32. Wahrscheinlichkeit

32.1 Grundbegriffe

Ohne einige grundlegenden Begriffe kommt man in der Wahrscheinlichkeitsrechnung nicht aus. Einige davon sind intuitiv einzusehen, andere sind etwas gewöhnungsbedürftig. Begriffen Beispiele angegeben, um die Unterscheidung zu anderen Begriffen klar herauszuheben.

32.1.1 Häufigkeit, Ergebnis, Ereignis, Wahrscheinlichkeit

Definition:
Ein **Zufallsexperiment** ist ein Vorgang mit zufälligem Ergebnis und durch folgende Eigenschaften gekennzeichnet:
1. Er besitzt mehrere mögliche Ergebnisse
2. Das Ergebnis kann vor Ablauf des Experimentes nicht vorhergesagt werden.

Den Begriff des Zufallsexperiments kann man sich leicht klar machen. Man betrachtet dafür genau das übliche Szenario, dass aus einer Urne verschiedenfarbige Kugeln gezogen werden.

Definition:
Anzahl des Auftretens des Ergebnisses e im Verlaufe einer Beobachtungsreihe bezeichnet man als **absolute Häufigkeit**.

Die absolute Häufigkeit tritt bei einer mehrfachen Durchführung eines Zufallsexperiments auf. Es bezeichnet die Anzahl, wie oft eine Ergebnis eingetreten ist. Diese Zahl ist also immer eine ganze Zahl und keine Wahrscheinlichkeit.

Definition:
Absolute Häufigkeit H(e) des Ergebnisses e geteilt durch Anzahl der Beobachtungen n insgesamt ist die **relative Häufigkeit**. Der Wert der relativen Häufigkeit h(e) liegt immer zwischen Null (Ergebnis trat nie auf) und 1 (Ergebnis trat bei allen Beobachtungen auf).

Das ist ein typischer Durchschnitt. Dieser Wert gibt an, wie oft bei einer Vielzahl von Zufallsversuchen das gewünschte Ergebnis eingetreten ist. Hier wird die absolute Häufigkeit auf die Anzahl der Versuche bezogen. das ist der erste Schritt zur Wahrscheinlichkeit. Die Wahrscheinlichkeit für das Eintreten eines Ergebnisses entsteht durch unendlich viele Wiederholungen des Zufallsexperiments, bei dem dann die relative Häufigkeit in die Wahrscheinlichkeit über geht.

Definition:
Die Menge alle möglichen Ergebnisse eines Zufallsvorgangs heißt **Ergebnismenge** sie wird mit Ω bezeichnet.

Die Ergebnismenge sind alle Ergebnisse, die bei einem Zufallsereignis eintreten können. Die Ergebnismenge listet nur die möglichen Ergebnisse auf, liefert aber keine Aussage, wie oft jedes Ergebnis eintritt.

Beispiel: Einmaliges würfeln mit einem normalen Würfel
$\Omega = \{ 1,2,3,4,5,6\}$

In einer Urne befinden sich 3rote, 4 gelbe und 4 blaue Kugeln. Es wird zweimal gezogen.
$\Omega = \{ rr, rg, rb, gr, gg, gb, br, bg, bb\}$

Die Anzahl der einzelnen Farben spielt keine Rolle. Dafür ist aber zu berücksichtigen, dass das Ergebnis rb vom Ergebnis br zu unterscheiden ist. Zunächst ist grundsätzlich zu unterscheiden, welche Farbe als erste und welche als zweite gezogen wurde. Das sind zwei verschiedene Ausgänge des Zufallsexperiments.

Definition:
Ein Element $\omega \in \Omega$ heißt **Ergebnis** oder **Ausgang**. Die Ergebnisse werden mit kleinen ω bezeichnet..

Bei dem ersten Beispiel mit dem Würfel ist ein Ereignis z.B. das Würfeln einer „4“ oder das Würfeln einer „2“. Beim zweiten Beispiel mit der Urne wäre ein Ereignis als erstes eine rote und dann eine blaue zu ziehen. Dieser Begriff ist scharf von dem folgenden zu trennen: Ein Ergebnis ist **immer** ein **Element** der Menge Ω, was beim nächsten Begriff „Ereignis“ nicht sein muss !

Definition:
Zusammenfassungen von Ergebnissen zu Teilmengen der Ergebnismenge heißen **Ereignisse** – d. h. der Begriff Ereignis steht für eine Menge.

Ein Ereignis ist eine Teilmenge von Ω . Ein Ereignis $A \subset \Omega$ tritt ein, wenn der Zufalls-vorgang ein Ergebnis ω liefert, das ein Element von A ist: $\omega \in A \subset \Omega$.
Bei den beiden angegebenen Beispielen würden sich Ereignisse etwa so darstellen:
Für das Würfelspiel wäre ein mögliches Ereignis: A = {es wird eine gerade Zahl gewürfelt }. Dieses **Ereignis** setzt sich aus mehreren **Ergebnissen** zusammen. Die Anzahl der zugehörigen Ergebnisse kann unterschiedlich sein. Aber es wird daraus anschaulich klar, dass Ereignisse Teilmengen Ω von sind. Teilmengen sind Mengen die aus Elementen der Obermenge bestehen. Also besteht A aus den Elementen {2,4,6}, damit eine Teilmenge von Ω. Ein weiteres mögliches Ereignis wäre B={die gewürfelte zahl ist kleiner als 3}. Dieses Ereignis setzt sich aus den Ergebnissen {1, 2} zusammen. Auch diese Menge aus zwei Elementen der Menge Ω ist damit eine Teilmenge von Ω.

Für das zweite Beispiel des Ziehen aus einer Urne könnte man folgende Ereignisse definieren: A = { die erste Kugel ist eine rote Kugel }. Dieses Ereignis setzt sich aus den Ergebnissen A ={ rr,rb,rg } zusammen. Ein weiteres mögliches Ereignis wäre B ={ die beiden gezogenen Kugeln sind verschiedenfarbig } . Dieses Ereignis setzt sich aus den Ergebnissen B = { rb, br, rg, gr, bg, gb } zusammen. Auch diese Mengen A und B sind Teilmengen der Menge Ω.

Definition:
Ist einem Ereignis nur ein einzelnes Ergebnis zugeordnet, spricht man von einem **Elementarereignis**. Die Ereignismenge besteht nur aus einem Element.

Ein Elementarereignis ist damit einem Ergebnis gleich zu setzen, es ist ein einzelnes Element aus Ω.

Definition:
Das Ereignis, welches keine Ergebnisse enthält, heißt **unmögliches Ereignis** – Bezeichnung mit dem Symbol Ø.

Bei den beiden angeführten Beispielen sind unmögliche Ereignisse etwas ungewohnt.
Für das Würfelspiel wäre ein unmögliches Ereignis: Es wird eine „7“ gewürfelt.
Für das Kugelspiel könnte man etwa ansehen: Es wird eine schwarze Kugel gezogen, oder es werden drei rote Kugel gezogen. Beide Ereignisse sind nicht in Ω enthalten ! das ist der Maßstab für unmögliche Ereignisse.

Definition:
Zwei Ereignisse E und F **unvereinbar** wenn die Durchschnittsmenge aus E und F leer ist.

Für das Würfelereignis könnte man dafür folgendes Beispiel angeben:
E : Es wird eine durch 3 teilbare zahl gewürfelt
F: Es wir eine durch 4 teilbare zahl gewürfelt
Das ist mit Zahlen von 1 bis 6 nicht realisierbar. Wäre das zweite Ereignis F: es wird eine durch 2 teilbare Zahl gewürfelt, dann würde das Ergebnis „6“ im Durchschnitt der beiden Ergebnisse liegen.

Definition:
Das Ereignis Ω, das alle Ergebnisse enthält, heißt das **sichere Ereignis**.

Ein sicheres Ereignis des Würfelspiels wäre etwa: C= { Es wird eine Zahl kleiner 7 gewürfelt } oder D = { es wird eine einstellige Zahl gewürfelt }.
Für das Ziehen der Kugeln könnte man folgendes Ereignis formulieren: E = { es wird keine schwarze Kugel gezogen }

Definition:
Die Menge aller Teilmengen von Ω heißt **Ereignisraum**.

Anzahl der Elemente eines Ereignisraums (Anzahl unterschiedlicher Ereignisse)
Besitzt eine Ergebnismenge Ω genau n Elemente ($|\Omega| = n$), so gibt es 2^n unterschiedliche Teilmengen von Ω. (Es gibt 2^n unterschiedliche Ereignisse. Der Ereignisraum hat 2^n unterschiedliche Elemente.)

Definition:
Ist jedem der Ergebnisse eines Zufallsexperiments mit $\Omega = \{e_1, e_2, \ldots, e_n\}$ eine reelle Zahl $P(e_i)$> so zugeordnet, dass

1. $\forall$ i: $0 < P(e_i)$
2. $P(e_1) + P(e_2) + \ldots + P(e_n) = 1$ gilt,

dann heißen die Zahlen $P(e_i)$ **Wahrscheinlichkeiten**.

32.1.2 Laplace – Versuche

Definition:
Die Wahrscheinlichkeitsverteilung, die jedem der möglichen Ereignisse e_1, $e_2, \ldots, e_n$ die gleiche Wahrscheinlichkeit zuordnet heißt **Gleichverteilung**

Definition:
Zufallsversuche, bei denen jedes mögliche Ereignis e_i mit der Wahrscheinlichkeit $P(e_i) = 1/n$ auftritt heißt **Laplace-Ereignis**.

32.1.3 Wahrscheinlichkeitsberechnungen ohne Wahrscheinlichkeitsfunktionen

Auch ohne Kenntnis von Wahrscheinlichkeitsverteilungen kann man die Wahrscheinlichkeit von zufälligen Ereignissen elementar berechnen. Dazu benötigt man lediglich einige Formeln aus der Kombinatorik.
Dazu wird folgendes Beispiel betrachtet:
In einer Urne befinden sich 10 weiße, 7 rote, 8 blaue und 5 grüne Kugeln

32.1.3.1 Mit Zurücklegen ; mit Reihenfolge

Aus der Urne werden 6 Kugeln mit zurücklegen gezogen. Gesucht ist die Wahrscheinlichkeit für das Ergebnis (w; w ; r; r; b; w) in genau dieser Reihenfolge.

$$P(\, w \;;\; w \;;\; r \;;\; r \;;\; b \;;\; w \,)$$
$$\frac{10}{40} \cdot \frac{10}{40} \cdot \frac{7}{40} \cdot \frac{7}{40} \cdot \frac{8}{40} \cdot \frac{10}{40}$$

Da mit Beachtung der Reihenfolge gezogen wird, gibt es nur einen Zweig im Baumdiagramm, in dem diese Reihenfolge auftritt. Es ist deshalb nicht mit einer Anzahl von Pfaden zu multiplizieren. Da mit Zurücklegen gezogen wird, sind bei jeder neuen Ziehung wieder alle Kugeln vorhanden, so daß die Gesamtzahl immer 40 ergibt und die Werte im Zähler immer allen vorhanden Kugeln in dieser Farbe entsprechen.

32.1.3.2 Mit Zurücklegen ; ohne Reihenfolge

Aus der Urne werden 6 Kugeln mit zurücklegen gezogen. Gesucht ist die Wahrscheinlichkeit daß am Ende 3 weiße , 2 rote und 1 blaue Kugel gezogen wurden. Die Reihenfolge spielt dabei keine Rolle, sondern nur das Ergebnis am Ende der Ziehung. Deshalb kann man hier die Reihenfolge so umändern, daß zuerst die weißen, dann die roten und zuletzt die blauen Kugeln erscheinen.

$$P(\ w\ ;\ w\ ;\ w\ ;\ r\ ;\ r\ ;\ b\ )$$
$$\frac{6!}{3!\cdot 2!\cdot 1!}\cdot\frac{10}{40}\cdot\frac{10}{40}\cdot\frac{10}{40}\cdot\frac{7}{40}\cdot\frac{7}{40}\cdot\frac{8}{40}$$

Die einzelnen Wahrscheinlichkeiten sind die gleichen wie bei „mit Reihenfolge“ . Da aber jetzt die Reihenfolge keine Rolle spielt sind alle Pfade zu addieren, bei denen 3 weiße, 2 rote und 1 blaue auftreten. In der Kombinatorik ist das Permutation mit Wiederholung. Die Anzahl der Pfade setzt sich zusammen aus der Anzahl der zu ziehenden Kugeln, nämlich 6! . Das ist die Anzahl 6 Kugeln in Reihenfolge anzuordnen. Aber von den 6 Kugeln sind 3 weiße nicht zu unterscheiden, 2 rote nicht zu unterscheiden und 1 blaue nicht zu unterscheiden. Deshalb ist durch die Permutationsanzahl diese Farben zu dividieren. Dadurch entsteht der Vorfaktor, der nichts anderes ausdrückt als die Anzahl der Pfade in dieser Farbkombination.

32.1.3.3 Ohne Zurücklegen ; mit Reihenfolge

Aus der Urne werden 6 Kugeln ohne zurücklegen gezogen. Gesucht ist die Wahrscheinlichkeit für das Ergebnis (w; w ; r; r; b; w) in genau dieser Reihenfolge.

$$P(\ w\ ;\ w\ ;\ r\ ;\ r\ ;\ b\ ;\ w\ )$$
$$\frac{10}{40}\cdot\frac{9}{39}\cdot\frac{7}{38}\cdot\frac{6}{37}\cdot\frac{8}{36}\cdot\frac{8}{35}$$

Da mit Beachtung der Reihenfolge gezogen wird, gibt es nur einen Zweig im Baumdiagramm, in dem diese Reihenfolge auftritt. Es ist deshalb nicht mit einer Anzahl von Pfaden zu multiplizieren. Da ohne Zurücklegen gezogen wird, sind es bei jeder neuen Ziehung eine Kugel weniger, und auch eine Kugel der jeweiligen Farbe weniger. Damit ändern sich sowohl Zähler als auch Nenner. Andererseits ist die Wahrscheinlichkeit an jedem Pfad die gleiche, wenn nur die Anzahl der Farben die gleiche ist. Man kann also auch die Reihenfolge ändern in (w ; w ; w ; r ; r ; b). Dieser Pfad hat die gleiche Wahrscheinlichkeit, wie der oben angegebene:

$$P(\ w\ ;\ w\ ;\ w\ ;\ r\ ;\ r\ ;\ b\ )$$
$$\frac{10}{40}\cdot\frac{9}{39}\cdot\frac{8}{38}\cdot\frac{7}{37}\cdot\frac{6}{36}\cdot\frac{8}{35}$$

Es handelt sich dabei um die gleichen Zähler und Nenner, die nur in ihrer Reihenfolge vertauscht sind. Da es sich aber im Zähler und Nenner um Produkte handelt, spielt das keine Rolle.

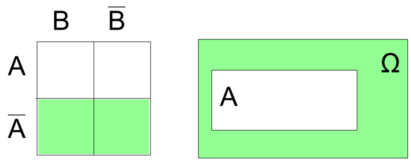
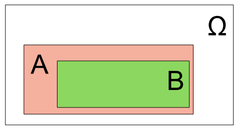
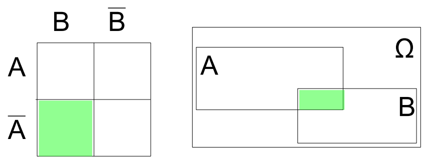
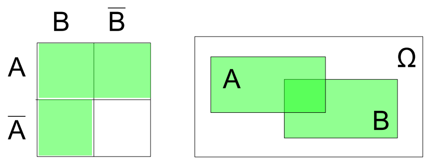
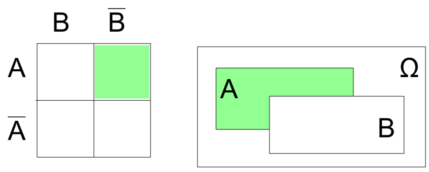
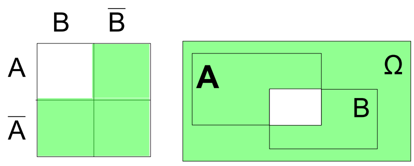
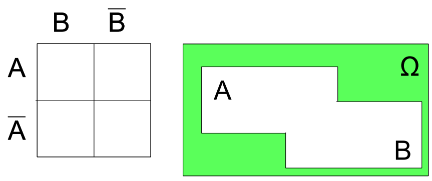
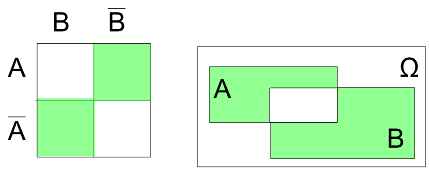
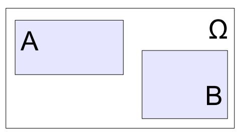
32.1.3.4 Ohne Zurücklegen ; ohne Reihenfolge

Bei einer Ziehung ohne Berücksichtigung der Reihenfolge ändert sich an den Einzelwahrscheinlichkeiten nichts, es kommt lediglich als Faktor noch die Anzahl der Pfade davor. Die Anzahl der Pfade ist aber die gleiche wie beim Ziehen mit Zurücklegen. Im Baum unterscheidet sich „ziehen mit Zurücklegen“ und „Ziehen ohne Zurücklegen“ nur durch die Wahrscheinlichkeiten an den Pfaden, aber nicht durch die Anzahl.

$$P(\ w\ ;\ w\ ;\ w\ ;\ r\ ;\ r\ ;\ b\ )$$
$$\frac{6!}{3!\cdot 2!\cdot 1!}\cdot\frac{10}{40}\cdot\frac{9}{39}\cdot\frac{8}{38}\cdot\frac{7}{37}\cdot\frac{6}{36}\cdot\frac{8}{35}$$

32.2 Das Venn – Diagramm

Eine Möglichkeit der Darstellung von Zufallsereignissen ist die Darstellung als Menge. Es hat sich gezeigt, das die Gesetze aus der Mengenlehre sehr gut auf die Wahrscheinlichkeiten zu übertragen sind. Deshalb sollen hier einige Grundsätze der Mengenlehre angegeben werden.

Ereignis	Beschreibung	Venn und Mengendarstellung
A	Das **Gegenereignis** $\overline{A}$ tritt genau dann ein, wenn A nicht eintritt.	
B ⊆ A	Das Ereignis B **zieht** das Ereignis A **nach sich** ; immer, wenn B eintritt, tritt auch A ein.	
A ∩ B	Das Ereignis **A und B** (A geschnitten B) tritt genau dann ein, wenn *sowohl A als auch B eintritt*	
A U B	Das Ereignis **A oder B** (A vereinigt B) tritt genau dann ein, wenn mindestens eines der Ereignisse A, B eintritt.	
A \ B	Das Ereignis **A und nicht B** (A minus B) tritt genau dann ein, wenn A eintritt und B nicht eintritt. A \ B = A ∩ B	
A ∩ B	**Höchsten eines** der Ereignisse A,B tritt ein, wenn entweder A oder B oder keines von beiden eintritt. A ∩ B = A U B	
A U B	Das Ereignis **Weder A noch B** tritt genau dann ein, wenn keines der beiden Ereignisse A,B eintritt. A U B = A ∩ B	
(A ∩ $\overline{B}$) U ($\overline{A}$ ∩ B)	Das Ereignis **Entweder A oder B** tritt genau dann ein, wenn das andere Ereignis **nicht** eintritt	
A ∩ B = Ø	Die Ereignisse A und B sind unvereinbar, d.h. sie können nicht gleichzeitig eintreten.	

Dabei stellt die 1. Grafik in der dritte Spalte das sogenannte Venn-Diagramm dar und die 2. Grafik die Mengendarstellung. In der folgenden Tabelle sollen einige grundlegende Rechenregeln zusammengestellt werden.

$A \cup B = B \cup A$	$A \cap B = B \cap A$	Kommutativgesetz
$(A \cup B) \cup C = A \cup (B \cup C)$	$(A \cap B) \cap C = A \cap (B \cap C)$	Assoziativgesetz
$(A \cup B) \cap C = (A \cup C) \cap (B \cup C)$ $(A \cap B) \cup C = (A \cap C) \cup (B \cap C)$		Distributivgesetz
$A \cup \varnothing = A$	$A \cap \Omega = A$	Neutrales Element
$A \cup \Omega = \Omega$	$A \cap \varnothing = \varnothing$	Dominantes Element
$A \cup \overline{A} = \Omega$	$A \cap \overline{A} = \varnothing$	Komplementäres Element
$A \cup A = A$	$A \cap A = A$	Idempotenzgesetz
$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	Absorptionsgesetz
$\overline{A \cup B} = \overline{A} \cap \overline{B}$	$\overline{A \cap B} = \overline{A} \cup \overline{B}$	de Morgan Gesetze

32.3 Baumdiagramme

Zufallsexperimente sind gekennzeichnet

- durch mehrere Ausgänge und
- durch Nichtvorhersagbarkeit der Ausgänge.

