- Gl. (7.1e): Um diese Form der Summenregel zu beweisen, schreiben wir die Ereignisse $a \sqcup b$ und b als Vereinigung ($\sqcup$) von sich gegenseitig ausschließenden Ereignissen

$$a \sqcup b = a \sqcup (\bar{a} \sqcap b); \qquad b = (a \sqcap b) \sqcup (\bar{a} \sqcap b) \qquad ,$$

auf die nun die Norm-Eigenschaft Gl. (6.6c) anwendbar ist

$$P(a \sqcup b) = P(a) + P(\bar{a} \sqcap b); \qquad P(b) = P(a \sqcap b) + P(\bar{a} \sqcap b) \qquad .$$

Wir lösen die zweite Gleichung nach $P(\bar{a} \sqcap b)$ auf und setzen das Ergebnis in die erste Gleichung ein und erhalten somit die gesuchte Summenregel für den Fall "überlappender" Ereignisse.

- Gl. (7.1f): Folgt leicht durch Induktion. Es gelte für $P(\sqcup_i^n a_i)$. Dann sind $\sqcup_i^n a_i$ und a_{n+1} sich gegenseitig ausschließende Ereignisse auf die Gl. (6.6c) anwendbar ist und somit gilt Gl. (7.1f) auch für $n+1$.

- Gl. (7.2a): Folgt sofort aus der Tatsache, dass $\mathcal{E} = \sqcup_i a_i$ und $P(\mathcal{E}) = 1$.

- Gl. (7.2b): Folgt ebenso aus der Tatsache, dass $\mathcal{E} = \sqcup_i a_i$ zusammen mit $b = b \sqcap \mathcal{E}$.

7.2 Bedingte Wahrscheinlichkeiten

Wir hatten bereits früher darauf hingewiesen, dass es nur bedingte Wahrscheinlichkeiten gibt. Im Rahmen der axiomatischen Wahrscheinlichkeitstheorie verbirgt sich der Bedingungskomplex hinter der Bedeutung und Struktur des Wahrscheinlichkeitsraumes, also der Angabe aller Experimente (Ω), der interessierenden Ereignisse (σ-Algebra) und der Prior-Wahrscheinlichkeiten. Insofern handelt es sich zwar auch hier um bedingte Wahrscheinlichkeiten, der Bedingungskomplex wird aber nicht explizit in die Notation aufgenommen, da er sich während der Analyse eines Problems nicht ändert. Das geht in den meisten Fällen gut, hat aber in der Literatur zu Paradoxa geführt, die Jahrzehnte lang unverstanden waren.
Neben der Abhängigkeit der Wahrscheinlichkeit vom Bedingungskomplex, kann man natürlich auch nach der Wahrscheinlichkeit $P(a|b)$ für ein Ereignis a fragen, unter der Annahme, dass das Ereignis b vorliegt. In Anlehnung an die Regeln der klassischen Wahrscheinlichkeitstheorie <u>definiert</u> man in der axiomatischen Theorie die

BEDINGTE WAHRSCHEINLICHKEIT

$$P(a|b) = \frac{P(a \sqcap b)}{P(b)} \qquad . \tag{7.3}$$

Man überzeugt sich leicht davon, dass bedingte Wahrscheinlichkeiten $P(a|x)$ ebenfalls alle Regeln der Wahrscheinlichkeitstheorie erfüllen; schließlich sind sie nichts anderes als Wahrscheinlichkeiten.
Aus der Definition der bedingten Wahrscheinlichkeit ergibt sich sofort die

PRODUKTREGEL

$$P(a \sqcap b) = P(a|b) * P(b) = P(b|a) * P(a) \qquad , \tag{7.4}$$

und durch Auflösen der beiden möglichen Darstellungen der Produktregel erhalten wir das Bayessche Theorem

BAYESSCHES THEOREM

$$P(a|b) = \frac{P(b|a) * P(a)}{P(b)} \qquad , \tag{7.5}$$

das es erlaubt, inverse Probleme zu lösen. Man interessiert sich in sehr vielen Fällen für die Posterior-Wahrscheinlichkeit $P(x|d)$ für ein Ereignis x, gegeben experimentelle Daten d und der Bedingungskomplex. Die Posterior-Wahrscheinlichkeit ist gemäß des Bayesschen Theorems verknüpft mit der Prior-Wahrscheinlichkeit $P(x)$, der sogenannte Daten-Evidenz $P(d)$[1] und der Likelihood-Funktion $P(d|x)$.
Man kann die Produkt-Regel verallgemeinern und in eine für die Anwendung sehr nützliche Gestalt bringen. Es sei $[a_1, \ldots, a_n]$ eine Partitionierung der Gesamtmenge Ω bzw. des sicheren Ereignisses und b ein beliebiges Element der σ-Algebra. Dann gilt wegen der Marginalisierungsregel Gl. (7.2b)

MARGINALISIERUNGSREGEL

$$P(b) = \sum_{i=1}^{n} P(a_i \sqcap b) = \sum_{i=1}^{n} P(b|a_i)\ P(a_i) \qquad . \tag{7.6}$$

Damit lässt sich das Bayessche Theorem auch in folgender Form schreiben:

[1]Die Daten-Evidenz spielt i.d.R. nur die Rolle des Normierungsfaktors.

$$P(a_i|b) = \frac{P(b|a_i)P(a_i)}{\sum\limits_j P(b|a_j)\ P(a_j)} \tag{7.7}$$

Def. 7.2 (Unabhängigkeit) *Man nennt zwei Ereignisse der σ-Algebra logisch unabhängig, wenn*

$$P(a \sqcap b) = P(a)P(b) \qquad \Longleftrightarrow \qquad P(a|b) = P(a)$$