3. Bedingte Wahrscheinlichkeit und Bayes

Yichen Han

9. November 2023

Schwierigkeitsumfrage

Entschuldigung fürs Durcheinander mit dem Zeitplan...

Abbildung: Quiz 6

Schlüsselfähigkeit

- Die Abstraktion von Wahrscheinlichkeitsausdrücken aus Textbeschreibungen unter realistischen Szenarien.
- Die Anwendung von den gelernten Formeln (bedingte W.keit, S.v.tot.W.keit, Bayes), um aus einer gegebenen bedingten Wahrscheinlichkeit $P(A|B)$ u.a. die umgekehrte bedingte Wahrscheinlichkeit $P(B|A)$ zu berechnen.

Bedingte Wahrscheinlichkeit

$$P(A \cap B) = P(A|B)P(B) \tag{1}$$

Satz von der totalen W.keit

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i) \tag{2}$$

Stochastische Unabhängigkeit

$$A \perp B \Leftrightarrow P(A \cap B) = P(A)P(B) \tag{3}$$

Satz von Bayes

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \tag{4}$$

Weiterführung: W.keitstheorie in Stat2

Was ihr bisher über Wahrscheinlichkeit gelernt habt, ist aber nicht unangreifbar...

[1]Credit: Prof. Dr. Volker Schmid, Prof. Dr. David Rügamer

Was ihr bisher über Wahrscheinlichkeit gelernt habt, ist aber nicht unangreifbar...

Überlegt euch die folgenden Fragen: [1]

- Kann eine Wahrscheinlichkeit immer auf der Potenzmenge $\mathcal{P} := \{A|A \in \Omega\}$ definiert werden? Was passiert, wenn $\mathcal{P}$ sehr groß ist? (S. Satz von Vitali)
- Laplace-Wahrscheinlichkeit: $P(A) = \frac{|A|}{|\Omega|}$ Was passiert bei $|A| = |\Omega| = \infty$? Was soll man dann machen?

[1] Credit: Prof. Dr. Volker Schmid, Prof. Dr. David Rügamer

Weiterführung: W.keitstheorie in Stat2

Was ihr bisher über Wahrscheinlichkeit gelernt habt, ist aber nicht unangreifbar...
Überlegt euch die folgenden Fragen: [1]

- Kann eine Wahrscheinlichkeit immer auf der Potenzmenge $\mathcal{P} := \{A | A \in \Omega\}$ definiert werden? Was passiert, wenn $\mathcal{P}$ sehr groß ist? (S. Satz von Vitali)
- Laplace-Wahrscheinlichkeit: $P(A) = \frac{|A|}{|\Omega|}$ Was passiert bei $|A| = |\Omega| = \infty$? Was soll man dann machen?

Zwar könntet ihr schon intuitive Ideen dazu haben, aber eine mathematische Erläuterung ist derzeit für euch noch unmöglich. (Maßtheorie)
Studiert weiter mit bestehender Zweifel. Alles wird in Stat2 geklärt.

[1] Credit: Prof. Dr. Volker Schmid, Prof. Dr. David Rügamer

Inhalt

1. Blatt 3
2. Wiederholung

Aufgabe 1

Gegeben seien zwei Ereignisse A und B mit $P(A) = P(B) = 1/2$ und $P(B|A) = 1/2$. Sind A und B unabhängig? Wie groß ist $P(A \cup B)$?

Aufgabe 1

Gegeben seien zwei Ereignisse A und B mit $P(A) = P(B) = 1/2$ und $P(B|A) = 1/2$. Sind A und B unabhängig? Wie groß ist $P(A \cup B)$?

$$
\begin{aligned}
P(B|A) &= \frac{P(B \cap A)}{P(A)} = \frac{1}{2} \\
\Rightarrow P(B \cap A) &= \frac{1}{4} = P(A) \cdot P(B) \Leftrightarrow A \perp B \\
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
&= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}
$$

Aufgabe 2.1

Die Eingänge eines Ladens sind mit einer Alarmanlage gegen Diebstahl gesichert. Wenn Diebe die Anlage passieren, wird mit W'keit 0.995 Alarm ausgelöst. Bei unbescholtenen Kund:innen beträgt die W'keit 0.006. Erfahrungswerte zeigen, daß auf 1000 Kund:innen zwei Dieb:innen kommen.

a Mit welcher W'keit alarmiert die Anlage zu Recht? Mit welcher Wahrscheinlichkeit werden harmlose Kund:innen erschreckt?

Die Eingänge eines Ladens sind mit einer Alarmanlage gegen Diebstahl gesichert. Wenn Diebe die Anlage passieren, wird mit W'keit 0.995 Alarm ausgelöst. Bei unbescholtenen Kund:innen beträgt die W'keit 0.006. Erfahrungswerte zeigen, daß auf 1000 Kund:innen zwei Dieb:innen kommen.

a Mit welcher W'keit alarmiert die Anlage zu Recht? Mit welcher Wahrscheinlichkeit werden harmlose Kund:innen erschreckt?