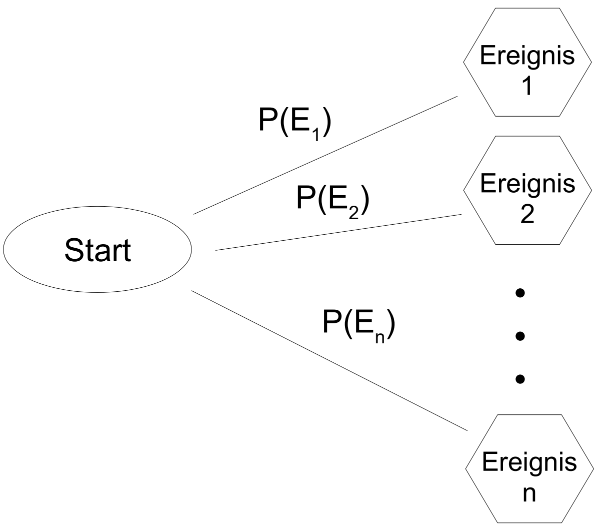
Es ist also grundsätzlich mit mehreren Ausgängen zu rechnen und es stellt sich das Problem, wie stellt man ein solche Ergebnisse dar. Bei dieser Umsetzung hat sich die die Darstellung als sogenannter Ereignisbaum als vorteilhaft erwiesen. Dieser Ereignisbaum besteht aus einem Startpunkt, der den Beginn des Experiments symbolisiert und aus verschiedenen Linien von diesem Startpunkt, Zweige genannt, die die möglichen Ergebnisse des Experiments symbolisieren. Für jedes Ergebnis existiert ein Zweig in diesem Baum. Diese Darstellungsweise bietet einige Vorteile. Es lassen sich theoretisch beliebig viele Zufallsexperimente hintereinander darstellen (die Begrenzung wird durch die Größe des Blattes gegeben). Es muss nicht jedem Ergebnis ein Zweig zugewiesen werden, es lassen sich auch mehrere Ergebnisse zu einem Ereignis zusammenfassen und als ein Zweig im Baum darstellen, wenn keine weiteren Unterscheidungen gewünscht sind.
Auf dieser Grundlage lassen sich verschiedene Baumtypen unterscheiden, die getrennt betrachtet werden können. Der Beginn eines jeden Experiments erzeugt einen Anfangsbaum, bei dem noch keine Unterscheidungen notwendig sind, da für alle Typen der Anfang gleich ist.
Alle Möglichen Ereignisse sind in ihrer Ausgangssituation – Anzahl – vorhanden

Für die weiteren Ereignisse ist entscheidend, wie wird nach dem ersten Ereignis weiter Verfahren. Diese Entscheidung hat Einfluß auf die Ergebnisse der zweiten Durchführung.

32.3.1 Einstufiges Zufallsexperiment

In einem beliebigen Zufallsexperiment gibt die Ergebnismenge, die wir mit dem großen griechischen Buchstaben Ω (Omega) bezeichnen, die Menge aller Ergebnisse (also alle Möglichkeiten, die auftreten können) an.

Am Ende eines jeden Zweiges wird das Ereignis eingetragen, das durch diesen Zweig interpretiert wird und an die Zweige selbst trägt man die Wahrscheinlichkeit ein, mit der dieses Ereignis zu erwarten ist. Diese Darstellung ist so allgemein, dass man alle auftretenden Situationen damit abdecken kann. Nach jedem Ereignis besteht die Möglichkeit wieder einen Baum anzuschließen, der auch völliges anderes Aussehen haben kann.



32.3.2 Mehrstufiges Zufallsexperiment

Einstufige Zufallsexperimente werden durch mehrmals hintereinander Ausführen zu mehrstufigen Zufallsexperimenten. Beispiele hierfür sind das mehrfache Würfeln oder das Ziehen von Kugeln aus Urnen. Die Ergebnisse werden dann als Zahlen- oder Buchstabenkolonnen geschrieben. Eine solche Kolonne hat genau so viele Einträge, wie das Zufallsexperiment wiederholt wurde.

Ist diesem Zusammenhang ist es interessant das Kreuzprodukt zweier Mengen zu kennen. Es wurde schon viel mit diesem Kreuzprodukt gearbeitet, ohne dass es als solches auch bezeichnet wurde. Eines der markantesten Gebiete ist die x-y-Koordinatenebene. Zeichnet man ein Koordinatensystem – wenn gar noch auf kariertem Papier – dann ist jeder Kreuzungspunkt der Linien ein Element aus des Kreuzproduktes, das aus der x-Achse und der y-Achse gebildet wird. Mathematisch kann man dieses Tatsache folgendermaßen formulieren:

$$X \times Y = \{ (x;y) \mid x \in X \text{ und } y \in Y\}$$

Das Kreuzprodukt von zwei Mengen umfasst alle Paare von Elementen, bei denen das erste Element aus der ersten Menge und das zweite Element aus der zweiten Menge sind. Solche Elemente eines Kreuzproduktes bezeichnet man auch als geordnete Paare, da zu unterscheiden ist, an welcher Stelle in der Reihenfolge ein Element steht. Für ein Koordinatensystem ist es sofort einleuchtend, dass das Paar (1;3) ein anderes Paar ist, wie (3;1). Die Reihenfolge der Anordnung spielt also eine entscheidende Rolle.

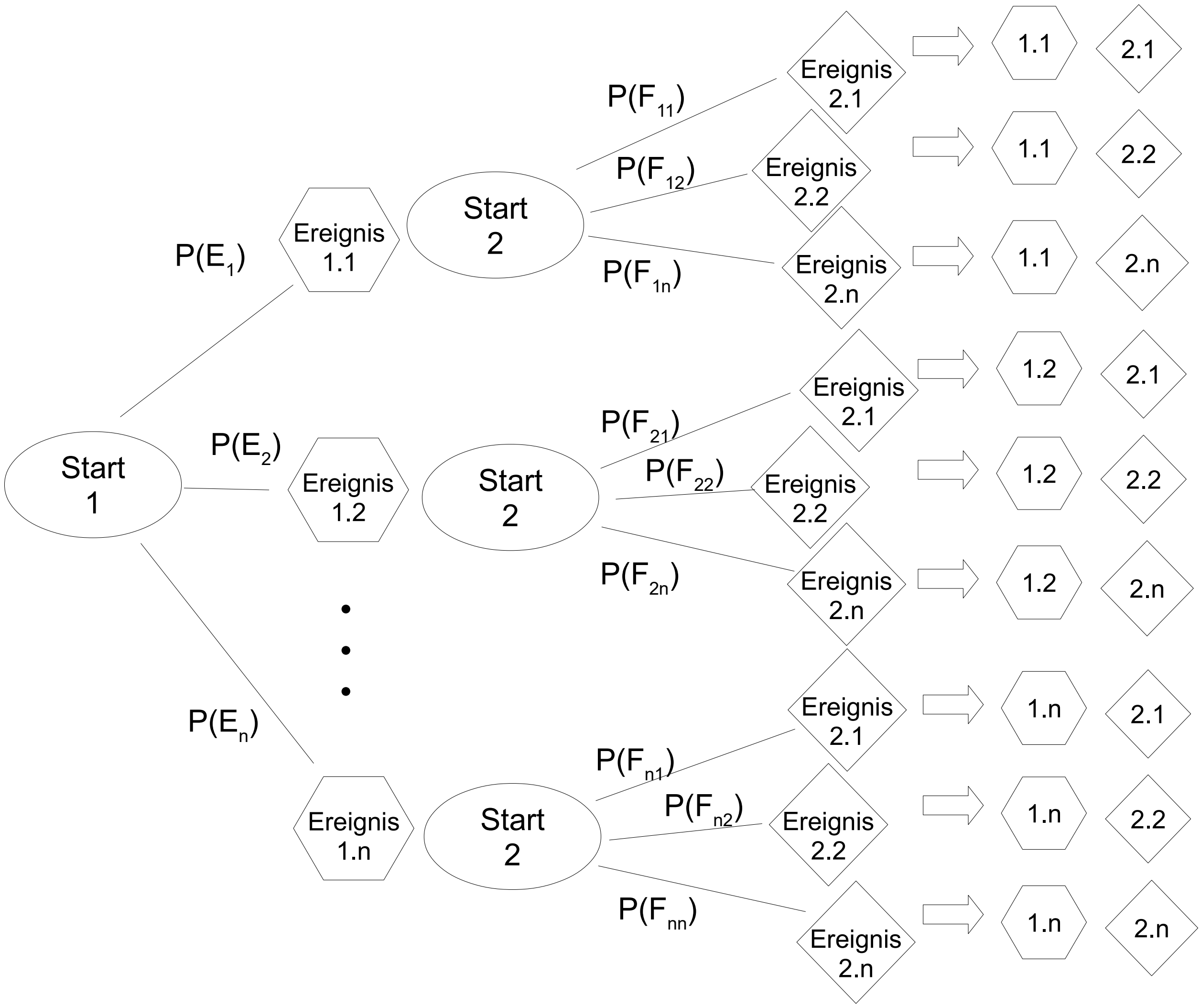
Diesen Konstrukt verwendet man in der Wahrscheinlichkeitsrechnung, um die Ergebnisse von Zufallsexperimenten darzustellen. Es ist eine weniger aufwendige Form, als das Baumdiagramm, dafür nicht so übersichtlich, aber mit beliebiger Länge zu versehen.
Man stellt sich jede Menge als ein Zufallsexperiment vor. Dann ist das Kreuzprodukt von X und Y die Menge aller Elemente, bei der an 1. Stelle ein Ereignis des 1. Zufallsexperiment X und an 2. Stelle ein Ereignis des 2. Zufallsexperiment Y steht. Wenn man diese Konstruktion auf beliebig viele nacheinander durchführbare Zufallsexperimente erweitert kommt man zum Begriff des n-Tupel:

Definition:
Ein **n-Tupel** $(x_1; x_2; \ldots x_n)$ ist eine geordnete Menge von Ereignissen, bei dem die 1. Position einem Ereignis bei der 1. Durchführung des Zufallsexperiments zugeordnet ist, die 2. Position einem Ereignis der 2. Durchführung des Zufallsereignisses usw.

Diese Definition ist so allgemein, dass alle möglichen Zufallsexperimente damit abgedeckt sind. Insbesondere ist es möglich, dass das folgende Zufallsexperiment mit dem ersten nichts zu tun hat. Als Beispiel könnte man folgendes Zufallsexperiment anführen: Es gibt 6 verschiedenen Urnen, in denen jeweils Kugeln in drei verschiedenen Farben enthalten sind. Das zweistufige Zufallsexperiment sieht nun folgendermaßen aus: Mit einem Würfel wir zuerst eine Urne ausgewählt und dann aus der Urne eine Kugel gezogen. Die Ereignisse des 2. Zufallsexperiments sind völlig andere als die des ersten, trotzdem kann man alle möglichen Kombinationen als ein solches n-Tupel darstellen. z. B (2; rot)

32.3.3 Darstellung des Baumes

Als nächstes sollen allgemeine Eigenschaften eines mehrfach aneinander gereihten Baumes untersucht werden. Dazu soll ein beliebiger zweistufiger Zufallsversuch dargestellt werden.



In dieser allgemeinen Baumstruktur ist alles enthalten, was bei einem mehrstufigen Zufallsversuch zu beachten ist.

1. Die Wahrscheinlichkeiten des 1. Versuchs haben mit den des zweiten Versuchs nichts zu tun! Der zweite Versuch startet auf seiner eigenen Grundlage, die eventuell vom Ausgang des ersten Versuchs beeinflußt sein kann.
2. Dass im zweiten Versuch alle zulässigen Ereignisse gleich sind, bedeutet nicht, dass auch die Wahrscheinlichkeiten gleich sind.

Diese beiden Erkenntnisse führen zu der Konsequenz bei mehrstufigen Versuchen:

- Für jeden Folgeversuch sind die Voraussetzungen des Versuchs genau zu untersuchen.

Damit kann man mehrstufige Versuche in zwei große Gruppen einteilen, die später noch weiter im einzelnen besprochen werden.

32.3.3.1 Unabhängige Versuche

Die erste Gruppe betrifft solche Zufallsexperimenten, bei denen die Durchführung des zweiten Experiments nicht vom Ausgang des ersten Experiments abhängt. Diese Versuche heißen unabhängige Versuche.

Merkmal:

Vor der Durchführung des folgenden Versuches ist die Ausgangssituation in allen Fällen die gleiche.
Die Wahrscheinlichkeiten an den Ästen des Baumes sind in allen Bäumen der 2. Stufe gleich.

Beispiele dafür sind alle Versuche, die mit Würfeln oder glücksradähnlichen Ereignissen erstellt werden. Nach jedem Versuch sind auf dem Würfel noch alle Zahlen vorhanden, bzw die die Farb- oder Zahlenverteilung auf einem Glücksrad sind die gleichen wie beim ersten Versuch.
Aber auch Versuche, die sich aus anderen gründen nicht gegenseitig beeinflussen: Wenn man aus einem Kleiderschrank ein Hose auswählt hat das keinen Einfluß auf die Auswahl eines Hemdes (vom Geschmack abgesehen). Es ist völlig uninteressant, ob man erst das Hemd oder erst die Hose auswählt, die Anzahl der Kombinationsmöglichkeiten sind gleich.
In diese Gruppe fallen alle Zufallsereignisse, die mit der Bezeichnung „.. mit Zurücklegen..“ bezeichnet werden.

32.3.3.2 Abhängige Versuche

Hier wird die Durchführung des zweiten Experiments **vom Ausgang** des ersten Experiments beeinflußt.

Merkmal:

Vor der Durchführung des folgenden Versuches findet man unterschiedliche Voraussetzung für den Versuch vor.
Die Wahrscheinlichkeiten an den Ästen der Bäume der 2. Stufe sind von Baum zu Baum verschieden.

Beispiele dafür sind alle Versuche, die von vorn herein unterschiedliche Voraussetzungen haben. dazu zählt das oben erwähnte Beispiel der 6 Urnen mit Kugeln in drei verschiedenen Farben. Je nachdem, welche Urne durch den Würfel ausgewählt wird, sind die Wahrscheinlichkeiten die Kugel einer Farbe zu ziehen unterschiedlich.
In diese Gruppe fallen alle Versuche, die mit der Bezeichnung „... ohne Zurücklegen...“ bezeichnet werden, aber auch die, die „ auf einmal gezogen ..“ werden, weil dann auch kein Zurücklegen erfolgen kann. Dazu zählen alle Mehrfachauswahlen aus einer Menge. Aus einer Menge von Personen werden 3 ausgewählt, wie viele Möglichkeiten der Auswahl gibt es. Es ist natürlich einleuchtend, dass eine Person nicht mehrfach ausgewählt werden kann, da sie nur einmal existiert. Nach jeder Auswahl existiert für den Folgeversuch eine andere Voraussetzung unter der dieser gestartet wird. Man kann die Personen auf einmal auswählen oder man kann dreimal durch die Reihe gehen, die Möglichkeiten der Auswahl ändern sich dadurch nicht.

In diesem Zusammenhang ist auch der Begriff der Bedingten Wahrscheinlichkeit zu sehen, der später erklärt wird.

32.3.4 1. Pfadregel

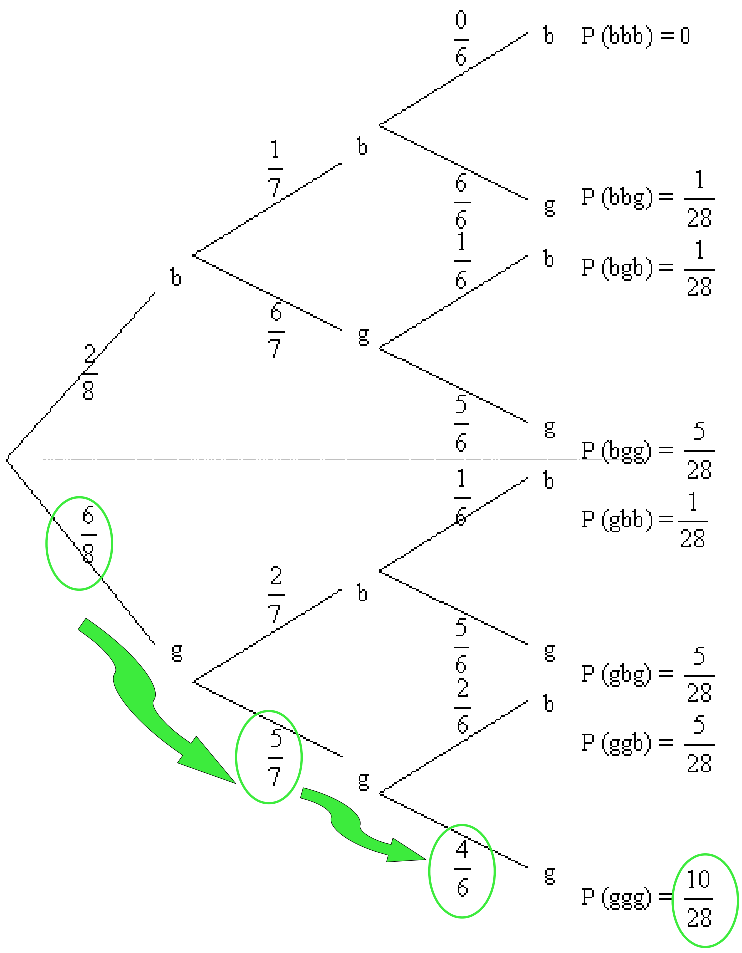
Die erste Frage, die sich bei einem solchen Baum stellt ist die, mit welcher Wahrscheinlichkeit tritt eine der möglichen Kombinationen aus erstem und zweitem Versuch auf.

Mit welcher Wahrscheinlichkeit kann man eine Kombination 1.k 2.m

(k;m) erwarten, wobei k ein Ereignis des ersten Versuchs und m ein Ereignis des zweiten Versuchs ist. Fest steht, dass man das Ereignis 1.k mit einer Wahrscheinlichkeit $P(E_k)$ erwarten kann. Das ist die Voraussetzung, dass man überhaupt eine solche Kombination erreichen kann. Nur wenn das eingetreten ist, kann man auf eine solche Kombination zweier Ereignisse hoffen. Wenn es eingetreten ist, kann man aber nur mit einer Wahrscheinlichkeit $P(F_{km})$ auch das zweite Ereignis erwarten. Also muss sowohl die Wahrscheinlichkeit $P(E_k)$ als auch die Wahrscheinlichkeit $P(F_{km})$ eingetreten sein. Dieses „sowohl – als – auch“ drückt sich durch das Produkt der Wahrscheinlichkeiten aus.

Multiplikationsregel: Bei einem mehrstufigen Zufalls-Versuch ist die **Wahrscheinlichkeit eines Ergebnisses** (eines Pfades im Baumdiagramm) gleich dem **Produkt der Wahrscheinlichkeiten** längs des zugehörigen Pfades im Baumdiagramm.

Aufgabe: In einer Urne Liegen 8 Kugeln 6 gelbe (g) und 2 blaue (b) Kugeln. Es wird nacheinander ohne Zurücklegen gezogen. Wie hoch ist die Wahrscheinlichkeit bei dreimaligem Ziehen dreimal eine gelbe Kugel zu ziehen.



Ergebnis:
Die Wahrscheinlichkeit für „ggg“ liegt bei

6/8 * 5/7 * 4/6 = 120/336 = 10/28 = 0,357

Die Wahrscheinlichkeitsangaben entlang der Pfade sind unterschiedlich. Für das Eintreten eines Ergebnisses b stehen am Baum des ersten Versuchs andere Wahrscheinlichkeiten, als an den Bäumen des zweiten Versuchs. Es muss also prinzipiell davon ausgegangen werden, dass die Wahrscheinlichkeiten späterer Versuche durch die Ergebnisse früherer versuche beeinflußt werden. Deshalb kann man nicht prinzipiell die Wahrscheinlichkeiten der Einzelereignisse multiplizieren, sondern muss folgende Formel zugrunde legen:

Multiplikationsregel: $P(A \cap B) = P(A)\ P(B|A) = P(B)\ P(A|B)$

Es ist das Produkt zu bilden aus der Wahrscheinlichkeit des Ersten Versuchs P(A) und der Wahrscheinlichkeit des Ergebnisses B **unter der Bedingung, dass im ersten Versuch eingetreten ist P(B|A)**. Es wird später noch zu klären sein, wie solche Fälle aussehen, bei denen P(B|A) = P(B) ist.

32.3.5 2. Pfadregel

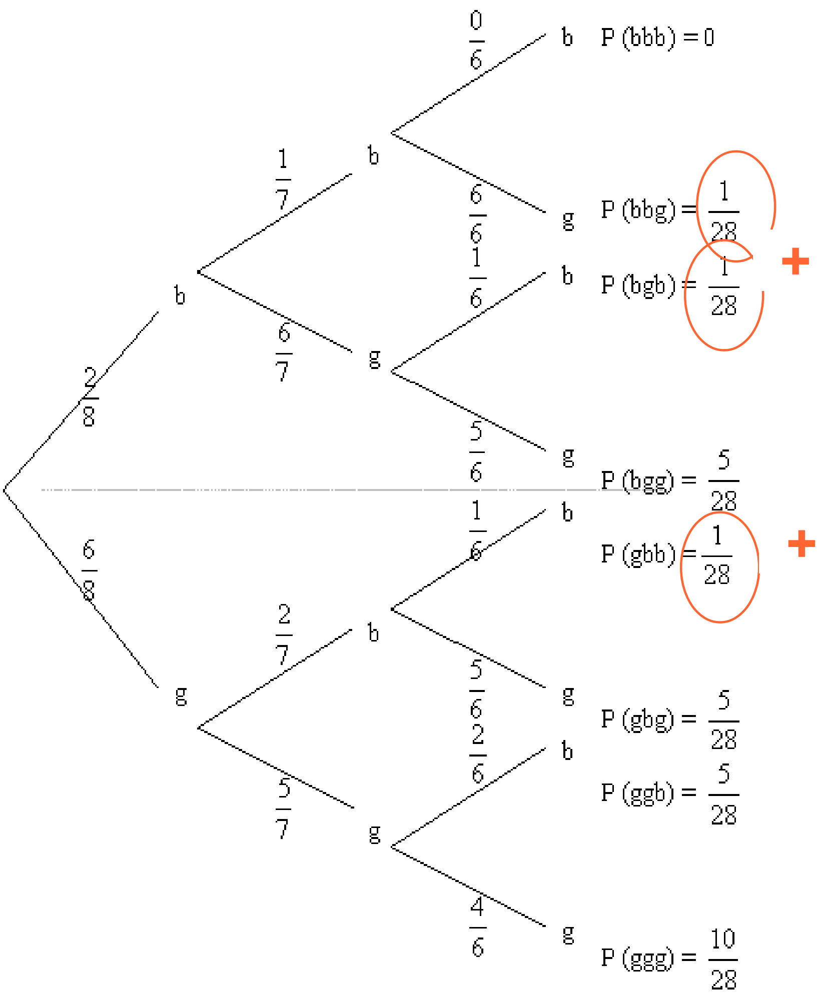
Die zweite Frage, die aus dem Baumdiagramm zu klären ist, ist die: Mit welcher

Wahrscheinlichkeit erreicht man im zweiten Versuch immer das Ereignis 2.m

Dazu muss man berücksichtigen, dass das Ereignis immer zu erreichen ist, gleichgültig, welchen Ausgang der erste Versuch genommen hat, Nur sind in Abhängigkeit vom ersten Versuch die Wahrscheinlichkeiten für das Ereignis jedes Mal anders: $P(F_{1m})$, $P(F_{2m})$; … $P(F_{nm})$. Alle diese Ereignisse führen dazu, dass im zweiten Versuch das Ereignis 2;m auftritt. Aber die einzelnen Ereignisse treten nur ein, wenn vorher das entsprechende Ereignis des ersten Versuchs eingetreten ist. Es sind also mehrere Wege möglich, dass im zweiten Versuch das Ereignis 2;m eintritt. Die Wahrscheinlichkeiten aller dieser Wege sind deshalb zu addieren.

