Satz von Bayes

Satz von Bayes

$A_1, \dots, A_K$ sei eine disjunkte Zerlegung von Ω mit $P(A_i) > 0$ für alle $i = 1, \dots, K$. Dann gilt für jedes Ereignis B mit $P(B) > 0$

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{K} P(B|A_k)P(A_k)}.$$

Diese Formel gilt sinngemäß auch für den Fall $K = \infty$.

Beispiel: Spam-Filter

Motivation

Ein Spam-Filter verschiebt E-Mails in den junk-Ordner, wenn gewisse Worte in der E-Mail vorhanden sind, z.B. *win*.
Durch Analysieren von alten E-Mails kann man die bedingten Wahrscheinlichkeiten der Form

$$P(\text{,,E-Mail enthält } \textit{Uni}\text{''} \,|\, \text{,,Email ist Spam''})$$

etc. gut schätzen.
Fragen:

1. Wie wahrscheinlich ist es, dass eine E-Mail Spam ist?
2. Wie groß ist die Wahrscheinlichkeit, dass eine E-Mail tatsächlich Spam ist, wenn das Wort *win* vorkommt?

Beispiel: Spam-Filter

Systematisch: Ereignisse definieren:

$$A = \text{„E-Mail ist Spam“},$$
$$B_1 = \text{„E-Mail enthält das Wort } \textit{Uni}\text{“},$$
$$B_2 = \text{„E-Mail enthält das Wort } \textit{win}\text{“}.$$

Bekannt seien: $P(A), P(B_1|A), P(B_1|\overline{A}), P(B_2|A)$ und $P(B_2|\overline{A})$.

1. Kann man hieraus

   $$P(B_i), \qquad i = 1, 2$$

   berechnen?
2. Kann man hieraus

   $$P(A|B_i)$$

   berechnen?

Beispiel: Spam-Filter

Mehrstufige Wahrscheinlichkeitsmodelle

Mehrstufige Zufallsexperimente, insbesondere n-malige Wiederholung

$$\Omega = \Omega_1 \times \cdots \times \Omega_n$$

Festlegung der Wahrscheinlichkeiten

$$p(\omega) = P(\{\omega\}) = ?$$

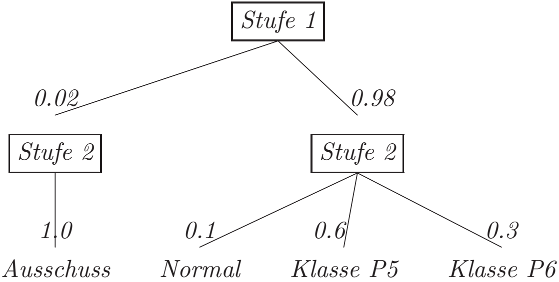
Mehrstufige Wahrscheinlichkeitsmodelle

Fallgestaltung: Produktion von Nadellagern bestehe aus zwei Stufen:

- Stufe 1: Vorbereitende Bearbeitung eines Rohling.
  Mit Wkeit 0.02 genügt ein Rohling nach Stufe 1 nicht den Qualitätsanforderungen.
- Stufe 2: Nachbearbeitung
  Entsprechend der Toleranzen Sortierung in drei Klassen (Normal/P5/P6).

Mit welcher Wkeit erhält man ein Nadellager der Klasse P5?

Mehrstufige Wahrscheinlichkeitsmodelle



Mehrstufige Wahrscheinlichkeitsmodelle

Oft: An verschiedenen Zeitpunkten bestimmen zufällige Ereignisse den Folgezustand.
Darstellung durch **Wahrscheinlichkeitsbaum**.
Modell:

$$\Omega = \Omega_1 \times \cdots \times \Omega_n$$

Startverteilung:

$$p(\omega_1), \qquad \omega_1 \in \Omega_1.$$

Bedingte Wahrscheinlichkeiten:

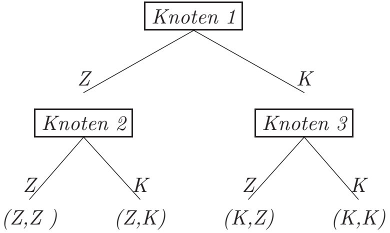
$$p(\omega_j | \omega_1, \dots, \omega_{j-1})$$

Pfadregel für $\omega = (\omega_1, \dots, \omega_n)$:

$$P(\{\omega\}) = p(\omega_1) p(\omega_2 | \omega_1) \cdots p(\omega_n | \omega_1, \dots, \omega_{n-1})$$

Mehrstufige Wahrscheinlichkeitsmodelle

Beispiel 2.3.3 *Eine faire Münze mit Kopf (K) und Zahl (Z) wird zweimal geworfen. Wir können auch dieses Zufallsexperiment als Wahrscheinlichkeitsbaum repräsentieren:*



Unabhängige Ereignisse

Heuristik: B nicht informativ für A, wenn $P(A|B) = P(A)$.

