# **29 Pravděpodobnost a statistika – met.**

**Stručný přehled teorie**

**Pravděpodobnost**

**Náhodný pokus** – pokus, jehož výsledek záleží i při dodržení předem stanovených podmínek na náhodě.

Ω={ω1; ω2;….; ωn} … množina všech možných výsledků náhodného pokusu

**Náhodný jev** A je podmnožinou množiny Ω ..... AΩ

**Klasická (Laplaceova) definice pravděpodobnosti:**

Nechť náhodný pokus splňuje předpoklady:

1. Všech možných výsledků je konečný počet
2. Všechny výsledky mají stejnou šanci na realizaci
3. Všechny výsledky se navzájem vylučují (tj. žádné dva nemohou nastat současně)
4. Jeden z výsledků jistě nastane.

Pak pravděpodobností jevu A se nazývá číslo $P(A)=\frac{m}{n}$**,** kde *n* je počet všech možných výsledků (počet prvků Ω) a *m* je počet výsledků příznivých jevu A (počet prvků A).

**Některé vlastnosti pravděpodobnosti:**

1. a) *Jistý jev* … A= Ω => 

b) *Nemožný jev* … A= Ø => 

 c) *Náhodný jev* … A Ω => 0 ≤ P(A) ≤ 1

1. **Opačný jev** k jevu A je takový jev A‘, který nastává právě tehdy, když nenastal jev A.

Tedy: A‘ = Ω – A (přesněji A ∩ A’ = Ø ∧ A ∪ A’ = Ω)

Pak ****

1. **Sjednocení jevů** A, B je jev A ∪ B, který nastane právě tehdy, když nastane aspoň jeden z jevů A nebo B.

 a) Platí-li, že se jevy A, B navzájem vylučují (tj. A ∩ B = Ø), pak P(A∪B) = P(A) + P(B)

 b) Pokud se jevy A, B navzájem nevylučují (tj. A ∩ B ≠ Ø), pak P(A∪B) = P(A) + P(B) – P(A∩B)

1. **Průnik jevů** A, B je jev A ∩ B, který nastane právě tehdy, když nastane jev A a zároveň jev B. Jestliže A, B jsou **nezávislé jevy**, pak P(A∩B) = P(A).P(B) . Jsou-li A, B závislé jevy, pak P(A∩B) $\ne $ P(A).P(B)

**Statistická (zobecněná) definice pravděpodobnosti:**

Pravděpodobnost P(A) jevu A je určena přibližně jeho relativní četností při dostatečně velkém počtu opakování náhodného pokusu.

 Nechť Ω = {ω1; ω2;….; ωn} je množina všech možných výsledků náhodného pokusu a *p1, p2, … ,pn* jsou jejich relativní četnosti (tzn. ).

Pak P(A) = , kde  značí relativní četnost výsledku  A (*k* je počet prvků A).

**Podmíněná pravděpodobnost** – pravděpodobnost jevu A podmíněnou jevem B určíme takto:

 **P(A|B) = **

**Bernoulliho schéma:**

Nechť při n-násobném opakování náhodného pokusu je stále stejná pravděpodobnost zdaru *p*

a pravděpodobnost nezdaru *q (*tedy *q = 1 – p)*. Pak pravděpodobnost jevu Ak, že zdar nastane

v těchto n pokusech právě k-krát je dána vztahem: $P\left(A\_{k}\right)=\left(\begin{matrix}n\\k\end{matrix}\right).p^{k}.q^{n-k}$

**Základy statistiky**

**Statistika** se zabývá zkoumáním a zpracováním velkého množství dat souvisejících s hromadnými jevy.

### **Základní pojmy**

* **statistický soubor** – konečná neprázdná množina objektů, které zkoumáme (např. obyvatelé Brna, obyvatelé ČR, rodinní příslušníci zaměstnanců určité firmy, dopravní nehody v určité oblasti za určité období, výrobky vyrobené v určité firmě za určité období ,…)
* **statistická jednotka** – prvek statistického souboru (např. jeden určitý obyvatel, jeden daný výrobek)
* **rozsah statistického souboru** – počet prvků statistického souboru
* **statistický znak** – společná vlastnost prvků statistického souboru, kterou zjišťujeme (např. věk, národnost, výše měsíčního příjmu, výška postavy, kvalita výrobku (vadný nebo bez vady), …);

