten

$$\binom{n+m-1}{m-1} = \binom{n+m-1}{n}\ .$$

Merke (Zusammenfassung) 1.13:

Ein **kombinatorische Modell** *bietet sich in Situationen an, in denen alle Elementarereignisse des Zufallsexperimentes* **gleichwahrscheinlich** *sind: das Modell besteht aus einer Menge Ω endlich vieler Elemente, denen jeweils die gleiche W'keit $1/|\Omega|$ zugeschrieben wird. Für ein Ereignis $E \subset \Omega$ gilt*

$$P(E) = \frac{|E|}{|\Omega|}.$$

Sobald Ω und E vom Modellierer festgelegt sind, besteht die Herausforderung darin, durch geschicktes Abzählen $|E|$ und $|\Omega|$ zu ermitteln.

1.2.2 Diskrete (nicht-kombinatorische) Modelle

Bei endlich vielen (oder abzählbar unendlich vielen) Elementarereignissen, die *nicht gleichwahrscheinlich* sind, bietet sich folgendes diskrete Modell an:

Definition 1.14: (Diskretes Modell)

Ein **diskretes Modell** $(\Omega, \mathcal{E}, P)$ *besteht aus*

i) einer endlichen oder abzählbar unendlichen Menge Menge $\Omega = \{\omega_1, \omega_2, \dots\}$ *(die* ω_i *heißen wieder „***Elementarereignisse***"),*

ii) den Ereignissen $\mathcal{E} = \mathcal{P}(\Omega) =$ *Potenzmenge von* Ω*,*

iii) dem folgenden W'keitsmaß: Es sind W'keitswerte $P(\{\omega_i\})$ *für die Elementarereignisse vorgegeben, für die* $\sum_{\omega \in \Omega} P(\omega) = 1$ *gilt. Dann wird für ein Ereignis* $E \subset \Omega$ *die folgende W'keit definiert:*

$$P(E) = \sum_{\omega \in E} P(\omega) \ .$$

Die kombinatorischen Modelle 1.7 sind der Spezialfall, wo für alle $\omega \in \Omega$ dieselbe W'keit $P(\{\omega\}) = 1/|\Omega|$ gewählt wird.

Beispiel 1.15: Ein Würfel ist manipuliert, die W'keiten p_i, die Zahl i zu würfeln, seien

i	1	2	3	4	5	6
p_i	0.1	0.1	0.1	0.2	0.2	0.3

Mit welcher W'keit würfle ich mindestens eine 4? Antwort:

$$\Omega = \{1, \dots, 6\} \ ; \quad E = \{4, 5, 6\} \ ; \quad P(E) = p_4 + p_5 + p_6 = 0.7 \ .$$

Beispiel 1.16: Ich werfe zweimal mit einem Würfel. Mit welcher W'keit erhalte ich die Augensumme 10?

a) Ich behandle dies kombinatorisch:

$$\Omega = \{(1,1), (1,2), \dots, (6,5), (6,6)\}, \quad E = \{(4,6), (5,5), (6,4)\}$$

$$\Rightarrow \quad P(\text{„Augensumme 10"}) = \frac{|E|}{|\Omega|} = \frac{3}{36} = \frac{1}{12}.$$

b) Alternativ kann ich die Augensumme (eine Zahl zwischen 2 und 12) direkt als Ergebnis des Experiments „Doppelwurf" ansehen, also

$$\Omega = \{2, 3, \dots, 12\}.$$

Dieser Stichprobenraum ist mit 11 Elementen wesentlich kleiner als der in a) gewählte Stichprobenraum mit 36 Elementen. Allerdings habe ich hier nun das Problem, dass die

Elementarereignisse nicht mehr gleichwahrscheinlich sind. Ich muss mir die W'keiten $p_2, \ldots, p_{12}$ aus der kombinatorischen Sicht a) startend konstruieren, damit das Modell realistisch ist:

$$\begin{aligned}
p_2 &= P(\text{„Augensumme 2“}) = P(\{(1,1)\}) = \tfrac{1}{36}, \\
p_3 &= P(\text{„Augensumme 3“}) = P(\{(1,2),(2,1)\}) = \tfrac{2}{36} = \tfrac{1}{18}, \\
p_4 &= \quad \ldots \quad , \\
&\vdots \\
p_{10} &= P(\text{„Augensumme 10“}) = P(\{(4,6),(5,5),(6,4)\}) = \tfrac{3}{36} = \tfrac{1}{12}, \\
&\vdots
\end{aligned}$$

Es folgt $P(\text{„Augensumme 10“}) = P(\{10\}) = p_{10} = \frac{1}{12}$.

Merke (Zusammenfassung) 1.17:

Ein **diskretes Modell** *entspricht einem Experiment mit endlich vielen oder abzählbar unendlich vielen Ausgängen, die unterschiedliche W'keiten haben können. Kennt man die W'keiten der Elementarereignisse (die möglichen Ausgänge des Experiments), so ergibt sich die W'keit eines Ereignisses als die Summe der W'keiten aller Elementarereignisse, aus denen das Ereignis zusammengesetzt ist:*

$$P(E) = \sum_{\omega \in E} P(\{\omega\}).$$

1.2.3 Kontinuierliche Modelle

Bei nicht diskreten Situationen werden W'keiten typischerweise über Integrale definiert:

Definition 1.18: (Kontinuierliches Modell)

Ein **kontinuierliches Modell** $(\Omega, \mathcal{E}, P)$ *besteht aus*

i) einer Teilmenge Ω von $\mathbb{R}$ oder $\mathbb{R}^2$ oder ...,

ii) den Ereignissen $\mathcal{E} = \{E \subset \Omega;$ „man kann über E integrieren“$\}$,

iii) dem folgenden W'keitsmaß: es ist eine Funktion $\rho : \Omega \mapsto [0, \infty)$ gegeben: die **W'keitsdichte**. *Sie muss $\int_{x \in \Omega} \rho(x)\, dx = 1$ erfüllen. Dann wird für ein Ereignis $E \in \mathcal{E}$ die folgende W'keit definiert:*

$$P(E) = \int_{x \in E} \rho(x)\, dx = \int_{x \in \Omega} \chi_E(x) \cdot \rho(x)\, dx \ ,$$

wobei

$$\chi_E(x) = \begin{cases} 1 & \text{für } x \in E \\ 0 & \text{für } x \notin E \end{cases}$$

die **charakteristische Funktion** *der Teilmenge $E \subset \Omega$ ist.*

$\boxed{16.4.07\downarrow}$ **Bemerkung 1.19:** *In dieser Situation darf man in der Tat nicht alle Teilmengen von Ω als Ereignisse zulassen, da es sonst mathematische Schwierigkeiten beim Integrieren gibt. Hierauf werden wir uns aber nicht einlassen. Man stelle sich für $\Omega = \mathbb{R}$ als zulässige Ereignisse die Intervalle sowie Vereinigungen bzw. Schnitte von Intervallen vor. Bei uns werden nur Vereinigungen bzw. Schnitte endlich vieler disjunkter Intervalle vorkommen. Dann gilt für $E = [a_1, b_1] \cup [a_2, b_2] \cup \ldots$:*

$$\int_{x \in E} \rho(x)\ dx \ = \ \int_{a_1}^{b_1} \rho(x)\ dx \ + \ \int_{a_2}^{b_2} \rho(x)\ dx \ + \ \cdots \ .$$

Bemerkung 1.20: *Damit über die Dichte ρ integriert werden kann, muss sie gewisse Glattheitseigenschaften erfüllen, worauf hier aber nicht eingegangen wird. Man beachte aber, dass in Anwendungen unstetige Dichten mit Sprüngen eine wichtige Rolle spielen. Es reicht hier, sich ρ als „stückweise stetig“ mit endlich vielen Sprungstellen vorzustellen.*

Beispiel 1.21: Die W'keit, eine Maus im Abstand $r \in \Omega = [0, \infty)$ von ihrem Mauseloch zu finden, sei durch die Dichte

$$\rho(r) = 2\,c\ r\,e^{-c\,r^2}$$

mit einem positiven Parameter $c > 0$ gegeben. Der konstante Faktor $2\,c$ sorgt dafür, dass die Normierung $\int_{r \in \Omega} \rho(r)\,dr = \int_0^\infty \rho(r)\,dr = 1$ gilt. Für das Ereignis $[r_1, r_2]$ („die Maus befindet sich in einem Abstand zwischen r_1 und r_2 von ihrem Loch“) ergibt sich die W'keit

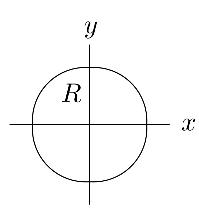
$$P([r_1, r_2]) = \int_{r_1}^{r_2} \rho(r)\,dr \ = \ e^{-c\,r_1^2} - e^{-c\,r_2^2} \ .$$

Kleine Werte von c entsprechen einer „mutigen Maus“, die mit großer W'keit weit weg von ihrem Loch zu finden ist: die W'keit

$$P([r, \infty)) = e^{-c\,r^2} \ ,$$

dass sie sich im Abstand $\geq r$ vom Loch aufhält, ist monoton fallend in c.

Beispiel 1.22: (Hier muss man wissen, wie man im $\mathbb{R}^2$ integriert).
Ein Schütze schießt auf eine im Ursprung von $\Omega = \mathbb{R}^2$ zentrierte Schießscheibe vom Radius R:



Die W'keit, dass ein kleines Flächenstück an der Stelle (x, y) getroffen wird, sei durch die Dichte

$$\rho(x,y) = \frac{c}{\pi}\ e^{-c\,(x^2+y^2)}$$

gegeben. Er trifft die Zielscheibe mit der W'keit

$$P(\text{Treffer}) = \int_{|(x,y)|\le R} \rho(x,y)\ dx\ dy$$

$$\stackrel{\text{(Polarkoordinaten)}}{=} \int_0^R \int_0^{2\pi} \frac{c}{\pi}\, e^{-c\,r^2}\ r\ dr\ d\phi = 1 - e^{-c\,R^2}.$$

Merke (Zusammenfassung) 1.23:

Ein **kontinuierliches Modell** *besteht aus einer Teilmenge Ω des $\mathbb{R}^n$ (in Zukunft bei uns i.W. $n = 1$), auf dem eine* **W'keitsdichte** $\rho(x) \ge 0$ *gegeben ist, die*

$$\int_{x\in\Omega} \rho(x)\ dx = 1$$

erfüllen muss. Die W'keit eines Ereignisses E ist dann

$$P(E) = \int_{x\in E} \rho(x)\ dx.$$

1.3 Bedingte Wahrscheinlichkeiten

1.3.1 Definitionen

Zur Motivation:

Beispiel 1.24: Ich werfe ein Münze: das Ergebnis ist „Kopf" oder „Zahl" mit den W'keiten $P(K) = P(Z) = 1/2$.

a) Mit welcher W'keit habe ich nach n Würfen n mal das Ergebnis „Kopf"?

b) Ich habe schon $n - 1$ mal „Kopf" geworfen. Mit welcher W'keit werfe ich bei nächsten mal „Kopf"?

Frage a) hat offensichtlich die kombinatorische Antwort $P(\text{n mal „Kopf"}) = 1/2^n$: es gibt bei n Würfen 2^n mögliche verschiedene Ergebnisse

$$\Omega = \{(\omega_1, \ldots, \omega_n)\ ;\ \omega_i \in \{K, Z\}\}\ ,$$

davon beschreibt nur ein einziges mein Ereignis $E = \{(K, \ldots, K)\}$. Diese W'keit wird für wachsendes n sehr schnell sehr klein.

Frage b) ist offensichtlich intuitiv unabhängig von n mit $P(K) = 1/2$ zu beantworten, denn die ersten $n-1$ Würfe haben nichts mit dem n-ten Wurf zu tun (Unabhängigkeit).