Unabhängige Ereignisse

Zwei Ereignisse A und B heißen **stochastisch unabhängig** (kurz: unabhängig), wenn

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

gilt. Diese Identität wird als **Produktsatz** bezeichnet.

Produktsatz

k Ereignisse $A_1, \dots, A_k \subset \Omega$ erfüllen den **Produktsatz**, wenn gilt:

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1) \cdot \ldots \cdot P(A_k).$$

Totale und paarweise Unabhängigkeit

$A_1, \dots, A_n \subset \Omega$ heißen **(total) stochastisch unabhängig**, wenn für jede Teilauswahl $A_{i_1}, \dots, A_{i_k}$ von $k \in \mathbb{N}$ Ereignissen der Produktsatz gilt.
$A_1, \dots, A_n$ heißen **paarweise stochastisch unabhängig**, wenn alle Paare A_i, A_j $(i \neq j)$ stochastisch unabhängig sind.

Totale und paarweise Unabhängigkeit

- $A_1, \dots, A_n \subset \Omega$ heißen **(total) stochastisch unabhängig**, wenn für jede Teilauswahl $A_{i_1}, \dots, A_{i_k}$ von $k \in \mathbb{N}$ Ereignissen der Produktsatz gilt.
- $A_1, \dots, A_n$ heißen **paarweise stochastisch unabhängig**, wenn alle Paare A_i, A_j $(i \neq j)$ stochastisch unabhängig sind.

Ersetzungsregel

Sind

$$A_1, \ldots, A_n \subset \Omega$$

unabhängig, dann auch

$$B_1, \ldots, B_k, \qquad k \leq n,$$

wobei jedes B_i entweder A_i oder $\overline{A}_i$ ist, für $i = 1, \ldots, k$.

Anwendung zur Unabhängigkeit

Beispiel: Parallelschaltung

- n Datenkabel sind parallel geschaltet. Sie können einzeln genutzt werden und fallen unabhängig voneinander aus.
- Die Übertragung fällt aus, wenn alle Kanäle versagen.
- Die Datenkabel sind durch Stecker verbunden, die unabhängig voneinander ausfallen.

Anwendung zur Unabhängigkeit

• Sei p die Wahrscheinlichkeit, dass ein Datenkabel ausfällt, und A_i das Ereignis, dass das i-te Kabel ausfällt.
Dann sind $A_1, \dots, A_n$ unabhängig mit $P(A_i) = p$, $i = 1, \dots, n$.
• Sei B das Ereignis $B =$ „Übertragung fällt aus". Dann ist

$$B = \bigcap_{i=1}^{n} A_i.$$

• Da $A_1, \dots, A_n$ unabhängig sind, ergibt sich die Ausfallwahrscheinlichkeit einer Übertragung zu

$$P(B) = P(A_1) \dots P(A_n) = p^n.$$

• Setzt man beispielsweise vier Datenkabel mit $p = 0.01$ ein, dann erhält man $P(B) = 0.01^4 = 10^{-8}$.

Anwendung zur Unabhängigkeit

Reihenschaltung: Das Datenkabel bestehe aus n Teilkabeln, die mit Steckern verbunden sind. Die Stecker versagen unabhängig voneinander mit Wahrscheinlichkeit q.

• Es bezeichne C_i das Ereignis, dass der i-te Stecker kaputt ist, und D das Ereignis $D =$ „Übertragung fällt aus“. Dann ist

$$D = \bigcup_{i=1}^{n} C_i, \qquad \overline{D} = \bigcap_{i=1}^{n} \overline{C}_i.$$

• Wir erhalten:

$$P(D) = 1 - P(\overline{D}) = 1 - P(\overline{C}_1 \cap \cdots \cap \overline{C}_n).$$

• Da $C_1, \ldots, C_n$ unabhängig sind, sind auch die komplementären Ereignisse $\overline{C}_1, \ldots, \overline{C}_n$ unabhängig. Somit ist:

$$P(\overline{C}_1 \cap \cdots \cap \overline{C}) = (1-q)^n.$$

• Die Übertragung fällt daher mit einer Wkeit von $P(D) = 1 - (1-q)^n$ aus. Für $q = 0.01$ und $n = 10$ ist: $P(D) = 0.0956$.