 **znak** může být - ***kvantitativní*** (číselný) – např. počet obyvatel daného věku, výše škody při nehodě, …

- ***kvalitativní*** (popsán slovy)– např. povolání, druh nemoci, příčina dopravní nehody, …

Pozn.: 1) *kvalitativní* znak může mít někdy více možností (např. příčin nehody může být víc) – pak se musí vybrat jedna, která je hlavní (ostatní mohou tvořit kategorii „jiné“);

Pozn.: 2) *nejjednodušší kvalitativní* znak je znak *alternativní* – dán jevem a jeho opakem – např. voják-nevoják, muž-žena, plavec-neplavec, prospěl-neprospěl, …

* **absolutní četnost** hodnoty znaku xi – číslo ni udávající počet prvků daného statistického souboru, které vykazují sledovanou hodnotu xi, neboli udávající, pro kolik prvků souboru nabývá statistický znak určité hodnoty nebo rozmezí hodnot (např. kolik nezaměstnaných osob je evidováno v dané oblasti, kolik osob má měsíční příjem ve vybraném rozmezí, …)
* **relativní četnost** znaku – poměr absolutní četnosti dané hodnoty a rozsahu souboru

Pozn.: Relativní četnost se nejčastěji uvádí v procentech .100 %

Statistické soubory rozdělujeme na - **základní** (mohou mít pro zkoumání příliš velký rozsah) - **výběrové** (část základního souboru, na němž se provádí zkoumání)

# **Charakteristiky statistického souboru**

1. **Charakteristiky polohy** hodnot znaku jsou číselné hodnoty, které určitým způsobem charakterizují typickou hodnotu sledovaného znaku
	* **Aritmetický průměr** – součet všech hodnot zjištěných znaků dělených jejich počtem.

  nebo tzv. vážený průměr $\overline{x}=\frac{1}{n}\sum\_{j=1}^{r}x\_{j}n\_{j}$

* **Modus Mod(x) – hodnota** znaku s největší četností.
* **Medián Med(x) – je** prostřední hodnota znaku, jsou-li hodnoty uspořádány podle velikosti.
 , je-li n liché, , je-li n sudé.
* **Geometrický průměr** 

## **Charakteristiky variability**

Každá charakteristika polohy je číslo, kolem něhož jednotlivé hodnoty znaku kolísají. Charakteristiky variability vyjadřují „velikost“ onoho kolísání.

* **Rozptyl** se definuje jako průměr druhých mocnin odchylek od aritmetického průměru 
* **Směrodatná odchylka** 

* **Variační koeficient** vx – podíl směrodatné odchylky a aritmetického průměru – udává se v procentech 
* **Koeficient korelace** r - souvisí s tím, že se velmi často zkoumá, zda a jak jsou na sobě závislé dva znaky *x* a *y*. Koeficient korelace vyjadřuje míru vzájemné závislosti těchto znaků *x* a *y*. , kde  ,  , 

 Pozn**.**: Vždy platí: . Čím víc se hodnota *r* blíží k **1**, tím považujeme závislost *x* a *y* za silnější. 

Met.: Když učitel probírá teorii k tématu Pravděpodobnost, přímo se mu nabízejí velmi jednoduché konkrétní situace, kterými by měl průběžně doplňovat svůj výklad, aby i nejslabší studenti okamžitě pochopili probírané pojmy, děje, jevy. Určitě k nejvhodnějším takovým situacím patří házení kostky ze hry Člověče, nezlob se.

Náhodný pokus Hod kostkou

Množina$Ω$ $Ω=\left\{1, 2, 3, 4, 5, 6\right\}$

Náhodný jev A $A⊆Ω, např. A=\left\{1, 2\right\}…padne číslo menší než 3$

- splněny předpoklady pro použití Laplaceovy definice pravděpodobnosti: 1) konečný počet prvků množiny $Ω$, 2) stejná šance pro všechny, 3) dva výsledky nemohou nastat současně, 4) jeden výsledek nastane jistě.