$D :=$ "Echter Diebstahl."
$A :=$ "Die Anlage alarmiert."
Wir wissen:

$$P(A|D) = 0.995 \qquad P(A|\overline{D}) = 0.006 \qquad P(D) = 0.002$$

Wir suchen: $P(D|A) = \frac{P(A|D) \cdot P(D)}{P(A)}$ und $P(\overline{D}|A) = 1 - P(D|A)$.
S. Tafel.

Aufgabe 2.2

$D :=$ "Echter Diebstahl."
$A :=$ "Die Anlage alarmiert."

b Wie groß müssten die W'keiten für korrekten und falschen Alarm seien, damit zumindest die Hälfte der Kompromittierten tatsächlich etwas geklaut haben?

Wir wissen:

$$P(D) = 0.002 \qquad P(\overline{D}) = 0.998$$

Wir lösen:

$$P(D|A) \overset{!}{=} 0.5$$

S. Tafel.

$D :=$ "Echter Diebstahl."
$A :=$ "Die Anlage alarmiert."

b Wie groß müssten die W'keiten für korrekten und falschen Alarm seien, damit zumindest die Hälfte der Kompromittierten tatsächlich etwas geklaut haben?

Wir wissen:

$$P(D) = 0.002 \qquad P(\overline{D}) = 0.998$$

Wir lösen:

$$P(D|A) \stackrel{!}{=} 0.5$$

S. Tafel.

$$P(A|D) = 499P(A|\overline{D})$$

Betrachten Sie die Ereignisse

- A: Anschnallpflicht befolgt
- K: Schwere Kopfverletzung nach Unfall

Aufgabe 3.1

Betrachten Sie die Ereignisse

- A: Anschnallpflicht befolgt
- K: Schwere Kopfverletzung nach Unfall

$$P(\bar{A}) = 0.15 \qquad P(K|A) = 0.08 \qquad P(\bar{K}|\bar{A}) = 0.62$$
$$P(A) = 0.85 \qquad P(\bar{K}|A) = 0.92 \qquad P(K|\bar{A}) = 0.38$$

a Interpretieren Sie das Ereignis $\bar{A} \cap K$ und berechnen Sie $P(\bar{A} \cap K)$.

Aufgabe 3.1

Betrachten Sie die Ereignisse

- A: Anschnallpflicht befolgt
- K: Schwere Kopfverletzung nach Unfall

$$P(\bar{A}) = 0.15 \qquad P(K|A) = 0.08 \qquad P(\bar{K}|\bar{A}) = 0.62$$
$$P(A) = 0.85 \qquad P(\bar{K}|A) = 0.92 \qquad P(K|\bar{A}) = 0.38$$

a Interpretieren Sie das Ereignis $\bar{A} \cap K$ und berechnen Sie $P(\bar{A} \cap K)$.

$\bar{A} \cap K$: Der Fahrer ist nicht angeschnallt und hat schwere Kopfverletzung.

$$P(\bar{A} \cap K) = P(K|\bar{A}) \cdot P(\bar{A}) = 0.057$$

$$P(\bar{A}) = 0.15 \qquad P(K|A) = 0.08 \qquad P(\bar{K}|\bar{A}) = 0.62$$
$$P(A) = 0.85 \qquad P(\bar{K}|A) = 0.92 \qquad P(K|\bar{A}) = 0.38$$

b Sind die Ereignisse $\bar{A}$ und K stochastisch unabhängig?

$$P(\bar{A}) = 0.15 \qquad P(K|A) = 0.08 \qquad P(\bar{K}|\bar{A}) = 0.62$$
$$P(A) = 0.85 \qquad P(\bar{K}|A) = 0.92 \qquad P(K|\bar{A}) = 0.38$$

b Sind die Ereignisse $\bar{A}$ und K stochastisch unabhängig?

$$P(K) = P(K|A)P(A) + P(K|\bar{A})P(\bar{A}) = 0.125$$
$$P(\bar{A})P(K) = 0.15 \times 0.125 \neq P(\bar{A} \cap K) \Rightarrow \text{Nein}$$

c Wie groß ist die Wahrscheinlichkeit, dass jemand bei einem Unfall nicht angegurtet war falls eine schwere Kopfverletzung diagnostiziert wurde?

Aufgabe 3.2

$$P(\bar{A}) = 0.15 \qquad P(K|A) = 0.08 \qquad P(\bar{K}|\bar{A}) = 0.62$$
$$P(A) = 0.85 \qquad P(\bar{K}|A) = 0.92 \qquad P(K|\bar{A}) = 0.38$$

b Sind die Ereignisse $\bar{A}$ und K stochastisch unabhängig?

$$P(K) = P(K|A)P(A) + P(K|\bar{A})P(\bar{A}) = 0.125$$

$$P(\bar{A})P(K) = 0.15 \times 0.125 \neq P(\bar{A} \cap K) \Rightarrow \text{Nein}$$

c Wie groß ist die Wahrscheinlichkeit, dass jemand bei einem Unfall nicht angegurtet war falls eine schwere Kopfverletzung diagnostiziert wurde?

