7.3.2 Modalwert

Eine Ausprägung a eines Merkmals X ist ein *Modalwert* (oder *Modus*) der Stichprobe $x_1, ..., x_n$, wenn es keine Ausprägung gibt, die in der Stichprobe häufiger vorkommt, d.h. wenn für die Häufigkeitsverteilung H_n der Stichprobe gilt:

$$H_n(a) \geq H_n(x_i) \quad \text{für alle } 1 \leq i \leq n.$$

Modalwerte sind Lagemaße, aber im Allgemeinen nicht eindeutig. Falls doch, werden sie mit x_{mod} bezeichnet und die Häufigkeitsverteilung heißt *unimodal*. Modalwerte lassen sich auch bei Stichproben von nominalen Merkmalen bestimmen.

7.3.3 Quantile

Zu $x_1, ..., x_n$ bezeichne $x_{(1)}, ..., x_{(n)}$ die *geordnete Stichprobe*, d.h. es gilt

$$\{x_1, ..., x_n\} = \{x_{(1)}, ..., x_{(n)}\} \text{ und } x_{(1)} \leq ... \leq x_{(n)}$$

Zur Beschreibung der Struktur einer Stichprobe zieht man dann sogenannte *Quantile* heran:

Definition
Sei $p \in [0, 1]$. Eine Ausprägung $\boxed{x_p}$ heißt p-**Quantil** der Stichprobe $x_1, ..., x_n$, wenn für die Häufigkeitsverteilung h_n gilt:

$$\sum_{\substack{1 \leq i \leq s \\ a_i \leq x_p}} h_n(a_i) \geq p \quad \text{und} \quad \sum_{\substack{1 \leq i \leq s \\ a_i \geq x_p}} h_n(a_i) \geq 1 - p, \qquad (\star)$$

d.h. mindestens $p \cdot 100\%$ der Stichprobenwerte sind $\leq x_p$ und mindestens $(1 - p) \cdot 100\%$ der Stichprobenwerte sind $\geq x_p$.

i	1	2	3	4	5
x_i	4	5	1	2	6
$x_{(i)}$	1	2	4	5	6

$x_{(3)} = 4$, $x_3 = 1$

$n = 5$, $h_5(1) = \frac{1}{5}$

$p = \frac{1}{2}$, $x_{\frac{1}{2}} = 4 = x_1$ ist $\frac{1}{2}$-Quantil:

$$\sum_{a_i \leq x_{\frac{1}{2}} = 4} h_5(a_i) = h_5(1) + h_5(2) + h_5(4) = \frac{3}{5} \geq \frac{1}{2}$$

$$\sum_{a_i \geq x_{\frac{1}{2}} = 4} h_5(a_i) = \frac{3}{5} \geq \frac{1}{2}$$

Bemerkung

- Für $p = \frac{1}{2}$ heißt das p-Quantil $x_{0,5}$ *Median* (oder *Zentralwert*),

  d.h. mindestens 50 % der Stichprobenwerte sind größer gleich und mindestens 50 % sind kleiner gleich $x_{0,5}$.
- Das $\frac{1}{4}$-Quantil $x_{0,25}$ heißt *unteres Quartil*, das $\frac{3}{4}$-Quantil $x_{0,75}$ *oberes Quartil*.
- $x_{j \cdot 0,1}$ heißt *j-tes Dezil*.
- Im Fall $p = 0$ erhält man den kleinsten Wert, das *Minimum* $x_{min} = x_{(1)}$ der Stichprobe, im Fall $p = 1$ den größten Wert, das *Maximum* $x_{max} = x_{(n)}$.

Alle p-Quantile sind Lagemaße. Verglichen mit dem arithmetischen Mittel ist der Median robuster (d.h. weniger anfällig) gegenüber „Ausreißern".

Satz 7.3.3 (Berechnung des p-Quantils). *Sei $0 < p < 1$. Eine Ausprägung a ist genau dann ein p-Quantil der Stichprobe $x_1, ..., x_n$, wenn gilt:*

$$a = x_{(\lceil np \rceil)}, \quad \text{falls } np \notin \mathbb{N} \quad \text{bzw.} \quad x_{(np)} \leq a \leq x_{(np+1)}, \quad \text{falls } np \in \mathbb{N}, \qquad (\star\star).$$

Hierbei steht $\lceil x \rceil$ für die kleinste ganze Zahl größer oder gleich x.

← obere Gaussklammer

Bsp $\lceil 2{,}1 \rceil = 3$, $\lceil -0{,}9 \rceil = 0$

Insbesondere gilt für den Median $x_{0,5}$ somit

$$x_{0,5} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{falls } n \text{ ungerade,} \\ a, & \text{falls } n \text{ gerade und } a \text{ eine Ausprägung} \\ & \text{mit } x_{\left(\frac{n}{2}\right)} \leq a \leq x_{\left(\frac{n}{2}+1\right)} \text{ ist.} \end{cases}$$

Einschub 7.3.4. ...

1)

i	1	2	3	4	5	6	7	8	9
x_i	-	-	-	-	-	-	-	-	-
$x_{(i)}$	2	4	7	8	10	13	15	18	20

$n = 9$

$p = \frac{1}{4}$: $\lceil np \rceil = \lceil 9 \cdot \frac{1}{4} \rceil = \lceil 2{,}25 \rceil = 3 \overset{\text{Satz}}{\Rightarrow} x_{(3)} = x_{1/4} = 7$

$p = 3/4$: $\lceil np \rceil = \lceil 9 \cdot 3/4 \rceil = \lceil \frac{27}{4} \rceil = 7 \overset{\text{Satz}}{\Rightarrow} x_{(7)} = x_{3/4} = 15$

$p = \frac{1}{2}$: $\lceil 9 \cdot \frac{1}{2} \rceil = \lceil 4{,}5 \rceil = 5 \Rightarrow x_{1/2} = x_{(5)} = 10$

2) In 1) nur bis $i = 8$, also $n = 8$

$p = \frac{1}{2}$: $np = 8 \cdot \frac{1}{2} = 4 \in \mathbb{N} \overset{\text{Satz}}{\Rightarrow} x_{(np)} = x_{(4)} \leq a \leq x_{(np+1)} = x_{(5)}$

$\Leftrightarrow 8 \leq a \leq 10$, d.h. jede Ausprägung a, die $8 \leq a \leq 10$ erfüllt ist ein $x_{\frac{1}{2}}$ - Quantil, also sowohl 8 als auch 10 ist ein Median (also nicht eindeutig).

Daher gibt es folgende Konvention: wähle

$a = \frac{1}{2}\left(x_{(np)} + x_{(np+1)}\right)$ hier dann $a = 9$

7.4 Streumaße

Definition 7.4.1. Sei $x_1, ...x_n$ eine Stichprobe eines quantitativen Merkmals X. Eine Funktion

$$s : \mathbb{R}^n \to \mathbb{R}, \qquad (x_1, ...x_n) \mapsto s(x_1, ...x_n)$$

heißt *Streumaß*, wenn für alle $a \in \mathbb{R}$ und alle $(x_1, ...x_n) \in \mathbb{R}^n$ gilt:

$$s(x_1 + a, ..., x_n + a) = s(x_1, ...x_n),$$

d.h. durch Verschieben der Stichprobe ändert sich die Streuung nicht. Oftmals fordert man für ein Streumaß s auch noch, dass $s(0, 0, ..., 0) = 0$ gilt.

Beispiele 7.4.2.

- *Spannweite*: $x_{max} - x_{min} = x_{(n)} - x_{(1)}$
- *p-Quantilsabstand*, $p \in (0, \frac{1}{2})$: $x_{1-p} - x_p$ = Differenz von $(1-p)$- und p-Quantil; im Fall $p = \frac{1}{4}$ heißt der p-Quantilsabstand kurz *Quartilsabstand*.
- *Medianabweichung*: Median der Abstände $|x_1 - x_{0,5}|, |x_2 - x_{0,5}|, ..., |x_n - x_{0,5}|$ der Stichprobenwerte vom Median $x_{0,5}$
- *Mittlere absolute Abweichung vom Median*: $\tilde{s} = \frac{1}{n}\sum_{i=1}^{n} |x_i - x_{0,5}|$
- *Mittlere quadratische Abweichung vom arithmetischen Mittel* (oder *Varianz*):

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

- *Standardabweichung*: $s = \sqrt{s^2}$

Es handelt sich hier in allen Fällen um Streumaße, da zur Berechung stets Differenzen von Stichprobenwerten und Lagemaßen bzw. von zwei Lagemaßen verwendet werden. Bei Verschiebung der Stichprobe um a, heben sich die Verschiebungen somit gegenseitig auf.

