

Rozpoznávanie Obrazcov

2. Cvičenie - Štatistika

Ing. Viktor Kocur, PhD.

22.2.2022

Klasické definície





Klasická

Pravdepodobnosť javu A je jeho relatívna početnosť v pokusoch:

$$P(A) = \frac{N_A}{N}$$

Limitná

Pravdepodobnosť javu A je jeho relatívna početnosť v nekonečne pokusov:

$$P(A) = \lim_{n \to \infty} \frac{N_A}{N}$$

Axiomatická definícia - sigma-algebra





Pole javov \mathcal{A} je σ -algebra na Ω teda preň platia:

$$\Omega \in \mathcal{A}$$
 (1)

$$A \in \mathcal{A} \implies A^c \in \mathcal{A}$$
 (2)

$$(\forall n \in \mathbb{N})(A_n \in \mathcal{A}) \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$$
 (3)

Elementárny jav

Elementárny jav je taký jav z Ω , ktorý sa nedá rozložiť na iné javy.

Axiomatická definícia - pravdepodobnosť



Pravdepodobnosť

Pravdepodobnosť je funkcia $P: A \rightarrow \langle 0, 1 \rangle$, taká že spĺňa:

$$P(\Omega) = 1 \tag{4}$$

$$P\left(\bigcup_{n\in\mathbb{N}}A_n\right)=\sum_{n\in\mathbb{N}}P(A_n),\tag{5}$$

ak A_n je postupnosť po dvoch disjunktných javov.

Pravdepodobnostný priestor

Pravdepodobnostný priestor je trojica (Ω, \mathcal{A}, P) .

Operácie





Zjednotenie

Zjednotenie $A \cup B$ nastane ak nastane aspoň jeden z javov.

Prienik

Prienik $A \cap B$ nastane ak nastanú oba javy.

Opačný jav

Opačný jav k A je A^c .

Nezlučiteľ né javy

Ak $A \cap B = \emptyset$ tak sú to nezlučiteľ né javy.



- Padne číslo X
 - počet priaznivých elementárnych javov
 - celkový počet elementárnych javov
- Padne nepárne číslo
 - počet priaznivých elementárnych javov
 - celkový počet elementárnych javov



- Padne číslo X
 - počet priaznivých elementárnych javov = 1
 - celkový počet elementárnych javov = 6
- Padne nepárne číslo
 - počet priaznivých elementárnych javov
 - celkový počet elementárnych javov



- Padne číslo $X = \frac{1}{6}$
 - počet priaznivých elementárnych javov = 1
 - celkový počet elementárnych javov = 6
- Padne nepárne číslo
 - počet priaznivých elementárnych javov
 - celkový počet elementárnych javov





- Padne číslo $X = \frac{1}{6}$
 - počet priaznivých elementárnych javov = 1
 - celkový počet elementárnych javov = 6
- Padne nepárne číslo
 - počet priaznivých elementárnych javov = 3
 - celkový počet elementárnych javov = 6





- Padne číslo $X = \frac{1}{6}$
 - počet priaznivých elementárnych javov = 1
 - celkový počet elementárnych javov = 6
- Padne nepárne číslo $=\frac{1}{2}$
 - počet priaznivých elementárnych javov = 3
 - celkový počet elementárnych javov = 6







- Označme A súčiastka je hrdzavá 100.
- Označme *B* súčiastka je matica 50.



- Označme A súčiastka je hrdzavá 100.
- Označme *B* súčiastka je matica 50.
- Označme $A \cap B$ súčiastka je hrdzavá matica 25.



- Označme A súčiastka je hrdzavá 100.
- Označme *B* súčiastka je matica 50.
- Označme $A \cap B$ súčiastka je hrdzavá matica 25.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{100}{200} + \frac{50}{200} - \frac{25}{200} = \frac{5}{8}$$

Podmienená pravdepodobnosť





Podmienená pravdepodobnosť

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

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



Ovca porodila dve jahniatka. Určite pravdepodobnosť, že obe sú barani, ak vieme, že jedno z detí je baran.





Ovca porodila dve jahniatka. Určite pravdepodobnosť, že obe sú barani, ak vieme, že jedno z detí je baran.

$$\blacksquare \Omega = \{(b,b),(b,o),(o,b),(o,o)\}$$





Ovca porodila dve jahniatka. Určite pravdepodobnosť, že obe sú barani, ak vieme, že jedno z detí je baran.