$$P(\bar{A}|K) = \frac{P(\bar{A} \cap K)}{P(K)} = 0.456$$

Alternativ: Bayes

Aufgabe 4.1

Die Software stuft eine tatsächlich bedrohliche Kommunikation mit einer sehr hohen Wahrscheinlichkeit von 99,5% richtig ein. Die Wahrscheinlichkeit dafür, dass eine harmlose E-Mail fälschlicherweise als potentielle Bedrohung klassifiziert wird, liegt dagegen nur bei 0,5%.
In Deutschland gibt es 71.000.000 Internetnutzer:innen. Nachfolgend gehen wir davon aus, dass

- jede:r Nutzer:in täglich 10 unverschlüsselte E-Mails verschickt,
- 10.000 Nutzer:innen das Internet für die Vorbereitung illegaler und terroristischer Aktivitäten nutzen
- und jede vierte E-Mail aus diesem Personenkreis direkte Kommunikation über eine bedrohliche Aktivität enthält.

Wie groß ist die Wahrscheinlichkeit dafür, dass an einem beliebigen Tag eine durch die Software als potentielle Bedrohung eingestufte E-Mail tatsächlich eine reale Bedrohung aufweist?

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

$$\begin{aligned}
P(B) &= \frac{|B|}{|\Omega|} = \frac{10.000 \cdot 10 \cdot 0.25}{71.000.000 \cdot 10} \approx 0.000035 \\
P(V) &= P(V|B) \cdot P(B) + P(V|\bar{B}) \cdot P(\bar{B}) \approx 0.005 \\
P(B \cap V) &= P(V|B) \cdot P(B) = 0.000033 \\
P(B|V) &= \frac{0.000033}{0.005} = 0.0066 = 0.66\%
\end{aligned}$$

Inhalt

Zusatzaufgabe 1: Abstraktion

Eine Fabrik produziert elektronische Bauteile mit drei Maschinen (A, B und C). Maschine A produziert 30% der Bauteile, Maschine B 50% und Maschine C die restlichen 20%. Die Wahrscheinlichkeit eines Produktionsfehlers liegt bei Maschine A bei 1%, bei Maschine B bei 2% und bei Maschine C bei 5%.

Ein zufällig ausgewähltes Bauteil weist einen Fehler auf.

Welche der folgenden Aussagen sind korrekt?

Abbildung: Quiz 7 (Aufgabe 1)

Zusatzaufgabe 1: Abstraktion

1. **Gesamtwahrscheinlichkeit eines Fehlers** $P(F)$:

$$P(F) = P(F|A)P(A) + P(F|B)P(B) + P(F|C)P(C) = 0.023$$

2. **Bedingte Wahrscheinlichkeiten:**

- Für Maschine A:

$$P(A|F) = \frac{P(F|A)P(A)}{P(F)} = 0.1304$$

- Für Maschine B:

$$P(B|F) = \frac{P(F|B)P(B)}{P(F)} = 0.4348$$

- Für Maschine C:

$$P(C|F) = \frac{P(F|C)P(C)}{P(F)} = 0.4348$$

Zusatzaufgabe 2: Kap.1 mit Bayes in Odds

In einer klinischen Studie werden Daten zur Diagnose einer seltenen Krankheit X gesammelt. Die Krankheit tritt bei 1 von 10.000 Personen in der Bevölkerung auf. Es gibt einen diagnostischen Test, dessen Sensitivität (Wahrscheinlichkeit, dass der Test positiv ist, wenn die Krankheit vorliegt) bei 99% liegt und dessen Spezifität (Wahrscheinlichkeit, dass der Test negativ ist, wenn die Krankheit nicht vorliegt) bei 95% liegt.

1. Bestimmen Sie für das Merkmal "Das Ergebnis des diagnostischen Tests (positiv/negativ)"das Skalenniveau.
2. Welche Erhebungsart ist für diese Studie geeignet?
3. Ein Patient erhält ein positives Testergebnis. Berechnen Sie unter Verwendung der Odds-Notation und des Satzes von Bayes die Odds, dass der Patient tatsächlich die Krankheit X hat.

Gegeben:

$$P(K) = 0.0001 \qquad P(T|K) = 0.99 \qquad P(\bar{T}|\bar{K}) = 0.95$$

$$\begin{aligned}
\text{Prior Odds} &= \gamma(K) = \frac{0.0001}{0.9999} = 1 : 9999 \\
\text{Likelihood Ratio} &= \frac{P(T|K)}{P(T|\bar{K})} = \frac{0.99}{1-0.95} = 99 : 5 \\
\text{Posterior Odds} &= \gamma(K|T) = \text{Prior Odds} \cdot \text{Likelihood Ratio} \\
&= 1 : 9999 \times 99 : 5 = 1 : 505
\end{aligned}$$

Die Odds, dass ein Patient mit einem positiven Testergebnis tatsächlich die Krankheit X hat, sind 1 zu 505.