Rozpoznávanie obrazcov
šk.r. 2017-18

Lies, Damned Lies, and Statistics

Zuzana Berger Haladová

Definícia pravdepodobnosti

- Definícia klasická $P(A) = \frac{N_A}{N}$.
- Definícia limitná $P(A) = \lim_{N \to \infty} \frac{N_A}{N}$.
- Výberový priestor
- Javy a ich zjednotenie, prienik,
- Rozklad výberového priestoru
- Bayes

Príklad 1

- Aká je pravdepodobnosť, že pri hode kockou:
  - Padne číslo X
  - Padne nepárne číslo
- A – padne číslo 4
- B – padne nepárne číslo
- m – počet priaznivých elementárnych javov
- n – celkový počet elementárnych javov

Príklad 1 pokračovanie

- Aká je pravdepodobnosť, že pri hode kockou:
  - Padne číslo X
  - Padne nepárne číslo
- P(A) = m/n = 1/6 = 0,167
- P(B) = m/n = 3/6 = 0,5

Príklad 2

- Dodávka obsahuje 50 matíc a 150 skrutiek.
- Polovica matíc a polovica skrutiek je hrdzavá.
- Ak náhodne vyberieme jednu súčiastku, aká je pravdepodobnosť, že to bude matica alebo, že to bude hrdzavá súčiastka?

Príklad 2 pokračovanie

- Pokus má 200 rovnako možných výsledkov.
- Ak označíme ako udalosť
  - A súčiastka je hrdzavá 100
  - B súčiastka je matica 50
  - $A \cap B$ súčiastka je hrdzavá matica 25
- potom

Príklad 2 pokračovanie II

- Potom pre udalosť $A \cup B$, že súčiastka je hrdzavá alebo matica platí
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 100/200 + 50/200 - 25/200 = 0.625$

Podmienená pravdepodobnosť

- Pravdepodobnosť, že nastane jav A za podmienky, že nastal jav B

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

Príklad 3

- V rodine sú dve deti.
- Určite pravdepodobnosť, že obe deti sú chlapci, ak vieme, že jedno z detí je chlapec.
- Výberový priestor

$$\Omega = \{(ch;ch);(ch;d);(d;ch);(d;d)\}$$

Príklad 3 pokračovanie

- A - náhodná udalosť „obe deti sú chlapci“
- A = {(ch;ch)}
- B - náhodná udalosť „aspoň jedno dieťa je chlapec“
- B ={(ch;ch);(ch;d);(d;ch)}

- P(A) =1/4 P(B)=3/4
- $P(A|B) = P(A \cap B)/P(B) =$

  $= P(A)/P(B) = ¼ / ¾ = 1/3$

Veta o úplnej pravdepodobnosti

- Majme systém disjunktných javov $B_1, B_2, \dots, B_n$, ktoré tvoria rozklad výberového priestoru
- Potom pravdepodobnosť $P(A)$ javu A sa dá určiť z pravdepodobností $P(B_k)$ a pod-mienených pravdepodobností $P(A|B_k)$ tak, že $P(A) = \ P(A|B_k).\ P(B_k)$

Príklad 1

- Elektrické žiarovky sú vyrábané 3 závodmi. Prvý dodáva 25%, druhý 40% a tretí 35% celej dodávky. Z celkovej produkcie 1. závodu je 88%, 2. závodu 75%, 3. závodu 85% štandardných. Aká je pravdepodob-nosť, že náhodne vybraná žiarovka je štandardná?

Príklad 1 - riešenie

- B_i - žiarovka bola vyrobená v *i*–tom závode, $P(B_1) = 0{,}25$, $P(B_2) = 0{,}4$ a $P(B_3) = 0{,}35$
- A je jav, že žiarovka je štandardná, $P(A|B_1) = 0{,}88, P(A|B_2) = 0{,}75, P(A|B_3) = 0{,}85$
- $P(A) = \sum_{i=1}^{3} P(A|B_i).P(B_i)$
- $P(A) = 0{,}88.0{,}25 + 0{,}75.0{,}4 + 0{,}85.0{,}35$
- $P(A) = 0{,}8175$

Príklad 2

- Elektrónka zapojená v televízore môže byť od 4 rôznych výrobcov s pravdepodobnosťami 0,2; 0,3; 0,35; 0,15 pre jednotlivých výrobcov. Pravdepodobnosti, že elektrónky od jednotlivých výrobcov vydržia predpísaný počet hodín sú postupne 0,45; 0,60; 0,75; 0,30. Aká je pravdepo-dobnosť, že náhodne vybratá elektrónka vydrží predpísaný počet hodín?