Satz 7.4.3 (Berechnung der Varianz, Steinersche Formel)**.** *Die Varianz s^2 einer Stichprobe $x_1, ..., x_n$ lässt sich wie folgt berechnen:*

$$s^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \overline{x}^2$$

Beweis
Mithilfe der Binomischen Formel gilt:

$$\begin{aligned} s^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2}{n}\overline{x}\sum_{i=1}^{n} x_i + \frac{1}{n}\cdot n\overline{x}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\overline{x}^2 + \overline{x}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2 \qquad \square \end{aligned}$$

$\square$

Beispiele 7.4.4. Urliste $n = 13$

$x_1 = 2, 1, 3, 1, 3, 5, 5, 5, 5, 5, 3, 5, 5 = x_{13}$

x_4

geordnet

$x_{(1)} = 1, 1, 2, 3, 3\ 3, 5, 5, 5, 5, 5, 5, 5 = x_{(13)}$

$x_{(4)}$ $x_{(7)}$

Median $x_{0,5} = x_{(\lceil 13 \cdot 0,5 \rceil)} = x_{(7)} = 5$

20% Quartilabstand $x_{0,2} = x_{(\lceil 13 \cdot 0,2 \rceil)} = x_{(\lceil 2,6 \rceil)} = x_{(3)} = 2$

$x_{0,8} = x_{(\lceil 13 \cdot 0,8 \rceil)} = x_{(11)} = 5$

$x_{1-p} - x_p = x_{0,8} - x_{0,2} = 5 - 2 = 3$

Medianabweichung $x_{0,5} = 5$

Abstände zum Median $|x_{(1)} - 5|, \ldots, |x_{(13)} - 5|$ ergibt

4, 4, 3, 2, 2, 2, 0, 0, 0, 0, 0, 0, 0

geordnet also

0, 0, 0, 0, 0, 0, 0, 2, 2, 2, 3, 4, 4

Medianabweichung (wie oben aus $\lceil 13 \cdot 0,5 \rceil$)

Mittlere absolute Abweichung vom Median $x_{0,5} = 5$

$$\tilde{s} = \frac{1}{13}\left(|x_1 - x_{0,5}| + \cdots + |x_{13} - x_{0,5}|\right) = \frac{17}{13}$$

$n = 13$

Mittelwert $\bar{x}$

$$\bar{x} = \frac{1}{13}\Big(\underbrace{1+1}_{=1 \cdot H_{13}(1)} + 2 + \underbrace{3 + 3 + 3}_{=3 \cdot H_{13}(3)} + \underbrace{5 + 5 + 5 + 5 + 5 + 5 + 5}_{=5 \cdot H_{13}(5)}\Big) = \frac{1}{13}\Big(1 \cdot \overbrace{H_{13}(1)}^{2} + 2 \cdot H_{13}(2) + 3 \cdot H_{13}(3) + 5 \cdot H_{13}(5)\Big) = \frac{48}{13}$$

← Tim

Varianz (Steiner)

$$s^2 = \frac{1}{13}\left(1^2 + 1^2 + 2^2 + 3^2 + 3^2 + 3^2 + 5^2 + \cdots + 5^2\right) - \bar{x}^2$$
$$= \frac{1}{13}\left(1^2 \cdot H_{13}(1) + 2^2 \cdot H_{13}(2) + 3^2 \cdot H_{13}(3) + 5^2 H_{13}(5)\right) - \bar{x}^2$$

$= \frac{1}{13}\left(2+4 + 27 + 7\cdot 25\right) - \left(\frac{48}{13}\right)^2 = \cancel{2} \ldots$

<u>Vortei</u> <u>Steiner bei Berechnung</u>:

$$\left.\begin{array}{l}\text{Mittelwert } x_i^2 \\ \text{Mittelwert } x_i\end{array}\right\} \xrightarrow[\text{Satz 7.4.4}]{\text{im}} \text{alle } H_n(a_i) \text{ gleich}$$

Satz 7.4.5 (Berechnung von $\overline{x}$, $\tilde{s}$ und s^2 mittels Häufigkeitsverteilungen))**.** *Seien $a_1, ..., a_s$ die Ausprägungen von X und H_n bzw. h_n die absolute bzw relative Häufigkeitsverteilung der Stichprobe $x_1, ..., x_n$. Dann gilt:*

$$\begin{aligned}
\overline{x} &= \frac{1}{n}\sum_{i=1}^{s} a_i \cdot H_n(a_i) = \sum_{i=1}^{s} a_i \cdot h_n(a_i) \\
\tilde{s} &= \frac{1}{n}\sum_{i=1}^{s} |a_i - x_{0,5}| \cdot H_n(a_i) = \sum_{i=1}^{s} |a_i - x_{0,5}| \cdot h_n(a_i) \\
s^2 &= \frac{1}{n}\sum_{i=1}^{s} (a_i - \overline{x})^2 \cdot H_n(a_i) = \sum_{i=1}^{s} (a_i - \overline{x})^2 \cdot h_n(a_i)
\end{aligned}$$

Beweis
Die Behauptungen folgen sofort aus den Definitionen durch Umordnen der Summanden sortiert nach den verschiedenen Ausprägungen. □

7.5 Vergleich von Stichproben

Einen schnellen visuellen Vergleich verschiedener Stichproben ermöglichen *Boxplots* (auch *Kistendiagramme* genannt). Hier werden Quantile und Streumaße zur graphischen Darstellung verschiedener Stichproben in geeigneter Weise herangezogen. Eine Variante, wie ein Boxplot erstellt werden kann, ist die folgende:

Lege eine geeignete Skala (für die darzustellenden Stichproben) fest und führe dann für jede Stichprobe die folgenden Schritte durch:

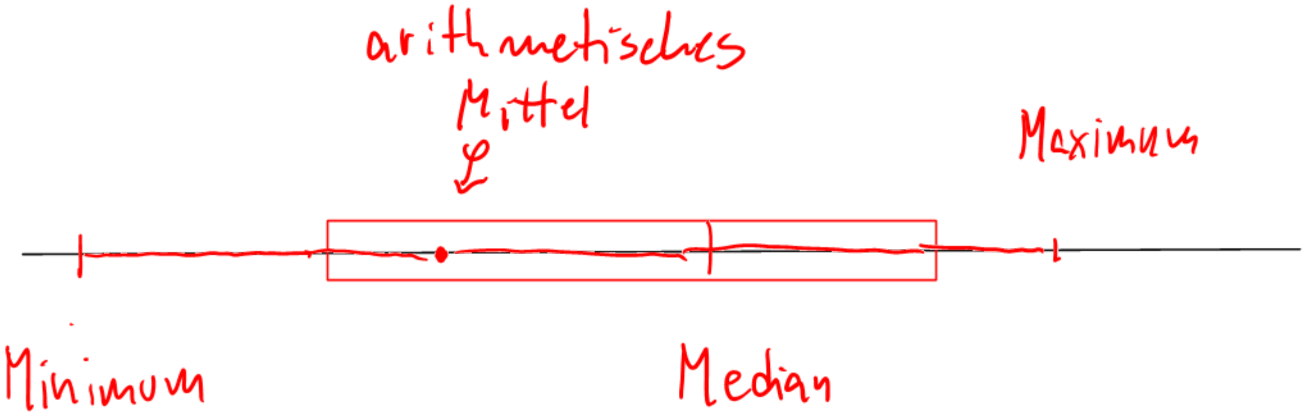
(i) Zeichne ein Rechteck ein, dessen eine Seitenlänge parallel zur Skala vom unteren Quartil bis zum oberen Quartil reicht. Die Länge der anderen Seite ist nach ästhetischen Gesichtspunkten zu wählen.

(ii) Unterteile das Rechteck an der Stelle des Medians senkrecht zur Skala in zwei Rechtecke.

(iii) Zeichne an der rechten und an der linken Seite des Rechtecks eine Strecke, die sogenannte *Antenne* bis zum größten bzw. kleinsten Wert der Stichprobe. Der größte und der kleinste Wert werden als Abschluss der Antennen durch einen kleinen senkrechten Strich oder einen Kreis gekennzeichnet.

Mitunter wird das arithmetische Mittel noch durch einen kleinen Vollkreis auf der mittleren Höhe des Rechtecks eingezeichnet. Varianten sehen vor, Ausreißer besonders zu kennzeichnen.

Einschub 7.5.1. ...



7.6 Grundlagen der Wahrscheinlichkeitstheorie

Die Stochastik ist paradigmatisch für den *Modellbildungskreislauf*. Hier sollen vom Zufall abhängige Vorgänge (wie z.B. Glücksspiele) modelliert werden. Der Zufall soll mit dem Begriff der Wahrscheinlichkeit erfasst werden. Als erste Näherung kann man den *frequentistischen Wahrscheinlichkeitsbegriff* heranziehen, der auf Richard von Mises (1883-1953) zurückgeht. Ihm liegt die Erfahrungstatsache zugrunde, dass sich die relative Häufigkeit eines Ergebnisses in einem vom Zufall abhängigen Experiment mit wachsender Anzahl an Wiederholungen des Experimentes stabilisiert (*Empirisches Gesetz großer Zahlen*). Den Grenzwert, gegen den die relative Häufigkeit mit wachsender Wiederholungszahl strebt, interpretiert man dann als *Wahrscheinlichkeit* des Ergebnisses. Dieser Erfahrungstatsache folgend kann man relative Häufigkeiten als Schätzwerte für Wahrscheinlichkeiten nehmen. Eine rigorose Definition des Wahrscheinlichkeitsbegriffs gelang von Mises nicht, da eine mathematische Präzisierung des Stabilisierens der relativen Häufigkeiten fehlschlug. Aus diesem Grund wählt man heute den auf Andrei Nikolajewitsch Kolmogorow (1903-1983) zurückgehenden *axiomatischen* Zugang zur Wahrscheinlichkeitstheorie.

