

Cvičení 4.: Pravidla pro počítání s pravděpodobnostmi, vzorec pro úplnou pravděpodobnost, Bayesův vzorec, nezávislé jevy

Pravidla pro počítání s pravděpodobnostmi

Pravděpodobnost nastoupení aspoň jednoho ze dvou jevů:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Jsou-li jevy A, B neslučitelné, pak $P(A \cup B) = P(A) + P(B)$

Pravděpodobnost společného nastoupení n nezávislých jevů:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \cdot \dots \cdot P(A_n)$$

Pravděpodobnost nastoupení opačného jevu:

$$P(\bar{A}) = 1 - P(A)$$

Vzorový příklad 1: V náhodně vybrané skupině 140 mužů ve věku 40 - 50 let ohrožených arteriální hypertenzí se vyskytl rizikový faktor "zvýšený cholesterol" ve 37 případech a rizikový faktor "kouření" v 98 případech. Ve 31 případech jsme zjistili současný výskyt obou rizikových faktorů. Odhadněte pomocí relativních četností pravděpodobnosti výskytu aspoň jednoho rizikového faktoru.

Řešení: Pravděpodobnost výskytu faktoru "zvýšený cholesterol" (označme jako jev A) je odhadnuta jako relativní četnost $p(A) = \frac{37}{140}$ a faktoru "kouření" (označme jako jev B) jako

relativní četnost $p(B) = \frac{98}{140}$. Odhad pravděpodobnosti současného výskytu obou faktorů je

relativní četnost $p(A \cap B) = \frac{31}{140}$. Odhad pravděpodobnosti výskytu "zvýšeného

cholesterolu" nebo "kouření" je $p(A \cup B) = \frac{37}{140} + \frac{98}{140} - \frac{31}{140} = \frac{104}{140} = 0,743$.

S pravděpodobností 74,3 % tedy můžeme očekávat, že u náhodně vybraného muže ve věku 40 - 50 let zjistíme, že má zvýšenou hladinu cholesterolu nebo kouří.

Vzorový příklad 2: Dva střelci střílejí nezávisle na sobě do terče. Pravděpodobnost zásahu je pro prvního střelce 0,9 a pro druhého střelce 0,8. Jaká je pravděpodobnost, že se netrefí ani jeden?

Řešení: A ... jev, že první střelec zasáhne terč, B ... jev, že druhý střelec zasáhne terč.

Počítáme $P(\bar{A} \cap \bar{B}) = P(\bar{A})P(\bar{B}) = [1 - P(A)][1 - P(B)] = (1 - 0,9)(1 - 0,8) = 0,02$

Vzorový příklad 3: Hodíme naráz 10 kostkami. Jaká je pravděpodobnost, že aspoň na jedné z nich padne šestka?

Řešení: Přejdeme k opačnému jevu, který spočívá v tom, že šestka nepadne vůbec.

$$P(A) = 1 - P(\bar{A}) = 1 - \left(\frac{5}{6}\right)^{10} = 0,8385$$

Vzorec pro úplnou pravděpodobnost a Bayesův vzorec

Jestliže H_1, \dots, H_n jsou jevy, které tvoří rozklad jistého jevu (tj. jsou neslučitelné a jejich sjednocením je celý základní prostor – říkáme, že tvoří úplný systém hypotéz), pak

pravděpodobnost libovolného jevu A lze vypočítat pomocí vzorce pro úplnou

$$\text{pravděpodobnost: } P(A) = \sum_{i=1}^n P(H_i)P(A/H_i)$$

Počítáme-li pravděpodobnost některé hypotézy za podmínky, že nastal jev A, použijme Bayesův vzorec:

$$P(H_k / A) = \frac{P(H_k)P(A/H_k)}{P(A)}, k = 1, \dots, n, \text{ kde } P(A) = \sum_{i=1}^n P(H_i)P(A/H_i).$$

Vzorový příklad 3.: Uvažme populaci mužů žijících v Praze. Označme jev H_1 ... muž má základní vzdělání, H_2 ... muž je vyučen, H_3 ... muž má středoškolské vzdělání, H_4 ... muž má vysokoškolské vzdělání. Pravděpodobnosti těchto jevů jsou postupně 0,12, 0,31, 0,32 a 0,25. Je A znamená, že muž kouří. Podmínené pravděpodobnosti kouření v závislosti na vzdělání jsou postupně 0,79, 0,72, 0,67 a 0,52. Zjistěte, jaká je pravděpodobnost výskytu kuřáka ve sledované populaci mužů.

