Zu einer Frage der bedingten Abhängigkeit

Prof. Dr. Hans-Jörg Starkloff

Technische Universität Bergakademie Freiberg
Fakultät für Mathematik und Informatik
Institut für Stochastik

17.01.2024

letzte Änderung: 18. Januar 2024

Behauptung/Vermutung

- aus Wikipedia-Artikel "Conditional dependence" (z.B. Version mit letztem Änderungsdatum 20.12.2023, aber auch schon davor)
- **Geg.:**
  - Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$
  - $A, B, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A), \mathbf{P}(B), \mathbf{P}(C) < 1$
  - A und B sind stochastisch unabhängig, $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$
  - C ist sowohl von A als auch von B "positiv abhängig", d.h.

$$\mathbf{P}(C|A) > \mathbf{P}(C) \quad \text{und} \quad \mathbf{P}(C|B) > \mathbf{P}(C) \qquad (*)$$

- **Dann gilt:** bedingt unter C ist A von B "negativ abhängig", d.h.

$$\mathbf{P}(A|B \cap C) < \mathbf{P}(A|C) \qquad (**)$$

- kein Beweis, Verweis auf Internetquelle, die als "[permanent dead link]" gekennzeichnet wird, eine gewisse heuristische Begründung und ein Beispiel (verbal und numerisch) werden gegeben

Zur positiven und negativen Abhängigkeit von Ereignissen

Beh.

- **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$ und $A, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A), \mathbf{P}(C) < 1$
- **Dann** sind alle folgenden 20 Beziehungen zueinander äquivalent:
  - $\mathbf{P}(C|A) > \mathbf{P}(C)$, $\mathbf{P}(A|C) > \mathbf{P}(A)$, $\mathbf{P}(A \cap C) > \mathbf{P}(A)\mathbf{P}(C)$
  - $\mathbf{P}(C|A^c) < \mathbf{P}(C)$, $\mathbf{P}(A^c|C) < \mathbf{P}(A^c)$, $\mathbf{P}(A^c \cap C) < \mathbf{P}(A^c)\mathbf{P}(C)$
  - $\mathbf{P}(C^c|A) < \mathbf{P}(C^c)$, $\mathbf{P}(A|C^c) < \mathbf{P}(A)$, $\mathbf{P}(A \cap C^c) < \mathbf{P}(A)\mathbf{P}(C^c)$
  - $\mathbf{P}(C^c|A^c) > \mathbf{P}(C^c)$, $\mathbf{P}(A^c|C^c) > \mathbf{P}(A^c)$, $\mathbf{P}(A^c \cap C^c) > \mathbf{P}(A^c)\mathbf{P}(C^c)$
  - $\mathbb{C}\mathbf{ov}[\mathbf{1}_A, \mathbf{1}_C] > 0$, $\mathbb{C}\mathbf{orr}[\mathbf{1}_A, \mathbf{1}_C] > 0$
  - $\mathbb{C}\mathbf{ov}[\mathbf{1}_{A^c}, \mathbf{1}_C] < 0$, $\mathbb{C}\mathbf{orr}[\mathbf{1}_{A^c}, \mathbf{1}_C] < 0$
  - $\mathbb{C}\mathbf{ov}[\mathbf{1}_A, \mathbf{1}_{C^c}] < 0$, $\mathbb{C}\mathbf{orr}[\mathbf{1}_A, \mathbf{1}_{C^c}] < 0$
  - $\mathbb{C}\mathbf{ov}[\mathbf{1}_{A^c}, \mathbf{1}_{C^c}] > 0$, $\mathbb{C}\mathbf{orr}[\mathbf{1}_{A^c}, \mathbf{1}_{C^c}] > 0$
- analoge Aussagen gelten für negativ abhängige und unabhängige Ereignisse A und C

Zusammenhang zum Begriff der positiven Assoziation

- In Cox, D.R., Causality: some statistical aspects, J. R. Statist. Soc. A (1992), v. 155, part 2, pp. 291–301, wird definiert (S. 292):

  "We may say that C is a candidate cause of E if C and E are positively associated, i.e. if
  $$\mathbf{P}(E|C) > \mathbf{P}(E|\text{not } C)."$$

- **Beh.** **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$ und $A, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A), \mathbf{P}(C) < 1$

  Dann sind die Beziehungen der Behauptung von Folie 3 auch äquivalent zu $\mathbf{P}(C|A) > \mathbf{P}(C|A^c)$, d.h. A ist eine mögliche Ursache von C im Sinne von Cox oder A und C sind positiv assoziiert im Sinne von Cox.