Thema der Stochastik sind Ereignisse (und die Wahrscheinlichkeit ihres Eintretens. Unsere Ausgangsfragen sind:

(i) Was kann passieren?

(ii) Mit welcher Wahrscheinlichkeit treten diese oder jene Ereignisse ein?

Unabhängig von der konkreten betrachteten Situation wird die *Menge aller möglichen Ereignisse* (also die Antwort auf die erste Frage) in der Wahrscheinlichkeitstheorie mit Ω bezeichnet.

Beispiele 7.6.1. Wir betrachten einige Situationen und modellieren das zugehörige Ω:

(i) *Ein Münzwurf.* Eine Münze wird geworfen, und als Ereignis tritt entweder "Kopf" oder "Zahl" ein. Wenn wir zur Vereinfachung das eine mit "0" notieren und das andere mit "1", ist die Menge aller möglichen Ereignisse in diesem Fall $\Omega = \{0, 1\}$.

(ii) *Würfelwurf.* $\Omega = \{1, 2, 3, 4, 5, 6\}$.

(iii) *Mehrere Münzwürfe.* In dieser Situation fassen wir jede n-elementige Folge von Nullen und Einsen (alias Köpfen und Zahlen) als ein Ereignis auf; es ist also

$$\Omega = \{(x_1, x_2, \dots, x_n) \mid x_i \in \{0,1\}\} =: \{0,1\}^n = \underbrace{\{0,1\} \times \{0,1\} \times \dots \times \{0,1\}}_{n\text{-mal}}$$

die Menge aller möglichen Ereignisse. Beim zweifachen ~~Münzwurf~~ also Würfelwurf

$$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

D.h. das Element $(2, 3) \in \Omega$ bedeutet: erst wurde eine 2 und dann eine 3 gewürfelt.

(iv) *Unendlich viele Münzwürfe.* Hier gilt

Folge

$$\Omega = \{(x_i)_{i \in \mathbb{N}} | x_i \in \{0, 1\}\} .$$

Im Gegensatz zum einfachen oder n-fachen Münzwurf ist Ω hier nicht endlich.

(v) *Zufallszahl zwischen* 0 *und* 1. $\Omega = [0, 1]$.

Die Fokussierung auf spezielle Fragestellungen erfolgt durch die Betrachtung sogenannter *Ereignisse*. Ein Ereignis A ist eine Teilmenge des Grundraums Ω, also $A \subset \Omega$.

Ein *Elementarereignis* ist ein Element $\omega \in \Omega$.

Man sagt „das Ereignis A tritt ein", wenn ein ω mit $\omega \in A$ beobachtet wird.

Da im Allgemeinen nicht jede Teilmenge von Ω als Ereignis sinnvoll sein muss (z.B. im Modell für den unendlichen Münzwurf), werden *gewisse* Teilmengen als Ereignisse ausgezeichnet, indem sie in einem *Ereignis-System*

$$\mathcal{A} := \{A \subseteq \Omega \,|\, A \text{ ist ein Ereignis }\} \subseteq \mathcal{P}(\Omega).$$

aufgelistet werden. Hier ist $\mathcal{P}(\Omega)$ die Potenzmenge von Ω. Wir verlangen, dass ein Ereignissystem $\mathcal{A}$ *abgeschlossen* gegenüber allen abzählbaren Mengenoperationen (etwa Durchschnitt, Vereinigung, Komplement, Mengedifferenzen,...) ist, d.h. führt man diese Operationen mit Ereignissen aus $\mathcal{A}$ durch, so erhält man wieder Ereignisse aus $\mathcal{A}$.

In den meisten Fällen wählen wir $\mathcal{A} = \mathcal{P}(\Omega)$.

Den Zusammenhang zwischen der Beschreibung von Ereignissen in der realen Welt und im Modell zeigt die folgende Tabelle:

Reale Welt	**Modell**
A und B treten (gleichzeitig) ein	$\omega \in A \cap B$
A oder B tritt ein (oder beide)	$\omega \in A \cup B$
Entweder A oder B tritt ein	$\omega \in (A \setminus B) \cup (B \setminus A)$
B tritt ein, A aber nicht	$\omega \in B \setminus A$
A tritt nicht ein	$\omega \in A^C = \Omega \setminus A$
A und B schließen sich aus	$A \cap B = \emptyset$ (A und B sind *disjunkt*.)
Mindestens ein A_i tritt ein	$\omega \in \bigcup_{i \geq 1} A_i$
Alle A_i treten ein	$\omega \in \bigcap_{i \geq 1} A_i$

Beispiele 7.6.2. Wir betrachten Situationen aus dem obigen Beispiel.

(i) *Ein Münzwurf.* "1 tritt ein": $A = \{1\}$.

(ii) *Würfelwurf.* $\Omega = \{1, 2, 3, 4, 5, 6\}$. Es wird eine gerade Zahl gewürfelt $A = \{2, 4, 6\}$.

(iii) *Mehrere Münzwürfe.* "Es treten genau k Einsen auf":

$$A = \left\{ (x_1, \ldots, x_n) \in \{0,1\}^n \,\middle|\, \sum_{i=1}^{n} x_i = k \right\}.$$

(iv) Unendlich viele Münzwürfe. "Die relative Häufigkeit der 1 ist p":

$$A = \{(x_i)_{i \in \mathbb{N}} \in \{0,1\}^{\mathbb{N}} | \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} x_i = p\}.$$

(v) *Zufallszahl zwischen* 0 *und* 1. "Es tritt eine Zahl aus $[a, b]$ auf": $A = [a, b] \subseteq \Omega = [0, 1]$.

7.7 Zufallsexperimente und Wahrscheinlichkeitsräume

Definition 7.7.1. Ein Paar (Ω, p) heißt *Zufallsexperiment*, falls

$\mathbb{N}, \mathbb{Q}, \mathbb{Z}$

(i) Ω eine *abzählbare* (d.h. endliche oder abzählbar unendliche), nicht-leere Menge (*Ergebnismenge, Grundraum, Stichprobenraum*) und

(ii) $p : \Omega \to [0, 1]$ eine Funktion mit

$$\sum_{\omega \in \Omega} p(\omega) = 1 \qquad (\textit{Wahrscheinlichkeitsfunktion, Wahrscheinlichkeitsdichte})$$

ist.

Ein Element $\omega \in \Omega$ heißt *Ergebnis bzw.Elementarereignis* und $p(\omega)$ heißt *Wahrscheinlichkeit* des Ergebnisses bzw. desElementarereignisses ω.

Beispiele 7.7.2. 1) Fairer Münzwurf : $\Omega = \{0,1\}$

$p(0) := \frac{1}{2}$, $p(1) := \frac{1}{2}$ $\Rightarrow$ $1 = p(0) + p(1)$

2) Gezinkte Münze : $q \in [0,1]$, $q \neq \frac{1}{2}$, $\Omega = \{0,1\}$

$p(0) := q$, $p(1) = 1-q$, $1 = p(0) + p(1) \left(= q + (1-q)\right)$

3) Fairer Würfelwurf : $\Omega = \{1,2,\dots,6\}$, $p(\omega) := \frac{1}{6}$

$$\sum_{\omega \in \Omega} p(\omega) = 6 \cdot \frac{1}{6} = \underline{1}$$

4) 10-facher Münzwurf: $\Omega = \{0,1\}^{10}$

$\underline{\text{zB}}$: $(1,0,1,1,0,0,0,1,0,1) \in \Omega$, $p(\omega) := \frac{1}{2^{10}}$ Kapitel 1

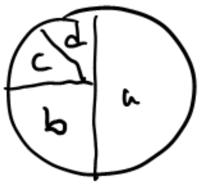
$$\sum_{\omega \in \Omega} p(\omega) = \sum_{\omega \in \Omega} \frac{1}{2^{10}} = 2^{10} \cdot \frac{1}{2^{10}} = 1 \quad \text{denn } |\Omega| = 2^{10}$$

5) $\Omega = \{0,1\} \times \{1,2,\dots,6\}$ zweistufiges Experiment aus fairem Münzwurf und fairem Würfelwurf, $p(\omega) := \frac{1}{2 \cdot 6} = \frac{1}{12}$

6) $\Omega = \{a,b,c,d\}$

ω	a	b	c	d
$p(\omega)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

$\underline{\text{zB}}$



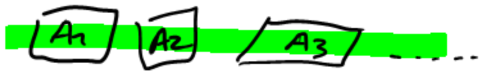
Glücksrad

Definition 7.7.3. Ein Tripel $(\Omega, \mathcal{A}, P)$ heißt *Wahrscheinlichkeitsraum*, falls gilt:

(1) Ω ist eine nicht-leere Menge (*Ergebnismenge*)

(2) $\mathcal{A}$ ist ein Ereignis-System (bei uns in der Regel $\mathcal{A} = \mathcal{P}(\Omega)$)

(3) P ist ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, d.h. eine Funktion $P : \mathcal{A} \to [0,1]$ mit folgenden Eigenschaften:



(N) *Normiertheit*: $P(\Omega) = 1$

(A) *σ-Additivität*:
Für paarweise disjunkte Ereignisse $A_1, A_2, \ldots \in \mathcal{A}$ (d.h. $A_i \cap A_j = \emptyset$ für $i \neq j$) gilt:

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Man sagt, dass $P(A)$ die *Wahrscheinlichkeit* von A ist (bzw. die Wahrscheinlichkeit ist, dass A eintritt).