> **Additionsregel:** Setzt sich bei einem mehrstufigen Zufalls-Versuch ein **Ereignis aus verschiedenen Pfaden** (im Baumdiagramm) zusammen, dann erhält man die Wahrscheinlichkeit des Ereignisses durch **Addition der einzelnen Pfadwahrscheinlichkeiten**.

Aufgabe: In einer Urne Liegen 8 Kugeln 6 gelbe (g) und 2 blaue (b) Kugeln. Es wird nacheinander ohne Zurücklegen gezogen. Wie hoch ist die Wahrscheinlichkeit bei dreimaligem Ziehen eine gelbe Kugel und zwei blaue Kugeln zu ziehen.



Ergebnis:

1/28

\+ 1/28

\+ 1/28 = 3/28

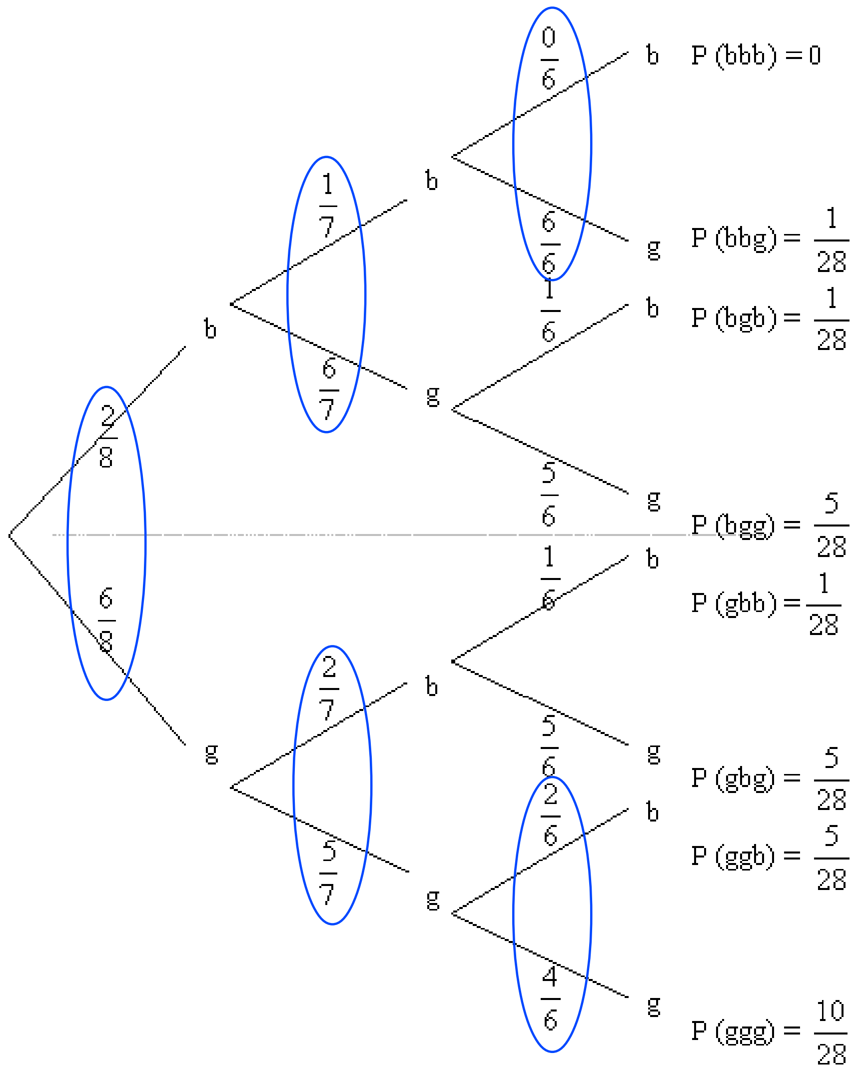
Jedes Ergebnis entlang eines Pfades ist durchschnittsfremd zu den anderen Ergebnissen. Deshalb ist hier eine einfache Addition der Wahrscheinlichkeiten der Pfade ausreichend. Es gibt aber auch Ereignisse, die nicht durchschnittsfremd sind und auf zwei verschiedene Weisen zu gleichen Ergebnis führen. Solche Werte würden beim einfachen Addieren doppelt erfasst, deshalb sind alle Ereignisse des Durchschnitts von dieser Summe wieder einmal zu subtrahieren. Die allgemeine Additionsformel sieht deshalb folgendermaßen aus:

Additionsregel: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

32.3.6 3. Pfadregel

Jeder Baum in jeder Stufe hat noch eine weitere Eigenschaft, die man zu Kontrollzwecken bei der Aufstellung eines Baumes gut nutzen kann. Jeder Baum, gleich welcher Stufe enthält natürlich alle Möglichkeiten, die ein Versuchsausgang haben kann. Diese Eigenschaft „alle Möglichkeiten" bedeutet mathematisch, dass ein Ereignis auf alle Fälle eintritt und deshalb die Summe der Wahrscheinlichkeiten aller Zweige eines Baumes 1 sein muß.

> **Verzweigungsregel:** Bei mehrstufigen Zufallsexperimenten ist die **Summe der Wahrscheinlichkeiten**, die von einem Knoten ausgehen immer **gleich 1**.



Damit sind die wichtigsten Kriterien für einen mehrstufigen Zufallsversuch durch Darstellung in einem Ereignisbaum zusammengetragen. Als nächstes sollen spezielle Bäume untersucht werden.

32.3.7 Zweistufige Bäume

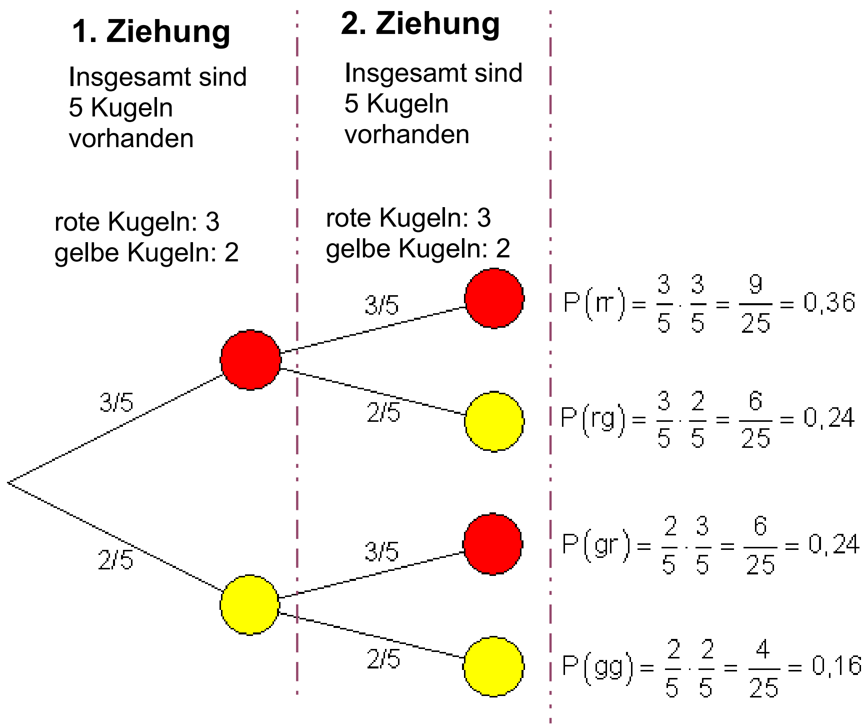
Als erstes sollen zweistufige Bäume mit ihren Wahrscheinlichkeiten untersucht werden, da die hier gewonnenen Erkenntnisse auf größere Bäume übertragbar sind. Dazu werden die Versuche als Ziehungen von verschiedenfarbigen Kugeln aus Urnen dargestellt und die Eigenschaften analysiert.

32.3.7.1 Zufallsversuche „ mit Zurücklegen"

Als erstes sollen hier Zufallsversuche betrachtet werden, die unter die Gruppe „mit Zurücklegen" fallen.

Es soll eine Urne mit 2 verschiedenfarbigen Kugeln betrachtet werden: 3 rote, 2 gelbe. Nach der ersten Ziehung wird die gezogene Kugel wieder zurückgelegt und es wird erneut gezogen.

Das Baumdiagramm für diesen Versuch hat folgendes Aussehen:



Betrachtet man im Baumdiagramm die erste Ziehung für die rote Kugel, dann ist die Wahrscheinlichkeit P(A) dafür:

- eine rote Kugel aus der Gesamtmenge Ω zu ziehen: $P_\Omega(A) = 3/5$

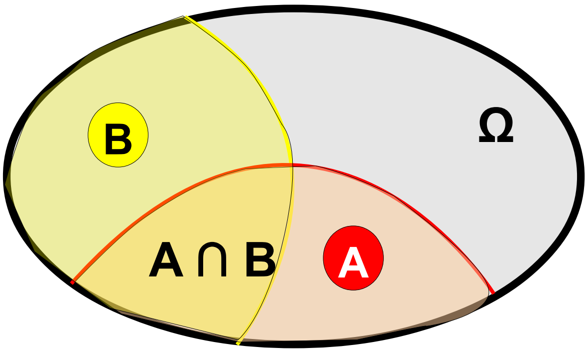
  (= die Wahrscheinlichkeit für das Eintreten des Ereignisses A unter der Bedingung der Grundmenge Ω)

Betrachtet man die zweite Ziehung für die rote Kugel, dann ist die Wahrscheinlichkeit P(A) dafür:

- eine rote Kugel aus der Gesamtmenge Ω zu ziehen: $P_\Omega(A) = 3/5$

  (= die Wahrscheinlichkeit für das Eintreten des Ereignisses A unter der Bedingung der Grundmenge Ω)

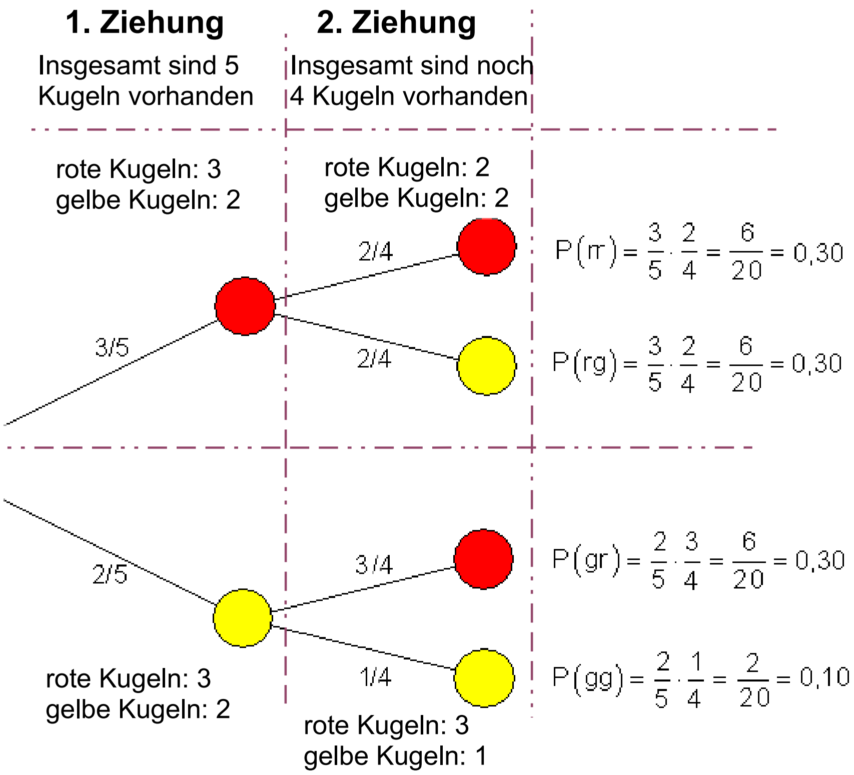
da durch das Zurücklegen die Ausgangsmenge wieder zurückgesetzt wurde.
Die Wahrscheinlichkeiten an den Bäumen der 2. Ziehung sind identisch mit denen der 1. Ziehung



Für das Ereignis, dass zuerst eine rote Kugel gezogen wird ist die Wahrscheinlichkeit P(A). Die Wahrscheinlichkeit, dass zuerst eine gelbe Kugel gezogen wird ist P(B).
ABER:
Die Wahrscheinlichkeit, dass als zweites eine gelbe Kugel gezogen wird, ist ebenfalls P(B), da sich die Versuche in keiner Weise unterscheiden: Es gibt 5 Kugel, 3 rote und 2 gelbe.

32.3.7.2 Zufallsversuche „ohne Zurücklegen“

Es soll jetzt die gleiche Versuchsanordnung gewählt werden, aber die Kugel nach dem ersten Ziehen nicht wieder zurückgelegt werden. Dann entsteht folgendes Baumdiagramm:



Bei diesem Versuch haben sich nicht nur die Wahrscheinlichkeiten am Ende der Äste verändert, sondern die Wahrscheinlichkeiten an den beiden Bäumen des zweiten Versuchs sind auch unterschiedlich. Das führt dazu, dass die Wahrscheinlichkeiten eine rote und eine gelbe Kugel zu ziehen sich völlig unterschiedlich zusammensetzen, je nachdem, ob erst eine rote Kugel oder erst eine gelbe Kugel gezogen wird.

Betrachtet man im Baumdiagramm die erste Ziehung für die rote Kugel, dann ist die Wahrscheinlichkeit P(A) dafür:

- eine rote Kugel aus der Gesamtmenge Ω zu ziehen: $P_\Omega(A) = 3/5$

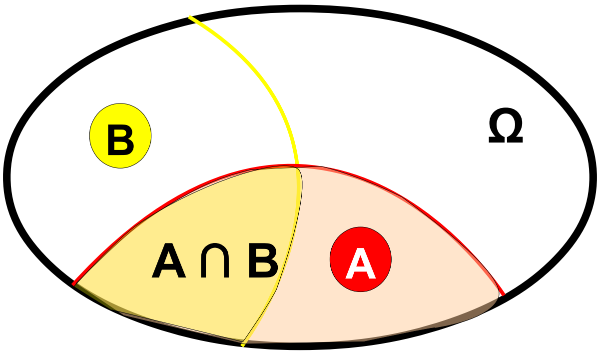
(= die Wahrscheinlichkeit für das Eintreten des Ereignisses A unter der Bedingung der Grundmenge Ω)

Betrachtet man die zweite Ziehung für die rote Kugel, dann ist die Wahrscheinlichkeit P(A) dafür:

- eine rote Kugel aus einer Menge zu ziehen, aus der bereits eine (rote oder gelbe) gezogen wurde:
  $P_A(B) = 2/4$, $P_A(\overline{B}) = 2/4$, $P_{\overline{A}}(B) = 3/4$, $P_{\overline{A}}(\overline{B}) = 1/4$.

(= die Wahrscheinlichkeit für das Eintreten des Ereignisses B (oder $\overline{B}$) unter der Bedingung, dass das Ereignis A (oder $\overline{A}$) eingetreten ist.)

Beim zweiten Versuch hat sich die Ausgangsmenge für alle Versuche geändert, alle Wahrscheinlichkeiten haben sich gegenüber dem ersten Versuch geändert.



Für das Ereignis, dass zuerst eine rote Kugel gezogen wird ist die Wahrscheinlichkeit P(A). Die Wahrscheinlichkeit, dass zuerst eine gelbe Kugel gezogen wird ist P(B).
ABER:
Die Wahrscheinlichkeit, dass als zweites eine gelbe Kugel gezogen wird, ist nicht mehr P(B), da sich die Anzahlen der Kugeln geändert haben. Nach dem ersten Versuch ist die Ausgangsmenge nicht mehr Ω.

Die Wahrscheinlichkeit, im zweiten Versuch eine gelbe zu ziehen ist nicht die gleiche, wie die Wahrscheinlichkeit im ersten Versuch eine gelbe zu ziehen, da im ersten Versuche entweder eine gelbe oder eine rote Kugel entfernt wurde. Damit hat sich nicht nur die Gesamtzahl der Kugeln geändert, je nachdem, was im ersten Versuch gezogen wurde, sondern es hat sich auch die Farbzusammensetzung der Kugeln geändert: 2 gelbe, 2 rote oder 1 gelbe, 3 rote.
Im Baumdiagramm wird deutlich, dass es sich um zwei getrennte Versuche handelt.

1. Versuch: 3 rote und 2 gelbe Kugeln, P1(A) ist die Wahrscheinlichkeit, dass eine rote Kugel gezogen wurde.
2. Versuch: 2 rote und 2 gelbe Kugeln, P1(B) ist die Wahrscheinlichkeit, dass eine gelbe Kugel gezogen wurde.

Das Baumdiagramm zeigt auch, dass sich die Summen der Zweige im 2. Baum wieder auf 1 addieren. Für P2(B) ist also die tatsächlich vorhandene Ausgangsmenge P1(A). P2(B) wird auf P1(A) normiert, P1(A) ist für den Versuch 2 die neue Ausgangsmenge Ω2.

$$P2(B) = P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Aus diesen Gründen bezeichnet man PA( ) auch als ein Wahrscheinlichkeitsmaß, das nicht wie P( ) auf dem Grundraum Ω, sondern auf dem neuen Grundraum der Ergebnismenge von A definiert ist.

32.4 Kontingenztafeln und Vier-Felder-Tafeln

Zweistufige Versuche lassen sich außer mit einem Baumdiagramm auch tabellarisch bearbeiten. Bei diesen Tabellen spricht man von **Kontingenztafeln**, bei denen alle möglichen Ergebnisse des ersten Versuchs in Zeilen eingetragen werden und alle möglichen Ergebnisse den zweiten Versuchs in Spalten.

	B_1	...	B_s	Σ
A_1	h_{11}	...	h_{1s}	$h_{1.}$
	h_{ij}			
A_r	h_{r1}	...	h_{rs}	$h_{r.}$
Σ	$h_{.1}$	...	$h_{.s}$	$h_{..} = 1$

Gibt es bei solchen Versuchen jeweils nur zwei mögliche Ereignisse, die untersucht werden sollen reduziert sich diese Tafel auf eine 4-Felder-Tafel:

	B	$\overline{B}$	Zeilensumme
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
Spalten-summe	$P(B)$	$P(\overline{B})$	1

Für die jeweils zwei möglichen Ereignisse eines jeden Versuchs kann man auch schreiben „*das Ereignis A tritt ein*“ und „*das Ereignis A tritt nicht ein*“. In der obigen Tabelle durch A und $\overline{A}$ gekennzeichnet, denn das Eintreten des zweiten möglichen Ereignisses ist gleichbedeutend damit, dass das erste nicht eingetreten ist, denn es gibt ja keine weiteren. Jeder solchen Tafel kann man eindeutig ein Baumdiagramm zuordnen und jedem zweistufigen Baumdiagramm eine solche Kontingenztafel. Wenn die Anzahl der Versuche nur 2 beträgt, dafür aber die Anzahl der möglichen Ausgänge eines Versuchs groß ist, kann man übersichtlicher mit diesen Tafeln an Stelle eines Baumdiagramms arbeiten. Das trifft bereits schon dann zu, wenn man Zufallsversuche analysiert, die durch zweimaliges Würfeln entstehen. Dann gibt es beim ersten Versuch sechs Möglichkeiten und beim zweiten Versuch für jeden Ausgang des ersten Versuchs wieder sechs Möglichkeiten. Dadurch entsteht schon ein recht großer Baum. Einen solche Versuchsdurchführung stellt man übersichtlicher in einer Vierfeldertafel dar, die dann allerdings aus 36 Feldern besteht.

32.4.1 Vierfeldertafel für Zufallsversuche „mit Zurücklegen“

Das im Kapitel über das Baumdiagramm angegebene Beispiel eines zweistufigen Versuchs hat mit der Vierfeldertafel folgendes Aussehen:

	B	$\overline{B}$	Zeilensumme
A	**P(rg)=0,24**	**P(rr)=0,36**	**P(A)=3/5 = 0,24+0,36**
$\overline{A}$	**P(gg)=0,16**	**P(gr)=0,24**	**P($\overline{A}$)=2/5 = 0,16+0,24**
Spalten-summe	**P(B) = 2/5 = 0,16+0,24**	**P($\overline{B}$)= 3/5 = 0,36 + 0,24**	**1**

In den Zeilen – mit A bezeichnet – stehen die Ausgänge des ersten Versuchs, in den Spalten – mit B bezeichnet – stehen die Ausgänge des zweiten Versuchs. Einige der eingetragenen Zahlen lassen sich unmittelbar im Baumdiagramm wiederfinden.
Die vier Werte in der Mitte der Tabelle sind genau die Werte, die im Baumdiagramm rechts die Wahrscheinlichkeiten der einzelnen Pfade als Produkt der Äste angeben. Die Zahlen in den Bereichen Spaltensumme und Zeilensumme geben die Wahrscheinlichkeiten an den Ästen wieder. Die Werte der Zeilensumme entsprechen den Werten des ersten Versuchs für die Ereignisse „rote Kugel“ und „gelbe Kugel“. Die Werte der Spaltensumme geben die Wahrscheinlichkeiten des zweiten Versuchs ebenfalls für „rote Kugel“ und „gelbe Kugel“ wieder.