- **Zum Beweis**: Nach der vorigen Behauptung gilt eine der folgenden Beziehungen:
  $$\mathbf{P}(C|A) > \mathbf{P}(C) > \mathbf{P}(C|A^c);$$
  $$\mathbf{P}(C|A) = \mathbf{P}(C) = \mathbf{P}(C|A^c);$$
  $$\mathbf{P}(C|A) < \mathbf{P}(C) < \mathbf{P}(C|A^c). \qquad \square$$

Symmetrie bezüglich der Ereignisse A und B

Beh.

- **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$ und $A, B, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A \cap C), 0 < \mathbf{P}(B \cap C)$
- **Dann gilt:**

$$(\mathbf{P}(A|B \cap C) < \mathbf{P}(A|C)) \quad \Leftrightarrow \quad (\mathbf{P}(B|A \cap C) < \mathbf{P}(B|C))$$

Eine allgemeine Beziehung

- **Beh.** **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$, $A, B, C \in \mathcal{F}$ mit $0 < \mathbf{P}(B \cap C)$

$$\Rightarrow \quad \mathbf{P}(A|B \cap C) \cdot \mathbf{P}(B|C) = \mathbf{P}(A \cap B|C) \qquad (***)$$

- dies ist auch ein Axiom in der Definition von bedingten Wahrscheinlichkeitsräumen von Rényi,

  "On a new axiomatic theory of probability", Acta Mathematica Hungarica, 1955, v. 6, nr. 3–4, p. 285–335

- damit gilt auch

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

  und so

$$(\mathbf{P}(A|B \cap C) < \mathbf{P}(A|C)) \Leftrightarrow (\mathbf{P}(A \cap B|C) < \mathbf{P}(A|C) \cdot \mathbf{P}(B|C)),$$

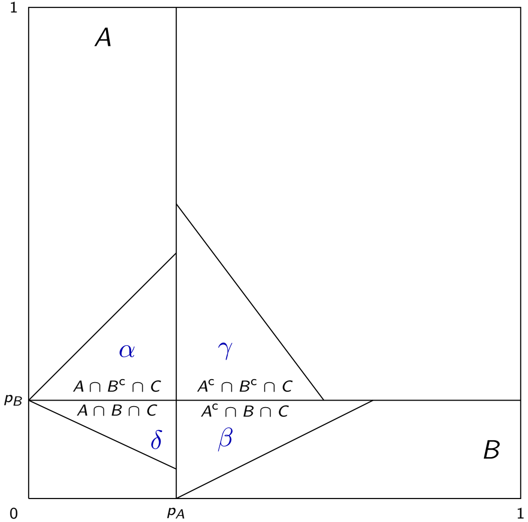
  d.h. bedingt unter C sind A und B negativ abhängig

Geometrische Veranschaulichung

- $\Omega = [0,1]^2$ mit $\sigma-$Algebra der Borelmengen und Flächenmaß als Wahrscheinlichkeitsmaß
- $0 < \mathbf{P}(A) =: p_A < 1$
- $0 < \mathbf{P}(B) =: p_B < 1$
- $0 < \mathbf{P}(C) =: p_C = \alpha + \beta + \gamma + \delta < 1$ mit
  - $0 \leq \alpha := \mathbf{P}(A \cap B^c \cap C) \leq p_A \cdot (1 - p_B) < 1$
  - $0 \leq \beta := \mathbf{P}(A^c \cap B \cap C) \leq (1 - p_A) \cdot p_B < 1$
  - $0 \leq \gamma := \mathbf{P}(A^c \cap B^c \cap C) \leq (1 - p_A) \cdot (1 - p_B) < 1$
  - $0 \leq \delta := \mathbf{P}(A \cap B \cap C) \leq p_A \cdot p_B < 1$

$\Rightarrow$ $0 < \mathbf{P}(A \cap C) = \alpha + \delta < 1$; $\qquad 0 < \mathbf{P}(B \cap C) = \beta + \delta < 1$