Satz 7.7.4. *Für ein Zufallsexperiment (Ω, p) wird durch $\mathcal{A} := \mathcal{P}(\Omega)$ und*

$$P(A) := \sum_{\omega\in A} p(\omega) \qquad \textit{für} \quad A \in \mathcal{A}$$

ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ definiert.

Satz 7.7.5. *(Eigenschaften von Wahrscheinlichkeitsmaßen) Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Dann gelten:*

2) $\cancel{(1)}$ Endliche Additivität*: Für paarweise disjunkte Ereignisse $A_1, A_2, \ldots A_n \in \mathcal{A}$ gilt:*

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

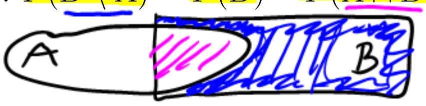
$\emptyset \cap \emptyset = \emptyset$

1) $\cancel{(2)}$ $P(\emptyset) = 0$ denn $P(\emptyset) = P(\emptyset \cup \emptyset) \overset{(A)}{=} P(\emptyset) + P(\emptyset) \Rightarrow P(\emptyset) = 0$

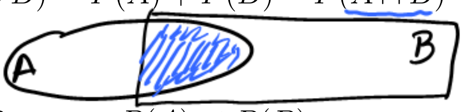
(3) Für alle $A \in \mathcal{A}$ gilt: $P(A^C) = 1 - P(A)$

$1 = P(\Omega) = P(\Omega \setminus A \cup A) \overset{(A)}{=} P(\overbrace{\Omega\setminus A}^{=A^C}) + P(A)$

(4) Für alle $A, B \in \mathcal{A}$ gilt: $P(B \setminus A) = P(B) - P(A \cap B)$



(5) Für alle $A, B \in \mathcal{A}$ gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



(6) Für alle $A, B \in \mathcal{A}$ gilt: $A \subset B \quad \Rightarrow \quad P(A) \leq P(B)$

(7) Endliche Subadditivität*: Für $A_1, \ldots, A_n \in \mathcal{A}$ gilt:*

müssen nicht disjunkt sein

$$P(A_1 \cup \ldots \cup A_n) \leq P(A_1) + \ldots + P(A_n)$$

(8) Subadditiviät*: Für $A_1, A_2, \ldots \in \mathcal{A}$ gilt: $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$*

Definition 7.7.6. Ein Zufallsexperiment (Ω, p) heißt *Laplace-Versuch* (oder *Laplace-Experiment*), wenn Ω endlich und $p : \Omega \to [0,1]$ die sogenannte *Gleichverteilung* (oder *Laplace'sche Wahrscheinlichkeit(sfunktion)*) ist, die jedem Ergebnis dieselbe Wahrscheinlichkeit, nämlich

$$\boxed{p(\omega) = \frac{1}{|\Omega|},}$$

zuordnet, wobei $|\Omega|$ gleich der Anzahl der Elemente in Ω bezeichnet. Der zugehörige Wahrscheinlichkeitsraum $(\Omega, \mathcal{P}(\Omega), P)$ heißt *Laplace'scher Wahrscheinlichkeitsraum* (oder *Laplace-Modell*). Im Laplace-Modell gilt für ein Ereignis $A \subset \Omega$:

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

aus Satz 7.7.4

In Laplace-Modellen ermittelt man Wahrscheinlichkeiten häufig durch geschicktes Abzählen.

Einschub 7.7.7. ... Im Bsp 7.7.2

Laplace 1), 3), 4), 5)   nicht Laplace 2) ~~5)~~ 6)

7.8 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Offenbar ist es möglich, dass durch das Eintreten eines Ereignisses B die Wahrscheinlichkeit eines Ereignisses A verändert werden kann. Gesucht ist in diesem Zusammenhang die Wahrscheinlichkeit von A unter der Bedingung, dass B eingetreten ist. Um dies zu modellieren führen wir bedingte Wahrscheinlichkeiten ein. Die bedingte Wahrscheinlichkeit kann als Neueinschätzung der Wahrscheinlichkeit von A interpretiert werden, wenn die Information vorliegt, dass das Ereignis B bereits eingetreten ist.

Definition 7.8.1. Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $B \in \mathcal{A}$ mit $P(B) > 0$. Dann heißt

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

die *bedingte Wahrscheinlichkeit von A gegeben B*.

Beispiele 7.8.2. Fairer Würfelwurf. $\Omega = \{1, \dots, 6\}$, $A := \{2,4,6\} = \{$"gerade"$\}$

$B := \{2,3,5\} = \{$"prim"$\}$, $P(A|B) = \frac{P(\{2\})}{P(B)} = \frac{1/6}{3/6} = \frac{1}{3}$

Ist B eingetreten, bleiben nur noch 3 gleich wahrscheinliche Ergebnisse übrig.

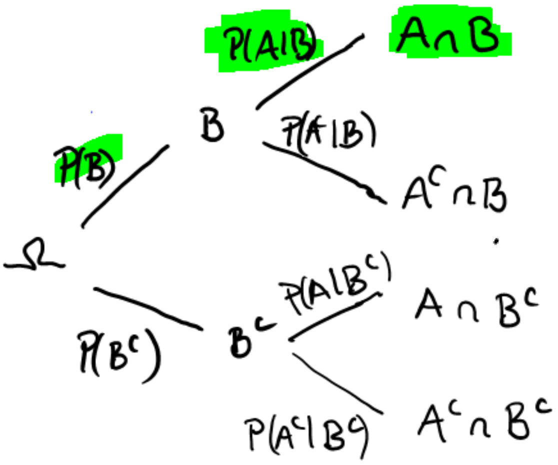
Bedingte Wahrscheinlichkeiten kann man wie folgt mithilfe von *Wahrscheinlichkeitsbäumen* visualisieren:

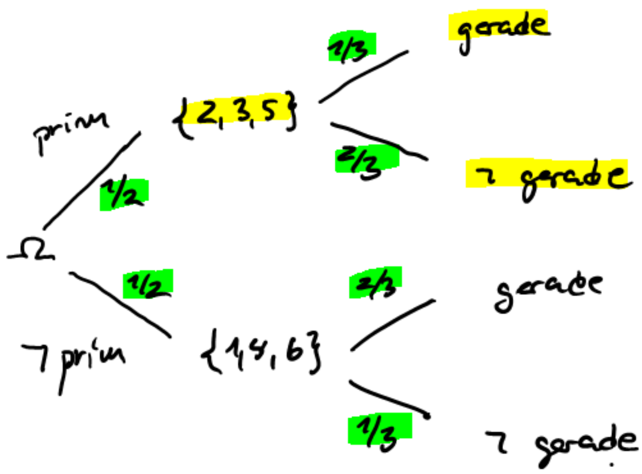
Einschub 7.8.3. ...

$$P(A|B) \overset{\text{Def}}{=} \frac{P(A\cap B)}{P(B)}$$

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

Im Beispiel oben





Für das Rechnen mit bedingten Wahrscheinlichkeiten gelten die folgenden Sätze:

Satz 7.8.4. *(Multiplikations- oder Pfadregel)*
Sei $(\Omega, \mathcal{A}, P)$ *ein Wahrscheinlichkeitsraum und* $A_1, ..., A_n \in \mathcal{A}$ *mit* $P(A_1 \cap ... \cap A_{n-1}) > 0$. *Dann gilt:*

$$P(A_1 \cap ... \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot ... \cdot P(A_n|A_1 \cap ... \cap A_{n-1}).$$

Einschub 7.8.5. ...

Satz 7.8.6. *(von der totalen Wahrscheinlichkeit)*
Sei $(\Omega, \mathcal{A}, P)$ *ein W.raum,* $A \in \mathcal{A}$ *und* $B_i \in \mathcal{A}$, $i \geq 1$, *eine* Partition *von* Ω, *d.h. es gelten:*

(i) $B_i \cap B_j = \emptyset$ *für alle* $i \neq j$,

(ii) $\bigcup_{i\geq 1} B_i = \Omega$.