32.4.2 Vierfeldertafel für Zufallsversuche „ohne Zurücklegen“

Die Darstellung des gleichen Versuchs aus dem Kapitel Baumdiagramm einer Vierfelderfafel hätte folgendes Aussehen:

	B	$\overline{B}$	Zeilensumme
A	**P(rg)=3/10**	**P(rr)=3/10**	**P(A)=3/5**
$\overline{A}$	**P(gg)=1/10**	**P(gr)=3/10**	**P($\overline{A}$)=2/5**
Spalten-summe	**P(B) = 2/5**	**P($\overline{B}$)= 3/5**	**1**

Die vier Werte in der Mitte der Tabelle entsprechen wieder den Wahrscheinlichkeiten am Ende der Äste im Baumdiagramm. Die Werte der Zeilensumme entsprechen auch den Werten an den Pfaden der ersten Versuchsdurchführung, für die die Ereignisse A und $\overline{A}$ stehen. Die Spaltensummen zu den Ereignissen B und $\overline{B}$ passen gar nicht zu den Werten an den Pfaden der zweiten Versuchsdurchführung, noch dazu, wo jeder Baum unterschiedliche Werte aufweist. Was zwangsläufig zu der Frage führt, was haben die Werte für eine Bedeutung, und wo sind sie im Baumdiagramm zu finden.

Dazu soll die Tabelle normiert werden. Das heißt, es wird die zweite Versuchsdurchführung gesondert betrachtet unter der Voraussetzung, dass das Ergebnis der ersten Versuchsdurchführung bekannt ist. Es sind dann für die zweite Versuchsdurchführung zwei Fälle zu unterscheiden. Die Normierung passiert so, dass man sagt, man geht davon aus, dass das Ereignis A oder $\overline{A}$ eingetreten ist. Das würde bedeuten, dass im ersten Fall P(A) = 1 ist und im zweiten Fall P($\overline{A}$) = 1 ist, da das jeweilige Ereignis mit Sicherheit eingetreten ist. In der Vierfeldertafel realisiert man das, indem man jede Zeile durch die Zeilen- oder Spaltensumme dividiert.

	B	$\overline{B}$	Zeilensumme
A	**P(rg)=1/2**	**P(rr)=1/2**	**P(A)=1**
$\overline{A}$	**P(gg)=1/4**	**P(gr)=3/4**	**P($\overline{A}$)=1**
Spalten-summe			

Spaltensummen machen jetzt keinen Sinn mehr, da zwei getrennte Versuche analysiert werden. Jetzt lassen sich die Zahlen in den vier Feldern im Baumdiagramm wiederfinden.
Die Werte der ersten Zeile sind die Wahrscheinlichkeiten beim ersten der beiden Bäume, der bei der 2. Durchführung entsteht. **Es sind die Wahrscheinlichkeiten, die auftreten, wenn beim ersten Versuch das Ereignis A eingetreten ist.** Das Eintreten des Ereignisses A reduziert die Anzahl der Kugeln insgesamt und reduziert die Anzahl der roten Kugeln. Damit ist für den zweiten Versuch eine völlig neue Situation gegenüber dem ersten entstanden.

Die Werte der zweiten Zeile sind die Wahrscheinlichkeiten beim zweiten der beiden Bäume, der bei der 2. Durchführung entsteht. **Es sind die Wahrscheinlichkeiten, die auftreten, wenn beim ersten Versuch das Ereignis $\overline{A}$ eingetreten ist.** Das Eintreten des Ereignisses reduziert ebenfalls die Anzahl der Kugeln insgesamt und reduziert die Anzahl der gelben Kugeln. Damit ist nicht nur eine Veränderung gegenüber dem ersten Versuch eingetreten, sondern auch eine Veränderung gegenüber dem ersten Baum des zweiten Versuchs. Die Ereignisse sind also völlig unterschiedlich zu bewerten.

Bei diesen Wahrscheinlichkeiten spricht man von „Bedingten Wahrscheinlichkeiten" die nur Auftreten „... unter der Bedingung, dass ein anderes Ereignis eingetreten ist .."
Bedingte Wahrscheinlichkeiten nehmen in der Wahrscheinlichkeitsrechnung eine großen Raum ein und sollen als nächstes eingehender betrachtet werden.

32.5 Bedingte Wahrscheinlichkeit

Mit dem Zählprinzip der Kombinatorik ist es in einfacher Weise möglich, die Wahrscheinlichkeiten mehrstufige Zufallsprozesse zu berechnen

32.5.1 Einführendes Beispiel

Eine Urne enthält 3 rote und 2 gelbe Kugeln. Es werden nacheinander 2 Kugeln ohne Zurücklegen gezogen. Jede Kugel soll die gleiche Wahrscheinlichkeit haben, gezogen zu werden, so dass ein Laplace-Experiment vorliegt. Wie groß ist die Wahrscheinlichkeit, beim ersten Zug eine rote und beim zweiten Zug eine grüne Kugel zu ziehen?

Nach dem Zählprinzip gibt es insgesamt $|\Omega| = 5 \cdot 4$ Möglichkeiten, zwei Kugeln ohne Zurücklegen aus der Urne mit 5 Kugeln zu ziehen (Variation ohne Wiederholung).

Weiter werden die beiden folgenden Ereignisse betrachtet:

A: „die erste Kugel ist gelb, die zweite Kugel ist beliebig“,
B: „die erste Kugel ist beliebig, die zweite Kugel ist rot“.

Dann ist $E = A \cap B$ das Ereignis

E: „die erste Kugel ist gelb und die zweite Kugel ist rot“.

- Für den **ersten Zug des Ereignisses *A*** sind zwei Möglichkeiten günstig: es kann eine der zwei gelben Kugeln gezogen werden.
- Für den **zweiten Zug** gibt es noch 4 Möglichkeiten, eine der verbliebenen 4 Kugeln zu ziehen, also ist.

$$|A| = 2 \cdot 4 = 8\,; \qquad P(A) = \frac{|A|}{|\Omega|} = \frac{8}{20} = \frac{2}{5}$$

- Beim **ersten Zug des Ereignisses *B*** kann zuerst eine von zwei gelben Kugeln.
- Für den **zweiten Zug** anschließend eine von 3 roten Kugeln gezogen werden;
- Wird im **ersten Zug** dagegen eine von den drei roten Kugeln gezogen,
- so gibt es im **zweiten Zug** noch 2 Möglichkeiten, eine der verbliebenen 2 roten Kugeln zu ziehen. Es ist daher

$$|B| = 2 \cdot 3 + 3 \cdot 2 = 12\,; \qquad P(B) = \frac{|B|}{|\Omega|} = \frac{12}{20} = \frac{3}{5}$$

- Für den **ersten Zug des Ereignisses *E*** sind zwei Möglichkeiten günstig und
- für den **zweiten Zug** sind 3 Möglichkeiten günstig, also:

$$|E| = 2 \cdot 3 = 6; \qquad P(E) = \frac{2 \cdot 3}{5 \cdot 4} = \frac{2}{20} \cdot \frac{3}{4}$$

Dabei ist $\frac{2}{5}$ die Wahrscheinlichkeit von *A*. Der zweite Faktor $\frac{3}{4}$ dagegen ist aber nicht die Wahrscheinlichkeit von *B*.

Er lässt sich aber ebenfalls als eine Wahrscheinlichkeit deuten, nämlich als $P_A(B)$

> die *Wahrscheinlichkeit des Ereignisses B unter der **Bedingung**, dass das Ereignis A eingetreten ist*. Man schreibt dafür und nennt dies die „*bedingte Wahrscheinlichkeit von B unter der Bedingung A*“.

Damit gilt also $\boxed{\mathbf{P(A \cap B) = P(A) \cdot P_A(B)}}$

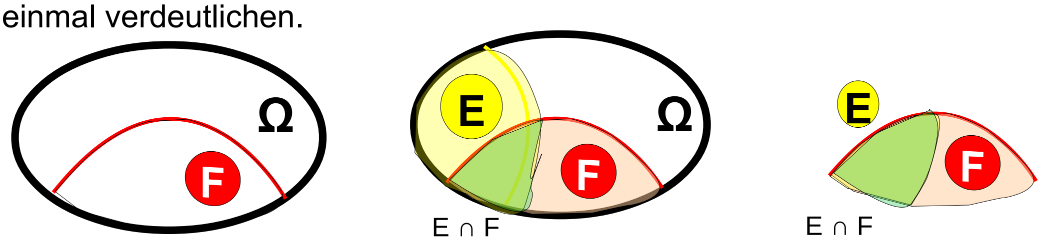
Diese Beziehung wird als „*Multiplikationssatz für das gleichzeitige Eintreten der Ereignisse A und B*“ bezeichnet.

Bedingte Wahrscheinlichkeiten treten als Pfadwahrscheinlichkeit ab der 2. Stufe auf. Den Ereignissen der 2.Stufe ist das Ereignis der 1. Stufe "vorangegangen" (zeitlich oder gedanklich). Die Wahrscheinlichkeiten $P_\Omega(B)=P(B)$ und $P_A(B)$ unterscheiden sich bei Laplace Experimenten eventuell (falls $A \subset \Omega$) durch eine gegenüber Ω eingeschränkte Ergebnismenge.

32.5.2 Was ist Bedingte Wahrscheinlichkeit

Ein vorgegebener Wahrscheinlichkeits-Raum (Ω,P) beschreibe zunächst ein stochastisches Experiment. Mit der Zusatzinformation, dass die Ergebnisse alle in einer nichtleeren Teilmenge $F \subset \Omega$ liegen, stellt sich die Frage nach der Modifikation des zunächst vorgegebenen Wahrscheinlichkeits-Raumes. Daher kann sinnvollerweise $P(F) > 0$ unterstellt werden, denn $P(F) = 0$ würde der hier betrachteten Situations-beschreibung widersprechen.

Schränkt man den Ausgangsraum Ω auf die Teilmenge F mit $P(F) > 0$ ein, so bietet es sich an, die durch (Ω,P) gelieferten Gewichte $P(\{\omega\})$ für $\omega \subset F$ zu übernehmen und das Gesamtgewicht durch Normierung, d.h. durch Division durch P(F), auf 1 zu bringen, wobei man von $P(. \mid F)$ als von der **(elementaren) bedingten Verteilung** spricht. Die Teilmenge F heißt Bedingung. Die folgende Skizze soll den Zusammenhang noch einmal verdeutlichen.



$$P(F) = \frac{|F|}{|\Omega|} \qquad P(E \cap F) = \frac{|F \cap E|}{|\Omega|} \qquad P_F(E) = \frac{|F \cap E|}{|F|}$$

Anteil der Fläche F in gesamt Ω

Anteil der Fläche E ∩ F in Ω

Anteil der Fläche E ∩ F in F

Für die bedingte Wahrscheinlichkeit gibt es deshalb folgende Definition:

Definition:
Von **bedingter Wahrscheinlichkeit** spricht man, wenn das Eintreten eines Ereignisses von dem Ausgang eines vorher durchgeführten Zufallsexperiment abhängt und durch dessen Ausgang der Ausgang des folgenden Ereignisses beeinflußt wird.
Die bedingte Wahrscheinlichkeit eines Ereignisses E unter der Voraussetzung, dass das Ereignis F eingetreten ist, berechnet sich durch:

$$P_F(E) = \frac{P(E \cap F)}{P(F)}$$

Nach dieser Definition ist es zunächst notwendig, sich klar zu machen, dass alle Wahrscheinlichkeiten in einem Baumdiagramm ab der zweiten Stufe bedingte Wahrscheinlichkeiten sind, da bis zu diesem Ereignis bereits andere Ereignisse eingetreten sind. Damit ist noch nicht festgelegt, dass der Ausgang der vorhergehenden Ereignisse einen Einfluß auf das folgende Ereignis hat. Ob ein solcher Einfluß besteht, oder nicht, hängt allein von den Wahrscheinlichkeiten an den Bäumen ab.
Wichtig in diesem Zusammenhang ist sind unvereinbarte Ereignisse, die sehr oft mit stochastisch unabhängigen Ereignissen verwechselt werden.

Definition:
Zwei Ereignisse E und F sind **unvereinbar**, wenn $E \cap F = \emptyset$ gilt

Für unvereinbare Ereignisse gilt die Formel

$$P(F \cup E) = P(F) + P(E)$$

während für zwei beliebige Ereignisse E und F gilt:

$$P(F \cup E) = P(F) + P(E) - P(F \cap E)$$

Gibt es Teile der Mengen E und F, die einander überdecken oder überlappen, so dürfen diese Teile für die Gesamtwahrscheinlichkeit nicht doppelt gezählt werden – das würde passieren, wenn man nur die Formel wie bei den unvereinbaren Ereignissen verwendet. Alle Ereignisse, die sowohl in E, als auch in F liegen, sind in P (F) + P(E) zweimal erfasst. Statt dessen ergibt sich die Gesamtmenge aus der Summe der beiden Teilmengen einmal reduziert um den überlappenden Bereich.
Versuche, bei denen die Unterscheidung z.B. „Junge“ oder „Mädchen“ eine Rolle spielt, sind unvereinbar, aber nicht unabhängig !
Merke: Unvereinbar ist eine Eigenschaft der Ereignisse, unabhängig ist eine Eigenschaft der Wahrscheinlichkeiten. Es gibt Ereignisse, die je nach Wahrscheinlichkeiten einmal abhängig und einmal unabhängig sein können.

32.5.2.1 Würfelbeispiel

Es werden zwei Würfel geworfen.

A sei das Ereignis: „Der erste Würfel zeigt eine Zahl < 5”.

B sei das Ereignis: „Die Augensumme ist > 6”.

Gesucht ist die Wahrscheinlichkeit für „A und B” .

Man kann die Wahrscheinlichkeit ermitteln, indem im Standardmodell für das Werfen mit zwei Würfeln (={(1,1),(1,2),...,(6,6)}) für alle Elementarereignisse geprüft wird, ob sie die Bedingung „A und B” erfüllen. Als Menge geschrieben, erhält man dann für das gesuchte Ereignis

E = { (1,6),(2,5),(2,6),(3,4),(3,5),(3,6),(4,3),(4,4),(4,5), (4,6) }

Da alle Elementarereignisse {(a,b)} gleichwahrscheinlich sind, kann man mit der Laplaceformel arbeiten.

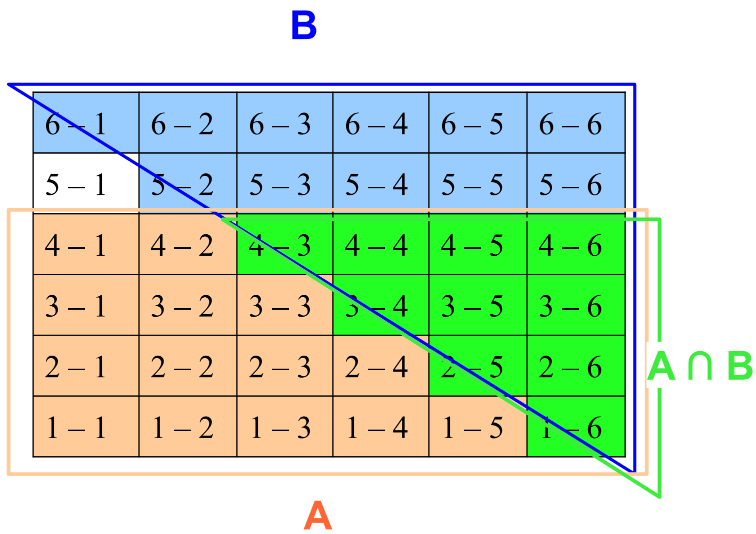
Das Ereignis „A und B” kann man als Durchschnitt zweier Mengen schreiben:

A = { (1,1),(1,2),...(1,6), (2,1),(2,2),...(2,6), .. ., (4,1),(4,2),...(4,6) },

B = { (1,6),(2,5),(2,6),(3,4),(3,5),(3,6),(4,3),(4,4),(4,5),(4,6),(5,2),...(6,6) }.

Gesucht ist P(A ∩ B) = P(E).

Es ist günstig, sich hier Wahrscheinlichkeiten als Größen von Teilflächen einer Gesamtfläche vorzustellen. Es wird also ein geometrisches Modell zur Veranschaulichung der Zusammenhänge gewählt. Dabei gelten folgende Beziehungen



Im geometrischen Modell bedeutet das:

- Die Wahrscheinlichkeit, die Elemente von A ∩ B aus der Gesamtmenge Ω zu ziehen beträgt **10/36** .
- Jetzt soll **das Ereignis A als erstes Ereignis** betrachtet werden. Nach dem ersten Schritt befindet man sich in der angegebenen Menge A. Die Wahrscheinlichkeit dafür ist 24/36. 24 Elemente von 36 gehören zur Menge A. Wenn man von dieser Teilmenge A als neuer ‘Menge’ ausgeht, ist die Wahrscheinlichkeit, den Durchschnitt A ∩ B = E zu erhalten, 10/24. Von 24 Elementen der Menge A gehören 10 Elemente auch zur Menge B. Für beide Schritte zusammen, also für P(A ∩ B), gilt damit: Es ist der Anteil (10/24) von ursprünglich (24/36) zu bestimmen. Das führt auf die übliche Multiplikation von Anteilen (10/24) • (24/36) = **10/36**.
- Jetzt soll **das Ereignis B als erstes Ereignis** betrachtet werden. Die Wahrscheinlichkeit für das Ereignis B beträgt 18/36 . Nach dem ersten Schritt befindet man sich in der angegebenen Menge B. Von den Elementen der Menge

B gehören 10 Elemente zu Menge $A \cap B$. Damit sind von 18/36 Ausgangswahrscheinlichkeit eine Wahrscheinlichkeit von 10/18 für Element aus $A \cap B$. Damit ergibt sich aus dem Produkt der beiden Wahrscheinlichkeiten (18/36) • (10/18) = **10/36.**

Also gilt $P(A \cap B) = P(A) \cdot$ P(„B unter der Bedingung, dass A eingetreten ist”).
Man schreibt kürzer $P(A \cap B) = P(A) \cdot \mathbf{P(B|A)} = P(A) \cdot \mathbf{P_A(B)}$. (zwei Schreibweisen für bedingte Wahrscheinlichkeit), aber es gilt auch $P(A \cap B) = P(A) \cdot$ P(„A unter der Bedingung, dass B eingetreten ist”). $P(A \cap B) = P(B) \cdot \mathbf{P(A|B)} = P(B) \cdot \mathbf{P_B(A)}$. Damit ist zunächst zu erkennen: Die Wahrscheinlichkeit $P(A \cap B)$ ist gleich, unabhängig, ob man zuerst das Ereignis A oder zuerst das Ereignis B betrachtet. Dies Erkenntnis wird später noch Bedeutung erlangen.

32.5.2.2 Ziehen von Kugeln mit zwei verschiedenen Eigenschaften

Beim mehrmaligen Würfeln hängt die Wahrscheinlichkeit eine bestimmte Zahl zwischen 1 und 6 zu werfen nicht von dem vorherigen Ergebnis ab. Jeder Wurf geschieht unabhängig von dem vorherigen. Werden hingegen aus einer Urne, die z.B. mehrere Kugeln mit zwei unterschiedlichen Farben enthält, nacheinander gezogen ohne sie wieder zurückzulegen, ist die Wahrscheinlichkeit des des zweiten Zuges vom ersten Ergebnis abhängig. In diesem Fall spricht man von einer **bedingten Wahrscheinlichkeit**. Bedingte Wahrscheinlichkeiten beim Ziehen mit Zurücklegen machen keinen Sinn, weil die zweite Ziehung nicht von der ersten abhängt.
Für die nächsten Abschnitte zur Erklärung der bedingten Wahrscheinlichkeit soll folgendes Beispiel betrachtet werden:

Eine Urne enthält 100 Kugeln.

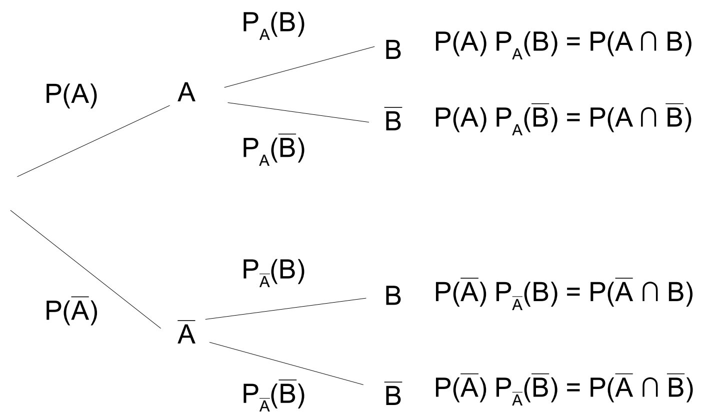
70 Kugeln bestehen aus dem Material Holz	und 30 Kugeln aus Kunststoff.
25 der Holzkugeln sind mit der Farbe rot gestrichen	und 45 sind grün.
10 der Kunststoffkugeln sind rot	und 20 sind grün.

Folgende Ereignisse werden definiert:

A: Die Kugel ist aus Holz	$\overline{A}$: Die Kugel ist aus Kunststoff
B: Die Kugel ist rot	$\overline{B}$: Die Kugel ist grün

32.5.3 Bedingte Wahrscheinlichkeiten im Baumdiagramm

Bedingte Wahrscheinlichkeiten sind eigentlich seit dem Baumdiagramm bekannt, auch, wenn sie dort nicht die Bezeichnung erhalten haben. Alle Wahrscheinlichkeiten, die an den zweiten Bäumen oder später auftreten sind bedingte Wahrscheinlichkeiten. Sie treten nur auf unter der Bedingung, daß das erste Ereignis bereits eingetreten ist.



Die Schreibweise $P_A(B)$ soll ausrücken:
Es ist die Wahrscheinlichkeit für das Eintreten des Ereignisses B, w**enn das Ereignis A eingetreten ist** und wird **bedingte Wahrscheinlichkeit** bezeichnet

Da die Indexschreibweise in gedruckte Dokumenten immer etwas problematisch ist, wird auch noch eine andere Schreibweise verwendet: $P(B\,|A)$
Dabei ist das Ereignis, was bereits eingetreten ist nach dem senkrechten Strich angegeben.