Es wird innerhalb desselben Experiments nach unterschiedlichen Dingen gefragt:

a) P(Ereignis: n mal „Kopf“)

b) P(Ereignis: n mal „Kopf“; Zusatzinformation:
„Kopf“ ist schon $n-1$ mal aufgetreten)

Frage: wie kann ich b) im Rahmen des für a) benutzten Modells Ω beschreiben? Das Experiment ist ja dasselbe, also sollte die Modellierung $(\Omega, \mathcal{E}, P)$ für beide Fälle benutzbar sein.

Wir betrachten die allgemeine Situation eines Modells $(\Omega, \mathcal{E}, P)$. Es geht darum: ich weiß, dass ein Ereignis eingetreten ist. Was nützt diese Information, wenn ich nach der W'keit für ein anderes Ereignis frage?

Definition 1.25: (bedingte W'keit)

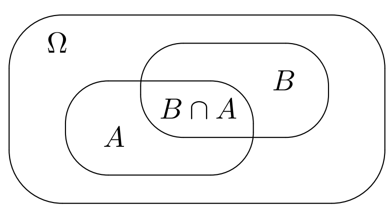
Seien $A, B \in \mathcal{E}$. Setze

$$P(B\,|A) \;=\; \begin{cases} \dfrac{P(A\cap B)}{P(A)} & \text{für } P(A) \neq 0\ , \\ 0 & \text{für } P(A) = 0 \end{cases}$$

und nenne es die **bedingte W'keit für das Ereignis B bei gegebenem Ereignis A**.

Interpretation (wichtige Modellierungshilfe) 1.26:

Dieser Begriff modelliert in der Tat die Situation, dass das Ereignis A eingetreten ist. Wenn ich nun nach der W'keit von B frage, so ist die Antwort das soeben definierte $P(B\,|A)$. Der Gedanke ist:



Ist A eingetreten, so befinden wir uns nicht mehr irgendwo in Ω, sondern in der Teilmenge A. Wenn B eintritt, so wissen wir, dass in Wirklichkeit das Ereignis $B \cap A$ eingetreten ist. Daher können wir A als neuen W'keitsraum interpretieren, welches jedem $B \subset \Omega$ die „reduzierte" W'keit $P(B \cap A)$ zuschreibt. Dabei müssen wir diese W'keit aber noch mit einem Faktor f multiplizieren, damit das neue W'keitsmaß $P_A(B) = f \cdot P(B \cap A)$ auf A wieder normiert ist. Da $P_A(A) = 1$ gelten muss, folgt $f = 1/P(A)$. Also:

$$P_A(B) = \frac{1}{P(A)} \cdot P(B \cap A) = P(B\,|A) \; .$$

Merke:

Die bedingte W'keit $P(B \mid A)$ entspricht einer „normalen" W'keit $P_A(B)$ auf einem geänderten (reduzierten) Modell $\Omega_A = A$, das nur aus Teilereignissen des „als eingetreten vorausgesetzten" Ereignisses A besteht.

Bemerkung 1.27: *Diese Interpretation ist auch mathematisch stichhaltig. Es gilt der leicht beweisbare Satz (Übungsaufgabe):*

> *Für einen W'keitsraum $(\Omega, \mathcal{E}, P)$ und gegebenes $A \in \mathcal{E}$ mit $P(A) > 0$ ist $(\Omega_A, \mathcal{E}_A, P_A)$ mit $\Omega_A = A$, $\mathcal{E}_A = \{E \cap A;\ E \in \mathcal{E}\}$ und $P_A(B) = P(B\,|A)$ wieder ein W'keitsraum nach Definition 1.2.*

Beispiel 1.28: Zurück zu Beispiel 1.24: n-facher Wurf einer Münze, $\Omega = \{(\omega_1, \ldots, \omega_n);\ \omega_i \in \{K, Z\}\}$ mit dem Zählmaß.

a) Das Ereignis $B =$ „es wurde n mal K geworfen" $= \{(\underbrace{K, \ldots, K}_{n})\}$ hat die W'keit $P(B) = |B|/|\Omega| = 1/2^n$.

b) Das Ereignis $A =$ „in den ersten $n-1$ Würfen wurde K geworfen" ist $A = \{(\underbrace{K, \ldots, K}_{n-1}, \omega_n);\ \omega_n \in \{K, Z\}\}$. In b) wird nach $P(B\,|A)$ gefragt:

$$P(B\,|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} = \frac{|B|/|\Omega|}{|A|/|\Omega|} = \frac{|B|}{|A|} = \frac{1}{2} \; .$$

Hier habe ich A als Teilmenge von Ω aufgefasst. Ich kann mich auch auf den Standpunkt stellen, dass ich, wenn A eingetreten ist, A als neuen Stichprobenraum auffasse, den ich vereinfachen würde:

$$\Omega_A = A = \{(\underbrace{K, \ldots, K}_{n-1}, \omega_n);\ \omega_n \in \{K, Z\}\} \quad \equiv \quad \tilde{\Omega} = \{\omega_n;\ \omega_n \in \{K, Z\}\}.$$

Hier wird nun auch formal deutlich, dass unter der Zusatzinformation „A ist eingetreten" das Gesamtexperiment „n-facher Münzwurf" äquivalent zu einem einzelnen Münzwurf $\tilde{\Omega}$ wird.

1.3.2 Folgerungen, „totale W'keit“ und der „Satz von Bayes“

Es gibt einige einfache Rechenregeln für bedingte W'keiten:

$\boxed{18.4.07\downarrow}$ **Satz 1.29:** (Eine Rechenregel für bedingte W'keiten)

Seien $E_1, E_2, \ldots$ Ereignisse in einem Modell $(\Omega, \mathcal{E}, P)$. Dann gilt

$$P(E_1 \cap \cdots \cap E_n) =$$
$$P(E_1)\, P(E_2|E_1)\, P(E_3|E_2 \cap E_1) \cdots P(E_n|E_{n-1} \cap \cdots \cap E_1)$$

Beweis: Einsetzen der Definition 1.25 liefert

$$P(E_1)\, P(E_2|E_1)\, P(E_3|E_2 \cap E_1) \cdots P(E_n|E_{n-1} \cap \cdots \cap E_1) =$$
$$P(E_1)\, \frac{P(E_2 \cap E_1)}{P(E_1)}\, \frac{P(E_3 \cap E_2 \cap E_1)}{P(E_2 \cap E_1)} \cdots \frac{P(E_n \cap E_{n-1} \cap \cdots \cap E_1)}{P(E_{n-1} \cap \cdots \cap E_1)}.$$

Fast alles kürzt sich heraus, es verbleibt $P(E_n \cap \cdots \cap E_1)$.

Q.E.D.

Beispiel 1.30: Mit welcher W'keit hat jeder der 3 Spieler beim Skat genau ein As?
Lösung: Beim Skat werden von 32 Karten jeweils 10 an die drei Spieler und 2 an den „Skat“ verteilt. Sei E_i das Ereignis „der i-te Spieler hat genau ein As“. Gefragt ist nach $E_1 \cap E_2 \cap E_3$. Die W'keit, dass der 1-te Spieler genau ein As hat, ist

$$P(E_1) = \frac{\binom{4}{1}\binom{28}{9}}{\binom{32}{10}}$$

(wähle 10 Karten aus den 32, davon eine aus den 4 Assen und 9 aus den restlichen 28 Karten). Die W'keit, dass der 2-te Spieler **ebenfalls** genau ein As hat, ist

$$P(E_2|E_1) = P(E_1 \cap E_2|E_1) = \frac{\binom{3}{1}\binom{19}{9}}{\binom{22}{10}}$$

(wähle 10 Karten aus den 22, die nicht an den ersten Spieler verteilt wurden. Davon eine aus den verbleibenden 3 Assen und 9 aus den restlichen 19 Karten). Die W'keit, dass der 3-te Spieler **ebenfalls** genau ein As hat, ist

$$P(E_3|E_2 \cap E_1) = P(E_1 \cap E_2 \cap E_3|E_2 \cap E_1) = \frac{\binom{2}{1}\binom{10}{9}}{\binom{12}{10}}$$

(wähle 10 Karten aus den 12, die nicht an die beiden ersten Spieler verteilt wurden. Davon eine aus den verbleibenden 2 Assen und 9 aus den restlichen 10 Karten). Damit folgt für $E_1 \cap E_2 \cap E_3 =$ „jeder der 3 Spieler hat genau ein As“:

$$P(E_1 \cap E_2 \cap E_3) = P(E_1)\, P(E_2|E_1)\, P(E_3|E_2 \cap E_1)$$
$$= \frac{\binom{4}{1}\binom{28}{9}}{\binom{32}{10}}\, \frac{\binom{3}{1}\binom{19}{9}}{\binom{22}{10}}\, \frac{\binom{2}{1}\binom{10}{9}}{\binom{12}{10}} = \frac{50}{899} \approx 0.0556\ .$$

Satz 1.31: („Formel der totalen W'keit")

Seien $U_1, \ldots, U_n \in \mathcal{E}$ disjunkte Ereignisse in einem Modell $(\Omega, \mathcal{E}, P)$, also $U_i \cap U_j = \emptyset$ für $i \neq j$. Dann gilt für jedes Ereignis $E \subset U_1 \cup \cdots \cup U_n$:

$$P(E) = \sum_{i=1}^{n} P(E\,|U_i)\,P(U_i).$$

Beweis:

$$P(E) = P\Big(E \cap \big(\bigcup_{i=1}^{n} U_i\big)\Big) = P\Big(\bigcup_{i=1}^{n} (E \cap U_i)\Big)$$

$$\stackrel{\text{(disjunkt)}}{=} \sum_{i=1}^{n} P(E \cap U_i) = \sum_{i=1}^{n} P(E\,|U_i)\; P(U_i)\;.$$

Hierbei wird benutzt, dass sich nach Folgerung 1.6.a) die W'keiten der disjunkten Mengen $E \cap U_i$ addieren.

Q.E.D.

Interpretation 1.32:

Für $E \subset U_1 \cup \cdots \cup U_n$ können die U_i als **vollständiges System disjunkter Ursachen** *interpretiert werden: das Ereignis E kann aus unterschiedlichen Gründen U_i eintreten. Ist bekannt, mit welcher W'keit das Ereignis E bei gegebenen Ursachen jeweils eintritt, so liefert die Formel der totalen W'keit die W'keit für E, wenn die W'keiten $P(U_i)$ bekannt sind, mit denen die Ursachen eintreten.*

Eine typische Anwendung dieses Satzes ist die Situation, wo mehrere Experimente hintereinandergeschaltet werden, wobei die möglichen Ausgänge U_i des ersten Experiments einen Einfluss auf das Ergebnis des zweiten Experiments haben.

Beispiel 1.33: Ich habe einen Topf mit 7 fairen Münzen und 3 Münzen mit 2 Köpfen.
Erstes Experiment: Ich entnehme dem Topf eine Münze.
Zweites Experiment: Ich nehme die gezogene Münze und werfe sie 3 Mal.
Frage: Mit welcher W'keit werfe ich im Gesamtexperiment 3 mal „Kopf"?

Lösung: Ich zerlege den Stichprobenraum Ω des zweiten Experiments in die beiden Teile U_1 und U_2, wobei U_1 die Situation beschreibt, dass ich im ersten Experiment eine faire Münze gezogen habe, während U_2 einer „Doppelkopf"-Münze entspricht. Offensichtlich gilt $P(U_1) = 7/10$, $P(U_2) = 3/10$.
Die W'keit, bei 3 Würfen 3 Köpfe zu werfen, ist für eine faire Münze 1/8 ($E = \{(K, K, K)\}$ bei insgesamt 8 Möglichkeiten $(K, K, K), (K, K, Z), \ldots, (Z, Z, Z)$), also

$$P(\text{„3 Köpfe in 3 Würfen"} \mid U_1) = \frac{1}{8}.$$

Für eine „Doppelkopf"-Münze ist die W'keit für 3 Köpfe bei 3 Würfen natürlich 1:

$$P(\text{„3 Köpfe in 3 Würfen"} \mid U_2) = 1.$$

Die Formel der totalen W'keit liefert für das Gesamtexperiment

$$\begin{aligned} & P(\text{„3 Köpfe in 3 Würfen"}) \\ = {} & P(\text{„3 Köpfe in 3 Würfen"} \mid U_1)\, P(U_1) + P(\text{„3 Köpfe in 3 Würfen"} \mid U_2)\, P(U_2) \\ = {} & \frac{1}{8} \cdot \frac{7}{10} + 1 \cdot \frac{3}{10} = \frac{31}{80} = 0.3875\ . \end{aligned}$$

Der folgende Satz beantwortet „die Schuldfrage". Ein Ereignis E kann aus unterschiedlichen Ursachen $U_1, U_2, \ldots$ heraus eintreten. Das Ereignis ist eingetreten. Mit welcher W'keit ist die Ursache U_i verantwortlich gewesen?