Ist weiterhin $P(B_i) > 0$ *für alle* $i \geq 1$, *so gilt:*

$$P(A) = \sum_{i\geq 1} P(A|B_i) \cdot P(B_i).$$

Einschub 7.8.7. ...

Beispiele 7.8.8. Satz von der totalen W.keit und Pfadregel im Spezialfall $\Omega = B \cup B^C$.

Beispiele 7.8.9. In einer Urne befinden sich zwei blaue und zwei rote Kugeln. Es wird zweimal aus gezogen. Wird im ersten Zug eine rote Kugel gezogen, so wird sie wieder zurückgelegt. Dann wird ein zweites Mal gezogen. Wird im ersten Zug eine blaue Kugel gezogen, so wird sie nicht zurückgelegt und dann ein zweites Mal gezogen.

Satz 7.8.10. *(von Bayes) Sei $(\Omega, \mathcal{A}, P)$ ein W.raum und $B_i \in \mathcal{A}$, $i \geq 1$ eine Partition von Ω mit $P(B_i) > 0$ für alle $i \geq 1$. Dann gilt für alle $j = 1, 2, ...$ und alle $A \in \mathcal{A}$:*

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i\geq 1} P(A|B_i) \cdot P(B_i)}.$$

Bemerkung 7.8.11. Speziell für die Partition (B, B^C) bedeutet der Satz von Bayes:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + P(A|B^C) \cdot P(B^C)}$$

Beispiele 7.8.12. In Deutschland sei eine Person von 1000 an einer bestimmten Krankheit erkrankt. Zur Diagnose der Krankheit wird ein Test verwendet, der in 99% der Fälle, in denen die Krankheit vorliegt, ein positives Testergebnis liefert. Bei einer nicht-erkrankten Person gibt der Test allerdings in 2% der Fälle fälschlicherweise ein positives Testergebnis aus.
Mit welcher Wahrscheinlichkeit ist eine Person, die ein positives Testergebnis erhalten hat, tatsächlich krank?

Bedingte Wahrscheinlichkeiten kann man wie folgt mithilfe von *Wahrscheinlichkeitsbäumen* visualisieren:

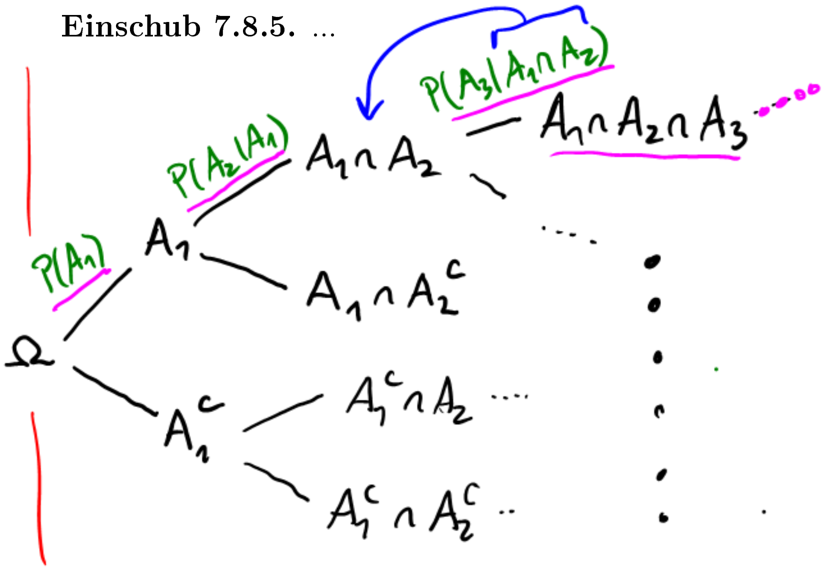
Einschub 7.8.3. ...

Für das Rechnen mit bedingten Wahrscheinlichkeiten gelten die folgenden Sätze:

Satz 7.8.4. *(Multiplikations- oder Pfadregel)*
Sei $(\Omega, \mathcal{A}, P)$ *ein Wahrscheinlichkeitsraum und* $A_1, ..., A_n \in \mathcal{A}$ *mit* $P(A_1 \cap ... \cap A_{n-1}) > 0$. *Dann gilt:*

$$P(A_1 \cap ... \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot ... \cdot P(A_n|A_1 \cap ... \cap A_{n-1}).$$

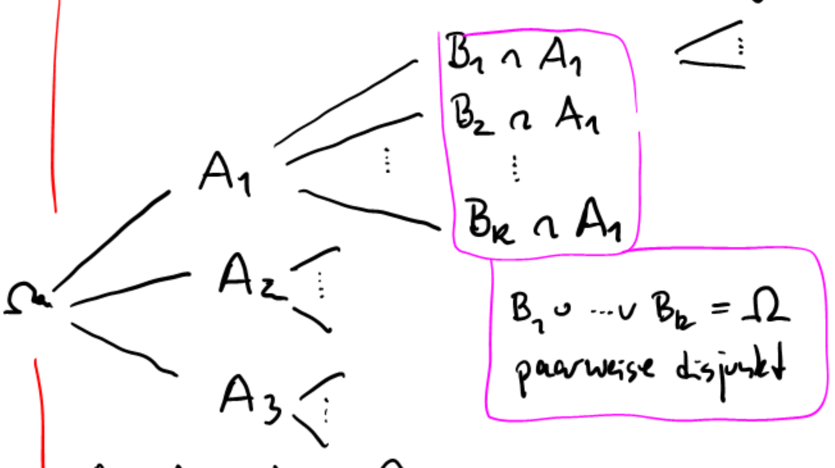
Einschub 7.8.5. ...



Beispiel von oben

$$P(\text{gerade } \underline{\text{und}} \text{ prim}) = P(\text{gerade} \mid \text{prim}) \cdot P(\text{prim}) = 1/3 \cdot 1/2$$

Einschub 7.8.6. ... „Mehrere Äste möglich"



$A_1 \cup A_2 \cup A_3 = \Omega$

$A_i \cap A_j = \emptyset$ (paarweise disjunkt)

Beispiele 7.8.7. (Mehrstufiges Experiment und Pfadregel) Wir werfen zweimal hintereinander einen fairen 3-seitigen Würfel. Dazu sei

$$\Omega = \{1,2,3\} \times \{1,2,3\}$$

$|\Omega| = 9$

die Ergebnismenge und $p(\omega_1, \omega_2) = \frac{1}{9}$ das zugehörige Laplace-Experiment. Das Ereignis $A \subset \Omega$ im ersten Wurf eine gerade Zahl zu werfen lautet

$$A = \{(2, s) \mid s \in \{1,2,3\}\} \subset \Omega.$$

$A^C = \{(1,1), (1,2), (1,3), (3,1), (3,2), (3,3)\}$

Das Ereignis $B \subset \Omega$ im zweiten Wurf eine gerade Zahl zu werfen lautet

$$B = \{(s, 2) \mid s \in \{1,2,3\}\} \subset \Omega.$$

Das Ereignis im zweiten Wurf eine gerade und im ersten eine ungerade Zahl zu werfen lautet

$$\{(1,2),(3,2)\} = B \cap A^C.$$

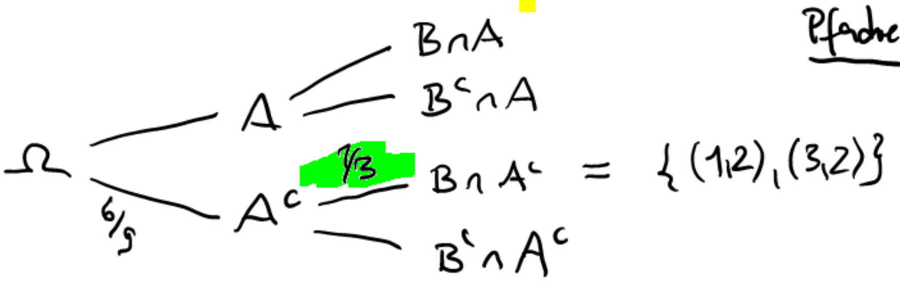
Also

$$\frac{2}{9} = P(B \cap A^C) \overset{\boxed{\text{Dfn}}}{=} P(B|A^C) \cdot P(A^C) = P(B|A^C) \cdot \frac{6}{9}$$

und somit

$$P(B|A^C) = \frac{1}{3}.$$

Der zugehörige Wahrscheinlichkeitbaum sieht so aus



Pfadregel $P(A^c \cap B) =$

$= \frac{1}{3} \cdot \frac{6}{9} = \frac{2}{9}$

Umgekehrt kann man die Wahrscheinlichkeitsfunktion p bestimmen (eventuell kein Laplace-Experiment), falls man ausreichend viele bedingte Wahrscheinlichkeiten kennt (siehe Beispiel unten).