 $P\_{\left(A\right)}=\frac{m}{n}=\frac{\left|A\right|}{\left|Ω\right|}=\frac{2}{6}=\frac{1}{3}$

Jistý jev B B …padne číslo menší než 8 $B=Ω=\left\{1, 2, 3, 4, 5, 6\right\}$ $P\_{\left(B\right)}=\frac{m\_{B}}{n}=\frac{\left|B\right|}{\left|Ω\right|}=\frac{6}{6}=1$

Nemožný jev C C … padne číslo větší než 6 $C=∅$ $P\_{\left(C\right)}=\frac{m\_{C}}{n}=\frac{\left|C\right|}{\left|Ω\right|}=\frac{0}{6}=0$

Náhodný jev D D … padne sudé číslo $D=\left\{2, 4, 6\right\}$

 $P\_{\left(D\right)}=\frac{m\_{D}}{n}=\frac{\left|D\right|}{\left|Ω\right|}=\frac{3}{6}=\frac{1}{2}$ $0\leq P\_{\left(D\right)}\leq 1$

Opačný jev A´ k jevu A A … $padne číslo menší než 3 $ $A=\left\{1, 2\right\}$ A´… $padne číslo větší nebo rovné 3 A=\left\{3, 4, 5, 6\right\}$ Platí: $(A\bigcap\_{}^{}A^{´}=∅) \bigwedge\_{}^{} (A\bigcup\_{}^{}A^{´}=Ω) .$ $P\_{\left(A\right)}=\frac{2}{6}=\frac{1}{3}; P\_{\left(A^{´}\right)}=\frac{4}{6}=\frac{2}{3}$ . $P\_{\left(A\right)}+ P\_{\left(A^{´}\right)}=1$

Sjednocení jevů a) disjunktních A a E A … padne číslo menší než 3 $A=\left\{1, 2\right\}$ E … padne číslo 5 $E=\left\{5\right\}$ $A\bigcup\_{}^{}E=\left\{1, 2, 5\right\}$ $P\_{\left(A\right)}=\frac{2}{6}; P\_{\left(E\right)}=\frac{1}{6}; P\_{\left(A\bigcup\_{}^{}E\right)}=\frac{m\_{A\bigcup\_{}^{}E}}{n}=\frac{3}{6}=\frac{2}{6}+\frac{1}{6}= P\_{\left(A\right)}+ P\_{\left(E\right)} $

 b) incidentních A a D A … padne číslo menší než 3 $A=\left\{1, 2\right\}$ D … padne sudé číslo $D=\left\{2, 4, 6\right\}$ $A\bigcup\_{}^{}D=\left\{1, 2, 4, 6\right\}$ $A\bigcap\_{}^{}D=\left\{2\right\}$ $P\_{\left(A\right)}=\frac{2}{6}; P\_{\left(D\right)}=\frac{3}{6}; P\_{\left(A\bigcup\_{}^{}D\right)}=\frac{m\_{A\bigcup\_{}^{}D}}{n}=\frac{4}{6}=\frac{2}{6}+\frac{3}{6}-\frac{1}{6}= P\_{\left(A\right)}+ P\_{\left(D\right)}-P\_{\left(A\bigcap\_{}^{}D\right)} $

Průnik jevů a nezávislé jevy A a B A … při prvním hodu padne 2 $P\_{\left(A\right)}=\frac{1}{6}$ B … při druhém hodu padne sudé číslo $P\_{\left(B\right)}=\frac{1}{2}$ $A\bigcap\_{}^{}B=\left\{\left[2,2\right],\left[2,4\right], \left[2,6\right]\right\}$ $m\_{A\bigcap\_{}^{}B}=3$ $n=V\_{\left(2,6\right)}^{´}=6^{2}=36$ $P\_{\left(A\bigcap\_{}^{}B\right)}=\frac{m\_{A\bigcap\_{}^{}B}}{n}=\frac{3}{36}=\frac{1}{12}$ $P\_{\left(A\right)}. P\_{\left(B\right)}=\frac{1}{6}.\frac{1}{2}=\frac{1}{12}$ $P\_{\left(A\bigcap\_{}^{}B\right)}=P\_{\left(A\right)}. P\_{\left(B\right)}$ Jevy A, B jsou nezávislé.