- \square $\Omega = \{(b,b),(b,o),(o,b),(o,o)\}$
- Označme A obe deti sú barani $\{(b,b)\}$.
- Označme B aspoň jedno dieť a je baran $\{(b,b),(b,o),(o,b)\}$





Ovca porodila dve jahniatka. Určite pravdepodobnosť, že obe sú barani, ak vieme, že jedno z detí je baran.

- \square $\Omega = \{(b,b),(b,o),(o,b),(o,o)\}$
- Označme A obe deti sú barani $\{(b,b)\}$.
- Označme B aspoň jedno dieť a je baran $\{(b,b),(b,o),(o,b)\}$
- $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$

Veta o úplnej pravdepodobnosti



Rozklad výberového priestoru

Množiná navzájom disjunktných javov $\{B_1,...,B_n\}$ z Ω tvorí rozklad výberového priestoru ak $\cup_{i\in\hat{n}}B_i=\Omega$.

Veta o úplnej pravdepodobnosti

Nech $\{B_1,...,B_n\}$ tvorí rozklad výberového priestoru Ω . Porom pre jav $A\in\Omega$ platí:

$$P(A) = \sum_{k \in \hat{n}} P(A|B_i)P(B_i)$$





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 pravdepodobnosť, že náhodne vybraná žiarovka je štandardná?

 \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tom závode



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tom závode
- $P(B_1) = 0.25, P(B_2) = 0.4, P(B_3) = 0.35$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tom závode
- $P(B_1) = 0.25, P(B_2) = 0.4, P(B_3) = 0.35$
- $P(A|B_1) = 0.88, P(A|B_2) = 0.75, P(A|B_3) = 0.85$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tom závode
- $P(B_1) = 0.25, P(B_2) = 0.4, P(B_3) = 0.35$
- $P(A|B_1) = 0.88, P(A|B_2) = 0.75, P(A|B_3) = 0.85$
- $P(A) = \sum_{k \in \hat{3}} P(A|B_i) P(B_i)$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tom závode
- $P(B_1) = 0.25, P(B_2) = 0.4, P(B_3) = 0.35$
- $P(A|B_1) = 0.88, P(A|B_2) = 0.75, P(A|B_3) = 0.85$
- $P(A) = \sum_{k \in \hat{3}} P(A|B_i) P(B_i)$
- $P(A) = 0.88 \cdot 0.25 + 0.75 \cdot 0.4 + 0.85 \cdot 0.35 = 0.8175$







 \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tym výrobcom



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tym výrobcom
- $P(B_1) = 0.2, P(B_2) = 0.3, P(B_3) = 0.35, P(B_4) = 0.15$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tym výrobcom
- $P(B_1) = 0.2, P(B_2) = 0.3, P(B_3) = 0.35, P(B_4) = 0.15$
- $P(A|B_1) = 0.45, P(A|B_2) = 0.6, P(A|B_3) = 0.75, P(A|B_4) = 0.3$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tym výrobcom
- $P(B_1) = 0.2, P(B_2) = 0.3, P(B_3) = 0.35, P(B_4) = 0.15$
- $P(A|B_1) = 0.45, P(A|B_2) = 0.6, P(A|B_3) = 0.75, P(A|B_4) = 0.3$
- $P(A) = \sum_{k \in \hat{A}} P(A|B_i) P(B_i)$



- \blacksquare B_i je pravdepodonosť že žiarovka bola vyrobená v i-tym výrobcom
- $P(B_1) = 0.2, P(B_2) = 0.3, P(B_3) = 0.35, P(B_4) = 0.15$
- $P(A|B_1) = 0.45, P(A|B_2) = 0.6, P(A|B_3) = 0.75, P(A|B_4) = 0.3$
- $P(A) = \sum_{k \in \hat{A}} P(A|B_i) P(B_i)$
- $P(A) = 0.45 \cdot 0.2 + 0.6 \cdot 0.3 + 0.75 \cdot 0.35 + 0.3 \cdot 0.15 = 0.5775$



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?



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?

Označíme ako A pravdepodobnosť že motorka dojde.



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

- Označíme ako A pravdepodobnosť že motorka dojde.
- \blacksquare B_i je spoľahlivosť motorky i-tej série.



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?

