

BAYESOVA VĚTA, DISKRÉTNÍ NÁHODNÁ VELIČINA

15. – 19. 10. 2012

BAYESOVA VĚTA

- Z pošty doručené na server je 80 % spamů. Spamový filtr úspěšně rozpozná 90% všech spamů, ale zároveň 15 % korektní pošty je označeno jako spam.
 - Jaké je pravděpodobnost, že email smazaný filtrem jste si chtěli přečíst?
 - Jaké je procento spamů ve vaší schránce?
- Máme tři truhly se dvěma mincemi. V truhle A jsou dvě zlaté mince, v truhle B dvě stříbrné mince a v truhle C zlatá a stříbrná mince. Náhodně vybereme truhlu a z ní vytáhneme náhodně minci. Ta je zlatá. Jaká je pravděpodobnost, že i druhá mince v této truhle je zlatá?
- Tři lovci vystřelili současně na divokého kance, který byl jednou střelou trefen. Určete pravděpodobnost toho, že kance zastřelil první, druhý nebo třetí střelec, jsou-li pravděpodobnosti zásahu po řadě rovny 0.2, 0.4 nebo 0.6.

DISKRÉTNÍ NÁHODNÁ VELIČINA

- V peněžence máte dvě pětistovky, jednu tisícikorunu a jednu dvoutisícovou bankovku. Zloděj Vám z peněženky náhodně vybere dvě bankovky. Označme jako X náhodnou veličinu, která udává, o kolik peněz jste právě přišli.
 - Určete rozdělení X .
 - Spočtěte Vaši očekávanou ztrátu.
 - Spočtěte rozptyl veličiny X .
 - Nakreslete distribuční funkci veličiny X . Jaké jsou obecné vlastnosti distribuční funkce?
- Při přenosu binárního souboru se náhodně vybraný znak zkreslí s pravděpodobností $p = 3/14$ a jednotlivé znaky se zkreslují nezávisle na sobě. Náhodná veličina X udává počet zkreslených znaků v binární posloupnosti délky n .
 - Jaké je rozdělení X ?
 - Určete očekávaný počet zkreslených znaků v posloupnosti délky n .
 - Spočtěte rozptyl veličiny X .
- Test obsahuje n otázek, ke každé z nich jsou uvedeny 4 možnosti a, b, c, d. U každé otázky je právě jedna odpověď správná. Předpokládejme, že student zaškrťává odpovědi zcela náhodně. Označme X počet správně zodpovězených otázek.
 - Odvoďte rozdělení veličiny X . Jak se toto rozdělení nazývá?
 - Jaký je střední (očekávaný) počet správně zodpovězených otázek?
 - Jaký je rozptyl počtu správně zodpovězených otázek?

OPAKOVÁNÍ

BAYESOVA VĚTA:

Nechť A, B_1, B_2, \dots jsou náhodné jevy takové, že $B_i \cap B_j = \emptyset$ pro všechna $i \neq j$, $\bigcup_i B_i = \Omega$, $P(B_i) > 0$ pro všechna i a necht' $P(A) > 0$. Pak

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}.$$

NÁHODNÁ VELIČINA: **Náhodná veličina** X je měřitelné zobrazení (funkce) z prostoru (Ω, \mathcal{A}) do $(\mathbb{R}, \mathcal{B})$. Jednotlivým prvkům $\omega \in \Omega$ tedy přiřazuje reálná čísla $X(\omega)$.

- **Rozdělení** náhodné veličiny X

- popisuje pravděpodobnosti $P(X \in B) = P(\{\omega : X(\omega) \in B\})$ pro všechny množiny $B \in \mathcal{B}$,
- je jednoznačně určeno **distribuční funkcí**, která je funkcí reálné proměnné $x \in \mathbb{R}$ a je definovaná jako

$$F(x) = P(X \leq x).$$

- Základní charakteristiky veličiny X jsou

- **střední hodnota** EX , která vyjadřuje „očekávanou hodnotu“ veličiny X ,
- **rozptyl** veličiny X , který popisuje její variabilitu kolem EX . Rozptyl je definovaný jako

$$\text{Var } X = E(X - EX)^2 = EX^2 - (EX)^2$$

a je to vždy **nezáporné** číslo!

DISKRÉTNÍ ROZDĚLENÍ: Nabývá-li náhodná veličina X s kladnou pravděpodobností **nejvýše spočetně** mnoha (tj. konečně nebo spočetně) hodnot x_1, x_2, \dots , říkáme, že má **diskrétní rozdělení**.

- Rozdělení X je charakterizováno pravděpodobnostmi $p_k = P(X = x_k)$, $k = 1, 2, \dots$ a platí $\sum_k p_k = 1$.
- **Distribuční funkce** je po částech konstantní, skokovitá se skoky o velikosti p_k v bodech x_k .
- **Střední hodnota** X se spočítá jako

$$EX = \sum_k x_k P(X = x_k) = \sum_k x_k p_k \quad (\text{existuje-li}).$$