Satz 7.8.8. *(von der totalen Wahrscheinlichkeit)*
Sei $(\Omega, \mathcal{A}, P)$ ein W.raum, $A \in \mathcal{A}$ und $B_i \in \mathcal{A}$, $i \geq 1$, eine Partition von Ω, d.h. es gelten:

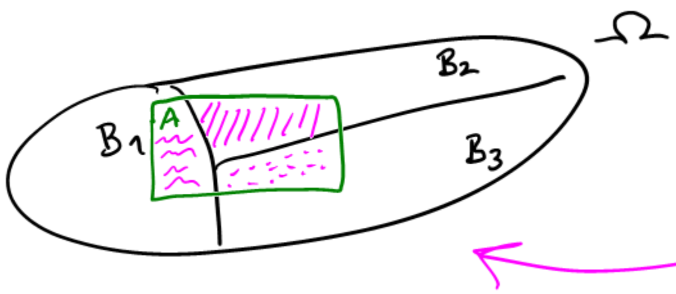
(i) $B_i \cap B_j = \emptyset$ *für alle* $i \neq j$,

(ii) $\bigcup_{i \geq 1} B_i = \Omega$.

Ist weiterhin $P(B_i) > 0$ für alle $i \geq 1$, so gilt:

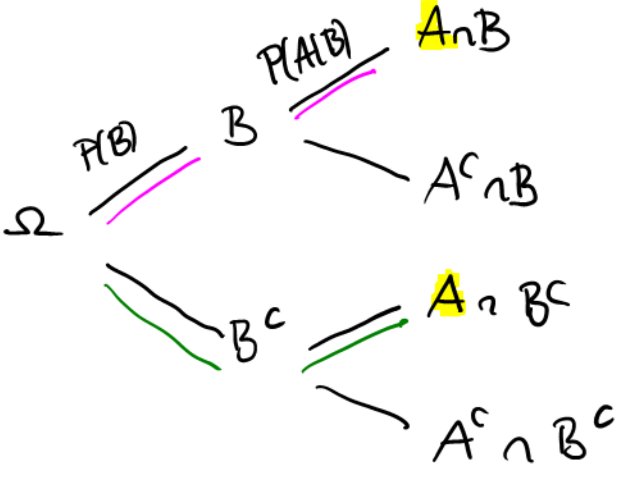
$$P(A) = \sum_{i \geq 1} \underbrace{P(A|B_i) \cdot P(B_i)}_{= P(A \cap B_i)}.$$

Einschub 7.8.9. ...



Beispiele 7.8.10. Satz von der totalen W.keit und Pfadregel im Spezialfall $\Omega = B \cup B^C$.

⚠ so merken



SvtW:

$$P(A) = P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c)$$

„Summe der Pfade zu A"

Beispiele 7.8.11. In einer Urne befinden sich zwei blaue und zwei rote Kugeln. Es wird zweimal aus gezogen. Wird im ersten Zug eine rote Kugel gezogen, so wird sie wieder zurückgelegt. Dann wird ein zweites Mal gezogen. Wird im ersten Zug eine blaue Kugel gezogen, so wird sie nicht zurückgelegt und dann ein zweites Mal gezogen.

Aufgabe: Bestimme Wkeit: im 2ten Zug rot; definieren Sie Ω und die Ereignisse formal

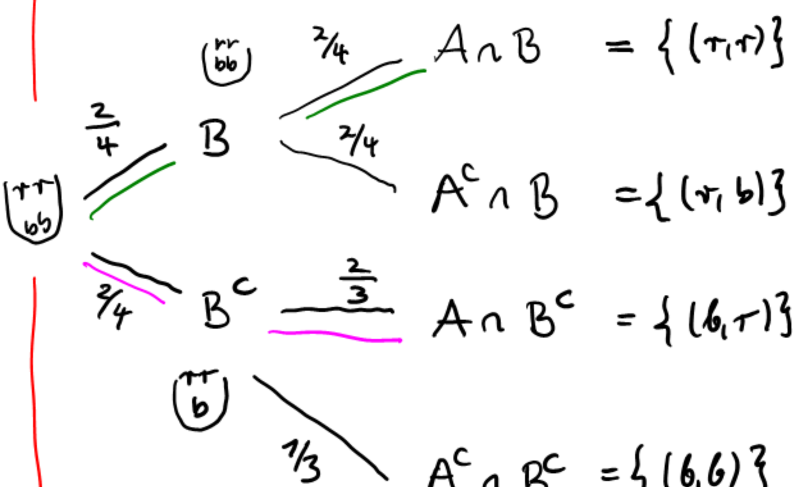
$$\Omega = \{(r,r), (r,b), (b,r), (b,b)\}$$

$$B := \{\text{im 1Z rot}\} = \{(r,r), (r,b)\} \subset \Omega$$

$$A := \{\text{im 2Z rot}\} = \{(r,r), (b,r)\} \subset \Omega$$

$$B^c := \{\text{im 1Z blau}\} = \{(b,r), (b,b)\} \subset \Omega$$

$$A^c := \{\text{im 2Z blau}\} = \{(r,b), (b,b)\} \subset \Omega$$



also $p : \Omega \to [0,1]$

$$p(r,r) = \frac{2}{4} \cdot \frac{2}{4} = \frac{1}{4}$$

$$p(r,b) = 1/4$$

$$p(b,r) = \frac{4}{12}$$

$$p(b,b) = \frac{2}{12}$$

$$P(A) = P(\text{„im 2ten Zug rot"}) \overset{\text{tW}}{=} \frac{2}{4} \cdot \frac{2}{4} + \frac{2}{4} \cdot \frac{2}{3} = \frac{7}{12}$$

Satz 7.8.12. *(von Bayes) Sei $(\Omega, \mathcal{A}, P)$ ein W.raum und $B_i \in \mathcal{A}$, $i \geq 1$ eine Partition von Ω mit $P(B_i) > 0$ für alle $i \geq 1$. Dann gilt für alle $j = 1, 2, \dots$ und alle $A \in \mathcal{A}$:*

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i\geq 1} P(A|B_i) \cdot P(B_i)}.$$

Bemerkung 7.8.13. Speziell für die Partition (B, B^C) bedeutet der Satz von Bayes:

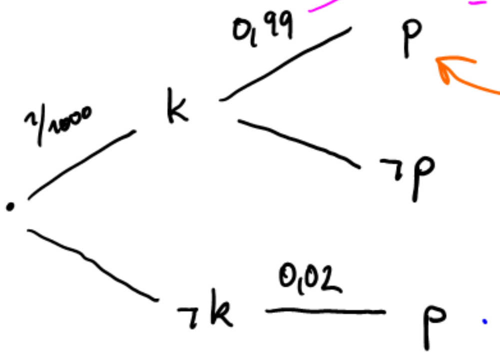
$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + P(A|B^C) \cdot P(B^C)}$$

Einschub 7.8.14. ...

Ist $P(A|B)$, $P(A|B^c)$ bekannt, kann man $P(B|A)$ ausrechnen

Beispiele 7.8.15. In Deutschland sei eine Person von 1000 an einer bestimmten Krankheit erkrankt. Zur Diagnose der Krankheit wird ein Test verwendet, der in 99% der Fälle, in denen die Krankheit vorliegt, ein positives Testergebnis liefert. Bei einer nicht-erkrankten Person gibt der Test allerdings in 2% der Fälle fälschlicherweise ein positives Testergebnis aus.
Mit welcher Wahrscheinlichkeit ist eine Person, die ein positives Testergebnis erhalten hat, tatsächlich krank?

$$P(\underbrace{\text{krank}}_{=:B} \mid \underbrace{\text{positiv}}_{=:A}) \underset{\text{SvB}}{=} \frac{P(p|k) \cdot P(k)}{P(p|k) \cdot P(k) + P(p|\neg k) P(\neg k)}$$



$0{,}99 = P(p|k)$

Sensitivität

$$= \frac{0{,}99 \cdot \frac{1}{1000}}{0.99 \cdot \frac{1}{1000} + 0.02 \cdot \frac{999}{1000}} \approx 0{,}0472$$

$\neg p$

Also etwa 4,7%

Spezifität des Tests

Definition 7.8.16. Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Zwei Ereignisse A und B aus $\mathcal{A}$ heißen *((P-)stochastisch))-unabhängig*, falls gilt:

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

Andernfalls heißen A und B *nicht unabhängig*. Allgemein heißt eine Familie $(A_i)_{i \in I}$, I eine beliebige Indexmenge, *((P-)stochastisch))-unab-hängig*, falls für jede endliche Teilfamilie $(A_i)_{i \in J}$ mit $J = \{i_1, ..., i_n\} \subset I$ gilt:

$$P(\bigcap_{i \in J} A_i) = P(A_{i_1} \cap ... \cap A_{i_n}) = P(A_{i_1}) \cdot ... \cdot P(A_{i_n}) = \prod_{i \in J} P(A_i).$$

Drei Ereignisse A_1, A_2, A_3 sind gemäß der Definition also unabhängig (hier $I = \{1, 2, 3\}$), falls gilt:

Einschub 7.8.17. ...