5

**1 + 4**

 **4 + 1**

 **2 + 3**

 **3 + 2**

Průnik jevů a závislé jevy A a C A … při prvním hodu padne 2 $P\_{\left(A\right)}=\frac{1}{6}$ C … součet při dvou hodech bude 5 $P\_{\left(C\right)}=\frac{4}{36}=\frac{1}{9}$ $A\bigcap\_{}^{}C=\left\{\left[2,3\right]\right\}$ $m\_{A\bigcap\_{}^{}C}=1$ $n=V\_{\left(2,6\right)}^{´}=6^{2}=36$ $P\_{\left(A\bigcap\_{}^{}B\right)}=\frac{m\_{A\bigcap\_{}^{}C}}{n}=\frac{1}{36}$ $P\_{\left(A\right)}. P\_{\left(B\right)}=\frac{1}{6}.\frac{1}{9}=\frac{1}{54}$ $P\_{\left(A\bigcap\_{}^{}B\right)}\ne P\_{\left(A\right)}. P\_{\left(B\right)}$ Jevy A, B jsou závislé.

Statistická (zobecněná) definice pravděpodobnosti - využívá se v případě, že všechny výsledky (prvky množiny $Ω$) nemají stejnou šanci, aby nastaly.

Př.: Házení nehomogenní kostkou na Člověče, nezlob se. S jakou pravděpodobností padne číslo větší než 3?

|  |  |  |
| --- | --- | --- |
| $$ω\_{i}$$ | $$n\_{i}$$ | $$p\_{i}$$ |
| 1 | 2 | $$\frac{2}{30}=\frac{1}{15}$$ |
| 2 | 4 | $$\frac{4}{30}=\frac{2}{15}$$ |
| 3 | 5 | $$\frac{5}{30}=\frac{1}{6}$$ |
| 4 | 6 | $$\frac{6}{30}=\frac{1}{5}$$ |
| 5 | 6 | $$\frac{6}{30}=\frac{1}{5}$$ |
| 6 | 7 | $$\frac{7}{30}$$ |
|  | $$Σn\_{i}=30$$ | $$Σp\_{i}=1$$ |

$$A=\left\{4, 5, 6\right\}$$

$$P\_{\left(A\right)}=\frac{1}{5}+\frac{1}{5}+\frac{7}{30}=\frac{19}{30}$$

Bernoulliho schéma

Př.: 20 hodů kostkou. Jev A … číslo 5 padne právě 3krát. Jaká je pravděpodobnost jevu A? $P\_{\left(A\right)}= \left(\genfrac{}{}{0pt}{}{20}{3}\right).\left(\frac{1}{6}\right)^{3}.\left(\frac{5}{6}\right)^{17}=\frac{20.19.18}{3.2.1}.\frac{5^{17}}{6^{20}}=190.\frac{5^{17}}{6^{19}}$

**Základy statistiky:**

Př. 1: V souboru A byl sledován údaj o počtu dětí v 13 rodinách. Rozsah souboru je n = 13 (liché):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| počet dětí | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| četnost | 2 | 4 | 3 | 2 | 1 | 1 | 0 | 0 |

Aritmetický průměr$\overline{x}=\frac{1}{13}\sum\_{i=1}^{13}x\_{i}$=$\frac{0+0+1+1+1+1+2+2+2+3+3+4+5}{13}=\frac{25}{13}=1,923$ (Vážený průměr $\overline{x}=\frac{1}{13}\sum\_{j=1}^{8}x\_{j}n\_{j}$=$\frac{0.2+1.4+2.3+3.2+4.1+5.1+6.0+7.0}{13}=\frac{25}{13}=1,923$)

Modus Mod(x) = 1 (má četnost 4, což je nejvíce)

Medián Med(x) = 2 (7. rodina).