Prinzipgrafik zur geometrischen Veranschaulichung



Der Spezialfall $C \subseteq A \cap B$, d.h. $\alpha = \beta = \gamma = 0$

- **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$; $A, B, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A), \mathbf{P}(B), \mathbf{P}(C) < 1$; $C \subseteq A \cap B$
- **Dann gelten:**
  - $\mathbf{P}(C|A) = \dfrac{\mathbf{P}(C)}{\mathbf{P}(A)} > \mathbf{P}(C), \quad \mathbf{P}(C|B) = \dfrac{\mathbf{P}(C)}{\mathbf{P}(B)} > \mathbf{P}(C)$
  - $\mathbf{P}(A|C) = \dfrac{\mathbf{P}(C)}{\mathbf{P}(C)} = 1$
  - $\mathbf{P}(A|B \cap C) = \dfrac{\mathbf{P}(C)}{\mathbf{P}(C)} = 1 \not< \mathbf{P}(A|C)$, sondern

$$\mathbf{P}(A|B \cap C) = \mathbf{P}(A|C)$$

- **Bem.** Allgemein gilt für $A, B, C \in \mathcal{F}$ mit $\mathbf{P}(C) > 0$

$$(\mathbf{P}(B|C) = 1) \quad \Rightarrow \quad (\mathbf{P}(A|B \cap C) = \mathbf{P}(A|C))$$

Der Spezialfall $\delta = 0$

- **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$; $A, B, C \in \mathcal{F}$ mit $0 < \mathbf{P}(A \cap C), \mathbf{P}(B \cap C)$
- **Dann** folgt wegen

$$\delta := \mathbf{P}(A \cap B \cap C) = 0 \Leftrightarrow \mathbf{P}(A|B \cap C) = 0 \Leftrightarrow \mathbf{P}(B|A \cap C) = 0\,,$$

- dass immer gilt:

$$0 = \mathbf{P}(A|B \cap C) < \mathbf{P}(A|C)\,;$$
$$0 = \mathbf{P}(B|A \cap C) < \mathbf{P}(B|C)$$

Der Spezialfall $C = A \cup B$, insbesondere $\gamma = 0$

- dies ist die Situation im Beispiel des Wikipedia-Artikels
- $0 < \delta = p_A \cdot p_B < 1\,; \quad p_A = \alpha + \delta\,; \quad p_B = \beta + \delta$
- $0 < \alpha = p_A \cdot (1 - p_B) < 1\,; \quad 0 < \beta = (1 - p_A) \cdot p_B < 1$
- $p_C = p_A + p_B - p_A \cdot p_B = \alpha + \beta + \delta$
- $\mathbf{P}(C|A) = \dfrac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1 > \mathbf{P}(C)\,;\; \mathbf{P}(C|B) = \dfrac{\mathbf{P}(B)}{\mathbf{P}(B)} = 1 > \mathbf{P}(C)$
- $0 < \mathbf{P}(A|C) = \dfrac{\mathbf{P}(A)}{\mathbf{P}(C)} = \dfrac{\alpha + \delta}{\alpha + \beta + \delta} < 1$
- $0 < \mathbf{P}(A|B \cap C) = \dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \dfrac{\delta}{\beta + \delta} < \mathbf{P}(A|C) = \dfrac{\alpha + \delta}{\alpha + \beta + \delta}\,,$

  da unter den gemachten Voraussetzungen von Folie 2

  $\mathbb{R} \ni \alpha \mapsto \dfrac{\alpha + \delta}{\alpha + \beta + \delta}$ streng monoton wachsend ist

Der Spezialfall $C \subseteq A \cup B$, d.h. $\gamma = 0$

- aus den Voraussetzungen von Folie 2 folgen
  - $\mathbf{P}(C|A) = \dfrac{\alpha+\delta}{p_A} > \mathbf{P}(C) = \alpha+\beta+\delta \;\Leftrightarrow\; \beta < (\alpha+\delta)\dfrac{1-p_A}{p_A}$
  - $\mathbf{P}(C|B) = \dfrac{\beta+\delta}{p_B} > \mathbf{P}(C) = \alpha+\beta+\delta \;\Leftrightarrow\; \alpha < (\beta+\delta)\dfrac{1-p_B}{p_B}$
- $\mathbf{P}(A|C) = \dfrac{\alpha+\delta}{\alpha+\beta+\delta}, \qquad \mathbf{P}(A|B\cap C) = \dfrac{\delta}{\beta+\delta} \qquad \Rightarrow$