- Označíme ako A pravdepodobnosť že motorka dojde.
- \blacksquare B_i je spoľahlivosť motorky i-tej série.
- $P(A) = \sum_{k \in \hat{\mathfrak{Z}}} P(A|B_i) P(B_i)$





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?

- Označíme ako A pravdepodobnosť že motorka dojde.
- \blacksquare B_i je spoľahlivosť motorky i-tej série.

$$P(A) = \sum_{k \in \hat{3}} P(A|B_i) P(B_i)$$

$$P(A) = 0.85 \cdot \frac{8}{23} + 0.75 \cdot \frac{8}{23} + 0.6 \cdot \frac{5}{23} = 0.24$$





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?

- Označíme ako A pravdepodobnosť že motorka dojde.
- \blacksquare B_i je spoľahlivosť motorky i-tej série.

$$P(A) = \sum_{k \in \hat{3}} P(A|B_i) P(B_i)$$

$$P(A) = 0.85 \cdot \frac{8}{23} + 0.75 \cdot \frac{8}{23} + 0.6 \cdot \frac{5}{23} = 0.24$$

■ Výsledok =
$$P(A^c) = 1 - P(A) = 0.76$$





- \blacksquare B_1 5 čučoriedkových a 10 malinových na prvej polici
- \blacksquare B_2 6 čučoriedkových a 9 malinových na prvej polici



- \blacksquare B_1 5 čučoriedkových a 10 malinových na prvej polici
- B₂ 6 čučoriedkových a 9 malinových na prvej polici

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



- B_1 5 čučoriedkových a 10 malinových na prvej polici
- \blacksquare B_2 6 čučoriedkových a 9 malinových na prvej polici

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

$$P(A) = \frac{10}{15} \cdot \frac{6}{16} + \frac{9}{15} \cdot \frac{10}{16} = 0.625$$





- \blacksquare B_1 katka vytiahne otázku, ktorú vie
- \blacksquare B_2 katka vytiahne otázku, ktorú nevie



- B₁ katka vytiahne otázku, ktorú vie
- B₂ katka vytiahne otázku, ktorú nevie
- $P(A|B_1) = \frac{20}{30}, P = \frac{10}{30}$



- B₁ katka vytiahne otázku, ktorú vie
- B₂ katka vytiahne otázku, ktorú nevie

$$P(A|B_1) = \frac{20}{30}, P = \frac{10}{30}$$

$$P(A) = \frac{19}{29} \cdot \frac{20}{30} + \frac{20}{29} \cdot \frac{10}{30} = \frac{2}{3}$$

Bayesov vzorec



Bayesov vzorec

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



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.





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 pozícij má rakovinu.

- $P(A_n) = 0.002$ chorý, $P(A_n) = 0.998$ zdravý
- B pozitívny test



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.

- $P(A_p) = 0.002$ chorý, $P(A_n) = 0.998$ zdravý
- *B* pozitívny test
- $P(B|A_p) = 0.6, P(B|A_n) = 0.05$



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.

- $P(A_p) = 0.002$ chorý, $P(A_n) = 0.998$ zdravý
- B pozitívny test
- $P(B|A_p) = 0.6, P(B|A_n) = 0.05$
- $P(A_p|B) = \frac{0.6 \cdot 0.002}{0.6 \cdot 0.002 + 0.05 \cdot 0.998} = 0.024$



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?



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?

- $lacksquare P(A_p)=0.15$ chorý, $P(A_n)=0.85$ zdravý
- *B* pozitívny test



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?

- $P(A_p) = 0.15$ chorý, $P(A_n) = 0.85$ zdravý
- *B* pozitívny test
- $P(B|A_p) = 0.6, P(B|A_n) = 0.05$



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?

- $P(A_n) = 0.15$ chorý, $P(A_n) = 0.85$ zdravý
- B pozitívny test
- $P(B|A_p) = 0.6, P(B|A_p) = 0.05$
- $P(A_p) = \frac{0.6 \cdot 0.15}{0.6 \cdot 0.15 + 0.05 \cdot 0.85} = 0.679$





- \blacksquare $P(A_c)=0.6$ chlapec, $P(A_d)=0.4$ dievča
- B osoba má nohavice



- $lacksquare P(A_c) = 0.6$ chlapec, $P(A_d) = 0.4$ dievča
- B osoba má nohavice
- $P(B|A_c) = 1, P(B|A_d) = 0.5$



- $lacksquare P(A_c) = 0.6$ chlapec, $P(A_d) = 0.4$ dievča
- B osoba má nohavice
- $P(B|A_c) = 1, P(B|A_d) = 0.5$
- $P(A_p) = \frac{0.5 \cdot 0.4}{0.5 \cdot 0.4 + 0.6} = 0.25$

Náhodná premenná



Náhodná premenná

Funkcia, ktorej hodnota je určená výsledkom náhodného pokusu. Priraďuje číselnú hodnotu každému javu.