Satz 1.34: (Thomas Bayes, 1702–1761)

Seien $U_1, \ldots, U_n \in \mathcal{E}$ disjunkte Ereignisse in einem Modell $(\Omega, \mathcal{E}, P)$, also $U_i \cap U_j = \emptyset$ für $i \neq j$. Dann gilt für jedes Ereignis $E \subset U_1 \cup \cdots \cup U_n$:

$$P(U_i\,|E) \;=\; \frac{P(E\,|\,U_i)\ P(U_i)}{P(E)} \;\stackrel{(1.31)}{=}\; \frac{P(E\,|\,U_i)\ P(U_i)}{\sum_{j=1}^{n} P(E\,|\,U_j)\, P(U_j)}\ .$$

Beweis: Es gilt

$$P(U_i|E) = \frac{P(U_i \cap E)}{P(E)} = \frac{P(U_i)\ P(U_i \cap E)/P(U_i)}{P(E)} = \frac{P(U_i)\ P(E\,|\,U_i)}{P(E)}\ .$$

$P(E)$ ist über die Formel der totalen W'keit 1.31 durch $P(E \mid U_i)$ und $P(U_i)$ ausdrückbar.

Q.E.D.

Interpretation 1.35:

Bayes beantwortet folgende Frage: das Ereignis E ist eingetreten. Mit welcher W'keit war die Ursache U_i für das Eintreten von E verantwortlich?

Bemerkung 1.36: *Es gilt:*

$$\sum_{i=1}^{n} P(U_i\,|\,E) = \sum_{i=1}^{n} \frac{P(U_i)\ P(E\,|U_i)}{P(E)} = \frac{1}{P(E)} \sum_{i=1}^{n} P(U_i)\ P(E\,|U_i)$$

$$\stackrel{(1.31)}{=} \frac{P(E)}{P(E)} = 1\ .$$

Intuitiv: eine der Ursachen war mit Sicherheit für das Eintreten von E verantwortlich. Obige Formel ist oft eine nette Kontrollmöglichkeit, ob man in konkreten Aufgaben richtig gerechnet hat.

Beispiel 1.37: Ein Test auf eine Krankheit ist mit der W'keit 0.9 positiv, wenn ein Kranker getestet wird. Bei Gesunden fällt der Test mit der W'keit 0.1 positiv aus. Es ist bekannt, dass jeder 1450-te der Bevölkerung diese Krankheit hat. Eine Person wird getestet.

a) Mit welcher W'keit ist der Test positiv?

b) Der Test ist positiv. Mit welcher W'keit ist die Person krank? Mit welcher W'keit ist sie gesund?

c) Wird die Krankenkasse für diesen Test zahlen?

Lösung: Sei $+$ das Ereignis „der Test ist positiv". Das Ursachensystem für $+$ ist k = „die Person ist krank" und g = „die Person ist gesund". Es gilt $P(k) = 1/1450$, $P(g) = 1449/1450$. Weiterhin ist bekannt $P(+\,|\,k) = 0.9$ und $P(+\,|\,g) = 0.1$.

a) Nach Satz 1.31 gilt

$$P(+) = P(+\,|\,k)\;P(k)\;\;+\;\;P(+\,|\,g)\;P(g) = 0.9 \cdot \frac{1}{1450} + 0.1 \cdot \frac{1449}{1450} \approx 0.10055...\;.$$

b) Nach Satz 1.34 gilt

$$\begin{aligned} P(k\,|\,+) &= \frac{P(k)\,P(+\,|\,k)}{P(+)} = \frac{1}{1450}\,\frac{0.9}{0.10055...} &&\approx\quad 0.0062... \\ P(g\,|\,+) &= \frac{P(g)\,P(+\,|\,g)}{P(+)} = \frac{1449}{1450}\,\frac{0.1}{0.10055...} &&\approx\quad \underline{0.9938...}\;. \\ & && \qquad\;\; 1.0000... \end{aligned}$$

Das wirkt überraschend. Die Tatsache, dass die Wahr'scheinlichkeit einer positiv getesteten Person, gesund zu sein, ungleich viel höher ist, als krank zu sein, liegt daran, dass ohne Test die W'keit, gesund zu sein, fast 1 ist.

c) Der Test erhöht die W'keit $P(k) = 1/1450 \approx 0.00069$ nur um den recht kleinen Faktor $P(+\,|\,k)/P(+) = 0.9/0.10055... \approx 9.0$ auf die W'keit $P(k\,|\,+) \approx 0.0062$.

Die W'keit $P(g) = 1449/1450 \approx 0.99931...$ wird bei einem positiven Tests nur um den dicht bei 1 liegenden Faktor $P(+\,|\,g)/P(+) = 0.1/0.10055 \approx 0.9945...$ auf die W'keit $P(g\,|\,+) \approx 0.9938$ erniedrigt.

Der Test ist damit völlig wertlos.

Beispiel 1.38: Folgende Tabelle ist eine fiktive Aufschlüsselung der 1980er Bundestagswahl nach Konfessionen:

Bevölkerung-anteil		CDU	SPD	FDP	Grüne	Sonstige
46%	katholisch	54%	41%	3%	1%	1%
44%	evangelisch	45%	45%	6%	3%	1%
10%	sonstig	30%	38%	15%	10%	7%

Welcher Anteil der CDU-Wähler ist katholisch?

Lösung: Die gegebenen Informationen sind

$$P(\text{kath}) = 0.46, \quad P(\text{ev}) = 0.44, \quad P(\text{sonst}) = 0.10$$

sowie

$$P(\text{CDU}\,|\,\text{kath}) = 0.54\ ,\ P(\text{CDU}\,|\,\text{ev}) = 0.45\ ,\ \text{etc.}$$

Nach Satz 1.34 gilt

$$P(\text{kath}\,|\text{CDU}) = \frac{P(\text{kath})\ P(\text{CDU}\,|\,\text{kath})}{P(CDU)} =$$

$$\frac{P(\text{kath})\ P(\text{CDU}\,|\,\text{kath})}{P(\text{kath})\,P(\text{CDU}\,|\,\text{kath}) + P(\text{ev})\,P(\text{CDU}\,|\,\text{ev}) + P(\text{sonst})\,P(\text{CDU}\,|\,\text{sonst})}$$

$$= \frac{0.46 \cdot 0.54}{0.46 \cdot 0.54 + 0.44 \cdot 0.45 + 0.1 \cdot 0.3} = \frac{0.248...}{0.476...} \approx 52\%\ .$$

Der Nenner $P(CDU) = 0.476...$ entspricht dem damaligen Wahlergebnis der CDU von 47.6%.

Merke (Zusammenfassung) 1.39:

Die **Formel der totalen W'keit** *und* **der Satz von Bayes** *führen Ereignisse E auf unterschiedliche (disjunkte) Ursachen U_i zurück. Kennt man $P(E\,|\,U_i)$ sowie die W'keiten $P(U_i)$ der einzelen Ursachen, so kann mit der Formel der totalen W'keit $P(E)$ berechnet werden. Bayes beantwortet die „Schuldfrage“: welche Ursache war mit welcher W'keit für das Eintreten eines beobachteten Ereignisses verantwortlich?*

Typischer Einsatz: Modellierung mehrerer <u>nicht unabhängiger</u> Einzelexperimente.

1.4 Irrfahrten auf Graphen

23.4.07↓

Man kann viele W'keitsexperimente, die aus der Hintereinanderschaltung von Einzelexperimenten bestehen, als Irrfahrt auf einem (gerichteten gewichteten) Graphen interpretieren. Dies liefert eine nützliche Visualisierungsmöglichkeit, mit der W'keiten oft leicht graphisch zusammengesetzt werden können.

1.4.1 Wahrscheinlichkeitsbäume

Zunächst zur formalen Definition der speziellen Graphen, um die es hier geht:

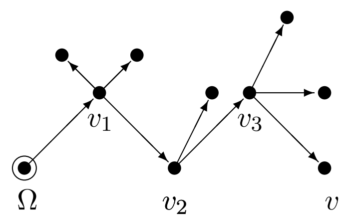
Definition 1.40: (Wurzelbäume)

Ein **nummerierter Wurzelbaum** (Ω, V, K) *mit* **Wurzel** Ω *besteht aus einer* **Knotenmenge** $V = \{\Omega, E_1, E_2, \dots\}$ *und einer* **Kantenmenge** $K \subset V \times V = \{(v,s);\ v,s \in V\}$, *so dass zu jedem* $v \in V \backslash \{\Omega\}$ *genau eine Kantenfolge* $k_1, k_2, \dots, k_k \in K$ *der Form*

$$k_1 = (\Omega, v_1), \dots, k_i = (v_{i-1}, v_i), \dots, k_k = (v_{k-1}, v)$$

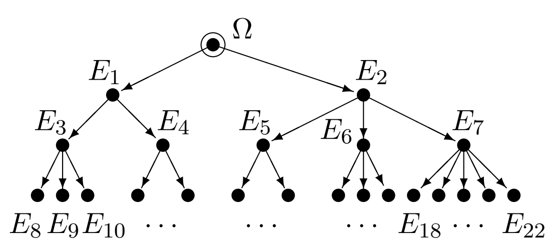
existiert (ein eindeutiger **Pfad** *der* **Länge** k *von der Wurzel* Ω *nach* v*).*

Beispiel:



Bemerkung 1.41: *Gewurzelte Bäume sind gerichtete (Kanten sind geordnete Paare) zusammenhängende (... es existiert ein Pfad ...) Graphen ohne geschlossene Wege (... genau ein Pfad ...). In einer Kante* $(v,s) \in K$ *heißt* v *„Vater“ und* s *„Sohn“ (der Abstand des Vaters zur Wurzel ist um 1 kleiner als der Abstand des Sohns zur Wurzel).*

Es bietet sich an, einen Wurzelbaum graphisch so darzustellen, dass alle Knoten mit dem selben Abstand zur Wurzel Ω jeweils eine „Schicht“ bilden. Die Wurzeln von Bäumen sind bei den Informatikern typischerweise oben `;-(`:



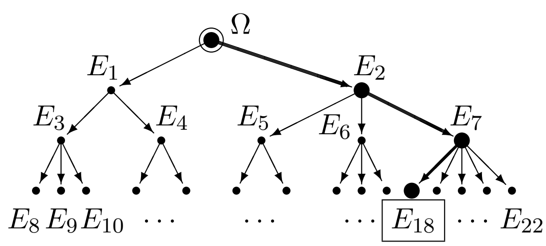
Eine wichtige Tatsache ist:

> Per Definition der Wurzelbäume gehört zu jedem Knoten genau ein Pfad von der Wurzel zum Knoten. Jeder Knoten kann also auch als Pfad auf dem Graphen interpretiert werden!