$$
\begin{aligned}
\mathbf{P}(A|B\cap C) \le \mathbf{P}(A|C) &\Leftrightarrow \frac{\delta}{\beta+\delta} \le \frac{\alpha+\delta}{\alpha+\beta+\delta} \\
\Leftrightarrow\; \delta(\alpha+\beta+\delta) &\le (\alpha+\delta)(\beta+\delta) \\
\Leftrightarrow\; \alpha\delta+\beta\delta+\delta^2 &\le \alpha\beta+\alpha\delta+\beta\delta+\delta^2 \\
\Leftrightarrow\; 0 &\le \alpha\beta
\end{aligned}
$$

- ist immer erfüllt, Gleichheitszeichen genau dann, wenn $\alpha = 0$ oder $\beta = 0$ (siehe auch den Fall $C \subseteq A \cap B$)

Ein konkretes Beispiel mit $\mathbf{P}(A|B \cap C) > \mathbf{P}(A|C)$

- $\mathbf{P}(A) = \mathbf{P}(B) = \frac{3}{8} = 0.375\,; \quad \mathbf{P}(A \cap B) = \frac{9}{64} = 0.140625$
- $\alpha = \beta = \delta = \frac{1}{16} = 0.0625\,; \quad \gamma = \frac{3}{32} = 0.09375$

$\Rightarrow \mathbf{P}(A \cap C) = \alpha + \delta = \mathbf{P}(B \cap C) = \beta + \delta = \frac{1}{8} = 0.125$

- $\mathbf{P}(C|A) = \frac{\mathbf{P}(A \cap C)}{\mathbf{P}(A)} = \frac{1}{3} = 0.3\overline{3} > \mathbf{P}(C) = \frac{9}{32} = 0.28125$
- $\mathbf{P}(C|B) = \frac{\mathbf{P}(B \cap C)}{\mathbf{P}(B)} = \frac{1}{3} = 0.3\overline{3} > \mathbf{P}(C) = \frac{9}{32} = 0.28125$
- $\mathbf{P}(A|C) = \frac{\mathbf{P}(A \cap C)}{\mathbf{P}(C)} = \frac{1/8}{9/32} = \frac{4}{9} = 0.4\overline{4}$
- $\mathbf{P}(A|B \cap C) = \frac{\mathbf{P}(A \cap B \cap C)}{\mathbf{P}(B \cap C)} = \frac{1/16}{1/8} = \frac{1}{2} = 0.5 > \mathbf{P}(A|C)$

Allgemeine Betrachtung

- aus (***) folgt unter den Voraussetzungen von Folie 2
$$(\mathbf{P}(A|C)) \overset{\geq}{\underset{<}{=}} \mathbf{P}(A|B\cap C)) \;\Leftrightarrow\; (\mathbf{P}(A|C)\cdot\mathbf{P}(B|C) \overset{\geq}{\underset{<}{=}} \mathbf{P}(A\cap B|C))$$
- mit den Bezeichnungen aus der geometrischen Veranschaulichung ist dies äquivalent zu
$$\frac{\mathbf{P}(A\cap C)}{\mathbf{P}(C)}\cdot\frac{\mathbf{P}(B\cap C)}{\mathbf{P}(C)} \overset{\geq}{\underset{<}{=}} \frac{\mathbf{P}(A\cap B\cap C)}{\mathbf{P}(C)}$$
$$\mathbf{P}(A\cap C)/\mathbf{P}(C) \overset{\geq}{\underset{<}{=}} \mathbf{P}(A\cap B\cap C)/\mathbf{P}(B\cap C)$$
oder $$\mathbf{P}(A\cap C)\cdot\mathbf{P}(B\cap C) \overset{\geq}{\underset{<}{=}} \mathbf{P}(A\cap B\cap C)\cdot\mathbf{P}(C)$$
$$(\alpha+\delta)\cdot(\beta+\delta) \overset{\geq}{\underset{<}{=}} \delta\cdot(\alpha+\beta+\gamma+\delta)$$
$$\alpha\beta+\alpha\delta+\beta\delta+\delta^2 \overset{\geq}{\underset{<}{=}} \alpha\delta+\beta\delta+\gamma\delta+\delta^2$$
$$\alpha\beta-\gamma\delta = \begin{vmatrix}\alpha & \gamma\\ \delta & \beta\end{vmatrix} \overset{\geq}{\underset{<}{=}} 0$$