$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2) P(A_3)$ und $P(A_i \cap A_j) = P(A_i) P(A_j)$

für $i \neq j$, $i, j \in \{1, 2, 3\}$

Beispiele 7.8.18.

(i) im Beispiel 7.8.7 gelten

$$P(A) = \tfrac{3}{9}, \quad P(B) = \tfrac{3}{9}, \quad P(A \cap B) = P(\{(2,2)\}) = \tfrac{1}{9}.$$

Also sind A und B unabhängig.

(ii) Falls $P(B) > 0$ gilt, sind die Ereignisse A und B genau dann unabhängig, wenn $P(A|B) = P(A)$ gilt.

$$P(A|B) = P(A) \overset{\text{Dfn}}{\Leftrightarrow} \frac{P(A \cap B)}{P(B)} = P(A) \Leftrightarrow \underbrace{P(A \cap B)}_{=} = P(A) P(B)$$

$$\overset{\text{unab}}{\Leftrightarrow} P(A) P(B) = P(A) P(B)$$

(iii) Falls A und B unabhängig sind, so auch A^C und B.

$$P(A^c \cap B) = P(B \setminus A) \overset{\text{Satz}}{=} P(B) - P(A \cap B) \overset{\text{unab}}{=} P(B) - P(A) P(B) = P(B)(1 - P(A)) \overset{\text{Satz}}{=} P(B) P(A^c)$$

(iv) Hat A die Wahrscheinlichkeit 0 oder 1, so sind A und B für jede Wahl von B unabhängig.

$\underline{P(A) = 0}$ $\quad P(A \cap B) \overset{\text{Satz}}{\leq} P(A) = 0 \Rightarrow P(A \cap B) = 0 = \underbrace{P(A)}_{=0} P(B)$

$\underline{P(A) = 1}$ ... ohne Beweis.

(v) Aus der Unabhängigkeit der Ereignisse $A_1, ..., A_n$ folgt die paarweise Unabhängigkeit (d.h. je zwei der Ereignisse $A_1, ..., A_n$ sind unabhängig), aber nicht umgekehrt. p.w. unab $\not\Rightarrow$ alle unabh.

(vi) (Un-)Abhängigkeit meint hier stets stochastische (Un-)Abhängigkeit und nicht zwingend eine kausale (un-)Abhängigkeit (d.h. (Nicht-)Existenz eines Ursache-Wirkungszusammenhang).

Beispiele 7.8.19. Ein zentrales Beispiel für Unabhängigkeit ist der n-fache unabhängige Wurf einer Münze, die mit Wahrscheinlichkeit $p \in [0, 1]$ auf 'Kopf' und Wahrscheinlichkeit $1 - p$ auf 'Zahl' fällt.

Wir betrachten den Spezialfall $n = 3$:

Einschub 7.8.20. ...

Die Anzahl der Pfade mit genau k Einsen kann man zählen, diesen Wert wollen wir vorerst $\alpha_{k,n}$ nennen. Daraus ergibt sich

$$P(\text{Es fallen genau } k \text{ Einsen}) = \alpha_{k,n} \cdot p^k (1-p)^{n-k} =: b_{n,p}(k)$$

Durch $b_{n,p}$ ist eine Wahrscheinlichkeitsfunktion auf $\Omega := \{0, 1, ..., n\}$ definiert, die sogenannte *Binomialverteilung zu den Parametern n und p.*

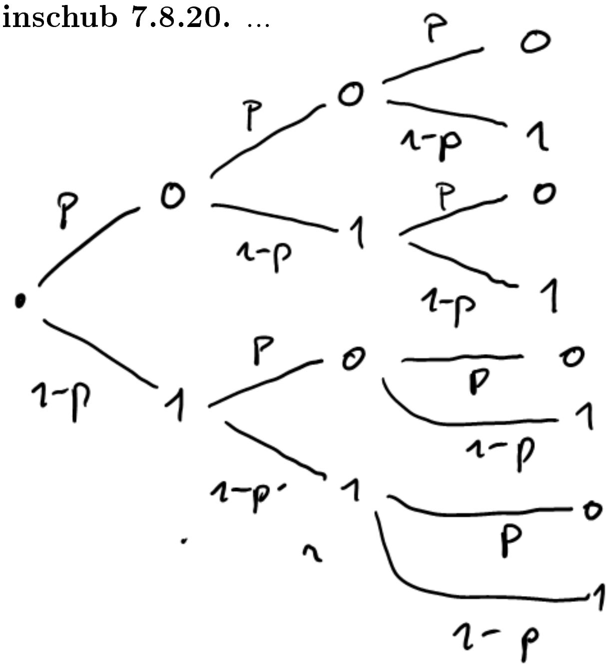
(v) Aus der Unabhängigkeit der Ereignisse $A_1, ..., A_n$ folgt die paarweise Unabhängigkeit (d.h. je zwei der Ereignisse $A_1, ..., A_n$ sind unabhängig), aber nicht umgekehrt.

(vi) (Un-)Abhängigkeit meint hier stets stochastische (Un-)Abhängigkeit und nicht zwingend eine kausale (un-)Abhängigkeit (d.h. (Nicht-)Existenz eines Ursache-Wirkungszusammenhang).

Beispiele 7.8.19. Ein zentrales Beispiel für Unabhängigkeit ist der n-fache unabhängige Wurf einer Münze, die mit Wahrscheinlichkeit $p \in [0,1]$ auf $\underbrace{\text{'Kopf'}}_{\hat{=}0}$ und Wahrscheinlichkeit $1-p$ auf $\underbrace{\text{'Zahl'}}_{\hat{=}1}$ fällt.

Wir betrachten den Spezialfall $n = 3$:

Einschub 7.8.20. ...



$\Omega_3 := \{0,1\}^3 \ni (0,1,0)$

allg. $\Omega_n := \{0,1\}^n$

Pfadregel $P(\underbrace{(x_1, \dots, x_n)}_{\in \Omega_n}) = p^{\text{Anzahl 0en}} \cdot (1-p)^{\text{Anzahl 1en}}$

$C \subset \Omega_3$, $C := \{\text{"2 Einsen"}\} = \{(0,1,1), (1,1,0), (1,0,1)\}$

$P(C) = p \cdot (1-p)^2 + p(1-p)^2 + p(1-p)^2 = \underbrace{3}_{\hat{=}\text{ Anzahl Pfade}} \cdot p(1-p)^2$

wobei $p =$ Wkeit für 1 fällt
$1-p =$ Wkeit für 0 fällt
wobei $\Omega_n := \{0,1\}^n$

hier anders herum (egal)

Die Anzahl der Pfade mit genau k Einsen kann man zählen, diesen Wert wollen wir vorerst $\alpha_{k,n}$ nennen. Daraus ergibt sich

$\alpha_{k,n} = \binom{n}{k}$ siehe Kapitel 7

$$P(\text{Es fallen genau } k \text{ Einsen}) = \alpha_{k,n} \cdot p^k (1-p)^{n-k} =: b_{n,p}(k)$$

Durch $b_{n,p}$ ist eine Wahrscheinlichkeitsfunktion auf $\Omega := \{0, 1, ..., n\}$ definiert, die sogenannte *Binomialverteilung zu den Parametern n und p.*

Einschub 7.8.21. ... Glücksrad

Wkeit bei 5 mal drehen 3 mal eine schwarz zu erhalten: $\alpha_{3,5} \cdot p^3 (1-p)^2$

7.9 Kombinatorik

In der Kombinatorik unterscheidet man bei der Auswahl von k aus n verschiedenen Objekten, ob

- *mit Beachtung der Reihenfolge* oder
- *ohne Beachtung der Reihenfolge* ausgewählt wird. $(1,2) \neq (2,1)$

Im ersten Fall erhält man eine *geordnete Menge*, die man als k-*Tupel* $(x_1, ..., x_k)$ notiert. Im zweiten Fall erhält man eine k-elementige Teilmenge $\{x_1, ..., x_k\}$. $\{1,2\} = \{2,1\}$

7.9.1 Allgemeines Zählprinzip

Es seien $M_1, ..., M_k$ endliche Mengen. Wählt man aus diesen Mengen nacheinander jeweils ein Element aus, so gibt es insgesamt

$$|M_1| \cdot |M_2| \cdot ... \cdot \overbrace{|M_k|}^{\text{Anzahl Elemente in } M_k}$$

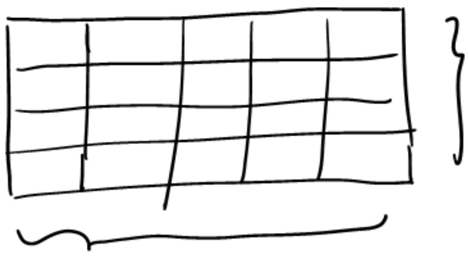
verschiedene Möglichkeiten, dies zu tun. Das Ergebnis dieser Auswahl wird mit

$$(x_1, ..., x_k) \in M_1 \times ... \times M_k$$

notiert und nennt man eine *Anordnung*, falls alle Mengen M_i gleich sind. Wir wissen bereits

$$|M_1 \times M_2 \times ... \times M_k| = |M_1| \cdot |M_2| \cdot ... \cdot |M_k|.$$

Beispiele 7.9.1. zum Abzählprinzip. Schokolade



n Stückchen

m Riegel

$M_1 := \{1, ..., n\}$

$M_2 := \{1, ..., m\}$

$\Rightarrow$ Anzahl Stücke $= n \cdot m = |M_1| \cdot |M_2|$

$= |M_1 \times M_2|$

Eine Anordnung von n verschiedenen Objekten heißt auch *Permutation*. Es gibt

$$n! := 1 \cdot 2 \cdot ... \cdot n$$

verschiedene Permutationen von n Objekten.