Aus den Pfadregeln des Baumdiagramms über Multiplikation und Addition erhält man folgende Beziehungen:

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \qquad P_A(\bar{B}) = \frac{P(A \cap \bar{B})}{P(A)} \qquad P_{\bar{A}}(B) = \frac{P(\bar{A} \cap B)}{P(\bar{A})} \qquad P_{\bar{A}}(\bar{B}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{A})}$$

Damit gibt es einen interessanten Zusammenhang mit der Wahrscheinlichkeit des Eintreten eines einzelnen Ereignisses wie A z.B. Die Wahrscheinlichkeit des Eintretens des Ereignisses A ist die Wahrscheinlichkeit von A unter der Bedingung des vollständigen Ereignisraumes Ω. So dass mit dem Formalismus der bedingten Wahrscheinlichkeit geschrieben werden kann:

$$P(A) = P_\Omega(A) = \frac{P(A)}{P(\Omega)}$$

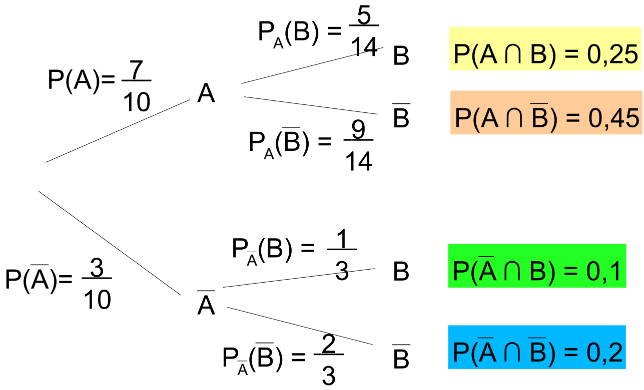
aber die Wahrscheinlichkeit $P(\Omega)$ ist natürlich 1, da ein Ereignis aus Ω mit Sicherheit eintrifft. Deshalb ist die Schreibweise $P(A)$ nur eine Kurzschreibweise für $P_\Omega(A)$.

Mit diesem Hintergrund kann man die bedingte Wahrscheinlichkeit $P_A(B)$ interpretieren als:

Wahrscheinlichkeit des Eintretens des Ereignisses B im Ereignisraum, der **nach dem Eintreten** des Ereignisses A entstanden ist.

Die Wahrscheinlichkeit wird also auf den Ereignisraum A reduziert und nicht über dem allgemeinen Ω Ereignisraum gebildet. Es ist die Wahrscheinlichkeit von B, **wenn man bereits weiß**, dass A eingetreten ist.

In dem vorher angegebenen Beispiel würde sich das Baumdiagramm folgendermaßen darstellen:



Die Wahrscheinlichkeit eine Kugel aus Holz zu ziehen beträgt 7/10 = P(A). Wenn man ein Kugel aus Holz gezogen hat ist die Wahrscheinlichkeit, daß die Kugel auch noch rot ist 25/70=5/14=$P_A(B)$, da von den 70 Holzkugeln 25 Kugeln die Farbe rot besitzen. In Worten ausgedrückt:

> Die Wahrscheinlichkeit, daß eine Kugel rot ist unter der Bedingung, daß sie aus Holz ist, beträgt 5/14.

Wichtig bei dieser Betrachtung ist, daß man den Unterschied zur Wahrscheinlichkeit des Durchschnitts erkennt. Die Wahrscheinlichkeit des Durchschnitts besagt:

> Wenn jetzt eine Kugel gezogen wird, dann wird mit einer Wahrscheinlichkeit von 7/10 * 5/14 = 0,25 eine rote Holzkugel gezogen.

Es ist noch nichts gemacht worden, es wird nur eine Aussage darüber getroffen, mit welcher Wahrscheinlichkeit eine bestimmtes Endergebnis eintreten kann. Es stehen noch alle 100 Kugeln zur Auswahl.
Die bedingte Wahrscheinlichkeit macht eine andere Aussage. Man zieht eine Kugel aus der Urne, hat sie in der Hand, sieht aber noch nicht, was das für eine Kugel ist. Durch befühlen der Kugel stellt man aber fest, man hat eine Holzkugel in der Hand, weil die Oberfläche vielleicht etwas rauer ist als bei einer Kunststoffkugel. Jetzt sind alle Kunststoffkugel ausgeschlossen. Das bedeutet, **ein Ereignis von den zweien ist schon eingetreten**, es ist nicht mehr unklar, was meine Kugel für ein Material haben wird. Damit hat man nicht mehr die Auswahl aus 100 Kugeln, sondern nur noch aus 70. Damit reduziert sich die Möglichkeit eine rote zu ziehen auch nur noch auf die 70. Geht man davon aus, es gebe gar keine roten Holzkugeln kann man noch so viele schöne rote Kunststoffkugeln haben, wenn man eine Holzkugel gezogen hat ist die Wahrscheinlichkeit für eine rote Kugel Null. Der Wahrscheinlichkeitsraum ist eingeschränkt auf das Ereignis A.

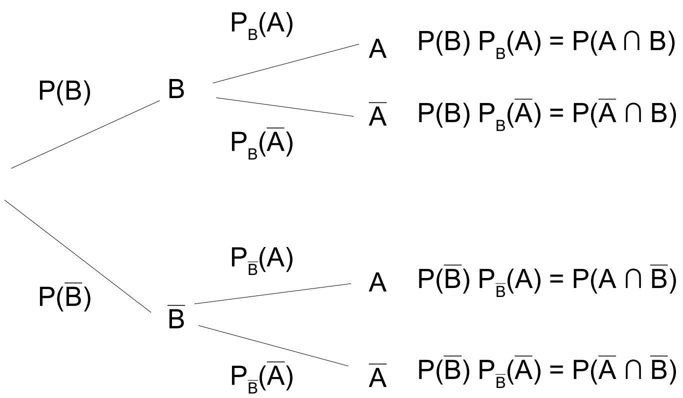
In diesem Baumdiagramm findet man aber nur die bedingten Wahrscheinlichkeiten zu A oder $\overline{A}$. Wie kann man im Baumdiagramm die Wahrscheinlichkeiten unter der Bedingung B oder $\overline{B}$ erstellen. Hier spielen jetzt die Wahrscheinlichkeiten des Durchschnitts eine Rolle. Die Wahrscheinlichkeit des Durchschnitts $P(A \cap B)$ muss die gleiche sein, unabhängig, ob erst das Ereignis A oder erste das Ereignis B eingetreten ist. Damit wird der **1. Pfadregel** oder des **Multiplikationssatzes** Rechnung getragen.

$$P(A) \bullet P_A(B) = P(A \mid B) = P_B(A) \bullet P(B)$$

Die Wahrscheinlichkeit des Durchschnitts zweier Ereignisse ist nicht generell das Produkt der Wahrscheinlichkeiten der Einzelereignisse, sondern das Produkt aus der Wahrscheinlichkeit des ersten Ereignisses multipliziert mit der Wahrscheinlichkeit für das zweite Ereignis unter der Bedingung, dass das erste Ereignis bereits eingetreten ist. Wenn das zweite Ereignis eintritt weiß man bereits, welches erste Ereignis eingetreten ist

32.5.4 Das umgekehrte Baumdiagramm

Damit stellt sich sofort die Frage, was passiert denn, wenn das Ereignis B als erstes eintritt und danach das Ereignis A. Am Ende der beiden Versuche ist wieder das Ereignis A und das Ereignis B eingetreten, also die gleiche Endsituation, wie im ersten Fall. Das zugehörige Baumdiagramm hätte dann folgendes Aussehen:



und für die bedingten Wahrscheinlichkeiten des umgekehrten Baumes ergeben sich folgende Gleichungen:

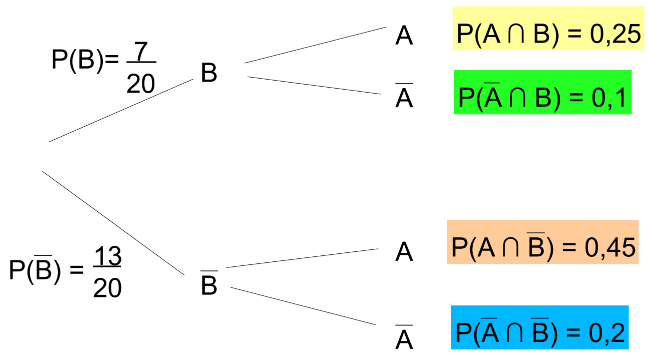
Ein Baumdiagramm kann man aber auch so aufbauen, dass das Ereignis B als erstes eintritt. Diese Wahrscheinlichkeiten kann man aus dem Baumdiagramm ermitteln. Betrachtet man die Durchschnittswahrscheinlichkeiten aller Ereignisse, an denen B beteiligt ist erhält man

$$P(A \cap B) + P(\overline{A} \cap B) = P(B)$$

Berücksichtigt man alle Wahrscheinlichkeiten, in denen das Ereignis B mit einem anderen Ereignis zusammen auftreten kann, dann ergibt diese Summe die Wahrscheinlichkeit für das Ereignis B selbst. Diesen Zusammenhang. der ganz allgemein gilt, bezeichnet man als **Satz von der totalen Wahrscheinlichkeit**, ist aber nichts anderes als die **2. Pfadregel**. Damit lässt sich ein sogenanntes umgekehrtes Baumdiagramm erstellen, bei dem das Ereignis B als erstes erscheint.

$$P(B) = P(A \cap B) + P(\overline{A} \cap B) = 0.35 \text{ und } P(\overline{B}) = P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B}) = 0.65$$

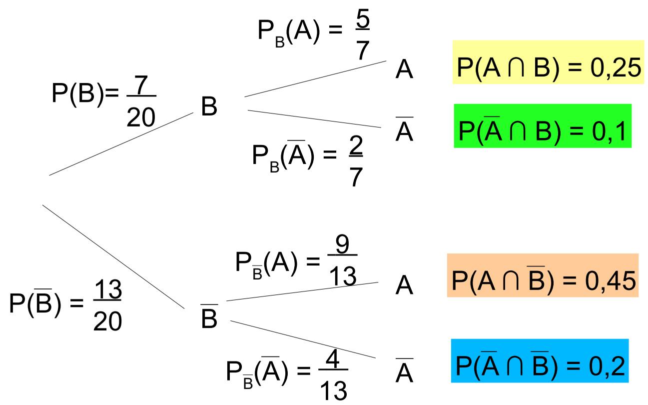
Das läßt sich im Beispiel auch unmittelbar nachprüfen. Es gibt von den 100 Kugeln 35 die rot sind und 65 die grün sind.



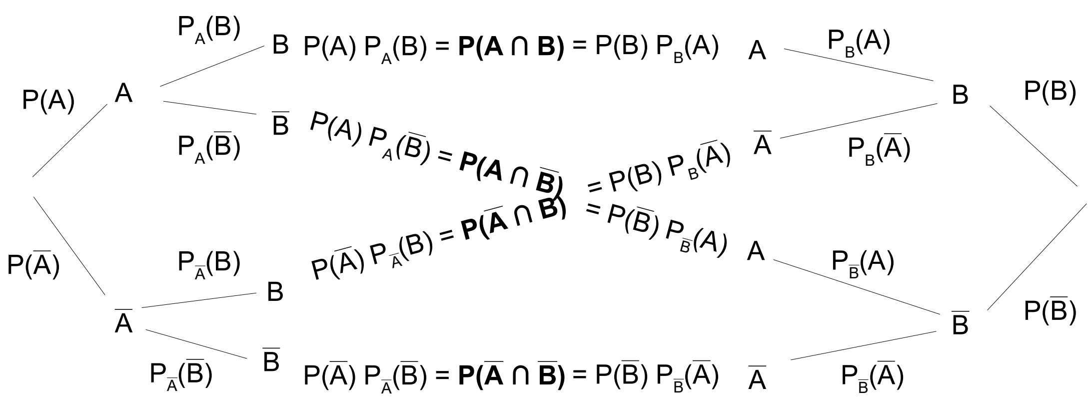
In dem so entstandenen rudimentären Baumprogramm fehlen zunächst wieder die Wahrscheinlichkeiten an den zweiten Bäumen. Andererseits müssen die Wahrscheinlichkeiten vom Durchschnitt zweier Ereignisse in beiden Fällen gleich sein, da die Wahrscheinlichkeit für das eintreten zweier Ereignisse nicht davon abhängt, welches Ereignis als erstes eingetreten ist Deshalb sind die farbig gestalteten Wahrscheinlichkeiten gleich, wobei zu beachten ist. dass die inneren beiden gegenüber dem ersten Baumdiagramm vertauscht wurden. P($\overline{A}$ I B) (grün) befindet sich jetzt am ersten Teilbaum der zweiten Ebene und P(A I B) (braun) am zweiten Teilbaum der zweiten Ebene. Nach den gleichen Rechenschritten wie im ersten Fall lassen sich die bedingten Wahrscheinlichkeiten für die zweiten Bäume berechnen.

$P_B(A) = \frac{P(A \cap B)}{P(B)}$ $\quad P_B(\overline{A}) = \frac{P(\overline{A} \cap B)}{P(B)}$ $\quad P_{\overline{B}}(A) = \frac{P(A \cap \overline{B})}{P(\overline{B})}$ $\quad P_{\overline{B}}(\overline{A}) = \frac{P(\overline{A} \cap \overline{B})}{P(\overline{B})}$

Damit hat das vollständige umgekehrte Baumdiagramm folgendes Aussehen



Jetzt lassen sich beide Bäume spiegelbildlich aneinander setzen, da in beiden Bäumen das Ende eines Zweiges durch das gleichzeitige Eintreten von zwei Ereignissen gekennzeichnet ist.



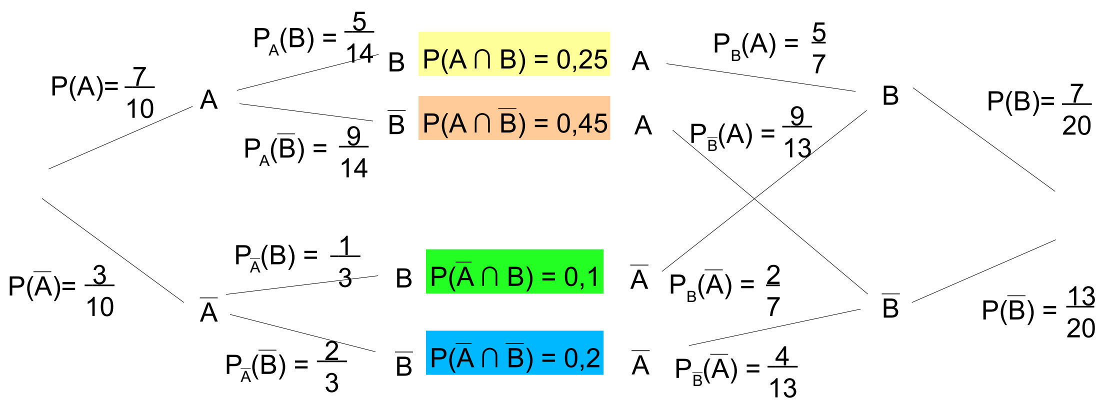
In beiden Bäumen sind die in der Mitte stehenden möglichen Ereignisse immer die gleichen, die aber auf zwei verschiedenen Wegen erreicht werden:

$P(A \cap B)$ =	$P(B)\,P_B(A) = P(A)\,P_A(B)$
$P(A \cap \overline{B})$ =	$P(\overline{B})\,P_{\overline{B}}(A) = P(A)\,P_A(\overline{B})$
$P(\overline{A} \cap B)$ =	$P(B)\,P_B(\overline{A}) = P(\overline{A})\,P_{\overline{A}}(B)$
$P(\overline{A} \cap \overline{B})$ =	$P(\overline{B})\,P_{\overline{B}}(\overline{A}) = P(\overline{A})\,P_{\overline{A}}(\overline{B})$

Diese Formeln ermöglichen das Berechnen der jeweils entgegengesetzten bedingten Wahrscheinlichkeit, wenn man eine bedingte Wahrscheinlichkeit kennt. Wenn man z.B aus $P_A(B)$ die Wahrscheinlichkeit $P_B(A)$ berechnen muss. Voraussetzung ist natürlich, dass man die beiden Einzelwahrscheinlichkeiten P(A) und P(B) kennt.
Für die Einzelwahrscheinlichkeiten erhält man:

$P(A)$ =	$P(A \cap B) + P(A \cap \overline{B})$
$P(\overline{A})$ =	$P(\overline{A} \cap B) + P(\overline{A} \cap \overline{B})$
$P(B)$ =	$P(A \cap B) + P(\overline{A} \cap B)$
$P(\overline{B})$ =	$P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B})$

Im angegebenen Beispiel hat das doppelte Baumdiagramm folgendes Aussehen:



32.5.5 Bedingte Wahrscheinlichkeit in der Vier-Felder-Tafel

Bei der Umsetzung des Baumdiagramms in eine Vier-Felder-Tafel wurde bereits geklärt, dass die Wahrscheinlichkeiten für das Eintreten von zwei nacheinander erfolgten Ereignissen in den mittleren vier Feldern auftauchen. Die Summen in den Zeilen und Spalten geben die Wahrscheinlichkeit der Einzelereignisse wieder.

	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	1

Aus der Anordnung der Vierfeldertafel kann man erkennen, dass die bedingten Wahrscheinlichkeiten aus der Vierfeldertafel nicht unmittelbar ablesbar sind, aber sie lassen sich aus ihr berechnen. Dazu benutzt man die sogenannten Randwahrscheinlichkeiten. Die Randwahrscheinlichkeiten sind die Zeilen- und Spaltensummen der Vierfeldertafel. Dividiert man die Wahrscheinlichkeiten in der Mitte der Tafel mit den Randwahrscheinlichkeiten der Zeilen- oder Spaltensummen, erhält man die jeweilige bedingte Wahrscheinlichkeit:

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

Damit erhält man die jeweiligen bedingten Wahrscheinlichkeiten, indem man die Zeilen oder Spalten auf die Zeilensumme normiert, soll heißen an Stelle von $P(A)$ oder $P(\overline{A})$ als Zeilensumme muss 1 stehen.

Vierfeldertafel

	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	1

Bedingte Wahrscheinlichkeit, daß A oder $\overline{A}$ eingetreten ist.

	B	$\overline{B}$	
A	$P_A(B)$	$P_A(\overline{B})$	1
$\overline{A}$	$P_{\overline{A}}(B)$	$P_{\overline{A}}(\overline{B})$	1
	╬	╬	╬

	B	$\overline{B}$	
A	$P_B(A)$	$P_{\overline{B}}(A)$	╬
$\overline{A}$	$P_B(\overline{A})$	$P_{\overline{B}}(\overline{A})$	╬
	1	1	╬

Bedingte Wahrscheinlichkeit, daß B oder $\overline{B}$ eingetreten ist.

Für manche Aufgaben hat es sich gezeigt, daß es nützlich ist für die Rechnung mit bedingten Wahrscheinlichkeiten mit einer solchen dreifachen Vier-Felder-Tafel zu starten. Die zweite Vier-Felder-Tafel in der Zeile dahinter enthält die bedingten Wahrscheinlichkeit bezogen auf die Ereignisse in der Zeile. Die dritte Vier-Felder-Tafel darunter enthält die bedingten Wahrscheinlichkeiten bezogen auf die Ereignisse in den Spalten. Nicht benötigte Felder bleiben leer. Der Vorteil, wenn man mit einer solchen Darstellung beginnt ist folgender: In der Aufgabenstellung sind mitunter schon bedingte Wahrscheinlichkeiten vorgegeben und man soll dann die jeweils umgekehrten bedingten Wahrscheinlichkeiten bestimmen. In einer einfachen Vier-Felder-Tafel hat man nicht die Möglichkeit die vorgegebenen Werte der bedingten Wahrscheinlichkeit einzutragen. In dieser dreifachen Tafel ist für jede mögliche angegebene Wahrscheinlichkeit Platz. Diese dreifache Vier-Felder-Tafel besitzt alle Einträge, die auch ein doppeltes Baumdiagramm besitzt. Durch Vorwärts- und Rückwärtsrechnen lassen sich dann alle 16 Felder bestimmen.
Umgesetzt auf das Beispiel liefert das folgende Vier-Felder-Tafeln:

Division durch $P(A)$ →
Division durch $P(\overline{A})$

(1)

	B	$\overline{B}$	Summe
A	$P(A \cap B) = 0{,}25$	$P(A \cap B) = 0{,}45$	$P(A) = 0{,}7$
$\overline{A}$	$P(\overline{A} \cap B) = 0{,}10$	$P(\overline{A} \cap \overline{B}) = 0{,}20$	$P(\overline{A}) = 0{,}3$
Summe	$P(B) = 0{,}35$	$P(\overline{B}) = 0{,}65$	1

(2)

	B	$\overline{B}$	Summe
A	$\frac{P(A \cap B)}{P(A)} = \frac{5}{14}$	$\frac{P(A \cap \overline{B})}{P(A)} = \frac{9}{14}$	1
$\overline{A}$	$\frac{P(\overline{A} \cap B)}{P(\overline{A})} = \frac{1}{3}$	$\frac{P(\overline{A} \cap \overline{B})}{P(\overline{A})} = \frac{2}{3}$	1
Summe			

Division durch $P(B)$ ↓
Division durch $P(\overline{B})$

	B	$\overline{B}$	Summe
A	$\frac{P(A \cap B)}{P(B)} = \frac{5}{7}$	$\frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{9}{13}$	
$\overline{A}$	$\frac{P(\overline{A} \cap B)}{P(B)} = \frac{2}{7}$	$\frac{P(\overline{A} \cap \overline{B})}{P(\overline{B})} = \frac{4}{13}$	
Summe	1	1	

(3)

Die drei 4- Felder-Tafeln werden so angeordnet, dass
1) oben links die Tafel der Wahrscheinlichkeiten des Durchschnitts.
2) oben rechts die Tafel für die bedingten Wahrscheinlichkeiten der Zeilensumme steht
3) unten links die Tafel für die bedingten Wahrscheinlichkeiten der Spaltensumme

In dieser Darstellung ist es relativ leicht, die einzelnen bedingten Wahrscheinlichkeiten zu berechnen. Der Wert der Zeilen oder Spaltensumme steht in der jeweiligen Zeile oder Spalte und man sehr schnell dividieren. Die jetzt in der Vier - Felder-Tafel entstehenden Wahrscheinlichkeiten sind die sogenannten bedingten Wahrscheinlichkeiten, das sind die Wahrscheinlichkeiten, die gelten, wenn man bereits weiß, dass das Ereignis A , $\overline{A}$, B oder $\overline{B}$ eingetreten ist.
Wenn man also weiß, dass die gezogene Kugel aus Kunststoff besteht, dann ist die Wahrscheinlichkeit dafür, dass sie die Farbe grün hat: 2/3. Die Wahrscheinlichkeit eine grüne Kunststoffkugel zu ziehen ist hingegen 0,2.

