Bayes' setning, Afsnit 7.3 [Ros]

March 13, 2021

Sandsynlighedsteori og lineær algebra (SLIAL)

Forår 2021

AALBORG UNIVERSITY
DENMARK

Part I

Repetition fra sidst

Sandsynlighedsfelt

Definition

- Et **udfaldsrum** er en *endelig* mængde $S = \{a_1, a_2, \ldots, a_n\}$.
- En **sandsynlighedsfunktion** på S er en funktion p med definitionsmængde S, som opfylder $p(a_i)$ er et reelt tal, $0 \leq p(a_i) \leq 1$, for ethvert a_i i S.
  Desuden er

$$\sum_{i=1}^{n} p(a_i) = p(a_1) + \ldots + p(a_n) = 1.$$

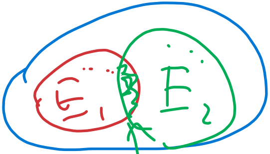
Sandsynlighedsfelt

Et udfaldsrum og en sandsynlighedsfunktion kaldes samlet et **sandsynlighedsfelt**.

En delmængde E af S kaldes en **hændelse**.
En hændelse tildeles også en sandsynlighed:

$$p(E) = \sum_{x \in E} p(x).$$

Sandsynlighedsfelt



- Hvis E_1 og E_2 er hændelser i udfaldsrummet S så er

$E_1 \cap E_2$

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

- Hvis

$$\boxed{p(E_1 \cap E_2) = p(E_1)p(E_2)}$$

så siges E_1 og E_2 at være **uafhængige**.

$p(E_1 | E_2) = p(E_1)$

Eksempel

$E_1 = \{2,4,6\}$ $E_2 = \{1,3,5\}$

Vi kaster med en fair 6-sidet terning. Lad E_1 være udfald med et lige antal øjne og E_2 være udfald med et ulige antal øjne.

1. Hvad er $p(E_1)$? $= p(2) + p(4) + p(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$
2. Hvad er $p(E_2)$? $= \frac{1}{2}$
3. Hvad er $p(E_1 \cap E_2)$? $= p(\emptyset) = 0$
4. Hvad er $p(E_1 \cup E_2)$? $= p(E_1) + p(E_2) - p(E_1 \cap E_2)$

$= \frac{1}{2} + \frac{1}{2} - 0 = 1$

$$p(E_1 \cup E_2) = p(S) = 1$$

Eksempel: 6-sidet terning

Vi kaster med en fair 6-sidet terning. Lad E_1 være udfald med et lige antal øjne og E_2 være udfald med et ulige antal øjne.

1. Hvad er $p(E_1)$?
2. Hvad er $p(E_2)$?
3. Hvad er $p(E_1 \cap E_2)$?
4. Hvad er $p(E_1 \cup E_2)$?

Eksempel: kort spil

Lad S være mængden af 52 forskellige kort.

Hvert kort tildeles sandsynlighed $\frac{1}{52}$.

Lad E_1 være de 13 hjertere.

Lad E_2 være de 4 damer.

1. Hvad er $p(E_1)$? $= \frac{13}{52} = \frac{1}{4}$
2. Hvad er $p(E_2)$? $= \frac{4}{52} = \frac{1}{13}$
3. Hvad er $p(E_1 \cap E_2)$? $= p(D\heartsuit) = \frac{1}{52} = \frac{1}{4} \cdot \frac{1}{13}$

E_1, E_2 - uafhængige!

Eksempel: kort spil

Vi fjerner en tilfældig kort: f.eks. 3♣
Lad S være mængden af de resterende 51 kort.
Hvert kort tildeles sandsynlighed $\frac{1}{51}$.
Lad E_1 være de 13 hjertere.
Lad E_2 være de 4 damer.

1. Hvad er $p(E_1)$? $= \frac{13}{51}$
2. Hvad er $p(E_2)$? $= \frac{4}{51}$
3. Hvad er $p(E_1 \cap E_2)$? $= p(D♡) = \frac{1}{51} \neq \frac{13}{51} \cdot \frac{4}{51} = \frac{52}{51 \cdot 51}$

E_1, E_2 - ikke uafhængige!

Part II

Bayes' sætning

Eksempel: skriftlig eksamen

Eksempel

Til en skriftlig eksamen er der en multiple choice opgave med fem svarmuligheder.
Lad udfaldsrummet S være mængden af eksaminander.

- *70% af eksaminanderne kender svaret på denne type og de sætter alle kryds ved det korrekte svar.*
- *De resterende 30% kender ikke svaret og de sætter et tilfældigt kryds.*

Givet at en eksaminand har sat kryds ved det rigtige svar, hvad er sandsynligheden for, at vedkommende kender svaret?