$3! = 1 \cdot 2 \cdot 3 = 6$

$4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$

$= 3! \cdot 4$

Wir wollen die Formel an folgendem Beispiel einsehen.

Beispiele 7.9.2. Zehn verschiedene Bücher sollen angeordnet werden.

M_1 = Menge 10 Bücher
M_2 = Menge 9 Bücher, nachdem eines aus M_1 genommen und ins Regal

M_3 = Menge 8 Büchern, ...

$\Rightarrow$ Anzahl der Anordnungen $= 10 \cdot 9 \cdot 8 \cdot 7 \cdot \ldots \cdot 1$
$= 10!$

(aus Binomialverteilung doch)

Eine k-elementige Teilmenge von einer n-elementigen Menge von Objekten wird auch *k-Kombination* genannt. Es gibt

$$d_{n,k} = \binom{n}{k} := \frac{n!}{k! \cdot (n-k)!} = \frac{n(n-1)...(n-k+1)}{k!}$$ „n über k"

verschiedene k-Kombinationen von n Objekten.

Wir wollen die Formel an folgendem Beispiel einsehen.

Beispiele 7.9.3. Aus 10 verschiedenen Büchern sollen 3 mit einem Griff ausgewählt/gezogen werden. Wie viele Ergebnisse sind möglich?

a) Wähle erst aus 10 B eines B_1 für Platz 1 aus : 10
Wähle dann aus 9 B —"— B_2 —"— 2 —"— : 9
—"— 8 B —"— B_3 —"— 3 —"— : 8

$\Rightarrow$ $10 \cdot 9 \cdot 8 = 720$ Möglichkeiten

b) (B_1, B_2, B_3), (B_2, B_1, B_3), ... liefert dasselbe Ergebnis (für die Aufgabe). Es gibt $3! = 6$ solche Anordnungen

c) Fazit es gibt $\frac{720}{6} = \frac{\overset{n}{10} \cdot \overset{n-1}{9} \cdot \overset{n-2}{8}}{3!}$ Ergebnisse
$k!$

allgemein $\frac{n\,(n-1) \cdot \ldots \cdot (n-k+1)}{k!}$

d) zum Gleichheitszeichen: Kosten

$$\frac{n!}{k!\,(n-k)!} \overset{\text{Hm.}}{=} \frac{1\cdot 2\cdot\ldots\cdot n}{(1\cdot 2\cdot\ldots\cdot k)\cdot(1\cdot 2\cdot\ldots\cdot(n-k))} = \frac{(n-k+1)\cdot\ldots\cdot n}{k!}$$

Definition 7.9.4. Die natürliche Zahl $\binom{n}{k}$ (sprich: „n über k”) heißt *Binomialkoeffizient*. Falls $k > n$ setzt man $\binom{n}{k} := 0$.

Satz 7.9.5. *(Binomischer Lehrsatz) Für alle $x, y \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt:*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Bemerkung 7.9.6.

$$(x+y)^2 = \binom{2}{0} x^0 y^{2-0} + \binom{2}{1} x^1 y^1 + \binom{2}{2} x^2 y^0$$

(i) Für $n = 2$ erhält man die übliche Binomische Formel.

(ii) Für $x = y = 1$ ergibt sich $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ und

(iii) für $x = -1, y = 1$ ergibt sich: $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0.$

Anzahl Teilmengen einer n-elementigen Menge ist 2^n

Beispiele 7.9.7. (zur Binomialverteilung)
Wir betrachten ein Hotel mit 200 Zimmern. Die Wahrscheinlickeit, dass ein Zimmer vor Anreise storniert wird sei 10%.

(i) Es seien alle Zimmer reserviert. Es sei Z, das Ereignis, dass weniger als 186 Zimmer vor Anreise nicht storniert werden. Dann gilt

$$P(Z) = \sum_{i=0}^{185} \binom{200}{i} \underbrace{\left(\frac{1}{10}\right)}_{\text{Wkeit für Stornierung}}^{200-i} \cdot \underbrace{\left(\frac{9}{10}\right)}_{\text{nicht storniert}}^{i}$$

(ii) Nun sollen mehr als 200 Reservierungen vorgenommen werden. Aber die Wahrscheilichkeit, dass das Hotel überbelegt ist soll maximal 1% betragen. Wie viele Reservierungen $n(\geq 200)$ dürfen maximal vorgenommen werden?

7.10 Urnenmodelle

Oftmals hilft beim strategischen Abzählen die Interpretation der Fragestellung als Ziehung aus einer Urne mit Kugeln. Hier sind im Wesentlichen vier Fälle zu unterscheiden:

Im Folgenden betrachten wir eine Urne mit n verschiedenen Kugeln. Es bezeichne k die Anzahl der Ziehungen von Kugeln aus der Urne.

i) Ziehen mit Zurücklegen und mit Beachtung der Reihenfolge:

Anzahl der möglichen Ergebnisse: n^k

Einschub 7.10.1. ... Würfel 3 mal werfen: $n = 6, \ k = 3$

$6 \cdot 6 \cdot 6 = 6^3 = 216$

ii) Ziehen ohne Zurücklegen und mit Beachtung der Reihenfolge:

Anzahl der möglichen Ergebnisse: $n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$

Spezialfall $k = n$:
Das Experiment liefert alle möglichen Anordnungen von n Elementen.

Anzahl möglicher Anordnungen von n Elementen: $n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!$

Einschub 7.10.2. ... 10 Bücher mit Platzierungen 1, 2, 3,

also $k = 3$, $n = 10$: $10 \cdot 9 \cdot 8 = 720$

„mit einem Griff" ziehen

iii) Ziehen **ohne** Zurücklegen und **ohne** Beachtung der Reihenfolge:

Anzahl der möglichen Ergebnisse: $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!}$

(= Anzahl der k-elementigen Teilmengen einer n-elementigen Menge)

gleichrangige

Einschub 7.10.3. ... Aus 10 Schüler:innen 3 Klassensprecher:innen wählen:

$$\binom{10}{3} = \frac{10!}{3!\,((10-3)!)} = \frac{10!}{3!\,7!} = \frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} = 4 \cdot 3 \cdot 10 = 120$$

iv) Ziehen **mit** Zurücklegen und **ohne** Beachtung der Reihenfolge:

Anzahl der möglichen Ergebnisse: $\binom{n+k-1}{k}$

Einschub 7.10.4. ... Aus 3 Eissorten 2 Kugeln wählen,

$k = 2$, $n = 3$ ergibt $\binom{3+2-1}{2} = \binom{4}{2}$

7.11 Hypergeometrische Verteilung

Sei

- N die Anzahl Kugeln in einer Urne, von denen

- M ($M \leq N$) eine bestimmte Eigenschaft $\mathcal{E}$ haben. Aus dieser Urne zieht man
- n ($n \leq N$) Kugeln.

Dann gilt:

$$P(\text{Anzahl der gezogenen Kugeln mit Eigenschaft } \mathcal{E} = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

Diese Verteilung heißt *hypergeometrische Verteilung zu den Parametern N, M und n.*

Beispiele 7.11.1. (Zahlenlotto) Berechnung der Wahrscheinlichkeit von k Richtigen im Lotto 6 aus 49, also $N = 49, M = 6$ und $n = 6$.

Vorstellung: getippte Zahlen sind rote eingefärbte Kugeln (Eigenschaft $\mathcal{E}$)

$$P(k \text{ richtige}) = P(k \text{ rote}) = \frac{\binom{6}{k} \cdot \binom{49-6}{6-k}}{\binom{49}{6}}$$

es müssen k Kugeln aus den 6 mit Eigenschaft $\mathcal{E}$ gezogen werden

(Anzahl = 49 − 6)

von denen ohne Eigenschaft $\mathcal{E}$ müssen 6−k gezogen werden

es gibt insgesamt $\binom{49}{6}$ 6-elementige Teilmengen von 1, ..., 49 Kugeln (also $\binom{49}{6}$ Ergebnisse)