Es besteht aber die Möglichkeit an Stelle der Zeilensummen auch auf die Spalten-summen zu normieren. Diese Wahrscheinlichkeiten, die man in der Vier Felder Tafel erst durch Erstellen einer neue Tafel finden kann, befinden sich die Wahrscheinlich-keiten am umgekehrten Baumdiagramm. Von den Vier- Felder-Tafeln ist es die untere Tafel.

32.5.6 Notwendige Unterscheidungen

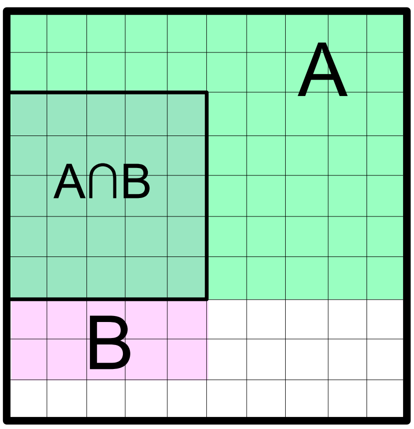
1. Die Voraussetzung P (B) > 0 ist nötig, damit die Definition überhaupt erst möglich ist.
2. P (A|B) ist die Wahrscheinlichkeit des Ereignisses A eingeschränkt auf die Menge B (als neue Ergebnismenge). Damit kommen natürlich nur die Elemente von A in Frage, die auch in B liegen, also solche, die in A ∩ B liegen. Die Schreibweise $P_B(A)$ in der Definition weist darauf hin, dass P_B als spezielles Wahrscheinlich-keitsmaß verstanden werden kann – eben dasjenige, das auf dem neuen Grundraum (der neuen Ergebnismenge B) definiert werden kann.
3. Vorsicht: Sehr oft werden die Wahrscheinlichkeiten P (A|B), P (B|A) oder P (A ∩ B) miteinander verwechselt, weil die umgangssprachlichen Beschreibungen dieser Wahrscheinlichkeiten sehr ähnlich sind. Hilfreich ist die genaue Überlegung, welche Menge man sich als neue Ergebnismenge denken kann – A oder B – oder ob diese überhaupt vonnöten ist.

Bsp.: M:”Die Person ist männlich.“, S:”Die Person ist mindestens 70 Jahre alt.“

– Wahrscheinlichkeit dafür, dass ein Mann mindestens 70 Jahre alt ist: P (S|M).
– Wahrscheinlichkeit dafür, dass jemand ein mindestens 70jähriger Mann ist: P (S ∩ M).
– Wahrscheinlichkeit dafür, dass jemand, der mindestens 70 Jahre alt ist, ein Mann ist: P (M|S).

32.5.7 Graphische Veranschaulichung bedingter Wahrscheinlichkeiten

Die Wahrscheinlichkeit von A ist P(A) = 70/100.
Die Wahrscheinlichkeit von B ist P(B) = 35/100.
Die Wahrscheinlichkeit von A∩B ist P(A∩B) = 25/100.



- Das Ereignis A nimmt innerhalb von Ω 70 Teile ein. Damit ist die Wahrscheinlichkeit von A : 70/100 = 7/10
- Das Ereignis B nimmt innerhalb des Ereignisses A 25 Teile ein, das Ereignis A selbst aber 70 Teile. Damit ist die Wahrscheinlichkeit von B innerhalb von A 25/70.
- Das Ereignis B nimmt innerhalb des Ereignisses $\overline{A}$ 10 Teile ein, das Ereignis $\overline{A}$ selbst aber 30 Teile. Damit ist die Wahrscheinlichkeit von B innerhalb von $\overline{A}$ 10/30.

- Betrachtet man das Ereignis des Durchschnitts nur bezogen auf A ergibt sich eine Wahrscheinlichkeit von 27/60 , dh: Wenn man das Ereignis A zugrunde legt, tritt das Ereignis B mit der Wahrscheinlichkeit 27/60 = 9/20 ein.
- Betrachtet man das Ereignis des Durchschnitts nur bezogen auf $\overline{A}$ ergibt sich eine Wahrscheinlichkeit von 10/30 , dh: Wenn man das Ereignis $\overline{A}$ zugrunde legt, tritt das Ereignis B mit der Wahrscheinlichkeit 10/30 = 1/3 ein.

Die Ereignisse A und B sind abhängig. Für das Eintreten des Ereignisses B ist es entscheidend, ob erst A oder erst $\overline{A}$ eingetreten ist.

32.5.8 Bedingte Wahrscheinlichkeit in der 3. Stufe

Für spätere Untersuchungen soll hier noch die die bedingte Wahrscheinlichkeit in der 3. Stufe hergeleitet werden. Danach ist eine weitere Verallgemeinerung für weitere Stufen leicht nachvollziehbar. Es soll das Ereignis A_3 betrachtet werden, das in der 3. Stufe auftritt, unter der Bedingung, dass es auch in den beiden vorhergehenden Stufen die Ereignisse A_2 und A_1 aufgetreten sind: $P(A_3 \mid A_1 \cap A_2)$. Die Formel der bedingten Wahrscheinlichkeit besagt, dass im Zähler der Durchschnitt der beiden Ereignisse steht und im Nenner die Wahrscheinlichkeit für die vorher eingetretenen Ereignisse:

$$P(A_3 \mid A_1 \cap A_2) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}$$

Das ist die Wahrscheinlichkeit, die am 3. Baum stehen würde, an dem Weg, der von A_1 beginnend über A_2 nach A_3 führt. Betrachtet man jetzt den kompletten Pfad von A_1 bis A_3 ergibt sich folgende Formel:

$$P(A_1)\, P(A_2|A_1)\, P(A_3 \mid A_1 \cap A_2) = P(A_1)\, \frac{P(A_1 \cap A_2)}{P(A_1)}\, \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} = P(A_1 \cap A_2 \cap A_3)$$

Zuerst tritt das Ereignis A_1 ein, danach das Ereignis A_2 unter der Bedingung, dass A_1 eingetreten ist und dann A_3 unter der Bedingung, dass A_1 und A_2 eingetreten sind. Das ist der Beweis für die erste Pfadregel, dass die Wahrscheinlichkeit des Durchschnitts von mehreren Ereignissen entlang eines Pfades (Wahrscheinlichkeit, dass genau diese Ereignisse gemeinsam aufgetreten sind) sich aus dem Produkt der Wahrscheinlichkeiten der Einzelpfade, die zu dem Gesamtereignis führen, bestimmt. Dabei spielt es keine Rolle, ob die Ereignisse abhängig oder unabhängig sind, da hier nur die bedingten Wahrscheinlichkeiten benutzt wurden. Ob die sich in den einzelnen Stufen ändern oder gleich bleiben ist uninteressant.

Zu diesem Thema gibt es ein eigenes spezielles Dokument mit entsprechenden Beispielen. Das Thema ist üblicherweise kein Schulstoff, allerdings wird in Abituraufgaben manchmal mit drei Ereignissen gearbeitet, bei denen bedingte Wahrscheinlichkeiten berechnet werden sollen.

32.6 Stochastisch unabhängige Ereignisse

Aus dem bisherigen stellt sich die Frage, wann ist denn die Wahrscheinlichkeit eines zweiten Versuchs unabhängig vom Ergebnis des ersten Versuchs, oder, wann ist eine Wahrscheinlichkeit eines zweiten Ereignisses B unabhängig vom Ausgang eines ersten Ereignisses A. In Formeln ausgedrückt:

$$P_A(B) = P(B)$$

Für zweistufiges Ziehen aus einer Urne kann man die Antwort noch relativ leicht geben: Beim Ziehen mit Zurücklegen sind die Wahrscheinlichkeit des zweiten Zuges identisch mit den Wahrscheinlichkeiten des ersten Ziehens, damit sind die Ereignisse unabhängig.

Bei anderen Ereignissen ist die Antwort nicht so einfach zu geben.

32.6.1 Formel für stochastische Unabhängigkeit

Als Erstes soll dargestellt werden, wie man zeigen kann, ob zwei Ereignisse stochastisch unabhängig sind. Dazu ist zunächst eine wichtige Bemerkung zu machen: Stochastische Unabhängigkeit hat nichts mit **Unvereinbarkeit** zu tun. Ereignisse wie z.B ob nach männlich oder weiblich unterschieden wird, oder nach Raucher und Nichtraucher sind unvereinbar, aber noch lange nicht stochastisch unabhängig. Man kann nur sagen: Stochastisch unabhängige Ereignisse sind auch unvereinbar, aber die Umkehrung gilt nicht. Oder: Ereignisse, die nicht unvereinbar sind können auch nicht stochastisch unabhängig sein.

Mit der eingangs erwähnten Formel kann man die stochastische Unabhängigkeit definieren. Aus der bedingten Wahrscheinlichkeit ist folgende Beziehung bekannt: $P(A \cap B) = P(A)\, P_A(B)$. Jetzt geht es darum, dass das Ereignis B nicht mehr vom Ereignis A abhängt, was eingangs schon einmal mit der Formel $P_A(B) = P(B)$ ausgedrückt wurde. Damit ergibt sich als Definition für die stochastische Unabhängigkeit:

Definition: Sei (Ω, P) ein Wahrscheinlichkeitsraum eines Zufalls-experiments und $A, B \subseteq \Omega$ und sei $P(A) > 0$. Das Ereignis B heißt **(stochastisch) unabhängig** vom Ereignis A, wenn gilt:

$$\mathbf{P_A(B) = P(B)}$$

In Worten: wenn die bedingte Wahrscheinlichkeit von B unter der Bedingung A gleich der Wahrscheinlichkeit von B ist.

Soll also nachgewiesen werden, ob zwei Ereignisse stochastisch unabhängig sind, dann ist für diese Ereignisse genau diese Gleichung nachzuweisen. Es sind die Einzelwahrscheinlichkeiten der beiden Ereignisse zu bestimmen und die Wahrscheinlichkeit, das beide Ereignisse eingetreten sind.

Satz: Sei (Ω, P) ein Wahrscheinlichkeitsraum eines Zufallsexperiments und gelte für $A, B \subseteq \Omega$: $P(A) > 0$ und $P(B) > 0$.
Dann gilt: **A, B (stochastisch) unabhängig** $\Leftrightarrow$

$$\mathbf{P(A \cap B) = P(A) \cdot P(B).}$$

32.6.2 Beziehungen zwischen unabhängigen Ereignissen

Sind zwei Ereignisse A und B stochastisch unabhängig, dann gelten folgende Aussagen:

- Ist A stochastisch unabhängig von B, dann ist auch B stochastisch unabhängig von A
- $\overline{A}$ und B sind stochastisch unabhängig
- A und $\overline{B}$ sind stochastisch unabhängig
- $\overline{A}$ und $\overline{B}$ sind stochastisch unabhängig
- $P_A(B) = P_{\overline{A}}(B)$

Für zwei Ereignisse ist dieser Nachweis noch relativ einfach zu erbringen, da es sich um eine Gleichung handelt. Stellt man sich die frage, wann sind den drei Ereignisse A, B und C stochastisch unabhängig, dann muss man schon die Gültigkeit aller folgenden Gleichungen nachweisen:

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$
$$P(\overline{A} \cap B \cap C) = P(\overline{A}) \cdot P(B) \cdot P(C)$$
$$P(A \cap \overline{B} \cap C) = P(A) \cdot P(\overline{B}) \cdot P(C)$$
$$P(A \cap B \cap \overline{C}) = P(A) \cdot P(B) \cdot P(\overline{C})$$
$$P(\overline{A} \cap \overline{B} \cap C) = P(\overline{A}) \cdot P(\overline{B}) \cdot P(C)$$
$$P(\overline{A} \cap B \cap \overline{C}) = P(\overline{A}) \cdot P(B) \cdot P(\overline{C})$$
$$P(A \cap \overline{B} \cap \overline{C}) = P(A) \cdot P(\overline{B}) \cdot P(\overline{C})$$
$$P(\overline{A} \cap \overline{B} \cap \overline{C}) = P(\overline{A}) \cdot P(\overline{B}) \cdot P(\overline{C})$$

Allgemein gilt, dass man für n Ereignisse 2^n Gleichungen überprüfen muss.

32.6.3 Unabhängige und unvereinbare Ereignisse

Die Unabhängigkeit zweier Ereignisse ist nicht zu verwechseln mit der Disjunktheit oder Unvereinbarkeit von Ereignissen !

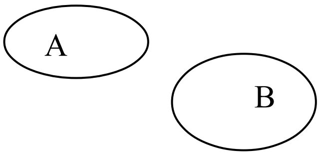
- Für unabhängige Ereignisse gilt eine Produktformel für den Durchschnitt der Ereignisse von A und B: $P(A \cap B) = P(A) \cdot P(B)$
- Für unvereinbare Ereignisse gilt eine Summenregel für die Vereinigung von A und B: $P(A \cup B) = P(A) + P(B)$, falls $A \cap B = \emptyset$.

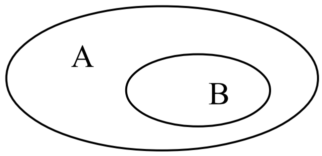
Es gilt hierbei sogar der folgende

Satz: Sei (Ω, P) ein Wahrscheinlichkeitsraum eines Zufallsexperiments und gelte für $A, B \subseteq \Omega$: $P(A) > 0$ und $P(B) > 0$. Dann gilt:
a) A, B unvereinbar $\Rightarrow$ A, B abhängig.
b) A, B unabhängig $\Rightarrow$ A, B vereinbar.

Bemerkungen:

1. Wie dieser Satz zeigt, darf die stochastische Unabhängigkeit nicht als völlige Beziehungslosigkeit der beiden Ereignisse gedeutet werden. Es handelt sich lediglich um eine Unabhängigkeit im statistischen Sinn. Auch wenn man von Unabhängigkeit der Ereignisse spricht, so ist die Unabhängigkeit genau genommen keine Eigenschaft der Ereignisse, sondern eher eine Eigenschaft ihrer Wahrscheinlichkeiten.
2. Viele Experimente sind so beschaffen, dass man aufgrund der Anordnung des (meist mehrstufigen) Experiments die Unabhängigkeit zweier Ereignisse postuliert. Dann dienen die bekannten Einzelwahrscheinlichkeiten P (A) und P (B) zur Ermittlung der Durchschnittswahrscheinlichkeit P $(A \cap B)$ mit Hilfe der Produktregel.





A und B sind unvereinbar
A und B sind abhängig
$A \cap B = \emptyset$.

Für unvereinbare Ereignisse gilt eine Summenregel für die Vereinigung von A und B:
$P(A \cup B) = P(A) + P(B)$

A und B sind vereinbar
A und B sind unabhängig, wenn $P(A \cap B) = P(A) \cdot P(B)$

Für unabhängige Ereignisse gilt eine *Produktformel für den Durchschnitt* der Ereignisse von A und B:
$P(A \cap B) = P(A) \cdot P(B)$

$B \subset A$
A und B sind abhängig

$P(A \cap B) = P(B)$
$P(A \cup B) = P(A)$

Zur Veranschaulichung unabhängiger Ereignisse wird die gleiche Aufgabenstellung wie in Kapitel 31.5.2.2 verwendet, das in Kapitel 31.6. benutzt wurde. Allerdings ist die Verteilung der Kugeln bezüglich Material und Farbe anders.

Eine Urne enthält 100 Kugeln.
70 Kugeln bestehen aus dem Material Holz und 30 Kugeln aus Kunststoff.
28 der Holzkugeln sind mit der Farbe rot gestrichen und 42 sind grün.
12 der Kunststoffkugeln sind rot und 18 sind grün.

Folgende Ereignisse werden definiert:

A: Die Kugel ist aus Holz — $\overline{A}$: Die Kugel ist aus Kunststoff
B: Die Kugel ist rot — $\overline{B}$: Die Kugel ist grün

$P(A) = 0{,}7$; $P(\overline{A}) = 0{,}3$; $P(B) = 0{,}4$; $P(\overline{B}) = 0{,}6$

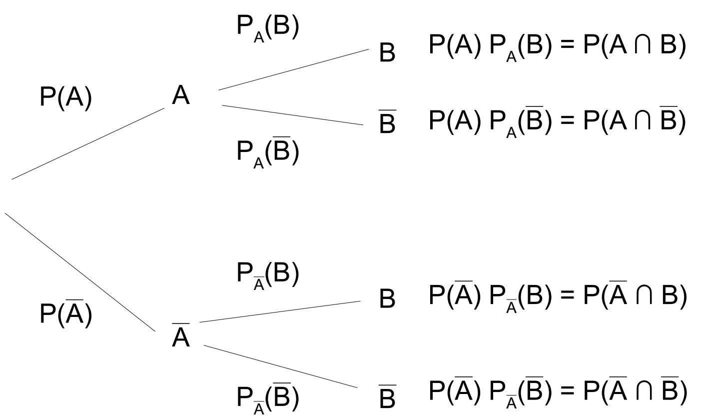
$P(A \cap B) = 0{,}28$; $P(\overline{A} \cap B) = 0{,}42$; $P(A \cap \overline{B}) = 0{,}12$; $P(\overline{A} \cap \overline{B}) = 0{,}18$;

32.6.4 Unabhängige Ereignisse im Baumdiagramm

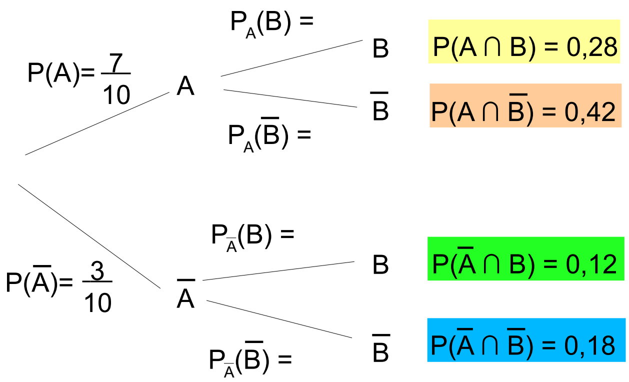
> Im Baumdiagramm stellen sich unabhängige Ereignisse so dar, dass jeder zu dem gleichen Ereignis zeigende Ast die gleiche Wahrscheinlichkeit trägt.

Besonders anschaulich kann man sich das klar machen beim „Ziehen mit Zurücklegen“. Da hat jeder Versuch die gleiche Wahrscheinlichkeit, da es keine Rückschlüsse auf die bereits erfolgten Versuche gibt. Solche Versuche sind stochastisch unabhängig. Im folgenden soll die mathematische Begründung dafür gegeben werden.

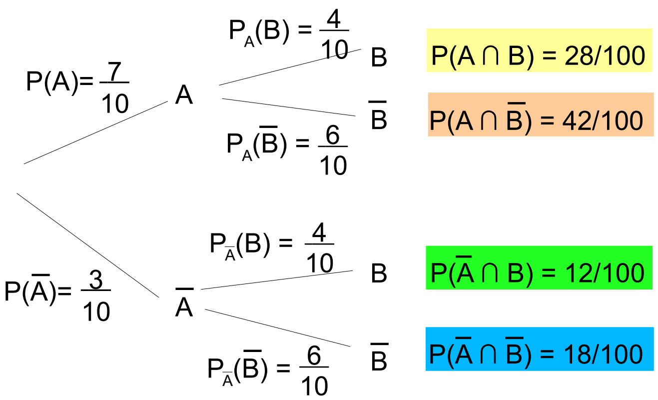
Es sollen zwei allgemeine Ereignisse A und B betrachtet werden.
Das erste Ereignis A hat natürlich auf die stochastische Unabhängigkeit keinen Einfluss, da es keinen Einfluss auf die erste Ziehung gibt.
Für den Eintritt des Ereignisses B sind die beiden Pfade für A und $\bar{A}$, die zum Ereignis B führen zu berücksichtigen.
$P(B) = P(A)\,P_A(B) + P(\bar{A})\,P_{\bar{A}}(B)$
Stochastische Unabhängigkeit heißt **$P_A(B) = P(B)$** damit gilt aber auch **$P_{\bar{A}}(B) = P(B)$**. Genau das bedeutet, dass an den beiden Pfaden, die zum Ereignis B führen die Wahrscheinlichkeit P(B) ist. Das heißt aber, dass die Wahrscheinlichkeit für B unabhängig davon ist, ob A oder $\bar{A}$ eingetreten ist. Setzt man die beiden bedingten Wahrscheinlichkeiten in der obigen Formel durch P(B), dann ergibt sich:
$P(B) = P(A)\,P(B) + P(\bar{A})\,P(B) = P(B)\,(P(A) + P(\bar{A})\,) = P(B)$.



Sind die Wahrscheinlichkeiten an allen Teilbäumen, die zu ein und demselben Ereignis führen gleich, dann kann man diese Wahrscheinlichkeit aus allen Zweigen, die zu diesem Ereignis führen ausklammern. Die Summe in der Klammer ergibt dann 1, da es die Summe aller vorhergehenden Ereignisse ist. Für abhängige Ereignisse, die unterschiedliche Wahrscheinlichkeiten besitzen, ist das Ausklammern nicht möglich.
Umgesetzt auf das geänderte Beispiel ergibt sich jetzt folgendes Baumdiagramm:



Da im Baum nach wie vor die 1. Pfadregel zu gelten hat, kann man die fehlenden Wahrscheinlichkeiten aus den gegebenen berechnen: 7/10 * $P_A(B)$ = 0,28 . Auf diese Art und Weise lassen sich die Wahrscheinlichkeiten am Baum ergänzen:



Die Wahrscheinlichkeit für das Eintreten des Ereignisses B ist immer 0,4 , gleichgültig, ob vorher das Ereignis A oder das Ereignis $\overline{A}$ eingetreten ist. Da es insgesamt 40 rote Kugeln gibt, ist auch durch diese Rechnung der Wahrscheinlichkeit für eine rote Kugel 0,4.