Beispiel: Der Knoten E_{18} im Baum oben kann als der Pfad

$$((\Omega, E_2), (E_2, E_7), (E_7, E_{18}))$$

mit den dicker eingezeichneten Kanten interpretiert werden:



Nun die Interpretation eines aus Teilexperimenten zusammengesetzten Gesamtexperiments durch einen solchen Baum:

Modellierungsregeln 1.42:

a) Ω ist der Stichprobenraum des Gesamtexperiments.

b) Alle weiteren Knoten sind Ereignisse des Gesamtexperiments, also Teilmengen von Ω.

c) Eine Kante von einem Knoten v („Vater“) zum Knoten s („Sohn“) wird als Durchführung eines Einzelexperiments interpretiert, das auf dem Stichprobenraum $v \subset \Omega$ ausgeführt wird und das Ergebnis (Ereignis) $s \subset v \subset \Omega$ liefert:

Jeder Sohn ist Teilmenge seines Vaters.

d) Die Kanten werden gewichtet mit den **Übergangsw'keiten**

$$P(s|v) = \frac{P(s \cap v)}{P(v)} \overset{(s \subset v)}{=} \frac{P(s)}{P(v)}.$$

wobei P das W'keitsmaß auf dem Stichprobenraum Ω des Gesamtexperiments ist. Nach Interpretation 1.26 dürfen wir diese bedingten W'keiten (in Ω) in der Tat als normale W'keiten im Stichprobenraum $v \subset \Omega$ ansehen, d.h., zur Berechnung der Übergangsw'keiten $P(s|v)$ können wir ein (in der Regel einfacher zu beschreibendes Einzelexperiment) v benutzen.

Die Durchführungen des Gesamtexperiments (Hintereinanderschaltung von n Teilexperimenten) sind die von der Wurzel des W'keitsbaums ausgehenden Pfade. Der Pfad $(\Omega, v_1, v_2, \ldots, v_{n-1}, v_n)$ über die Kanten (Ω, v_1), (v_1, v_2), …, (v_{n-1}, v_n) hat die Interpretation:

> *„Das auf dem Stichprobenraum Ω durchgeführte erste Einzelexperiment liefert das Ergebnis $v_1 \subset \Omega$. Das dann auf dem Stichprobenraum v_1 durchgeführte zweite Einzelxperiment liefert das Ergebnis $v_2 \subset v_1 \subset \Omega$. Usw.“*

In der Interpretation der Knoten des W'keitsbaums als Ereignisse im Gesamtexperiment gelten für jeden Pfad $(\Omega, v_1, v_2, \ldots, v_n)$ von der Wurzel Ω zum Knoten v_n die Inklusionen

$$v_n \subset v_{n-1} \subset \cdots \subset v_1 \subset \Omega,$$

also

$$v_n = v_n \cap v_{n-1} = v_n \cap v_{n-1} \cap v_{n-2} = \cdots = v_n \cap v_{n-1} \cap \cdots \cap v_1.$$

Damit berechnet sich nach Satz 1.29 die W'keit des Ereignisses $v_n \subset \Omega$ durch

$$P(v_n) = P(v_n \cap v_{n-1} \cap \ldots \cap v_1) =$$

$$\underbrace{P(v_1)}_{=P(v_1|\Omega)} \cdot P(v_2|v_1) \cdot P(v_3|\underbrace{v_2 \cap v_1}_{=v_2}) \cdot \ldots \cdot P(v_n|\underbrace{v_{n-1} \cap \cdots \cap v_1}_{=v_{n-1}}),$$

also:

$$P(v_n) = \underbrace{P(v_1|\Omega)}_{=P(v_1)} P(v_2|v_1)\, P(v_3|v_2) \, \cdots \, P(v_n|v_{n-1}).$$

Die Berechnung der W'keiten $P(v_n)$ geschieht also rein über die Übergangsw'keiten aller Kanten längs des Pfads von Ω nach v_n. Dies ergibt die erste „Pfadregel“:

Satz 1.43: (Pfadregel 1)

Die W'keit eines Knoten in einem W'keitsbaum, der den Modellierungsregeln 1.42 genügt, ist das Produkt aller Übergangsw'keiten der Kanten des Pfades, der von der Wurzel zu dem Knoten führt.

Die Erleichterung in der Modellierung von W'keitsexperimenten ist, dass die W'keit eines Ereignisses im Gesamtexperiment vollständig durch die W'keiten von Ereignissen in den (einfacher zu beschreibenden) Teilexperimenten berechenbar ist.

Beispiel 1.44: *Gabriel und Neumann* betrachteten im *Quarterly Journal of the Royal Meteorological Society*, vol. **83**, pp 375–380 (1957), folgendes einfache stochastische Modell für das Wetter in Israel in den Wintermonaten Dezember, Januar, Februar: ↓23.4.07

a) Ist es in einem Monat überwiegend trocken (T), so ist der Folgemonat mit der W'keit 3/4 überwiegend trocken und mit der W'keit 1/4 überwiegend nass (N).

b) Ist es in einem Monat nass, so ist der Folgemonat mit der W'keit 1/3 überwiegend trocken und mit der W'keit 2/3 überwiegend nass.

Wir haben einen trockenen Dezember. Mit welcher W'keit ist das Wetter im Februar trocken?

Der Stichprobenraum sei $\Omega = \{(Dez, Jan, Feb); Dez = T; Jan, Feb \in \{T, N\}\}$. Er beschreibt einen trockenen Dezember und die vier unterschiedlichen Wetterkombinationen für Januar und Februar. Wir betrachten den Baum

$$\Omega = \{(T, Jan, Feb); Jan, Feb \in \{T, N\}\}$$

$$\xleftarrow{3/4} \qquad \xrightarrow{1/4}$$

$$T_{Jan} = \{(T, T, Feb); Feb \in \{T, N\}\} \qquad N_{Jan} = \{(T, N, Feb); Feb \in \{T, N\}\}$$

$$\xleftarrow{3/4} \quad \xrightarrow{1/4} \qquad\qquad \xleftarrow{1/3} \quad \xrightarrow{2/3}$$

$$T^{(1)}_{Feb} = \{(T, T, T)\} \quad N^{(1)}_{Feb} = \{(T, T, N)\} \quad T^{(2)}_{Feb} = \{(T, N, T)\} \quad N^{(2)}_{Feb} = \{(T, N, N)\}$$

Die eingetragenen Übergangswahrscheinlichkeiten sind dabei die vorgegebenen Daten. Aus der Pfadregel 1 ergeben sich sofort die Wahrscheinlichkeiten für die Februar-Knoten, die die Pfade kodieren:

$$P(\{(T, T, T)\}) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}, \quad P(\{(T, T, N)\}) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16},$$

$$P(\{(T, N, T)\}) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}, \quad P(\{(T, N, N)\}) = \frac{1}{4} \cdot \frac{2}{3} = \frac{1}{6}.$$

Das Ereignis „der Februar ist trocken“ besteht aus den Ereignissen $\{(T, T, T)\}$ und $\{(T, N, T)\}$. Da diese disjunkt sind, sind die Wahrscheinlichkeiten für die Vereinigungsmenge zu addieren:

$$P(\text{„der Februar ist trocken“}) = P(\{(T, T, T\} \cup \{T, N, T\})$$

$$= P(\{(T, T, T\}) + P(\{T, N, T\}) = \frac{9}{16} + \frac{1}{12} = \frac{31}{48}.$$

Bei einem nassen Dezember würde sich ergeben:

$$\Omega = \{(N, Jan, Feb); Jan, Feb \in \{T, N\}\}$$

$$\xleftarrow{1/3} \qquad \xrightarrow{2/3}$$

$$T_{Jan} = \{(N, T, Feb); Feb \in \{T, N\}\} \qquad N_{Jan} = \{(N, N, Feb); Feb \in \{T, N\}\}$$

$$\xleftarrow{3/4} \quad \xrightarrow{1/4} \qquad\qquad \xleftarrow{1/3} \quad \xrightarrow{2/3}$$

$$T^{(1)}_{Feb} = \{(N, T, T)\} \quad N^{(1)}_{Feb} = \{(N, T, N)\} \quad T^{(2)}_{Feb} = \{(N, N, T)\} \quad N^{(2)}_{Feb} = \{(N, N, N)\}$$

Hiermit folgt

$$P(\{(N, T, T)\}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}, \quad P(\{(N, T, N)\}) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12},$$

$$P(\{(N, N, T)\}) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}, \quad P(\{(N, N, N)\}) = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

und somit:

$$P(\text{„der Februar ist trocken"}) = P(\{(N,T,T\} \cup \{N,N,T\})$$
$$= P(\{(N,T,T\}) + P(\{N,N,T\}) = \frac{1}{4} + \frac{2}{9} = \frac{17}{36}.$$

Beispiel 1.45: Aus der Urne |SAVE OUR SOULS| (= |AELOORSSSUUV|) werden ohne Zurücklegen drei Buchstaben gezogen. Mit welcher W'keit bilden diese Buchstaben in der gezogenen Reihenfolge das Wort SOS?

Wir betrachten als Einzelexperimente die drei einzelnen Züge. Der entsprechende W'keitsbaum ist:

|AELOORSSSUUV|

3/12

|AELOORSSUUV| <u>S</u> „erster gezogener Buchstabe kein S"

2/11

|AELORSSUUV| S<u>O</u> „zweiter gezogener Buchstabe kein O"

2/10

|AELORSUUV| SO<u>S</u> „dritter gezogener Buchstabe kein S"

Die Übergangsw'keiten ergeben sich dabei leicht folgendermaßen aus den Einzelexperimenten:

Zug 1: Die Wahrscheinlichkeit ist 3/12, aus einer Urne mit 12 Buchstaben, von denen 3 ein S sind, ein S zu ziehen.

Zug 2: Die Wahrscheinlichkeit ist 2/11, aus einer Urne mit 11 Buchstaben, von denen 2 ein O sind, ein O zu ziehen.

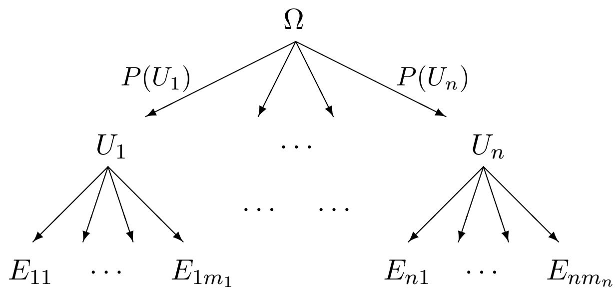
Zug 3: Die Wahrscheinlichkeit ist 2/10, aus einer Urne mit 10 Buchstaben, von denen 2 ein S sind, ein S zu ziehen.

Als Endergebnis erhält man über die Pfadregel 1:

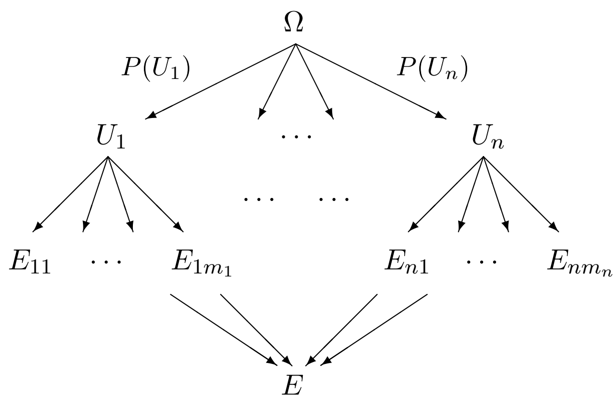
$$P(\text{„SOS"}) = \frac{3}{12} \cdot \frac{2}{11} \cdot \frac{2}{10} = \frac{1}{110}.$$

1.4.2 Allgemeinere Wahrscheinlichkeitsgraphen

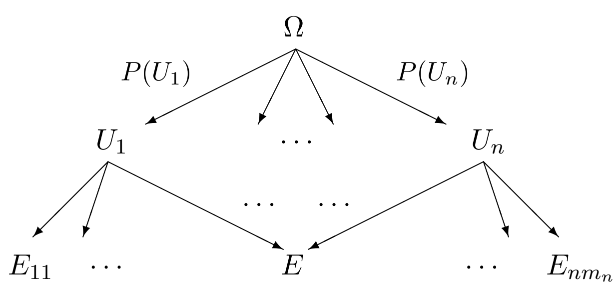
Wir verallgemeinern nun das graphisches Konzept des letzten Abschnitts, indem wir allgemeinere Graphen zulassen, die nicht unbedingt Bäume zu sein brauchen. Als Motivation betrachten wir zunächst ein zweistufiges Experiment:



Das interessierende Ereignis E, dessen W'keit berechnet werden soll, sei die Vereinigung von mehreren der Ereignisse E_{ij} der unteren Schicht des Baums:



Wir wollen das Ereignis E direkt von den U_i anfahren, ohne über die „Teilergebnisse“ E_{ij} laufen zu müssen:



Wir können nun nicht mehr verlangen, dass ein Ereignis auf nur einem Pfad von der Wurzel aus erreicht werden kann:

Definition 1.46: (W'keitsgraph)

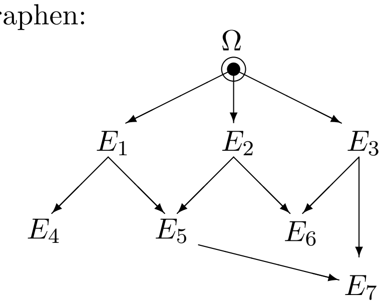
Ein **W'keitsgraph** (Ω, V, K) *mit* **Wurzel** Ω *besteht aus einer* **Knotenmenge** $V = \{\Omega, E_1, E_2, \dots\}$ *und einer* **Kantenmenge** $K \subset V \times V = \{(v,s);\ v,s \in V\}$*, so dass zu jedem* $s \in V \backslash \{\Omega\}$ *mindestens eine und maximal endlich viele Kantenfolgen* $k_1, k_2, \dots, k_k \in K$ *der Form*

$$k_1 = (\Omega, v_1), \dots, k_i = (v_{i-1}, v_i), \dots, k_k = (v_{k-1}, s)$$

existiert (Pfade von der Wurzel Ω *nach* s*).*

Die Bedingung „maximal endlich viele“ Pfade von Ω nach s impliziert, dass es keine Zyklen (geschlossene Wege) von einem Punkt s zurück zu sich selbst geben darf. Wäre ein solcher Zyklus vorhanden, könnte man auf beliebig vielen Pfade von Ω nach s laufen, indem man unterwegs den Zyklus beliebig oft durchläuft.

Beispiel eines W'keitsgraphen:



Hier die **Modellierungsanleitung** zur Darstellung eines aus der Hintereinanderausführung von Teilexperimenten zusammengesetzten Gesamtexperiments:

Modellierungsregeln 1.47:

a) *Ω ist der Stichprobenraum des Gesamtexperiments.*

b) *Alle weiteren Knoten sind Ereignisse des Gesamtexperiments, also Teilmengen von Ω.*

c) *Eine Kante von einem Knoten v („Vater“) zum Knoten s („Sohn“) wird als Durchführung eines Einzelexperiments interpretiert, das auf dem Stichprobenraum $v \subset \Omega$ ausgeführt wird und das Ergebnis (Ereignis) $s \subset \Omega$ liefert:*

Ein Sohn ist dabei Teilmenge der Vereinigung aller seiner Väter.

d) *Die Kanten werden gewichtet mit den* **Übergangsw'keiten**

$$P(s|v) = \frac{P(s \cap v)}{P(v)},$$

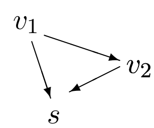
wobei P das W'keitsmaß auf dem Stichprobenraum Ω des Gesamtexperiments ist.

e) *Wir setzen weiterhin voraus:*

Väter mit einem gemeinsamen Sohn sind disjunkt.

Hierbei ist 1.47.c) die Verallgemeinerung der Modellierungsregel 1.42.c) („Jeder Sohn ist Teilmenge seines Vaters“) für Bäume. Beim Punkt d) ist zu beachten, dass wir (bei Söhnen mit mehreren Vätern) nicht mehr die Inklusion $s \subset v$ haben, also nicht $P(s \cap v)$ zu $P(s)$ vereinfachen können. Regel 1.47.e) ist neu hinzugekommen (für Bäume, wo jeder Sohn nur genau einen Vater besitzt, ist dieser Punkt inhaltslos).

Die Regeln c) und e) implizieren gewisse Einschränkungen an die Graphen bzw. an die Übergangsw'keiten. Beispielsweise kann es keine „Dreiecksbeziehungen“ im Graphen geben (ein Sohn eines Vaters kann nicht gleichzeitig ein Enkel dieses Vaters sein):



Im obigen Beispiel müssen v_1 und v_2 als gemeinsame Väter von s disjunkte Teilmengen von Ω sein. Die Übergangsw'keit längs der Kante von v_1 nach v_2 wäre damit

$$P(v_2|v_1) = \frac{P(v_1 \cap v_2)}{P(v_1)} = \frac{P(\emptyset)}{P(v_1)} = 0.$$

Die Kante (v_1, v_2) ist damit überflüssig und kann entfallen (man kann sie natürlich auch formal mit der Übergangsw'keit 0 stehen lassen).

Definition 1.48: (Pfadw'keit)

Für einen Pfad $(\Omega, v_1, v_2, \ldots, v_n)$ von der Wurzel zu einem Knoten definieren wir die **Pfadw'keit***:*

$$P_{(\Omega, v_1, \ldots, v_n)} = P(v_1) \cdot P(v_2|v_1) \cdot P(v_3|v_2) \cdot \ldots \cdot P(v_n|v_{n-1})$$

als das Produkt der Übergangsw'keiten längs der Kanten des Pfades.

Bemerkung 1.49: *Liegen auf einem Pfad $(\Omega, v_1, \ldots, v_n)$ nur Knoten, die nur jeweils einen Vater haben, so gelten wie bei Bäumen die Inklusionen*

$$v_n \subset v_{n-1} \subset \cdots \subset v_1 \subset \Omega$$

und damit

$$P_{(\Omega, v_1, \ldots, v_n)} = P(v_1 \cap v_2 \cap \cdots \cap v_n) = P(v_n).$$

Hat aber mindestens einer der Knoten im Pfad mehrere Väter, so gilt diese Interpretation der Pfadw'keit nicht mehr! Es gilt weder $P_{(\Omega, v_1, \ldots, v_n)} = P(v_n)$ noch $P_{(\Omega, v_1, \ldots, v_n)} = P(v_1 \cap v_2 \cap \cdots \cap v_n)$.

Satz 1.50: (Pfadregel 2)

In einem W'keitsgraphen, der den Modellierungsregeln 1.47 genügt, gilt für jeden Knoten v:

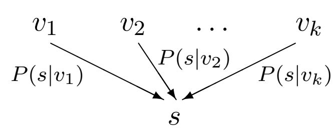
$$P(v) = \sum_{\text{Pfad} \in \text{Pfade}(v)} P_{\text{Pfad}},$$

wobei Pfade(v) die Menge aller Pfade von Ω nach v ist und P_{Pfad} die Pfadw'keit nach Definition 1.48.

Beweis: Induktion nach der Länge der Pfade.

Induktionsstart: Für einen direkt an der Wurzel Ω hängenden Knoten v kann es neben Ω keinen weiteren Vater geben, denn dieser müsste wegen 1.47.e) disjunkt von Ω sein. Die Kante (Ω, v) (mit der Übergangsw'keit ($P(v)$) ist damit der einzige Pfad von Ω nach v und liefert die korrekte W'keit $P(v)$.

Induktionsschritt: Die Behauptung gelte für alle Knoten, die von Ω aus mit einem Pfade der Länge $\leq n$ erreichbar sind. Sei nun s ein Knoten, der über einen Pfad der maximalen Länge $n+1$ von Ω aus erreichbar ist. Seien die Knoten $v_1, \ldots, v_k$ die Väter von s:



Gemäß der Regel 1.47.e) sind die Ereignisse $v_1, \ldots, v_k$ disjunkt und überdecken gemäß Regel 1.47.c) das Ereignis s. Damit ist der Satz von der totalen W'keit 1.31 anwendbar, es gilt:

$$P(s) = \sum_{i=1}^{k} P(s \mid v_i)\ P(v_i).$$

Da die Väter v_i alle über Pfade der maximalen Länge n mit der Wurzel verbunden sind, gilt nach Induktionsvoraussetzung

$$P(v_i) = \sum_{\text{Pfad} \in \text{Pfade}(v_i)} P_{\text{Pfad}},$$

also

$$P(s) = \sum_{i=1}^{k} P(s \mid v_i) \cdot \sum_{\text{Pfad} \in \text{Pfade}(v_i)} P_{\text{Pfad}} = \sum_{i=1}^{k} \sum_{\text{Pfad} \in \text{Pfade}(v_i)} P(s \mid v_i) \cdot P_{\text{Pfad}}.$$

Für jeden Pfad in Pfade(v_i) ist das Produkt $P(s \mid v_i) \cdot P_{\text{Pfad}}$ die Pfadw'keit des erweiterterten Pfades, der von Ω über v_i nach s führt. Da jeder der Pfade von Ω nach s über einen der Väter v_i führt, erstreckt sich die Doppelsumme $\sum_i \sum_{\text{Pfad}\in\text{Pfade}(v_i)}$ in der Tat über alle Pfade von Ω nach s, womit die Pfadregel 2 für den Knoten s bewiesen ist.

Q.E.D.

Beispiel 1.51: Wir betrachten erneut das Wetter in Israel nach Beispiel 1.44. Wir benutzen wieder das schon vorher verwendete Modell

$$\Omega = \{(Dez, Jan, Feb); Dez = T; Jan, Feb \in \{T, N\}\} = \text{„Dez = T"},$$

betrachten aber nun die Ereignisse

$$\begin{aligned}
\text{„Jan = T"} &= \{(T, T, Feb); Feb \in \{T, N\}\},\\
\text{„Jan = N"} &= \{(T, N, Feb); Feb \in \{T, N\}\},\\
\text{„Feb = T"} &= \{(T, Jan, T); Jan \in \{T, N\}\},\\
\text{„Feb = N"} &= \{(T, Jan, N); Jan \in \{T, N\}\}
\end{aligned}$$

im folgenden Graphen:

$$\begin{tikzcd}
 & \text{Dez = T} \arrow[dl, "3/4"'] \arrow[dr, "1/4"] & \\
\text{Jan = T} \arrow[dd, "3/4"'] \arrow[ddrr, "1/4" near start] & & \text{Jan = N} \arrow[dd, "2/3"] \arrow[ddll, "1/3"' near start] \\
 & & \\
\text{Feb = T} & & \text{Feb = N}
\end{tikzcd}$$

Dieser Graph erfüllt in der Tat die Modellierungsregeln 1.47. Die W'keit für einen trockenen Februar ergibt sich über die Pfadregel 2 unmittelbar zu

$$P(\text{„Feb = T"}) = \frac{3}{4}\cdot\frac{3}{4} + \frac{1}{4}\cdot\frac{1}{3} = \frac{31}{48}$$

als Summe der Pfadw'keiten der beiden Pfade (Dez = T, Jan = T, Feb = T) und (Dez = T, Jan = N, Feb = T). Man beachte, dass sich der W'keitsgraph hier gegenüber dem in Beispiel 1.44 verwendeten W'keitsbaum deutlich vereinfacht hat.

1.5 Unabhängigkeit von Ereignissen

$\boxed{30.4.07\downarrow}$

Nun kommt die Formalisierung eines ganz zentralen Begriffs: Unabhängigkeit.

1.5.1 Definition

Intuitiv sind zwei Ereignisse A, B voneinander unabhängig, wenn das Eintreten von A keinerlei Informationen über B liefert, d.h., wenn $P(B\,|A) = P(B)$ gilt, also $P(A\cap B)/P(A) = P(B)$, also $P(A\cap B) = P(A)\,P(B)$:

Definition 1.52: (Unabhängigkeit von Ereignissen)

Zwei Ereignisse $A, B \in \mathcal{E}$ in einem Modell $(\Omega, \mathcal{E}, P)$ heißen **unabhängig**, *wenn gilt*

$$P(A \cap B) = P(A) \cdot P(B) \ ,$$

also $P(B\,|A) = P(B)$ bzw. $P(A\,|B) = P(A)$.