Distribučná funkcia

Opisuje rozdelenie pravdepodobnosti náhodnej premennej definovanej na pravdepodobnostnom priestore.



Máme 4 stroje (nezávislé na sebe): 1. sa pokazí s pravdepodobnosťou 10%, 2. s pravdepodobnosťou 15%, 3. s pravdepodobnosťou 30%, 4. stroj spravdepodobnosťou 50%. Náhodná premenná X označuje počet pokazených strojov. Popíšte túto náhodnú premennú. T.j. popíšte s akými pravdepodobnosťami bude mať jednotlivé hodtnoty $\{0,...,4\}$.



Máme 4 stroje (nezávislé na sebe): 1. sa pokazí s pravdepodobnosťou 10%, 2. s pravdepodobnosťou 15%, 3. s pravdepodobnosťou 30%, 4. stroj spravdepodobnosťou 50%. Náhodná premenná X označuje počet pokazených strojov. Popíšte túto náhodnú premennú. T.j. popíšte s akými pravdepodobnosťami bude mať jednotlivé hodtnoty $\{0,...,4\}$.

 $P(X=4) = 0.1 \cdot 0.15 \cdot 0.3 \cdot 0.5 = 0.00225$



Máme 4 stroje (nezávislé na sebe): 1. sa pokazí s pravdepodobnosťou 10%, 2. s pravdepodobnosťou 15%, 3. s pravdepodobnosťou 30%, 4. stroj spravdepodobnosťou 50%. Náhodná premenná X označuje počet pokazených strojov. Popíšte túto náhodnú premennú. T.j. popíšte s akými pravdepodobnosťami bude mať jednotlivé hodtnoty $\{0,...,4\}$.

$$P(X = 4) = 0.1 \cdot 0.15 \cdot 0.3 \cdot 0.5 = 0.00225$$

$$P(X = 0) = 0.9 \cdot 0.85 \cdot 0.7 \cdot 0.5 = 0.26775$$



Máme 4 stroje (nezávislé na sebe): 1. sa pokazí s pravdepodobnosťou 10%, 2. s pravdepodobnosťou 15%, 3. s pravdepodobnosťou 30%, 4. stroj spravdepodobnosťou 50%. Náhodná premenná X označuje počet pokazených strojov. Popíšte túto náhodnú premennú. T.j. popíšte s akými pravdepodobnosťami bude mať jednotlivé hodtnoty $\{0, ..., 4\}$.

- $P(X = 4) = 0.1 \cdot 0.15 \cdot 0.3 \cdot 0.5 = 0.00225$
- $P(X = 0) = 0.9 \cdot 0.85 \cdot 0.7 \cdot 0.5 = 0.26775$
- $P(X = 1) = P(S_1 \cap S_2^c \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2 \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2^c \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2^c \cap S_3^c \cap S_4^c) = 0.4595$



Máme 4 stroje (nezávislé na sebe): 1. sa pokazí s pravdepodobnosťou 10%, 2. s pravdepodobnosťou 15%, 3. s pravdepodobnosťou 30%, 4. stroj spravdepodobnosťou 50%. Náhodná premenná X označuje počet pokazených strojov. Popíšte túto náhodnú premennú. T.j. popíšte s akými pravdepodobnosťami bude mať jednotlivé hodtnoty $\{0,...,4\}$.

- $P(X = 4) = 0.1 \cdot 0.15 \cdot 0.3 \cdot 0.5 = 0.00225$
- $P(X=0) = 0.9 \cdot 0.85 \cdot 0.7 \cdot 0.5 = 0.26775$
- $P(X = 1) = P(S_1 \cap S_2^c \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2 \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2^c \cap S_3^c \cap S_4^c) + P(S_1^c \cap S_2^c \cap S_3^c \cap S_4^c) = 0.4595$
- P(X = 2) = 0.23, P(X = 3) = 0.0405