Die Ereignisse A und B sind **stochastisch unabhängig**.

Am Baumdiagramm sind unabhängige Ereignisse daran zu erkennen, daß die Wahrscheinlichkeiten an den zweiten Bäumen gleich sind. Der zweite Baum, der bei A beginnt hat die gleichen Wahrscheinlichkeiten, wie der Baum, der bei $\overline{A}$ beginnt.

32.6.5 Unabhängige Ereignisse in der Vierfeldertafel

In der Vierfeldertafel kann man die unabhängigen Ereignisse über die sogenannten Randwahrscheinlichkeiten nachweisen. Die Struktur einer Vierfeldertafel hat etwa folgendes Aussehen:

	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	1

Wenn sich jetzt für **jedes** der vier Mittelfelder die Produktbeziehung nachweisen lässt, dann sind die beiden Ereignisse stochastisch unabhängig.

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

$P(A \cap \overline{B}) = P(A)\ P(\overline{B})$

$P(\overline{A} \cap B) = P(\overline{A})\ P(B)$

$P(\overline{A} \cap \overline{B}) = P(\overline{A})\ P(\overline{B})$

dabei sind die Wahrscheinlichkeiten P(A), P(B) usw. die sogenannten Randwahrscheinlichkeiten, da sie außerhalb der 2x2 Tafel stehen. Die Erklärungen dazu sind im Kapitel 1.5 zu finden.

In einer Vierfeldertafel lässt sich die stochastische Unabhängigkeit recht einfach über die Zeilen- und Spaltensummen prüfen.

	B	$\overline{B}$	
A	$P(A) \cdot P(B)$	$P(A) \cdot P(\overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A}) \cdot P(B)$	$P(\overline{A}) \cdot P(\overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	

Ergeben sich die Werte der vier mittleren Felder genau aus dem Produkt der Zeilen– und Spaltensummen, dann sind die Ereignisse stochastisch unabhängig.

Wenn für die neue Aufgabenstellung die Vier-Felder-Tafel erstellt wir, hat sie folgendes Aussehen:

	B	$\overline{B}$	Summe
A	$P(A \cap B) = 0{,}28$	$P(A \cap \overline{B}) = 0{,}42$	$P(A) = 0{,}7$
$\overline{A}$	$P(\overline{A} \cap B) = 0{,}12$	$P(\overline{A} \cap \overline{B}) = 0{,}18$	$P(\overline{A}) = 0{,}3$
Summe	$P(B) = 0{,}40$	$P(\overline{B}) = 0{,}60$	1

Für unabhängige Ereignisse ist die Wahrscheinlichkeit des Durchschnitts von zwei Ereignissen gleich dem Produkt der Wahrscheinlichkeiten der beiden Einzelereignisse:

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

Für die Vier-Felder-Tafel bedeutet das, die vier Felder in der Mitte ergeben sich aus dem Produkt der Randwahrscheinlichkeiten:

	B	$\overline{B}$	Summe
A	$0{,}4 \cdot 0{,}7 = 0{,}28$	$0{,}6 \cdot 0{,}7 = 0{,}42$	$P(A) = 0{,}7$
$\overline{A}$	$0{,}4 \cdot 0{,}3 = 0{,}12$	$0{,}6 \cdot 0{,}3 = 0{,}18$	$P(\overline{A}) = 0{,}3$
Summe	$P(B) = 0{,}40$	$P(\overline{B}) = 0{,}60$	1

Die Wahrscheinlichkeiten in den 4 Feldern entsprechen genau den Produkten der Zeilen- und Spalten-summen. Damit sind die Ereignisse stochastisch unabhängig.
Die 4-Felder-Tafel zeigt auch:
Ist B unabhängig von A, dann ist auch A unabhängig von B.
Das gleiche gilt für die Gegenwahrscheinlichkeiten: B ist auch unabhängig von $\overline{A}$.

Aus diesem Beispiel ist folgende wichtige Erkenntnis zu ziehen:

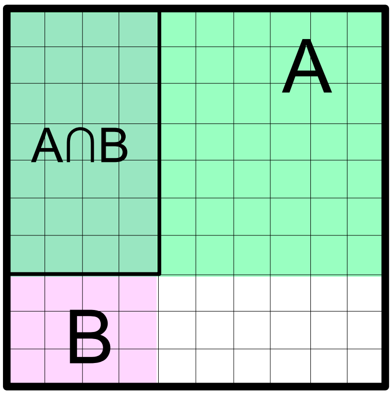
Stochastische Unabhängigkeit kann man nicht am Ereignis erkennen.

Stochastische Unabhängigkeit ist **keine Eigenschaft des Ereignisses** (Versuchs), stochastische Unabhängigkeit ist **eine Eigenschaft der Wahrscheinlichkeiten**.

Damit ist stochastische Unabhängigkeit nur über Rechnung, aber nicht über den Versuchsaufbau und dessen Zusammenhänge zu bestimmen.

32.6.6 Graphische Veranschaulichung unabhängiger Ereignisse

Die Wahrscheinlichkeit von A ist P(A) = 70/100.
Die Wahrscheinlichkeit von B ist P(B) = 40/100.
Die Wahrscheinlichkeit von A∩B ist P(A∩B) = 28/100.



- Das Ereignis A nimmt innerhalb von Ω 70 Teile ein. Damit ist die Wahrscheinlichkeit von A : 60/100
- Das Ereignis B nimmt innerhalb des Ereignisses A 28 Teile ein, das Ereignis A selbst aber 70 Teile. Damit ist die Wahrscheinlichkeit von B innerhalb von A
  P_A (B) = 28/70 = 2/5 .
- Das Ereignis B nimmt innerhalb des Ereignisses $\overline{A}$ 12 Teile ein, das Ereignis $\overline{A}$ selbst aber 30 Teile. Damit ist die Wahrscheinlichkeit von B innerhalb von $\overline{A}$
  $P_{\overline{A}}$ (B) = 12/30 = 2/5 .

32.7 Totale Wahrscheinlichkeit

Beim umgekehrten Baumdiagramm im Kapitel 1.4 wurde folgende Formel hergeleitet:

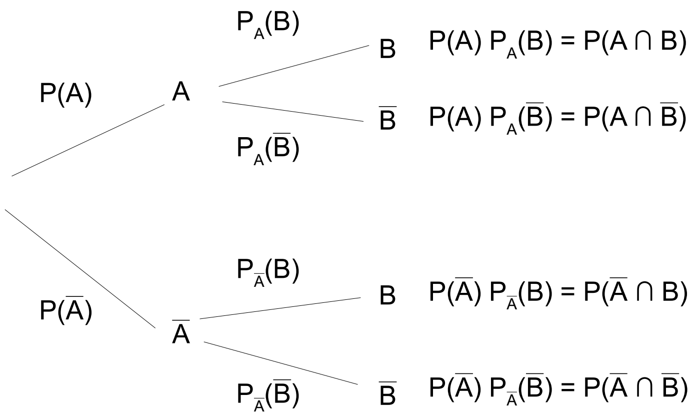
$$P(B)\,P_B(A) = P(A \cap B) = P(A)\,P_A(B)$$

Unabhängig, welchen Baum man benutzt, ist die Wahrscheinlichkeit für das gleichzeitige Eintreten zweier Ereignisse immer gleich. Es resultiert aus dem Produkt der Wahrscheinlichkeit des ersten Ereignisses mit der Wahrscheinlichkeit des zweiten Ereignisses unter der Bedingung, dass das erste Ereignis eingetreten ist. Diese Formel ermöglicht die Bestimmung einer bedingten Wahrscheinlichkeit, wenn die umgekehrte bedingte Wahrscheinlichkeit bekannt ist.

$$\boxed{P_B(A)=\frac{P(A)\cdot P_A(B)}{P(B)}}$$

Dies ist nichts anderes, als eine zweite Formel der bedingten Wahrscheinlichkeit, da $P(A)\,P_B(A) = P(A \cap B)$ ist. Das ist insofern eine interessante Formel, als hier Ursache und Wirkung vertauscht wird. Nach dem Eintreten des Ereignisses A tritt das Ereignis B ein. Dafür kann man die Wahrscheinlichkeit $P_A(B)$ bestimmen. Es wird aber eine Aussage gemacht über das Ereignis A, wenn das Ereignis B als erstes eingetreten wäre $P_B(A)$. Mitunter lässt sich diese Reihenfolge gar nicht ermitteln, aber die Wahrscheinlichkeit dafür berechnen.

Aus dem Baumdiagramm ist bekannt, dass die Wahrscheinlichkeit des gleichzeitigen Eintretens zweier Ereignisse gleich dem Produkt der Pfadwahrscheinlichkeiten ist, wobei der zweite Faktor eine bedingte Wahrscheinlichkeit ist, unter der Bedingung, dass das erste Ereignis eingetreten ist. Besonders deutlich sieht man das bei „Ziehen ohne Zurücklegen“. Die Wahrscheinlichkeiten am zweiten Baum ändern sich im Verhältnis zu ersten Baum.



Will man außerdem die Wahrscheinlichkeit für das einzelne Ereignis B im zweiten Zug bestimmen, sind von allen Pfaden, die zu dem Ereignis B führen, die Wahrscheinlichkeiten zu addieren.

$$\boxed{P(A)=P(B)\cdot P_B(A)+P(\bar{B})P_{\bar{B}}(A)}$$

Wenn $P(A) > 0$ ist, alles andere macht keinen Sinn, dann kann man die bedingte Wahrscheinlichkeit $P_B(A)$ auch wie folgt berechnen, indem man im Nenner den Wert von P(A) ersetzt:

$$\boxed{P_A(B)=\frac{P(B)\cdot P_B(A)}{P(A)}=\frac{P(B)\cdot P_B(A)}{P(B)\cdot P_B(A)+P(\bar{B})\cdot P_{\bar{B}}(A)}}$$

Diese Formel spielt bei der Untersuchung von Gesundheitstests eine Rolle, die nicht 100% Sicherheit bringen. Bei solchen Tests werden zu einem Prozentsatz die Krankheit bei Kranken nicht erkannt und bei Gesunden Menschen bringen die Test ein Ergebnis, dass dieser von der Krankheit betroffen sei.

Eine bestimmte Krankheit tritt bei 1% aller Personen auf.
Ein Testverfahren für diese Krankheit liefert ein positives Ergebnis bei 98% aller Kranken und bei 5% aller gesunden Personen (5% der gesunden Personen werden durch diesen Test fälschlicherweise als krank eingestuft).

Frage: Wie groß ist die Wahrscheinlichkeit, dass eine Person mit positivem Testergebnis wirklich krank ist ?

32.7.1 Umgekehrte bedingte Wahrscheinlichkeit mit Formel

Analyse der Lösung:
K : Die Person ist krank – $P(K) = 0{,}01$ und $P(\overline{K}) = 0{,}99$
P: Test positiv ausgefallen – $P_K(P) = 0{,}98$ und $P_{\overline{K}}(P) = 0{,}05$

$P_K(P) = 0{,}98$: 98 % der tatsächlich Erkrankten werden als solche erkannt.
$P_{\overline{K}}(P) = 0{,}05$: 5 % der nicht Erkrankten werden als krank erkannt.
Diese bedingten Wahrscheinlichkeiten beschreiben die Verläßlichkeit des Labortests. Ihnen entsprechen die Fehlerwahrscheinlichkeiten.
Aus der Sicht eines Patienten ist es nicht so sehr interessant, wie viele Kranke oder Gesunde der Labortest als solche erkennt; es geht ihm vielmehr darum, was aus einer einmal getroffenen Entscheidung geschlossen werden kann. Es geht also um die folgenden bedingten Wahrscheinlichkeiten, die man auch **a posteriori Wahrscheinlichkeiten** nennt:
$P_P(K)$ =? Wie viele der als krank eingestuften Untersuchungspersonen sind tatsächlich krank?
$P_P(\overline{K})$ =? Wie viele der als krank eingestuften Untersuchungspersonen sind tatsächlich gesund?

$$\mathbf{P_P(K)} = \frac{\mathbf{P(K) \cdot P_K(P)}}{\mathbf{P(K) \cdot P_K(P) + P(\overline{K}) \cdot P_{\overline{K}}(P)}} = \frac{0{,}01 \cdot 0{,}98}{0{,}01 \cdot 0{,}98 + 0{,}99 \cdot 0{,}05} \approx 0{,}16526$$

Die Gegenwahrscheinlichkeit, dass ein positiver Test einen gesunden Menschen betrifft ist:

$$\mathbf{P_P(\overline{K})} = \frac{\mathbf{P(\overline{K}) \cdot P_{\overline{K}}(P)}}{\mathbf{P(K) \cdot P_K(P) + P(\overline{K}) \cdot P_{\overline{K}}(P)}} = \frac{0{,}05 \cdot 0{,}99}{0{,}01 \cdot 0{,}98 + 0{,}99 \cdot 0{,}05} \approx 0{,}834738$$

Warum die Formel des Nenners in dieser Weise entsteht, ist noch einmal im Kapitel 3.3 über die Berechnung im Baumdiagramm zu finden. Offenbar genügt es nicht, die Verläßlichkeit bzw. die Fehlerwahrscheinlichkeiten des Labortests zu kennen. Es müssen darüber hinaus die a priori Wahrscheinlichkeiten $P(K)$ und $P(\overline{K})$ der mögliche Zustände des Patienten bekannt sein. Die Wahrscheinlichkeit, dass ein positiver Test einen gesunden Menschen betrifft ist größer, als dass ein solcher Test einen kranken Menschen betrifft. Test scheint nicht sehr geeignet zu sein.

32.7.2 Umgekehrte bedingte Wahrscheinlichkeit mit Vier-Felder-Tafel

In Kapitel 1.5 wurden die Bedingten Wahrscheinlichkeiten in einer Vier-Felder-Tafel angezeigt. Markanter Ausdruck dafür ist die Zeilen oder Spaltensumme von „1“.
Hier würde für die Aufgabestellung folgende Vier-Felder-Tafel Anwendung finden:

	P	$\overline{P}$	
K	$P_K(P)$	$P_K(\overline{P})$	1
$\overline{K}$	$P_{\overline{K}}(P)$	$P_{\overline{K}}(\overline{P})$	1

Da aus der Aufgabenstellung nur die bedingten Wahrscheinlichkeiten zur „Positiver Befund unter Bedingung Krank“ und „Positiver Befund unter Bedingung gesund“ bekannt sind, muss über die Zeilen normiert werden.“ In der ersten Spalte findet man die beiden bedingten Wahrscheinlichkeiten, die in der Aufgabenstellung angegeben sind. da die Wahrscheinlichkeiten der Zeilensummen „1“ ergeben muss, hat die Vier-Felder-Tafel zunächst folgendes Aussehen:

	P	$\overline{P}$	
K	0,98	*0,02*	1
$\overline{K}$	0,05	*0,95*	1

Die Wahrscheinlichkeiten der Spalte $\overline{P}$ ergibt sich nur rechnerisch daraus, dass die Zeilensumme „1“ sein muss. Der Wert in der Spalte P und Zeile K entsteht als bedingte Wahrscheinlichkeit aus $P_K(P) = P(P \cap K) / P(K)$. Für den Wert in der normalen Vier-Felder-Tafel ist also dieser Wert mit $P(K)$ zu multiplizieren, um die übliche Vier-Felder-Tafel zu erhalten. Die zweite Zeile ist entsprechenden mit $P(\overline{K})$ zu multiplizieren. Diese Wahrscheinlichkeiten sind aus der Aufgabestellung bekannt: $P(K) = 0{,}01$ und $P(\overline{K}) = 0{,}99$. Durch Multiplikation erhält man die nicht normierte Vier-Felder Tafel:

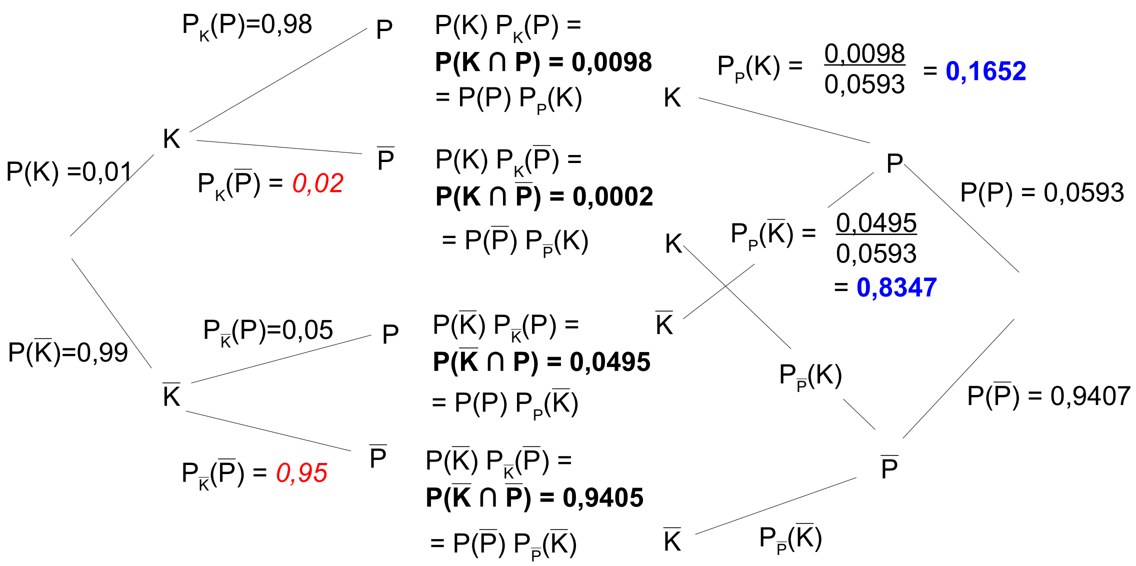
	P	$\overline{P}$	
K	0,0098	0,0002	0,01
$\overline{K}$	0,0495	0,9405	0,99
	0,0593	0,9407	**1**

Die Summen der Spalten P und $\overline{P}$ ergeben sich aus Addition der beiden drüber liegenden Zeilen. Die Summe der Spaltensummen und die Summe der Zeilensummen liefern beide den Wert „1“, was bei einer normalen Vier-Felder-Tafel sein muss. Gesucht ist die bedingte Wahrscheinlichkeit eines „Kranken unter der Bedingung eines positiven Tests“ und eines „Gesunden unter der Bedingung eines positiven Tests“. Das ist die Umkehrung und bedeutet für die Vier-Felder-Tafel das Normieren auf die Spaltensummen. Es ist also jede Spalte durch ihre Spaltensumme zu dividieren.

	P	$\overline{P}$
K	0,1652	0,000212
$\overline{K}$	0,8347	0,99978
	1	1

In der ersten Spalte erkennt man genau die Werte, die im vorherigen Abschnitt über die Formeln berechnet wurden.

32.7.3 Umgekehrte bedingte Wahrscheinlichkeit mit doppeltes Baumdiagramm



Der Aufbau des Baumes beginnt auf der linken Seite. Die Wahrscheinlichkeiten $P(K) = 0,01$ und $P(\overline{K}) = 0,99$ sind der Aufgabenstellung entnommen. $P(\overline{K})$ ergibt sich aus $1 - P(K)$. Die bedingten Wahrscheinlichkeiten $P_K(P) = 0,98$ und $P_{\overline{K}}(P) = 0,05$ sind ebenfalls in der Aufgabenstellung angegeben.
Die ersten zu berechnenden Zahlen sind die Wahrscheinlichkeiten der zweiten Zweige an den Bäumen der zweiten Stufe (rot gekennzeichnet). Diese ergeben sich jeweils aus der Bedingung, dass alle Wahrscheinlichkeiten die von einem Knoten ausgehen die Summe 1 haben müssen.
Die Produkte der Wahrscheinlichkeiten der vollständige Zweige eines Baumes liefern die Wahrscheinlichkeit für das gleichzeitige Eintreten beider Ereignisse, oder den Durchschnitt der Ereignisse. Diese Wahrscheinlichkeiten sind in der mittleren Spalte berechnet (fett geschrieben). Die Wahrscheinlichkeiten dieses Durchschnitts sind aber gleichzeitig die Wahrscheinlichkeiten des zweiten Baumes, wenn die Reihenfolge der Ereignisse vertauscht werden. Die Wahrscheinlichkeit des Durchschnitts der Ereignisse kann nicht von der Reihenfolge, welches Ereignis zuerst eintritt bestimmt werden. Diese Vertauschungen sind jeweils die dritten Zeilen in der mittleren Spalte angezeigt.

Damit beleibt die Frage, wie kommt man zu den Wahrscheinlichkeiten im rechten Baum. Für die erste Stufe sind die Wahrscheinlichkeiten für das Eintreten eines „positiven Ergebnisses“ und die Wahrscheinlichkeiten für das eintreten einen „negativen Ergebnisses“ gefragt, die beide nicht direkt bekannt sind. Aber es gilt im Baumdiagramm der Grundsatz: Die Wahrscheinlichkeit eines Ereignisses der zweiten Stufe ist die Summe aller möglichen Ereignisse, die zu diesem Ereignis führen. Damit gilt für:

$P(P) = P(K)\ P_K(P) + P(\overline{K})\ P_{\overline{K}}(P) = 0{,}01\ 0{,}98 + 0{,}99\ 0{,}05 = 0{,}0593$

$P(\overline{P}) = P(K)\ P_K(\overline{P}) + P(\overline{K})\ P_{\overline{K}}(\overline{P}) = 0{,}01\ 0{,}02 + 0{,}99\ 0{,}95 = 0{,}9407$

Die Formel zu P(P) entspricht genau dem Nenner im Kapitel 3.1, bei der die Wahrscheinlichkeit $P_P(K)$ über die Formel ermittelt wurde.