Príklad 2 - riešenie

- B_i - elektrónka od i–teho výrob., $P(B_1) = 0{,}2$ $P(B_2) = 0{,}3, P(B_3) = 0{,}35$ a $P(B_4) = 0{,}15$
- A – elektrónka vydrží predpísaný počet hodín $P(A|B_1) = 0{,}45, P(A|B_2) = 0{,}6, P(A|B_3) = 0{,}75$ a $P(A|B_4) = 0{,}3$
- $P(A) = \sum_{i=1}^{4} P(A|B_i).P(B_i)$
- $P(A) = 0{,}45.0{,}2 + 0{,}6.0{,}3 + 0{,}75.0{,}35 + 0{,}3.0{,}15$
- $P(A) = 0{,}5775$

Príklad 3

- V terénnej súťaži spoľahlivosti zostalo 10 motocyklov prvej série, ktoré vykazovali spoľahlivosť jazdy 85%, 8 motocyklov druhej série, ktoré vykazovali spoľahli-vosť 75%, 5 motocyklov tretej série, ktoré vykazovali spoľahlivosť 60%. Aká je pravdepodobnosť, že nasledujúci deň náhodne vybratý motocykel nedošiel?

Príklad 3 - riešenie

- B_i - motocykle *i*–tej série, $P(B_1) = 10/23$, $P(B_2) = 8/23$ a $P(B_3) = 5/23$
- A – spoľahlivosť motocykla, $P(A|B_1) = 0{,}85, P(A|B_2) = 0{,}75, P(A|B_3) = 0{,}60$
- $P(A) = \sum_{i=1}^{3} P(A|B_i).P(B_i)$
- $P(A) = 0{,}85.\frac{10}{23} + 0{,}75.\frac{8}{23} + 0{,}6.\frac{5}{23}$
- $P(A) = 0{,}76$
- $1 - P(A) = 0{,}24$

Príklad 4

- V pivnici je tma. Kompóty sú poukladané na policiach. Na prvej polici je 6 čučoriedkových a 10 malinových kompótov, na druhej polici je 7 čučoriedkových a 5 malinových kompótov. Marienka preložila jeden kompót z prvej police na druhú. Janko si potom vybral jeden kompót z a) prvej police, b) druhej police. Aká je pravdepodobnosť, že

Príklad 4 – riešenie a)

- B_1 - 6 resp. 5 Č kompótov a 10 M kompótov na 1. polici
- B_2 - 6 Č kompótov a 10 resp. 9 M kompótov na 1. polici

$$P(B_1) = \frac{6}{16} \text{ a } P(B_2) = \frac{10}{16},$$

- A – bol vybraný malinový kompót z prvej police,

$$P(A|B_1) = \frac{10}{15}, \; P(A|B_2) = \frac{9}{15}$$

- $P(A) = \sum_{i=1}^{2} P(A|B_i).P(B_i)$
- $P(A) = \frac{10}{15} . \frac{6}{16} + \frac{9}{15} . \frac{10}{16}$
- $P(A) = 0{,}625$

Príklad 4 – riešenie b)

- B_1 - 6 resp. 5 Č kompótov a 10 M kompótov na 1. polici
- B_2 - 6 Č kompótov a 10 resp. 9 M kompótov na 1. polici

$$P(B_1) = \frac{6}{16} \text{ a } P(B_2) = \frac{10}{16},$$

- A – bol vybraný malinový kompót z druhej police,

$$P(A|B_1) = \frac{5}{13}, \; P(A|B_2) = \frac{6}{13}$$

- $P(A) = \sum_{i=1}^{2} P(A|B_i).P(B_i)$
- $P(A) = \frac{5}{13}.\frac{6}{16} + \frac{6}{13}.\frac{10}{16}$
- $P(A) = 0{,}432$

Príklad 5

- Katarína A., Lenka B. a Monika C. sa spolu učia na maturitu. Z 30 maturitných otázok z angličtiny sa však stihli naučiť len prvých 20. Aká je pravdepodobnosť, že a) Lenka b) Monika si vytiahne otázku, ktorú vie, ak odpovedajú v poradí podľa abecedy a každá musí mať inú otázku?

Príklad 5 – riešenie a)

- B_1 - Katka vytiahne otázku, ktorú vie
- B_2 - Katka vytiahne otázku, ktorú nevie

$$P(B_1) = \frac{20}{30} \text{ a } P(B_2) = \frac{10}{30},$$

- A – Lenka si vytiahne otázku, ktorú vie

$$P(A|B_1) = \frac{19}{29},\ P(A|B_2) = \frac{20}{29}$$

- $P(A) = \sum_{i=1}^{2} P(A|B_i).P(B_i)$
- $P(A) = \frac{19}{29}.\frac{2}{3} + \frac{20}{29}.\frac{1}{3}$
- $P(A) = 0{,}666$

Príklad 5 – riešenie b)

- B_1 - ani jedna si nevytiahne otázku, ktorú vie
- B_2 - práve jedna si vytiahne otázku, ktorú vie
- B_3 - obe si vytiahnu otázku, ktorú vedia

$$P(B_1) = \frac{10}{30} . \frac{9}{29} = \frac{90}{870}$$

$$P(B_2) = \frac{20}{30} . \frac{10}{29} + \frac{10}{30} . \frac{20}{29} = \frac{400}{870}$$

$$P(B_3) = \frac{20}{30} . \frac{19}{29} = \frac{380}{870}$$

Príklad 5 – riešenie b) pokrač.