Př. 2: V souboru B byl sledován údaj o počtu dětí ve 14 rodinách. Rozsah souboru je n = 14 (sudé):

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| počet dětí | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| četnost | 2 | 5 | 3 | 1 | 1 | 0 | 0 | 1 | 1 |

Aritmetický průměr $\overline{x}=\frac{1}{14}\sum\_{i=1}^{14}x\_{i}$=$\frac{0+0+1+1+1+1+1+2+2+2+3+4+7+8}{14}=\frac{33}{14}=2,357$ (Vážený průměr $\overline{x}=\frac{1}{14}\sum\_{j=1}^{9}x\_{j}n\_{j}=\frac{0.2+1.5+2.3+3.1+4.1+5.0+6.0+7.1+8.1}{14}=\frac{33}{14}=2,357$)

Modus Mod(x) = 1 (má četnost 5, což je nejvíce)

Medián Med(x) = 1,5 (aritmetický průměr ze 7. a 8.rodiny).

Základní poznatky:

1. Jaká je pravděpodobnost, že při hodu třemi kostkami padne součet 12?

 [0,116]

1. 55% populace tvoří ženy, 45% muži. Určitou chorobou trpí 1% žen a 5% mužů. Jaká je p., že náhodně vybraná osoba z populace trpí touto chorobou?

 [0,028]

1. Dvanáct studentů, mezi kterými je Pavel a Tomáš, mají ze svého středu vylosovat 4-člennou skupinu.
Jaká je pravděpodobnost, že ve skupině bude:

a) Tomáš b) Tomáš, ale Pavel ne c) Tomáš a Pavel d)Tomáš nebo Pavel

 [0,333 ; 0,242 ; 0,091 ; 0,576]

1. Jaká je pravděpodobnost, že se Jana a Tomáš narodili ve stejný měsíc?
(Počítejte, že 1 měsíc je 1/12 roku) [0,0833]
2. **Státní maturita 2017**

Jaká je pravděpodobnost, že oba vybraní žáci budou mít úkol vypracován správně? [2/5]
3. **Státní maturita 2015**

Kolik písemných prací bylo oznámkováno? [16]

Typové příklady standardní náročnosti

1. Z 10 studentů, mezi nimiž jsou Adam a Petr vybíráme tříčlennou komisi. Jaká je pravděpodobnost, že Adam nebo Petr budou mezi nimi? [8/15]
2. Hodíme 3-krát kostkou. Jaká je pravděpodobnost, že při 1. hodu nebo při 2. nebo při 3. padne sudé číslo? [7/8]
3. Každý ze spínačů je náhodně v poloze zapnuto nebo vypnuto nezávisle na druhých. Jaká je pravděpodobnost, že součástkou bude protékat el. proud?

 [15/64]

1. Žárovka svítí se spolehlivostí 85%. Jaká je spolehlivost systému (alespoň část svítí), jsou-li zapojeny:
a) dvě žárovky sériově, b) dvě žárovky paralelně, c) dvě žárovky sériově a třetí k nim paralelně

 [0,723 ; 0,978 ; 0,958]

1. Test obsahuje 10 otázek, čtyři možné odpovědi, z nichž jedna je správná. Jaká je pravděpodobnost, že náhodným volením odpovědí vybereme alespoň 5 správných? [0,078]
2. Lék úspěšně léčí 90% případů onemocnění. Vypočítejte pravděpodobnost, že vyléčí alespoň 18 pacientů z 20, kterým je lék podán. [0,677]
3. Tabulka zaznamenává známky z matematiky pro 16 testů. Vypočítejte:
	1. Relativní četnost známky 2.
	2. Modus a medián známek z testu.
	3. Aritmetický průměr známek z testu.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Číslo testu | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Známka | 1 | 5 | 2 | 3 | 5 | 4 | 3 | 4 | 2 | 2 | 4 | 4 | 5 | 2 | 3 | 4 |

[a) 25%, b) modus 4, medián 3,5, c) 53/16]

1. Průměrná cena 1 kg jablek je 27 Kč, průměrná cena 1 kg hrušek je 30 Kč. Nakoupili jsme jablka, hrušky a pomeranče. Čtvrtinu nakoupeného množství tvořily jablka, třetinu hrušky a zbytek tvořily pomeranče. Jaká je průměrná cena 1 kg pomerančů, jestliže byla průměrná cena 1 kg smíšeného zboží 32 Kč? [39,6 Kč]

Rozšiřující cvičení

1. **Státní maturita Matematika+ 2017** [A, N, A]