Eksempel: skriftlig eksamen

- Lad $E \subseteq S$ være mængden af eksaminander der kender svaret. Så er $p(E) = 0.7$.
- $\overline{E}$ er da mængden af eksaminander der ikke kender svaret. Så $p(\overline{E}) = 0.3$.
- Lad $F \subseteq S$ være mængden af eksaminander der sætter rigtigt kryds.

Da er

$$p(F \mid E) = 1, \quad \text{og} \quad p(F \mid \overline{E}) = \frac{1}{5} = 0.2$$

Bestem $p(E \mid F)$.

Betingede sandsynlighed

Fra betingede sandsynlighed ved vi at

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} \quad \Rightarrow \quad p(E \mid F)p(F) = p(E \cap F)$$

Betingede sandsynlighed

Fra betingede sandsynlighed ved vi at

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} \quad \Rightarrow \quad p(E \mid F)p(F) = p(E \cap F)$$

Tilsvarende ved vi at

$$p(F \mid E) = \frac{p(F \cap E)}{p(E)} \quad \Rightarrow \quad \underline{p(F \mid E)p(E)} = \underline{p(F \cap E)} = \underline{p(E \cap F)}$$

Betingede sandsynlighed

Fra betingede sandsynlighed ved vi at

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} \quad \Rightarrow \quad p(E \mid F)p(F) = p(E \cap F)$$

Tilsvarende ved vi at

$$p(F \mid E) = \frac{p(F \cap E)}{p(E)} \quad \Rightarrow \quad p(F \mid E)p(E) = p(F \cap E) = p(E \cap F)$$

Derfor er

$$p(E \mid F)p(F) = p(F \mid E)p(E).$$

Betingede sandsynlighed

Fra betingede sandsynlighed ved vi at

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} \quad \Rightarrow \quad p(E \mid F)p(F) = p(E \cap F)$$

Tilsvarende ved vi at

$$p(F \mid E) = \frac{p(F \cap E)}{p(E)} \quad \Rightarrow \quad p(F \mid E)p(E) = p(F \cap E) = p(E \cap F)$$

Derfor er

$$p(E \mid F)p(F) = p(F \mid E)p(E).$$

Ved at isolere $p(F \mid E)$ fås:

$$\boxed{p(F \mid E) = \frac{p(E \mid F)\, p(F)}{p(E)}}$$

Dette udtryk er en version af Bayes' sætning.

Betingede sandsynlighed

Fra betingede sandsynlighed har vi også at

$$p(E \cap F) = p(E \mid F)p(F)$$

for enhver hændelse F.

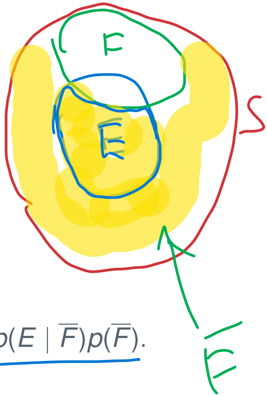
Betingede sandsynlighed

Fra betingede sandsynlighed har vi også at

$$p(E \cap F) = p(E \mid F)p(F)$$

for enhver hændelse F.
Derfor tilsvarende at $p(E \cap \overline{F}) = p(E \mid \overline{F})p(\overline{F})$.

Betingede sandsynlighed



Fra betingede sandsynlighed har vi også at

$$p(E \cap F) = p(E \mid F)p(F)$$

for enhver hændelse F.
Derfor tilsvarende at $p(E \cap \overline{F}) = p(E \mid \overline{F})p(\overline{F})$.
Desuden gælder at

$$E = (E \cap F) \cup (E \cap \overline{F})$$

fra forening af disjunkte mængder, og derfor at

$$p(E) = p(E \cap F) + p(E \cap \overline{F}) = p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F}).$$

$$(E \cap F) \cap (E \cap \bar{F}) = \emptyset$$

Betingede sandsynlighed

Fra betingede sandsynlighed har vi også at

$$p(E \cap F) = p(E \mid F)p(F)$$

for enhver hændelse F.
Derfor tilsvarende at $p(E \cap \overline{F}) = p(E \mid \overline{F})p(\overline{F})$.
Desuden gælder at

$$E = (E \cap F) \cup (E \cap \overline{F})$$

fra forening af disjunkte mængder, og derfor at

$$p(E) = p(E \cap F) + p(E \cap \overline{F}) = p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F}).$$

Indsættes dette i ovenstående “Bayes’ sætning” fås

$$\boxed{p(F \mid E) = \frac{p(E \mid F)\, p(F)}{p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F})}}$$

Som er Bayes’ sætning.