- A – Monika si vytiahne otázku, ktorú vie

$$P(A|B_1) = \frac{20}{28},\; P(A|B_2) = \frac{19}{28},$$
$$P(A|B_3) = \frac{18}{28}$$

- $P(A) = \sum_{i=1}^{3} P(A|B_i).P(B_i)$
- $P(A) = \frac{20}{28}.\frac{9}{87} + \frac{19}{28}.\frac{40}{87} + \frac{18}{28}.\frac{38}{87}$
- $P(A) = 0{,}666$

Bayesov vzorec

- $\mathrm{P(A_k)}$ – apriórna pravdepodobnosť objavenia sa udalosti $\mathrm{A_k}$
- $\mathrm{P(B|A_k)}$ - podmienená pravdepodob-nosť B za podmienky, že $\mathrm{A_k}$
- Nás zaujíma opačný vzťah

$$P(A_k|B) = \frac{P(B|A_k).P(A_k)}{P(B)}$$

Príklad 8

- Karolovi urobili v rámci prehliadky Rtg hrudníku, ktorý sa vrátil s pozitívnym nálezom na rakovinu pľúc. Aká je pravdepodobnosť, že Karol má rakovinu, ak vieme, že test je falošne pozitívny v 5% prípadov a falošne negatívny v 40% prípadov. A tiež vieme, že iba jeden z 500 zamestnancov v Karolovej pozícii má rakovinu.

Príklad 8 - riešenie

Snímka/Skutočnosť	Zdravý pacient	Chorý pacient
Pozitívna snímka (tvrdí, že je chorý)	0,05	0,6
Negatívna snímka (tvrdí, že je zdravý)	0,95	0,4

Príklad 8 - riešenie

- A_1 - chorý, A_2 - zdravý
- $P(A_1) = 0{,}002$ a $P(A_2) = 0{,}998$
- B – snímok bol pozitívny
- $P(B|A_1) = 0{,}6$ a $P(B|A_2) = 0{,}05$
- $P(A_1|B) = \dfrac{0{,}6.0{,}002}{0{,}6.002+0{,}998.0{,}05} = 0{,}024$
- To je 2,4%

Príklad 9

- Rovnaký test ako Karol podstúpil aj Peter a tiež mu vyšiel pozitívny test. Peter však 20 rokov pracoval ako baník v banskom dole. Peter vie, že 15% jeho bývalých kolegov má rakovinu pľúc. Ak sa ostatné podmienky nezmenili, aká je pravdepo-dobnosť, že rakovinu má aj Peter?

Príklad 9 - riešenie

-
- A_1- chorý, A_2- zdravý
- $P(A_1) = 0{,}15$ a $P(A_2) = 0{,}85$
- B – snímok bol pozitívny
- $P(B|A_1) = 0{,}6$ a $P(B|A_2) = 0{,}05$
- $P(A_1|B) = \dfrac{0{,}6.0{,}15}{0{,}6.0{,}15+0{,}05.0{,}85} = 0{,}679$
- To je 67,9%

Príklad 10

- Predpokladajme, že máme školu so 60% chlapcami a 40% dievčatami. Všetci chlapci nosia nohavice a z dievčat nosí nohavice polovica. Pozorovateľ vidí z diaľky študenta v nohaviciach. Aká je pravdepodobnosť, že ten náhodný študent je dievča?

Príklad 10 - riešenie

-
- A_1 - chlapci, A_2 - dievčatá
- $P(A_1) = 0{,}6$ a $P(A_2) = 0{,}4$
- B – študent/študentka má nohavice
- $P(B|A_1) = 1$ a $P(B|A_2) = 0{,}5$
- $P(A_2|B) = \dfrac{0{,}5.0{,}4}{0{,}5.0{,}4+0{,}6} = 0{,}25$

Náhodná premenná

- Funkcia, ktorej hodnota je určená výsledkom náhodného pokusu
- Priraďuje číselnú hodnotu každému elementárnemu javu
- Distribučná funkcia - opisuje rozdelenie pravdepodobnosti náhodnej premennej definovanej na pravdepodobnostnom priestore

Príklad 11

- Máme 4 stroje (nezávislé na sebe)
  - 1. sa pokazí s pravdepodobnosťou 10%   P(S1)
  - 2. s pravdepodobnosťou 15%   P(S2)
  - 3. s pravdepodobnosťou 30%   P(S3)
  - 4. s pravdepodobnosťou 50%   P(S4)
- Náhodná premenná X označuje počet pokazených strojov.
- Popíšte túto náhodnú premennú

Príklad 11 pokračovanie

- P(X=0) = P(S1 C ∩ S2 C∩ S3 C∩ S4 C) =
  = 0.9*0.85*0.7*0.5 = 0.26775
- P(X = 1) = P(S1 ∩ S2 C ∩ S3 C ∩ S4 C) ∪
  P(S1 C ∩ S2 ∩ S3 C ∩ S4 C) ∪
  P(S1 C ∩ S2 C ∩ S3 ∩ S4 C) ∪
  P(S1 C ∩ S2 C ∩ S3 C ∩ S4) = 0.4595
- P(X=2) = 0.23
- P(X=3) = 0.0405
- P(X=4) = 0.00225