Řešení:

Podle vzorce pro úplnou pravděpodobnost máme:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3) + P(H_4) \cdot P(A/H_4) =$$

$$= 0,12 \cdot 0,79 + 0,31 \cdot 0,72 + 0,32 \cdot 0,67 + 0,25 \cdot 0,52 = 0,66$$

Znamená to, že kuřáci tvoří asi 66 % sledované populace.

Vzorový příklad 4: U jistého druhu elektrického spotřebiče se s pravděpodobností 0,01 vyskytuje výrobní vada. U spotřebiče s touto výrobní vadou dochází v záruční lhůtě k poruše s pravděpodobností 0,5. Výrobky, které tuto vadu nemají, se v záruční lhůtě porouchají s pravděpodobností 0,01. Jaká je pravděpodobnost, že výrobek, který se v záruční lhůtě porouchá, bude mít dotýčnou výrobní vadu?

Řešení:

H_1 - výrobek má dotýčnou výrobní vadu

H_2 - výrobek nemá tuto výrobní vadu

A - výrobek se v záruční době porouchá

Pak je: $P(H_1) = 0,01$, $P(H_2) = 0,99$, $P(A/H_1) = 0,5$, $P(A/H_2) = 0,01$

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) = 0,01 \cdot 0,5 + 0,99 \cdot 0,01 = 0,0149$$

$$P(H_1 / A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{0,01 \cdot 0,5}{0,0149} = 0,3386$$

Příklady k samostatnému řešení:

1. Ve společnosti je 45% mužů a 55% žen. Výšku nad 190 cm má 5% mužů a 1% žen. Náhodně vybraná osoba je vyšší než 190 cm. Jaká je pravděpodobnost, že je to žena?

Návod: A ... osoba měří více než 190 cm, H_1 ... osoba je žena, H_2 ... osoba je muž.

Výsledek: 0,1964

2. Potřebu smrkových sazenic kryje lesní závod produkcí dvou školky. První školka kryje 75% výsadby, přičemž ze 100 sazenic je 80 první jakosti. Druhá školka kryje výsadbu z 25%, přičemž na 100 sazenic připadá 60 první jakosti. Jaká je pravděpodobnost, že náhodně vybraná sazenice první jakosti pochází z produkce první školky?

Návod: A ... sazenice je 1. jakosti, H_1 ... sazenice pochází z 1. školky, H_2 ... sazenice pochází z 2. školky.

Výsledek: 0,8

3. Uvažujeme soubor 10 195 dětí z několika okresů, které se narodily ve stejném roce. Během prvního roku života mělo zánět středního ucha 2 925 dětí, 7 941 dětí mělo infekční zánět horních cest dýchacích a u 2 761 dětí se vyskytly obě nemoci současně. Odhadněte pravděpodobnost, že náhodně vybrané dítě mělo zánět středního ucha nebo infekční zánět horních cest dýchacích.

Výsledek: 0,8.

4. V loterii je 1 000 losů, z nichž jeden vyhrává 1. cenu, pět 2. cenu a dvacet 3. cenu. Jaká je pravděpodobnost, že na zakoupený los vyhraje?

Výsledek: 0,026

5. Populace Kypřanů se skládá ze 75% Řeků a 25% Turků. Víme, že 20% Řeků a 10% Turků mluví anglicky. Jaká je pravděpodobnost, že náhodně vybraný Kypřan mluví anglicky?

Výsledek: 0,175

6. Přístroj se skládá ze tří částí, které pracují nezávisle na sobě. Ve sledovaném časovém intervalu je pravděpodobnost poruchy každé části 0,1. Jaká je pravděpodobnost, že

a) ani jedna součástka nebude mít poruchu

b) všechny části budou mít poruchu

c) právě jedna část bude mít poruchu

d) aspoň jedna část bude mít poruchu

Výsledek: ad a) 0,729, ad b) 0,001, ad c) 0,243, a d) 0,271