Die Definition ist offensichtlich symmetrisch, wie es intuitiv auch sein sollte: A ist genau dann unabhängig von B, wenn B unabhängig von A ist. Daher spricht man von der Unabhängigkeit eines Paares A, B.

Beispiel 1.53: Ich werfe 2 mal mit einem fairen Würfel.

a) Sind die Ereignisse A = „ich habe im ersten Wurf eine 6“ und B_1 = „ich habe im zweiten Wurf eine 4“ unabhängig?

b) Sind die Ereignisse A = „ich habe im ersten Wurf eine 6“ und B_2 = „meine Augensumme ist 10“ unabhängig?

Lösung: Seien

$$\begin{array}{lcl} \Omega & = & \text{„alle Wurfkombinationen“} \ = \ \{(1,1), \ldots, (6,6)\} \ , \\ A & = & \text{„im ersten Wurf eine 6“} \ = \ \{(6,1), (6,2), \ldots, (6,6)\} \ , \\ B_1 & = & \text{„im zweiten Wurf eine 4“} \ = \ \{(1,4), (2,4) \ldots, (6,4)\} \ , \\ B_2 & = & \text{„Augensumme ist 10“} \ = \ \{(4,6), (5,5), (6,4)\} \ . \end{array}$$

a) Mit $|\Omega| = 36$, $P(A) = |A|/|\Omega| = 6/36 = 1/6$, $P(B_1) = |B_1|/|\Omega| = 6/36 = 1/6$ und

$$A \cap B_1 = \{(6,4)\} \quad \Rightarrow \quad P(A \cap B_1) = \frac{|A \cap B_1|}{|\Omega|} = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(A)\ P(B_1)$$

sind A und B_1 unabhängig (was intitutiv klar war).

b) Mit $P(B_2) = |B_2|/|\Omega| = 3/36 = 1/12$ und

$$A \cap B_2 = \{(6,4)\} \quad \Rightarrow \quad P(A \cap B_2) = \frac{|A \cap B_2|}{|\Omega|} = \frac{1}{36} \neq P(A)\ P(B_2) = \frac{1}{72}$$

sind A und B_2 nicht unabhängig. Intuitiv: die Chance, eine hohe Augensumme zu erhalten, hat sich durch die Tatsache erhöht (hier: verdoppelt), dass schon der erste Wurf ein hohes Ergebnis brachte.

Aus technischen Gründen brauchen wir noch den Begriff der Unabängigkeit von mehr als zwei Ereignissen, der später in die Definition der Unabhängigkeit von Zufallsvariablen eingeht:

Definition 1.54: (Technische Definition: Unabhängigkeit von Ereignisfamilien)
Sei $\mathcal{A} = \{A_i;\ i \in I\} \subset \mathcal{E}$ eine Familie von Ereignissen (I eine beliebige Indexmenge) in einem Modell $(\Omega, \mathcal{E}, P)$.

a) *Ein Ereignis $B \in \mathcal{E}$ heißt* **unabhängig von der Familie** *$\mathcal{A}$, wenn B unabhängig von jedem Schnitt $A_{i_1} \cap \cdots \cap A_{i_k}$ beliebig ausgewählter endlich vieler verschiedener Elemente dieser Familie ist:*

$$P(B \cap A_{i_1} \cap \cdots \cap A_{i_k}) = P(B) \cdot P(A_{i_1} \cap \cdots \cap A_{i_k}) \ .$$

b) *Die Familie $\mathcal{A}$ selbst heißt* **unabhängig**, *wenn jedes Ereignis $A_i \in \mathcal{A}$ unabhängig von der Restfamilie $\mathcal{A} \setminus \{A_i\}$ ist. Dies ist der Fall, wenn für jede beliebige Auswahl von endlich vielen verschiedenen Elementen $A_{i_1}, \ldots, A_{i_k} \in \mathcal{A}$ gilt:*

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot \ \cdots \ \cdot P(A_{i_k}) \ .$$

Bemerkung 1.55: (Warnung) *Ist B unabhängig von jedem $A_i \in \mathcal{A}$, so ist B damit nicht automatisch unabhängig von der Familie $\mathcal{A}$. Entsprechend ist eine Familie $\mathcal{A}$ nicht notwendigerweise unabhängig, wenn jedes Paar $A_i, A_j \in \mathcal{A}$ (mit $i \neq j$) unabhängig ist. Gegenbeispiel:*

$$\Omega = \{1,2,3,4\} \text{ mit dem Zählmaß}, \ A_1 = \{1,3\}, \ A_2 = \{2,3\}, \ A_3 = \{3,4\} \ .$$

Es gilt die paarweise Unabhängigkeit der Ereignisse in $\mathcal{A} = \{A_1, A_2, A_3\}$:

$$P(A_1 \cap A_2) = \frac{1}{4} = P(A_1)\, P(A_2) \ , P(A_1 \cap A_3) = \frac{1}{4} = P(A_1)\, P(A_3) \ ,$$

$$P(A_2 \cap A_3) = \frac{1}{4} = P(A_2)\, P(A_3) \ ,$$

aber

$$P(A_1 \cap A_2 \cap A_3) = P(\{3\}) = \frac{1}{4} \neq P(A_1)\, P(A_2)\, P(A_3) = \frac{1}{8} \ .$$

Merke (Zusammenfassung) 1.56:
Unabhängigkeit zweier Ereignisse A, B wird modelliert durch

$$P(\text{„sowohl } A \text{ als auch } B \text{ treten ein“}) \equiv P(A \cap B) = P(A) \cdot P(B).$$

1.5.2 Modellierung unabhängiger Experimente: Produktmodelle

Ein typisches Beispiel für unabhängige Ereignisse ist die „unabhängige Wiederholung eines Experiments“ (mehrere Male würfeln, mehrere Roulette-Versuche etc). Wie baut man sich systematisch ein Modell für *unabhängige* Durchführungen einzelner Zufallsexperimente?

Wir haben schon mehrfach in den Beispielen unabhängige Wiederholungen von Würfelexperimenten, Münzwürfen etc. durch Modelle der Form

$$\Omega = \{(\text{erstes Ergebnis}, \text{zweites Ergebnis}, \dots)\}$$

behandelt. Dies war mehr intuitiv. Die Tatsache, dass Unabhängigkeit in der Tat durch geordnete Tupel von Einzelexperimenten beschrieben werden kann, soll nun formalisiert werden. Zur Erinnerung: das kartesische Produkt zweier Mengen M_1, M_2 ist die Menge aller geordneten Tupel

$$M_1 \times M_2 = \{(m_1, m_2)\,;\; m_1 \in M_1\,,\; m_2 \in M_2\}\ .$$

Im Sinne von Definition 1.14 wird ein W'keitsmaß für ein diskretes Produktmodell festgelegt:

Definition 1.57: (Produktmodelle[2])

Seien $(\Omega_1, \mathcal{E}_1, P_1), \dots, (\Omega_n, \mathcal{E}_n, P_n)$ diskrete Modelle („Einzelexperimente“). Das **Produktmodell** *ist das diskrete Modell $(\Omega, \mathcal{E}, P)$ mit*

a) $\Omega = \Omega_1 \times \dots \times \Omega_n = \{(\omega_1, \dots, \omega_n);\; \omega_1 \in \Omega_1;\; \dots\; ;\; \omega_n \in \Omega_n\}$.

b) $\mathcal{E} = \mathcal{P}(\Omega)$ *= Potenzmenge von* Ω.

c) Für die Elementarereignisse $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ *wird definiert:*

$$P(\{\underbrace{(\omega_1, \dots, \omega_n)}_{\omega}\}) = P_1(\{\omega_1\}) \cdot \ \dots\ \cdot P_n(\{\omega_n\})$$

*(das „**Produktw'keitsmaß**“). Hiermit wird für beliebige Ereignisse* $E \subset \Omega$ *definiert:*

$$P(E) = \sum_{\omega \in E} P(\{\omega\})\ .$$

[2] Da die Einzelexperimente hier als diskret vorausgesetzt sind (also: die Ω_i sind abzählbar und $\mathcal{E}_i = \mathcal{P}(\Omega_i)$), ist der Produktraum wieder diskret. Man kann auch Produkte allgemeiner (nicht notwendigerweise diskreter) Einzelexperimente $(\Omega_i, \mathcal{E}_i, P_i)$ definieren, was aber etwas technischer ist. Dann setzt man z.B.

$$\mathcal{E} = \text{ die von } \{E_1 \times \dots \times E_n;\; E_1 \in \mathcal{E}_1, \dots, E_n \in \mathcal{E}_n\} \text{ erzeugte Sigma-Algebra.}$$

Hierbei wird von einem Mengensystem eine Sigma-Algebra dadurch erzeugt, dass man zum Mengensystem alle Komplemente und abzählbaren Vereinigungen hinzunimmt, dann wiederum alle Komplemente und abzählbaren Vereinigungen hinzunimmt, usw.

2.5.07↓ **Interpretation und Folgerungen 1.58:**

Die Elementarereignisse $(\omega_1, \ldots, \omega_n)$ *sind folgendermaßen zu interpretieren: Das Experiment* $(\Omega_1, \mathcal{E}_1, P_1)$ *wird durchgeführt und liefert den Wert* ω_1*. Unabhängig davon wird das Experiment* $(\Omega_2, \mathcal{E}_2, P_2)$ *durchgeführt und liefert den Wert* ω_2*. Usw.*

Das Ereignis $E =$ *„im* i*-ten Schritt tritt das Ereignis* $E_i \subset \Omega_i$ *ein" (und es ist egal, was in den anderen Schritten passiert) ist dann als das Ereignis*

$$E = \Omega_1 \times \cdots \times \Omega_{i-1} \times E_i \times \Omega_{i+1} \times \cdots \times \Omega_n \subset \Omega$$

innerhalb des Gesamtexperiments zu interpretieren. Es gilt

$$P(E) = \sum_{\omega_1 \in \Omega_1} \cdots \sum_{\omega_i \in E_i} \cdots \sum_{\omega_n \in \Omega_n} P_1(\omega_1) \cdot \ldots \cdot P_j(\omega_j) \cdot \ldots \cdot P_n(\omega_n)$$

$$= \underbrace{\Big(\sum_{\omega_1 \in \Omega_1} P_1(\{\omega_1\}) \Big)}_{P_1(\Omega_1)=1} \cdots \underbrace{\Big(\sum_{\omega_i \in E_i} P_i(\{\omega_i\}) \Big)}_{=P_i(E_i)} \cdots \underbrace{\Big(\sum_{\omega_n \in \Omega_n} P_n(\{\omega_n\}) \Big)}_{P_n(\Omega_n)=1} = P_i(E_i)\,.$$

Mit einem weiteren Ereignis $\tilde{E} =$ *„im* j*-ten Schritt tritt das Ereignis* $\tilde{E}_j \subset \Omega_j$ *ein"*

$$\tilde{E} = \Omega_1 \times \cdots \times \Omega_{j-1} \times \tilde{E}_j \times \Omega_{j+1} \times \cdots \times \Omega_n \subset \Omega$$

folgt dann

$$E \cap \tilde{E} = \begin{cases} \Omega_1 \times \cdots \times E_i \times \cdots \times \tilde{E}_j \times \cdots \times \Omega_n & \text{für } i \neq j \\ \Omega_1 \times \cdots \times \big(E_i \cap \tilde{E}_j\big) \times \cdots \times \Omega_n & \text{für } i = j \end{cases}$$

und

$$P(E \cap \tilde{E}) = \begin{cases} \Big(\sum_{\omega_i \in E_i} P_i(\{\omega_i\}) \Big) \Big(\sum_{\omega_j \in \tilde{E}_j} P_j(\{\omega_j\}) \Big) & \text{für } i \neq j \\ \sum_{\omega \in E_i \cap \tilde{E}_j} P_i(\{\omega\}) & \text{für } i = j \end{cases}$$

$$= \left\{ \begin{array}{ll} P_i(E_i)\; P_j(\tilde{E}_j) & \text{für } i \neq j \\ P_i(E_i \cap \tilde{E}_j) & \text{für } i = j \end{array} \right\} = \begin{cases} P(E)\; P(\tilde{E}) & \text{für } i \neq j\,, \\ P_i(E_i \cap \tilde{E}_j) & \text{für } i = j\,. \end{cases}$$

Resultat: zwei Ereignisse E *und* $\tilde{E}$ *sind* <u>*automatisch*</u> *unabhängig, wenn sie Ereignisse in unterschiedlichen Einzelexperimenten beschreiben. Das Produktmodell modelliert damit in der Tat* <u>*unabhängige*</u> *Ausführungen von Einzelexperimenten.*

Bemerkung 1.59: *Eine n-fache unabhängige Wiederholung eines Experiments $(\Omega_1, \mathcal{E}_1, P_1)$ ist der Spezialfall $(\Omega_1, \mathcal{E}_1, P_1) = \cdots = (\Omega_n, \mathcal{E}_n, P_n)$.*

Beispiel 1.60: Eine manipulierte Münze $\Omega_1 = \{K, Z\}$ mit $P(K) = 0.6$, $P(Z) = 0.4$ wird 3 mal geworfen. Mit welcher W'keit wird höchstens 1 mal Kopf geworfen?