Analytische Herangehensweise

- Bestimmung des Wertebereichs (des Minimums und Maximums)

$$\frac{\mathbf{P}(A \cap C)}{\mathbf{P}(C)} \cdot \frac{\mathbf{P}(B \cap C)}{\mathbf{P}(C)} - \frac{\mathbf{P}(A \cap B \cap C)}{\mathbf{P}(C)}$$
$$= \frac{(\alpha + \delta) \cdot (\beta + \delta)}{(\alpha + \beta + \gamma + \delta)^2} - \frac{\delta}{\alpha + \beta + \gamma + \delta}$$

  als Funktion von $p_A, p_B, \alpha, \beta, \gamma, \delta$ mit allen Nebenbedingungen, z.B.

- $0 < p_A, p_B < 1\,;\ 0 < \alpha + \beta + \gamma + \delta < 1\,;$
- $\dfrac{\alpha + \delta}{p_A}, \dfrac{\beta + \delta}{p_B} > \alpha + \beta + \gamma + \delta\,;$
- $0 \le \alpha, \beta, \gamma, \delta < 1\,;$
- $0 < \alpha + \delta, \beta + \delta < 1\,;$
- $0 \le \alpha \le p_A(1 - p_B)\,;\ 0 \le \beta \le p_B(1 - p_A)\,;\ 0 \le \delta \le p_A p_B\,;$
  $0 \le \gamma \le (1 - p_A)(1 - p_B)$

Das Mengensystem von mit C unabhängigen Ereignissen

- **Geg.:** Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbf{P})$, $C \in \mathcal{F}$ mit $0 < \mathbf{P}(C) < 1$
- **Def.** Mengensystem der mit C unabhängigen Ereignisse:

$$\begin{aligned} \mathcal{UA}_C &:= \{A \in \mathcal{F} : \mathbf{P}(A|C) = \mathbf{P}(A)\} \\ &= \{A \in \mathcal{F} : \mathbf{P}(A \cap C) = \mathbf{P}(A)\mathbf{P}(C)\} \end{aligned}$$

- **Beh.** Unter obigen Voraussetzungen ist $\mathcal{UA}_C$ ein Dynkin-System ($\lambda-$System), d.h. es enthält Ω und ist stabil bezüglich eigentlicher Differenzen und abzählbarer disjunkter Vereinigungen.
- **Beh.** Unter obigen Voraussetzungen enthält $\mathcal{UA}_C$ die $\sigma-$Algebra $\{A \in \mathcal{F} : \mathbf{P}(A) \in \{0,1\}\} = \{A \in \mathcal{F} : \mathbf{P}(A)(1 - \mathbf{P}(A)) = 0\}$.
- **Beh.** Unter obigen Voraussetzungen ist $\mathcal{F}$ die disjunkte Vereinigung von $\{A \in \mathcal{F} : \mathbf{P}(A|C) = \mathbf{P}(A)\}$, $\{A \in \mathcal{F} : \mathbf{P}(A|C) > \mathbf{P}(A)\}$ und $\{A \in \mathcal{F} : \mathbf{P}(A|C) < \mathbf{P}(A)\}$.

$\Rightarrow$ erinnert an Situation in Linearer Algebra mit Halbräumen

Weitere mögliche Fragen

- Rolle der Unabhängigkeit der Ereignisse A und B ?
- Rolle der positiven Abhängigkeit des Ereignisses C von A und B ?
- Beziehung zur paarweisen Unabhängigkeit bzw. vollständigen Unabhängigkeit der Ereignisse A, B, C ?
- Beziehung zu Paradoxien der Wahrscheinlichkeitsrechnung ? (Vgl. z.B. das Paradox von Berkson.)
- Untersuchungen zu Mengensystemen unabhängiger/positiv abhängiger/negativ abhängiger Ereignisse
- ...