Betingede sandsynlighed

Eksempel (Forsat)

For at bestemme sandsynligheden for at en eksaminand kender svaret, givet at vedkommende har sat kryds ved det rigtige svar, udregnes

$$
\begin{aligned}
p(E \mid F) &= \frac{p(F \mid E)\, p(E)}{p(F \mid E)p(E) + p(F \mid \overline{E})p(\overline{E})} \\
&= \frac{1 \cdot 0.7}{1 \cdot 0.7 + 0.2 \cdot 0.3} = \frac{0.7}{0.76} \approx 0.921
\end{aligned}
$$

Bayes' sætning

Den betingede sandsynlighed for E givet F, $p(E \mid F)$, er sandsynligheden for at hændelsen E indtræffer givet at vi at hændelsen F indtræffer.

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Sætning (Bayes')

$$p(F \mid E) = \frac{p(E \mid F)\, p(F)}{p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F})}$$

Bayesiansk spamfilter

- Lad $p(w)$ være sandsynligheden for at en spam mail indeholder ordet w.
- Lad $q(w)$ være sandsynligheden for at en ikke-spam mail indeholder ordet w.
- Lad S være hændelsen: en indkommende mail er spam.
- Lad E være hændelsen: en indkommende mail indholder ordet w.

Så er $p(w) \approx p(E \mid S)$ og $q(w) \approx p(E \mid \overline{S})$.

Bayesiansk spamfilter

Fra Bayes' sætning (hvor F erstattes af S):

$$p(S \mid E) = \frac{p(E \mid S)\, p(S)}{p(E \mid S)p(S) + p(E \mid \overline{S})p(\overline{S})}.$$

På baggrund af (manglende) erfaring antager vi $p(S) = p(\overline{S}) = \frac{1}{2}$.

Bayesiansk spamfilter

Fra Bayes' sætning (hvor F erstattes af S):

$$p(S \mid E) = \frac{p(E \mid S)\, p(S)}{p(E \mid S)p(S) + p(E \mid \overline{S})p(\overline{S})}.$$

På baggrund af (manglende) erfaring antager vi $p(S) = p(\overline{S}) = \frac{1}{2}$. Indsættes dette får vi

$$\begin{aligned} \boxed{p(S \mid E)} &= \frac{p(E \mid S) \cdot \frac{1}{2}}{p(E \mid S) \cdot \frac{1}{2} + p(E \mid \overline{S}) \cdot \frac{1}{2}} \\ &= \frac{p(E \mid S)}{p(E \mid S) + p(E \mid \overline{S})} \approx \boxed{\frac{p(w)}{p(w) + q(w)}} \end{aligned}$$

Bayesiansk spamfilter

Fra Bayes' sætning (hvor F erstattes af S):

$$p(S \mid E) = \frac{p(E \mid S)\, p(S)}{p(E \mid S)p(S) + p(E \mid \overline{S})p(\overline{S})}.$$

På baggrund af (manglende) erfaring antager vi $p(S) = p(\overline{S}) = \frac{1}{2}$. Indsættes dette får vi

$$\begin{aligned} p(S \mid E) &= \frac{p(E \mid S) \cdot \frac{1}{2}}{p(E \mid S) \cdot \frac{1}{2} + p(E \mid \overline{S}) \cdot \frac{1}{2}} \\ &= \frac{p(E \mid S)}{p(E \mid S) + p(E \mid \overline{S})} \approx \frac{p(w)}{p(w) + q(w)} \end{aligned}$$

Sandsynligheden $\frac{p(w)}{p(w)+q(w)}$, estimerer sandsynligheden for, at en besked er spam givet at beskeden indeholder ordet w.

Bayesiansk spamfilter

Eksempel

Lad w=tilbud. Vi laver en undersøgelse, som viser at

- *ud af* 1000 *spam mails forekommer ordet "tilbud"* 200 *gange,*
- *ud af* 1000 *ikke-spam mails forekommer ordet "tilbud"* 20 *gange.*

så er

$$p(w) = \frac{200}{1000} = 0.2 \quad og \quad q(w) = \frac{20}{1000} = 0.02$$

Bayesiansk spamfilter

Eksempel

Lad w=tilbud. Vi laver en undersøgelse, som viser at

- *ud af* 1000 *spam mails forekommer ordet "tilbud"* 200 *gange,*
- *ud af* 1000 *ikke-spam mails forekommer ordet "tilbud"* 20 *gange.*

så er

$$p(w) = \frac{200}{1000} = 0.2 \quad og \quad q(w) = \frac{20}{1000} = 0.02$$

Heraf; når vi modtager en mail med ordet "tilbud", så er sandsynligheden for at det er spam

$$\frac{p(w)}{p(w) + q(w)} = \frac{0.2}{0.2 + 0.02} = \frac{0.2}{0.22} \approx 0.9091$$