Lösung:

$$\Omega = \{(\omega_1, \omega_2, \omega_3);\ \omega_i \in \{K, Z\}\}\ ,$$

$$E = \{(K, Z, Z), (Z, K, Z), (Z, Z, K), (Z, Z, Z)\}.$$

Die Elementarw'keiten der Elemente in E sind:

$$P(\{(K, Z, Z)\} = P(\{(Z, K, Z)\} = P(\{(Z, Z, K)\} = 0.6 \cdot 0.4 \cdot 0.4\ ,$$

$$P(\{(Z, Z, Z)\} = 0.4 \cdot 0.4 \cdot 0.4\ .$$

Damit folgt:

$$P(E) = 3 \cdot 0.6 \cdot 0.4 \cdot 0.4\ +\ 0.4 \cdot 0.4 \cdot 0.4 = 0.352\ .$$

Bemerkung 1.61: *Ein Produktmodell $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ kann natürlich auch dann verwendet werden, wenn die Einzelexperimente $\Omega_1, \ldots, \Omega_n$ nicht unabhängig sind. Allerdings ist das Produktmodell dann nicht mit dem Produkt-W'keitsmaß*

$$P(E_1 \times \cdots \times E_n) = P_1(E_1) \cdot\ \cdots\ \cdot P_n(E_n)$$

mit Einzelereignissen $E_i \subset \Omega_i$ auszustatten.

Eine sehr einfache, in Anwendungen aber sehr wichtige Situation ist das Bernoulli-Experiment bzw. unabhängige Wiederholungen eines solchen Experiments:

Definition 1.62:

Ein **„Bernoulli-Experiment“** *ist ein diskretes Experiment (= Modell) mit nur 2 möglichen Ausgängen 1 („Erfolg“) und 0 („Misserfolg“). Meist werden die Elementarw'keiten mit p (***„Erfolgsw'keit“***) und $q = 1 - p$ (***„Misserfolgsw'keit“***) bezeichnet:*

$$\Omega = \{0, 1\}\ , \quad P(\{1\}) = p\ , \quad P(\{0\}) = q = 1 - p\ .$$

In praktischen Aufgaben tauchen oft unabhängige Wiederholungen von Bernoulli-Experimenten auf:

Satz 1.63: (Erfolgsw'keiten in wiederholten Bernoulli-Experimenten)

Ein Bernoulli-Experiment mit Erfolgsw'keit p wird n mal wiederholt. Die W'keit, dabei insgesamt genau k mal Erfolg zu haben ($k \in \{0, 1, \ldots, n\}$), beträgt

$$P(\text{„genau } k \text{ Erfolge“}) = \binom{n}{k} p^k \, q^{n-k} \ .$$

Beweis: Benutze den Stichprobenraum

$$\Omega = \{(\omega_1, \ldots, \omega_n); \ \omega_i \in \{0, 1\}\}$$

(ein Produktmodell) zur Modellierung der n unabhängigen Wiederholungen. Jedes Elementarereignis $(\omega_1, \ldots, \omega_n)$ mit genau k Erfolgen hat die W'keit $p^k \, q^{n-k}$ (dies ist die Unabhängigkeit). Ein solches Elementarereignis unterscheidet jedoch die Reihenfolgen, in denen die Erfolge und Misserfolge auftreten. Gefragt ist nach dem Ereignis

$$\text{„genau } k \text{ Erfolge“} \equiv \Big\{(\omega_1, \ldots, \omega_n); \ \sum_{i=1}^{n} \omega_i = k\Big\} \ ,$$

in dem die Reihenfolge der Erfolge nicht eingeht. Nach Hilfssatz 1.9.b) gibt es $\binom{n}{k}$ verschiedene Reihenfolgen mit genau k Erfolgen und $n-k$ Misserfolgen. Damit ist die W'keit für genau k Erfolge ohne Berücksichtigung der Reihenfolge $\binom{n}{k} \, p^k \, q^{n-k}$.

Q.E.D.

Bezeichnung 1.64:

Die Abbildung

$$k \in \{0, 1, \ldots, n\} \quad \to \quad P(\text{„genau } k \text{ Erfolge“}) = \binom{n}{k} p^k \, q^{n-k}$$

heißt **Binomialverteilung** *zu den gegebenen Parametern n (Anzahl der Wiederholungen) und p (Erfolgsw'keit).*

Beispiel 1.65: Zurück zur manipulierten Münze mit $P(K) = p = 0.6$, $P(Z) = q = 0.4$ aus Beispiel 1.60. Ein Münzwurf kann als Bernoulli-Experiment aufgefasst werden. Bei n Würfen ist die W'keit, höchsten 1 mal „Kopf“ (= „Erfolg“) zu werfen, gegeben durch

$$\begin{aligned}
P(\text{„höchstens 1 Erfolg“}) &= P(\text{„genau 0 Erfolge“} \ \cup \ \text{„genau 1 Erfolg“}) \\
&\overset{(*)}{=} P(\text{„genau 0 Erfolge“}) + P(\text{„genau 1 Erfolg“}) \\
&= \binom{n}{0} p^0 \, q^n + \binom{n}{1} p^1 \, q^{n-1} \\
&= q^n + n \, p \, q^{n-1} \ .
\end{aligned}$$

Hierbei wird in ($*$) die Additivität von W'keiten bei disjunkten Ereignissen benutzt (Folgerung 1.6.a). Für $n = 3$ ergibt sich (vergleiche mit Beispiel 1.60):

$$P(\text{„höchstens 1 Erfolg“}) = 0.4^3 + 3 \cdot 0.6 \cdot 0.4^2 = 0.352 \ .$$

Merke (Zusammenfassung) 1.66:

Ein **Produktmodell** *bietet sich in Situationen an, in denen das Gesamtexperiment aus mehreren Einzelexperimenten besteht, die* <u>*unabhängig*</u> *voneinander durchgeführt werden. Es besteht aus geordneten Tupeln der Ergebnisse der Einzelexperimente, das W'keitsmaß ist auf den Elementarereignissen durch*

$$P(\{(\omega_1, \ldots, \omega_n)\}) = P(\{\omega_1\}) \cdot P(\{\omega_2\}) \cdot \ldots \cdot P(\{\omega_n\})$$

gegeben, was die Unabhängigkeit der Einzelexperimente wiederspiegelt. In diesem Modell sind Ereignisse automatisch unabhängig, wenn sie sich auf unterschiedliche Teilexperimente beziehen.

1.6 Einige Beispiele

In diesem Abschnitt werden einige etwas komplexere Beispiele diskutiert, in denen die meisten der bisherigen Begriffsbildungen und Aussagen zum Einsatz kommen:

Beispiel 1.67: (Verknüpfung mehrerer Experimente) Man hat 2 Münzen, eine faire ($P(K) = P(Z) = 1/2$) und eine manipulierte ($P(K) = 0.3$, $P(Z) = 0.7$), die man allerdings nicht unterscheiden kann. Man nimmt eine und möchte herausbekommen, ob es die manipulierte ist. Man wirft sie hierzu n mal und hat k mal „Zahl“. Mit welcher W'keit handelt es sich um die manipulierte Münze? (Die Idee ist: wenn bei n Würfen etwa $n/2$ mal „Zahl“ eintritt, so dürfte es sich wohl um die faire Münzen handeln. Wird hingegen etwa $0.7 \cdot n$ mal „Zahl“ beobachtet, so ist die W'keit groß, dass es sich um die manipulierte Münze handelt.)
Lösung: Zerlege das Experiment des n-fachen Münzwurfs in das Ursachensystem F = „die Münze ist fair“ und U = „die Münze ist unfair“. Eine der beiden Münzen wird in einem ersten Bernoulli-Experiment mit der W'keit $1/2$ gezogen:

$$\Omega = F \cup U \ , \quad P(F) = P(U) = \frac{1}{2} \ .$$

Für gegebenes F bzw. U betrachten wir den einfachen Münzwurf jeweils als Bernoulli-Experiment mit den Erfolgsw'keiten $P(Z) = P(K) = 1/2$ bzw. $p = P(Z) = 0.7$, $q = P(K) = 0.3$. Sei beim n-fachen Wurf E_k das Ereignis „es wird k mal „Zahl“ geworfen“. Mit Satz 1.63 gilt:

$$P(E_k \,|\, F) \ = \ \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \binom{n}{k} \frac{1}{2^n} \ , \quad P(E_k \,|\, U) \ = \ \binom{n}{k} p^k \, q^{n-k} \ .$$

Nach Bayes 1.34 gilt:

$$P(U \,|\, E_k) = \frac{P(E_k \,|\, U)\, P(U)}{P(E_k \,|\, U)\, P(U) + P(E_k \,|\, F)\, P(F)} = \frac{1}{1 + \dfrac{P(E_k \,|\, F)}{P(E_k \,|\, U)}\, \dfrac{P(F)}{P(U)}}$$

$$= \frac{1}{1 + \dfrac{1/2^n}{p^k\, q^{n-k}}\, \dfrac{1/2}{1/2}} = \frac{1}{1 + \dfrac{1}{(2\,p)^k\, (2\,q)^{n-k}}} \, .$$

Für $p = 0.7$, $q = 0.3$, $n = 1000$ ergeben sich beispielsweise folgende Werte. Mit Bemerkung 1.36 gilt $P(F \,|\, E_k) = 1 - P(U \,|\, E_k)$:

k	596	598	600	602	603	604	606	608
$P(U\|E_k)$	0.00291..	0.0156..	0.0796..	0.320..	0.523..	0.719..	0.939..	0.987..
$P(F\|E_k)$	0.997..	0.984..	0.920..	0.679..	0.476..	0.208..	0.0668..	0.0129..

Also: wird 600 mal „Zahl“ geworfen, handelt es sich mit großer W'keit um die faire Münze. Ab 603 mal „Zahl“ ist es wahrscheinlicher, dass es sich um die manipulierte Münze handelt. Ab 606 mal „Zahl“ kann man fast sicher davon ausgehen, dass es sich um die manipulierte Münze handelt. Bayes liefert damit ein extrem scharfes Kriterium, auf „manipuliert“ versus „fair“ schließen zu können. (Die extreme Schärfe liegt in diesem Beispiel daran, dass wir die Münzen sehr genau kennen: für die unfaire Münze ist $P(K) = 0.7$, $P(Z) = 0.3$ bekannt).

7.5.07↓ **Beispiel 1.68:** Ein System besteht aus 4 Komponenten vom Typ A und 16 Komponenten vom Typ B.
Typ A ist unzuverlässig: eine Komponente fällt mit der W'keit 0.1 pro Arbeitszyklus aus. Typ B ist zuverlässiger: eine Komponente fällt mit der W'keit 0.01 pro Arbeitszyklus aus.
Das System fällt sicher aus, wenn mindestens 2 Teile vom Typ A ausfallen oder mindestens 1 Teil vom Typ B ausfällt. Wenn genau 1 Teil vom Typ A ausfällt und kein Teil vom Typ B ausfällt, so fällt das Gesamtsystem mit der W'keit 1/2 aus.

a) Mit welcher W'keit fällt das Gesamtsystem in einem Arbeitszyklus aus?

b) Ich will das Gesamtsystem sicherer machen. Ist es sinnvoller, Typ A oder Typ B zu verbessern, oder müssen beide Typen verbessert werden?

