Das Theorem von Bayes

Ziel: Vertauschen von „Bedingung und Ereignis“
Also: gegeben $P(B|A)$, gesucht $P(A|B)$

Satz von Bayes

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

Allgemeiner nicht nur bei Dichotomie A und $\bar{A}$, sondern bei beliebiger vollständiger Zerlegung $A_1, \ldots, A_k$ anwendbar:

Satz von Bayes (allgemeine Formulierung)

Sei $A_1, \ldots A_k$ eine vollständige Zerlegung von Ω (wobei $P(A_i) > 0$, $P(B|A_i) > 0$, $i = 1, \ldots k$ und $P(B) > 0$ erfüllt seien). Dann gilt

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{k} P(B|A_i) \cdot P(A_i)}.$$

Satz von Bayes: Beispiel

T+: Test positiv
T-: Test negativ

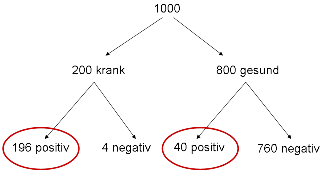
K+: krank
K-: gesund

- Sensitivität $P(T + |K+) = 0.98$
- Spezifität $P(T - |K-) = 0.95$
- Prävalenz $P(K+) = 0.2$ $\Rightarrow$ P(K-) = 0.8

Wie hoch ist die Wahrscheinlichkeit, dass eine positiv getestete Person wirklich krank ist?

Prädiktiver Wert $= P(K + |T+)$

Lösung durch hypothetische Population



Positiver prädiktiver Wert: $\frac{196}{196+40} = 0.83$

Lösung durch Satz von Bayes

- Sensitivität $P(T + |K+) = 0.98$
- Spezifität $P(T - |K-) = 0.95$
- Prävalenz $P(K+) = 0.2$

Dann gilt für den positiven prädiktiven Wert

$$\begin{aligned} P(K + |T+) &= \frac{P(K + \cap T+)}{P(T+)} \\ &= \frac{P(K+) \cdot P(T + |K+)}{P(K+) \cdot P(T + |K+) + P(K-) \cdot P(T + |K-)} \\ &= \frac{0.2 \cdot 0.98}{0.2 \cdot 0.98 + 0.8 \cdot 0.05} \\ &= \frac{0.196}{0.196 \cdot 0.040} \end{aligned}$$

Inhaltliche Bemerkungen

- Problematik: Flächendeckendes Screening nicht unumstritten, da viele falsch-positive Ergebnisse. Gegenposition: Anwendung nur auf Risikopatienten.
- Bei Mammographie oder PSA-Test auf Prostatakrebs teilweise sogar noch viel geringere Spezifität.
- Wert der mathematischen Theorie: Wenn es etwas komplexer wird, verlässt einen sofort der „gesunde Menschenverstand". Untersuchungen (u.a. von Gigerenzer) haben gezeigt, dass viele Ärzte sich dieser Problematik nicht bewusst sind.

Bayes Inferenz

Das Theorem von Bayes liefert ein Schema für das probabilistische Lernen aus Beobachtungen („Aufdatieren von Wahrscheinlichkeiten").

$$\left.\begin{array}{r}\text{priori}\\ +\ \text{Daten}\end{array}\right\} \longrightarrow \text{posteriori}$$

Es dient als Grundlage der sog. *Bayesianischen Inferenz*, einer bestimmten Schule der statistischen Methodologie. Dabei geht es darum, aus Daten zu lernen, indem man die subjektiven Wahrscheinlichkeiten $P(A_i)$ für bestimmte Modellparameter mit Hilfe der Daten (B) aufdatiert, und somit zur besseren Wahrscheinlichkeitsaussagen für die Modellparameter kommt.

Beispiel zur Bayes Inferenz I

Frage: Mit welcher Wahrscheinlichkeit trifft Schweinsteiger beim Elfmeter. Gesucht ist also ein der Parameter P_S. Es werden dazu 3 Experten gefragt:

$$\begin{array}{rl} \text{FB: } M_1 & P_S = 1 \\ \text{JH: } M_2 & P_S = 0.6 \\ \text{JL: } M_3 & P_S = 0.9 \end{array}$$

Priori - Wahrscheinlichkeiten: $P(T_1) = P(T_2) = P(T_3) = \frac{1}{3}$

Daten D: 1 Elfmeter 1 Treffer

Gesucht: $P(M_i|D)$

Lösung mit Satz von Bayes:

$$P(M_1|D) = \frac{P(M_1) \cdot P(D|M_1)}{P(D|M_1) \cdot P(M_1) + P(D|M_2) \cdot P(M_2) + P(D|M_3) \cdot P(M_3)}$$

Es ergibt sich: $P(M_1|D) = 0.4; P(M_2|D) = 0.24; P(M_3|D) = 0.36$

Beispiel zur Bayes InferenzII: 8 von 10

Weitere Berechnung für 10 Elfmeter 8 Treffer (D2)
Es ergibt sich: $P(M_1|D2) = 0; P(M_2|D2) = 0.13; P(M_3|D2) = 0.87$

- Daten widerlegen Modell 1
- Daten sprechen für Modell 3
- Inferenz durch priori und Daten

Bayes Inferenz

- Modelle M_i mit Priori Wahrscheinlichkeiten $P(M_i)$
- Daten D, die mit den Modellen auf verschiedene Weisen beschrieben werden
- Modellspezifikation liefert für jedes Modell $P(D|M_i)$ Diese wird auch als likelihood bezeichnet
- Satz von Bayes liefert:

$$P(M_i|D) = P(D|M_i) * P(M_i) * \frac{1}{P(D)} = \text{Likelihood} * \text{Priori} * \text{Konstante}$$

- Verallgemeinerung auf komplexe Situationen (z.B. Regressionsmodelle) möglich
- Expertenwissen und frühere Analysen können durch priori Verteilung eingebracht werden