Bleibt die Berechnung der bedingten Wahrscheinlichkeiten, die wieder zur gleichen Wahrscheinlichkeit des Durchschnitts führen muss. Diese Wahrscheinlichkeiten entstehen wieder durch Division $P_P(K) = P(P \cap K) / P(P)$. Im obigen Baumdiagramm ist das Ergebnis (als blaue Zahlen) eingetragen. Es sind die gleichen Wahrscheinlichkeiten, die in den beiden vorhergehenden Abschnitte ermittelt wurden. Es ist darauf zu achten, dass sich beim rechten Baum die Pfade einmal kreuzen. Das ist der Tatsache geschuldet, dass immer die gleichen Ereignisse zusammengebracht werden müssen. Im Kapitel 1.4 ist das bei der theoretische Herleitung bereits zu erkennen.
Setzt man für die Wahrscheinlichkeit $P(P \cap K)$ die Berechnung aus dem linken Baum ein, erhält man die Formel $P(K)\ P_K(P)$, was genau dem Zähler in der Formel des Kapitels 3.1 entspricht.

Es ist auf alle Fälle interessant zu vergleichen, wo die einzelnen Wahrscheinlichkeiten diese Baumdiagramms in der Vierfeldertafel wieder auftauchen. Man kann sie dort alle wiederfinden.

32.8 Zufallsvariable

Eine Zufallsvariable ist weder ein zufälliges Ereignis, noch eine Variable, sondern eine unglückliche Bezeichnung für eine Funktion !

Nicht immer sind bei einem Zufallsversuch die einzelnen Ereignisse von Bedeutung. Meistens interessiert nur eine Zusammenfassung von mehreren Ereignissen und dann die Wahrscheinlichkeit, mit der diese Gesamtmenge eintritt. Ein Beispiel dafür wäre: Es wird aus einer Urne mit 3 verschieden farbigen Kugeln 10 mal gezogen, wie groß ist die Wahrscheinlichkeit bei diesem 10-maligen Ziehen 4 rote Kugel zu ziehen.
Aus dem Baumdiagramm ist bekannt, dass es mehrere Möglichkeiten gibt bei 10-maligem Ziehen 4 rote Kugeln zu ziehen. Man kann die 4 roten Kugeln zuerst ziehen, man kann aber auch erst zwei andere Farben ziehen, dann eine rote, wieder eine andere Farbe und dann zwei rote. Es gibt also mehrere Möglichkeiten. Aber eines ist allen Möglichkeiten gleich: Am Ende sind 4 rote Kugeln auf dem Tisch, in welcher Reihenfolge auch immer. Diese Zusammenfassung drückt man einer Zufallsvariablen, meist als X, Y Z oder ähnlichen Großbuchstaben bezeichnet, aus:
X : Anzahl der gezogenen roten Kugeln
Von diesen möglichen Anzahlen interessiert im angegebenen Fall, wann sind es 4 rote Kugeln. Diese Fragestellung wird durch die Formulierung **X = 4** zu Ausdruck gebracht. Da natürlich nicht immer 4 rote Kugeln am Ende auf dem Tisch liegen ist die Frage nach der Wahrscheinlichkeit gestellt, oder wie oft ist damit zu rechnen, dass 4 rote Kugeln auf dem Tisch liegen. Diese Aufgabenstellung wird formuliert durch den Ausdruck
P(X=4)
Eine mathematische Schreibweise für diesen Sachverhalt wäre etwa folgende:

$$\{ X = k \} := \{\omega \in \Omega \mid X(\omega) = k, k \in \mathbb{R}\}.$$

Der Wert der Zufallsfunktion X nimmt den Wert k an für diejenigen $\omega \in \Omega$, für die die Anzahl des Eintretens des Ereignisses X gleich k ist.
Dazu soll folgendes Beispiel betrachtet werden:
Es wird mit zwei Würfeln gleichzeitig gewürfelt, dann sind folgende Zufallsvariable möglich:

X : Die Summe der beiden Augenzahlen
Damit wird folgende Abbildung von Ω auf $\mathbb{R}$ durchgeführt:

$\omega \in \Omega$	$X = k$
(1;1)	2
(1;2);(2;1)	3
(1,3);(2,2);(3,1)	4
(1,4);(2,3);(3,2);(4,1)	5
(1,5);(2,4);(3,3);(4,2);(5,1)	6
(1,6);(2,5);(3,4);(4,3);(5,2);(6,1)	7
(2,6);(3,5);(4,4);(5,3);(6,2)	8
(3,6);(4,5);(5,4);(6,3)	9
(4,6);(5,5);(6,4)	10
(5,6);(6,5)	11
(6,6)	12

Y: Das Maximum der beiden Augenzahlen

$\omega \in \Omega$	$Y = k$
(1;1)	1
(1;2);(2;1);(2,2)	2
(1,3);(3,1);(2,3);(3,2),(3,3)	3
(1,4);(2,4);(3,4);(4,1);(4,2);(4,3);(4,4)	5
(1,5);(2,5);(3,5);(4,5);(5,5);(5,4);(5,3);(5,2),(5,1)	6
(1,6);(2,6);(3,6);(4,6);(5,6);(6,6);(6,5);(6,4);(6,3);(6,2);(6,1)	7

Bei der ersten Funktion X ist die Bildmenge der Funktion die Zahlen von 2 bis 12. Der Wert 1 ist nicht in der Bildmenge enthalten. Bei der zweiten Funktion Y ist die Bildmenge die Zahlen von 1 bis 6.
Damit treten die Werte von X als Wertebereich der Abbildung X auf. Der Definitionsbereich sind die möglichen $\omega \in \Omega$ des (tatsächlichen) Zufallsexperiments. Wichtig ist hier vielleicht festzuhalten, dass die Werte des Definitionsbereichs in beiden Fällen gleich sind, aber die Werte de Wertevorrats verschieden. Eine Tatsache, die auch bei normalen Funktionen bekannt ist. Der Definitionsbereich von $y = x^2$ und der Definitionsbereich von $y = x^3$ sind der gleiche, aber die y -Werte sind verschieden.

Hier besteht jetzt noch der Zusatz, dass die Werte von X auch noch Definitionsbereich einer weiteren Funktion sind, nämlich der Funktion, die die Werte von X auf ihre Wahrscheinlichkeit abbildet, mit der diese Werte eintreten. Diese Wahrscheinlichkeit ergibt sich aus den summierten Wahrscheinlichkeiten aller $\omega \in \Omega$, die auf das gleiche k in X abgebildet werden. Diese Abbildung ist diejenige, die mittels Histogramm dargestellt wird.

Was heißt das mathematisch
Es ist bekannt , das es mehrere Ereignisse (Reihenfolgen) gibt, die zu dem Ergebnis führen, dass 4 rote Kugeln auf dem Tisch liegen. Jede einzelne Reihenfolge ist ein Element $\omega \in \Omega$. Es interessiert aber nicht, in welcher Reihenfolge die roten Kugeln gezogen werden, sondern nur die 4 als Anzahl der gezogenen Kugeln. Alle diese Elementarereignisse werden zu einem Gesamtereignis zusammengefasst, das besagt: 4 rote Kugeln wurden gezogen. Mathematisch bedeutet das, dass man eine Abbildung der Ereignisse auf die Zahl 4 erzeugt. Man kann sich das vorstellen, wie eine normale Funktion bei der die übliche x-Werte die $\omega \in \Omega$ sind und die y-Werte sind ganze Zahlen, hier die Zahl 4. Daß das keine solch schöne Kurve gibt, wie z.B $y = x^2$ oder $y = \sin(x)$ hat mathematisch keine Bedeutung. Von der sin – Funktion ist ja auch bekannt, dass mehrere x-Werte durchaus den gleichen y -Wert haben können. Bei den trigonometrischen Funktionen liegt das an der Periodizität der Funktionen. Aber selbst bei der Funktion $y =x^2$ gibt es zu zwei x-Werten den gleichen y-Wert.

Definition:
Eine **Zufallsvariable** ist eine Funktion, die jedem Ergebnis eines Zufallsversuchs $\omega \in \Omega$ eine reelle Zahl $X(\omega)$ zuordnet.

Letztenendes werden durch eine Zufallsvariable die Ergebnisse eines Zufallsversuchs vergröbert, indem nur die Anzahl interessiert, aber nicht, wie diese zustande gekommen ist. Gemäß der Definition einer Funktion ist es durchaus möglich einem $X(\omega)$ mehrere ω zuzuordnen. Andererseits gibt diese Zusammenfassung die Möglichkeit andere Unterscheidungen in den Vordergrund zu schieben, wie z.B Gesucht ist die

Wahrscheinlichkeit 6 rote Kugeln zu ziehen, oder die Wahrscheinlichkeit 2 rote Kugeln zu ziehen.

Definition:
Eine **Verteilung** P(X) ist eine Funktion, die jedem Ergebnis $x \in X(\Omega)$ eines Zufallsversuchs $\omega \in \Omega$ eine Wahrscheinlichkeit $P(X(\omega))$ zuordnet.

Damit wird jeder Zusammenfassung von ω zu einem x(ω) eine Wahrscheinlichkeit ihres Erscheinens zugeordnet. Das Bestimmen der Wahrscheinlichkeit führt dann in der Ausdrucksweise eines Baumdiagramms dazu: Bestimme alle Pfade, bei denen 4 rote Kugeln auftreten multipliziere die Wahrscheinlichkeit entlang der Pfade und addiere schließlich alle betroffenen Pfade mit ihren Wahrscheinlichkeiten auf. (1. und 2. Pfadregel). Alle diese Pfade werden auf die gleiche reelle Zahl abgebildet.

Beispiele:	für Zufallsvariable Werte für $X = x_i$:
X = „Augenzahl eines Würfels“	$x_1 = 1, x_2 = 2, \ldots x_6 = 6$
X = „Zahl der 6er bei 5maligem Würfeln“	$x_1 = 0, x_2 = 1, \ldots x_6 = 5$
X = „Anzahl der entdeckten Schmuggler an der Grenze bei einer Stichprobe vom Umfang 4“	$x_1 = 0, x_2 = 1, \ldots x_5 = 4$
X = „Füllmenge einer 3kg Waschmittelpackung“ (ACHTUNG! stetige Verteilung, keine Einzelwerte !)	$xi \in [\,2{,}95\,;\,3{,}05\,]$
X = „Anzahl der verdorbenen Äpfel bei zufälliger Entnahme von 4 Stück aus einem Sack mit 10 Äpfeln, von denen 2 verdorben sind“ (gleichgültig, wie oft man zieht, man kann nur 2 schlechte Äpfel finden, da es laut Voraussetzung nicht mehr gibt, aber die Möglichkeit die beiden schlechten Äpfel zu finden erhöht sich mit der Anzahl der Ziehungen !)	$x_1 = 0, x_2 = 1, x_3 = 2$

Für das erste Beispiel kann man die Wahrscheinlichkeiten ohne Probleme angeben. Die Beispiele 2 und 3 unterliegen der Binomialverteilung und Beispiel 5 der hypergeometrischen Verteilung, die in späteren Kapitel besprochen und deshalb hier noch nicht weiter betrachtet werden. Die Zufallsvariable 4 ist eine stetige Verteilung, die in der Schule nicht behandelt wird.

Für die erste Zufallsvariable „Augenzahl eines Würfels“ kann die Zufallsvariable den Wert 1 annehmen (aber niemals den Wert 0) . Also ist möglich X = 1 mit einer Wahrscheinlichkeit p = 1/6 . Damit ergibt sich P(X=1) = 1/6 „Die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert 1 annimmt ist 1/6“.
„Die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert 5 annimmt ist 1/6“
P(X=5)=1/6.

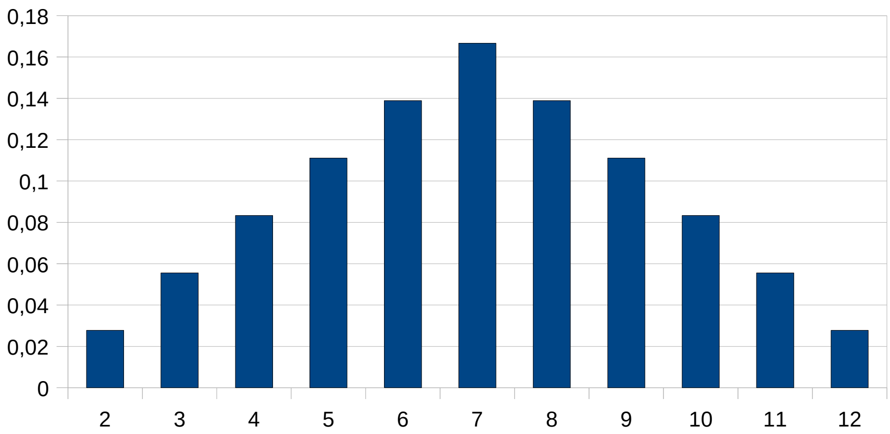
Als nächstes soll ein zweimaliges Würfeln betrachtet werden und es interessiert die Summe der beiden Augenzahlen.

Zufallsvariable
X = „ Summe der Augenzahlen bei zweimaligen Würfeln“

Die möglichen Summen sind 2 bis 12, aber mit unterschiedlichen Wahrscheinlichkeiten:

P(X=2) =	1/36	P(X=8) =	5/36
P(X=3) =	2/36	P(X=9) =	4/36
P(X=4) =	3/36	P(X=10) =	3/36
P(X=5) =	4/36	P(X=11) =	2/36
P(X=6) =	5/36	P(X=12) =	1/36
P(X=7) =	6/36		

Die Abbildungen der einzelnen ω, die zu dem gleichen Wert der Zufallsvariablen führen werden nicht graphisch als Funktion dargestellt, aber die Werte der Zufallsfunktion mit ihren zugehörigen Wahrscheinlichkeiten stellt man oft in Balkendiagrammen dar. Diesen Zusammenhang zwischen den Werten der Zufallsvariablen und ihren zugehörigen Wahrscheinlichkeiten nennt man Wahrscheinlichkeitsfunktion. Diese Funktion zeigt an, wie oft ein Wert der Zufallsvariablen eintreten kann.



Das ist eine Darstellung der Zufallsvariablen „Summe der Augenzahlen beim zweimaligen Würfeln“ in der Form eines Histogramms. Die y-Werte sind die Wahrscheinlichkeiten der jeweiligen Realisierungen der Zufallsvariablen.

32.8.1 Erwartungswert

Ganz eng mit dem Begriff der Zufallsvariablen ist der Begriff des Erwartungswertes verbunden. Der Erwartungswert drückt einen Mittelwert in der Wahrscheinlichkeitsrechnung aus. Dabei werden aber nicht die möglichen Werte für die Zufallsvariable zusammenaddiert und durch die Anzahl geteilt, was wenig Sinn macht, sondern man benutzt einen gewichteten Mittelwert. Man berücksichtigt bei der Berechnung des Mittelwertes, mit welcher Wahrscheinlichkeit die jeweilige Anzahl auftritt.
Am einfachsten und einleuchtendsten ist immer der Erwartungswert eines Spiels. Man stelle sich eine Drehscheibe – Glücksrad – vor, das aus drei unterschiedlich großen Sektoren besteht. Dies unterschiedlichen Sektoren repräsentieren unterschiedliche Wahrscheinlichkeiten des Eintretens des Ereignisses. Ein Sektor sei 1/2 der Scheibe,

was zu einer Wahrscheinlichkeit von 0,5 führt, ein zweiter Sektor 1/3 und ein dritter dann 1/6 der Scheibe. Die beiden Wahrscheinlichkeiten dazu sind 0,33 und 0,16.
Je größer der Sektor, desto kleiner die Auszahlung, wobei die Auszahlung nicht proportional zur Wahrscheinlichkeit sein muss. Für den ersten Sektor soll der Gewinn 1€ betragen, für den zweiten Sektor 2€ und für den dritten Sektor 5€.
Dann berechnet sich der Erwartungswert aus den Gewinnmöglichkeiten mit den zugehörigen Wahrscheinlichkeiten ihres Eintretens. Das führt in diesem Fall zu folgendem Ergebnis:

$$1€\ 0{,}5 + 2€\ 0{,}33 + 5\ €\ 0{,}166 = 1{,}99€$$

Bei dieser Drehscheibe gewinnt man also im Durchschnitt 2€, wenn man lange genug das Spiel betreibt. Ob für das Spiel ein Einsatz verlangt wird oder nicht, spielt zunächst für den Erwartungswert keine Rolle. Wird ein Einsatz verlangt, ist der Betrag des Einsatzes vom Erwartungswert zu subtrahieren.

Definition:
Unter dem **Erwartungswert** E(X) einer Zufallsvariablen X versteht man das gewichtete Mittel der Funktion X über Ω, wobei jeder Wert x(ω) mit seiner Wahrscheinlichkeit p(x) gewichtet wird.

Die Formel für den Erwartungswert ist deshalb:

$$\mu = \mathbf{E}(\mathbf{X}) = \sum_{\mathbf{k}=0}^{\mathbf{n}} (\mathbf{x_k} \cdot \mathbf{p}(\mathbf{x_k}))$$

dabei ist x_k der numerische wert, der durch die Zufallsvariable festgelegt wird und $p(x_k)$ die Wahrscheinlichkeit, mit der dieser Wert auftritt. Der Erwartungswert ist damit keine Wahrscheinlichkeit, sondern ein Element aus der Grundmenge X(ω). Wobei der errechnete Wert nicht wirklich in dieser Grundmenge enthalten sein muß. Der Erwartungswert kann eine reelle Zahl sein, obwohl nur ganzzahlige Werte x(ω) möglich sind. Dieser Effekt entsteht dadurch, daß die Wahrscheinlichkeiten in den meisten Fällen reelle Zahlen sind.
Ungewöhnlicher ist dabei schon folgende Aufgabenstellung: Eine Münze wird dreimal geworfen. Gesucht ist das Ereignis, dass „Wappen" geworfen wird. Wir groß ist der Erwartungswert dafür, dass beim dreimaligen Werfen Wappen auftritt. Hier kommt es auf die Definition der Zufallsvariablen an. Die Zufallsvariable wird wie folgt definiert:
X : Beim dreimaligen Werfen einer Münze liegt auf der Oberseite Wappen

Klar ist, dass für diese Zufallsvariable die Realisierungen 0, 1, 2 und 3 auftreten können, aber nicht mit der gleichen Wahrscheinlichkeit. Für die Wahrscheinlichkeiten der einzelnen Realisierungen soll als erstes die Ergebnismenge Ω bestimmt werden. Ω enthält alle Möglichkeiten, in jeder möglichen Reihenfolge:

Ω = {ZZZ, ZZW, ZWZ, WZZ, ZWW, WZW, WWZ, WWW }

Jedes einzelne Ereignis W und Z treten mit der Wahrscheinlichkeit ½ auf, so dass jedes dieser aufgeführten Ereignisse mit der Wahrscheinlichkeit ⅛ auftritt. Wie sieht es aber bei dem Ereignis „Wappen" aus:
P(X = 0) = ⅛ (0 mal Wappen tritt nur einmal auf)
P(X = 1) = ⅜ (1 mal Wappen tritt dreimal auf)
P(X = 2) = ⅜ (2 mal Wappen tritt dreimal auf)

$P(X = 3) = \frac{1}{8}$ (3 mal Wappen tritt nur einmal auf)

Damit ergibt sich als Erwartungswert für diese Zufallsvariable:
$0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1{,}5$

Man kann also davon ausgehen, dass bei drei Würfen im Durchschnitt 1,5 mal Wappen auftritt, das bedeutet, dass in den meisten Fällen 1 mal Wappen oder zweimal Wappen auftritt. Ein Ergebnis, dass nicht wirklich verwundert.

Der Erwartungswert E(X) einer Zufallsvariablen X gibt an, welchen Wert sie bei einer unbegrenzten Wiederholung des Zufallsexperiments im Durchschnitt annehmen würde. Praktisch lässt er sich als Durchschnittswert bei einer großen Anzahl von Wiederholungen des Zufallsvorgangs interpretieren. Der Erwartungswert E(X) von X wird auch als arithmetisches Mittel μ der Grundgesamtheit bezeichnet.

32.8.2 Varianz

In vielen Fällen ist man beim zufälligen Verhalten eines Experiments nicht nur daran interessiert, bestimmte Merkmale zu beobachten und den zu erwartenden Wert dieser Merkmale zu berechnen; oft ist auch von Bedeutung, wie stark ein Merkmal “streut”, d.h., wie stark es vom zu erwartenden Wert abweicht. Für diese zu erwartende Abweichung wird der Begriff der Varianz wie folgt eingeführt.

Definition:
Unter der **Varianz** V(X) einer Zufallsvariablen X mit dem Erwartungswert E(X) versteht man den Erwartungswert von (X – E(X))² .

Um Abweichungen vom Mittelwert zu untersuchen, könnte man auch einfach die Differenz zum Mittelwert nehmen. Da hätte dann aber den Nachteil, daß sich negative und positive Abweichungen gegenseitig aufheben würden. Um dem zu entgehen könnte man wiederum den Betrag nehmen. Die Betragsfunktion hat aber genau an der Stelle 0, also hier die Stelle des Erwartungswertes, einen entscheidenden Nachteil. Sie besitzt dort eine Spitze und ist nicht differenzierbar. Mit dem Quadrat hat man beides: Vorzeichen heben sich nicht gegenseitig auf und an der Stelle 0 ist die Funktion differenzierbar. Dazu kommt noch ein positiver Effekt hinzu: Kleine Werte werden durch das Quadrieren kleiner, fallen also weniger ins Gewicht und große Werte werden größer, so daß starke Abweichungen größere Auswirkungen haben. Hier spielt die Problematik der extremen Ausreißer mit rein.

Varianz und Standardabweichung einer Zufallsvariablen X sind Streuungsmaße, die angeben, wie stark ihre Ergebnisse um den Erwartungswert streuen. Die Varianz Var(X) gibt die durchschnittliche quadrierte Abweichung wieder. Sie wird auch durch das Symbol σ² gekennzeichnet.

$$\sigma^2 = \mathbf{Var}(\mathbf{X}) = \mathbf{E}(\mathbf{X} - \mu)^2 = \sum_{\mathbf{k}=0}^{\mathbf{n}} (\mathbf{x_k} - \mu)^2 \cdot \mathbf{p}(\mathbf{x_k})$$

Die Standardabweichung σ gibt als Wurzel aus der Varianz an, wie stark die Werte der Zufallsvariablen X durchschnittlich von ihrem Erwartungswert E(X) abweichen.