Lösung: Der Stichprobenraum für einen Arbeitszyklus sei

$$\begin{aligned} \Omega &= A_1 \times \cdots \times A_4 \times B_1 \times \cdots \times B_{16} \\ &= \{(a_1, \ldots, a_4, b_1, \ldots, b_{16}); a_1, \ldots, b_{16} \in \{\text{defekt}, \text{ok}\}\} \, , \end{aligned}$$

wobei $A_1, \ldots, B_{16}$ jeweils Bernoulli-Experimente sind mit

$$\begin{aligned} p_A &= P_{A_i}(\text{defekt}) = 0.1 \, , \quad & q_A &= P_{A_i}(\text{ok}) = 0.9 \, , \\ p_B &= P_{B_i}(\text{defekt}) = 0.01 \, , \quad & q_B &= P_{B_i}(\text{ok}) = 0.99 \, . \end{aligned}$$

Frage a): Mit welcher W'keit fällt das Gesamtsystem in einem Arbeitszyklus aus?

Bilde zunächst ein disjunktes System von Ursachen für das Ereignis $E =$ „das Gesamtsystem fällt aus“:

U_1 : „kein B fällt aus, kein A fällt aus“,
U_2 : „kein B fällt aus, genau ein A fällt aus“,
U_3 : „kein B fällt aus, mindestens zwei A fallen aus“,
U_4 : „mindestens ein B fällt aus“.

Das Kriterium für die Wahl dieser Zerlegung ist:

i) alle U_i müssen disjunkt sein,

ii) alle Ausfallmöglichkeiten müssen abgedeckt sein: $E \subset \cup_i U_i$,

iii) wir müssen $P(E \,|\, U_i)$ kennen,

iv) $P(U_i)$ sollte ohne allzu großen kombinatorischen Aufwand berechenbar sein.

Die obige Zerlegung erfüllt diese Kriterien:

i): Ist offensichtlich erfüllt.
ii): Es gilt sogar $\cup_i U_i = \Omega$.
iii):

$$P(E \,|\, U_1) = 0 \ , \ P(E \,|\, U_2) = \frac{1}{2} \ , \ P(E \,|\, U_3) = 1 \ , \ P(E \,|\, U_4) = 1 \ .$$

Fragen wir nach der W'keit des Ausfalls des Gesamtsystems, so gilt nach der Formel der totalen W'keit 1.31:

$$P(E) = \sum_i P(E \,|\, U_i) \cdot P(U_i) = \frac{1}{2} \cdot P(U_2) + P(U_3) + P(U_4) \ . \qquad (\#)$$

iv) Es verbleibt, $P(U_2)$, $P(U_3)$, $P(U_4)$ zu bestimmen. Dies ist nicht schwierig, da es sich um unabhängige Wiederholungen von 2 verschiedenen Bernoulli-Experimenten $A_1 = \cdots = A_4$, $B_1 = \cdots = B_{16}$ handelt. Betrachten wir zunächst die beiden Typen getrennt. Mit Satz 1.63 gilt:

$$\begin{aligned}
&P(\text{„kein } A \text{ defekt“}) &&= \binom{4}{0} \cdot p_A^0 \, q_A^4 = q_A^4 \ , \\
&P(\text{„genau ein } A \text{ defekt“}) &&= \binom{4}{1} \cdot p_A^1 \, q_A^3 = 4 \, p_A \, q_A^3 \ , \\
&P(\text{„mindestens zwei } A \text{ defekt“}) &&= 1 - P(\text{„kein } A \text{ defekt“}) - P(\text{„genau ein } A \text{ defekt“}) \\
& &&= 1 - q_A^4 - 4 \, p_A \, q_A^3 \ , \\
&P(\text{„kein } B \text{ defekt“}) &&= \binom{16}{0} \cdot p_B^0 \, q_B^{16} = q_B^{16} \ , \\
&P(\text{„mindestens ein } B \text{ defekt“}) &&= 1 - P(\text{„kein } B \text{ defekt“}) = 1 - q_B^{16} \ .
\end{aligned}$$

Da die A- und B-Teile unabhängig voneinander ausfallen können, werden die Bernoulli-Experimentwiederholungen für den Ausfall der A- und B-Teile unabhängig durchgeführt. Damit gilt

$$\begin{aligned} P(U_2) &= P(\text{„kein } B \text{ defekt, genau ein } A \text{ defekt“}) \\ &= P(\text{„kein } B \text{ defekt}) \, P(\text{genau ein } A \text{ defekt“}) \\ &= q_B^{16} \, 4 \, p_A \, q_A^3 \ \approx\ 0.2483 \ , \end{aligned}$$

$$\begin{aligned} P(U_3) &= P(\text{„kein } B \text{ defekt, mindestens zwei } A \text{ defekt“}) \\ &= P(\text{„kein } B \text{ defekt“}) \, P(\text{„mindestens zwei } A \text{ defekt“}) \\ &= q_B^{16} \, (1 - q_A^4 - 4 \, p_A \, q_A^3) \ \approx\ 0.04453 \ , \end{aligned}$$

$$P(U_4) = P(\text{„mindestens ein } B \text{ defekt“}) = 1 - q_B^{16} \approx\ 0.1485 \ .$$

Mit Gleichung (#) folgt die Ausfallw'keit des Gesamtsystems:

$$\begin{aligned} P(E) &= \frac{1}{2} \, q_B^{16} \, 4 \, p_A \, q_A^3 \ + q_B^{16} \, (1 - q_A^4 - 4 \, p_A \, q_A^3) + 1 - q_B^{16} \\ &= 1 - q_A^4 \, q_B^{16} - 2 \, p_A \, q_A^3 \, q_B^{16} \ . \end{aligned}$$

Für $p_A = 0.1$, $q_A = 0.9$, $p_B = 0.01$, $q_B = 0.99$ ergibt sich der konkrete Wert

$$P(E) = 0.3172.. \ .$$

Frage b): Ich will das Gesamtsystem sicherer machen. Ist es sinnvoller, Typ A oder Typ B zu verbessern, oder müssen beide Typen verbessert werden?

Wir berechnen die W'keiten der Ereignisse

U_A = „mindestens ein A-Teil fällt aus, kein B-Teil fällt aus“,
U_B = „mindestens ein B-Teil fällt aus, kein A-Teil fällt aus“,
U_{AB} = „mindestens ein A-Teil fällt aus und mindestes ein B-Teil fällt aus“

unter dem vorgegebenen Ereignis E = „das Gesamtsystem ist ausgefallen“. Dann interpretieren wir

$P(U_A|\, E) = P($„nur die A-Teile waren Schuld am Ausfall“$)$,
$P(U_B|\, E) = P($„nur die B-Teile waren Schuld am Ausfall“$)$,
$P(U_{AB}|\, E) = P($„sowohl A- als auch B-Teile waren Schuld am Ausfall“$)$.

Hierbei bilden U_A, U_B, U_{AB} ein neues disjunktes Ursachensystem für das Ereignis E. In Analogie zu den Rechnungen in a) folgt:

$$\begin{aligned} P(U_A) &= P(\text{„mindestens ein } A \text{ defekt“}) \, P(\text{„kein } B \text{ defekt“}) \\ &= \Big(1 - P(\text{„kein } A \text{ defekt“})\Big) \, P(\text{„kein } B \text{ defekt“}) \\ &= (1 - q_A^4) \, q_B^{16} \ \approx\ 0.2928 \ , \end{aligned}$$

$$\begin{aligned} P(U_B) &= P(\text{„mindestens ein } B \text{ defekt“}) \, P(\text{„kein } A \text{ defekt“}) \\ &= \Big(1 - P(\text{„kein } B \text{ defekt“})\Big) \, P(\text{„kein } A \text{ defekt“}) \\ &= (1 - q_B^{16}) \, q_A^4 \ \approx\ 0.09746 \ , \end{aligned}$$

$$\begin{aligned} P(U_{AB}) &= P(\text{„mindestens ein } B \text{ defekt“}) \, P(\text{„mindestens ein } A \text{ defekt“})) \\ &= \Big(1 - P(\text{„kein } B \text{ defekt“})\Big) \, \Big(1 - P(\text{„kein } A \text{ defekt“})\Big) \\ &= (1 - q_B^{16}) \, (1 - q_A^4) \ \approx\ 0.05108 \ . \end{aligned}$$

Wir benutzen den Satz von Bayes 1.34:

$$P(U_i \,|\, E) = \frac{P(E \,|\, U_i)\, P(U_i)}{P(E)} \ .$$

Die bedingten W'keiten $P(E \,|\, U_B) = P(E \,|\, U_{AB}) = 1$ sind bekannt (siehe Formulierung der Aufgabe). Damit folgt

$$P(U_B \,|\, E) = \frac{1 \cdot P(U_B)}{P(E)} \approx 0.3072 \ , \quad P(U_{AB} \,|\, E) = \frac{1 \cdot P(U_{AB})}{P(E)} \approx 0.1610 \ .$$

Es fehlt noch $P(U_A \,|\, E)$. Nach Bemerkung 1.36 gilt

$$P(U_A \,|\, E) = 1 - P(U_B \,|\, E) - P(U_{AB} \,|\, E) \ \approx \ 0.5318 \ .$$

Zu Übungszwecken berechnen wir diesen Wert auch noch anders. Da $P(E \,|\, U_A)$ nicht direkt gegeben ist, können wir den Satz von Bayes nicht unmittelbar heranziehen. Zur Berechnung benutzen wir $U_A = U_2 \cup U_3$ mit den Ursachen U_2, U_3 aus a). Es gilt

$$\begin{aligned} P(U_A \,|\, E) &= P(U_2 \cup U_3 \,|\, E) = \frac{P(E \cap (U_2 \cup U_3))}{P(E)} \\ &= \frac{P((E \cap U_2) \cup (E \cap U_3))}{P(E)} \overset{\text{(disjunkt)}}{=} \frac{P(E \cap U_2) + P(E \cap U_3)}{P(E)} \\ &= P(U_2 \,|\, E) + P(U_3 \,|\, E) \ . \end{aligned}$$

Die bedingten W'keiten $P(E \,|\, U_2) = 1/2$, $P(E \,|\, U_3) = 1$ sind aus der Aufgabenstellung bekannt. Damit folgt per Bayes

$$P(U_2 \,|\, E) = \frac{1/2 \cdot P(U_2)}{P(E)} \ , \ P(U_3 \,|\, E) = \frac{1 \cdot P(U_3)}{P(E)}$$

und mit den Werten aus a):

$$P(U_A \,|\, E) = P(U_2 \,|\, E) + P(U_3 \,|\, E) = \frac{1/2 \cdot P(U_2) + P(U_3)}{P(E)}$$

$$\approx \frac{0.5 \cdot 0.2483 + 0.04453}{0.3172} \approx 0.5318 \ .$$

Zusammenfassung:

$$\begin{array}{rcccl} P(\text{„nur } A\text{-Teile waren schuld“}) & = & P(U_A \,|\, E) & \approx & 0.5318 \ , \\ P(\text{„nur } B\text{-Teile waren schuld“}) & = & P(U_B \,|\, E) & \approx & 0.3072 \ , \\ P(\text{„sowohl } A\text{- als auch } B\text{-Teile waren schuld“}) & = & P(U_{AB} \,|\, E) & \approx & \underline{0.1610} \ . \\ & & & & 1.0000 \end{array}$$

Wenn ich es mir nur leisten kann, einen Typ zu verbessern, so sollte dies der Typ A sein. Dies wird aber keine dramatische Verbesserung erbringen: man muss beide Typen verbessern, um das Gesamtsystem deutlich sicherer zu machen.