STATISTIKA 1

RNDr. K. Hrach, Ph.D.

Zápočet:

odevzdání seminární práce (úkoly na PC)

Zkouška: písemná (bez kalkulačky, bez vzorců)

STATISTIKA

Činnost vedoucí k získávání dat

Instituce zajišťující tuto činnost

Jakákoli shromažďovaná data

Údaje získané výpočtem z dat

Matematická teorie o chování dat

STATISTIKA

DESKRIPCE (popis)

ANALÝZA (modely, odhady, testy)

Základní pojmy

- STATISTICKÁ JEDNOTKA =

  = na kom (čem) zjišťujeme

- STATISTICKÉ ŠETŘENÍ =

  = jak zjišťujeme

- STATISTICKÁ VELIČINA =

  = co zjišťujeme

STATISTICKÁ JEDNOTKA

s.j. = např. každý(á/é)

ptačí vejce (hmotnost v g?);

snůška (počet vajec?);

územní celek - obec; region;…; záměrně selektovaný čtverec (% zalesnění?);

respondent (dotazníkové šetření)

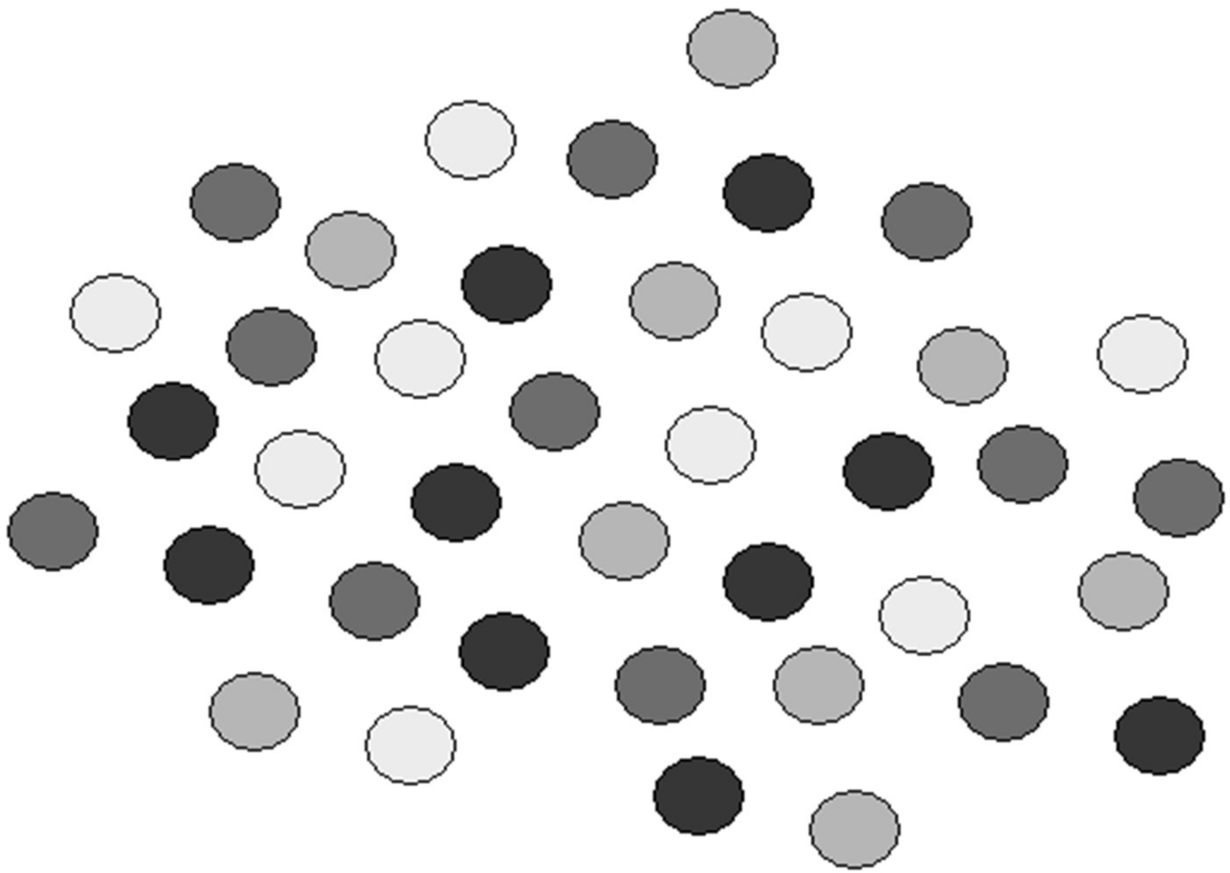
…

STATISTICKÉ ŠETŘENÍ

- <u>ÚPLNÉ</u> ⇔
  informace od všech stat.jednotek
  (od celé populace)
- <u>VÝBĚROVÉ</u> ⇔
  informace od vybraných stat. jednotek
  (od „výběru“)
  nevýhody versus výhody výběru?
  * neúplnost informace
  * rychlejší a levnější informace

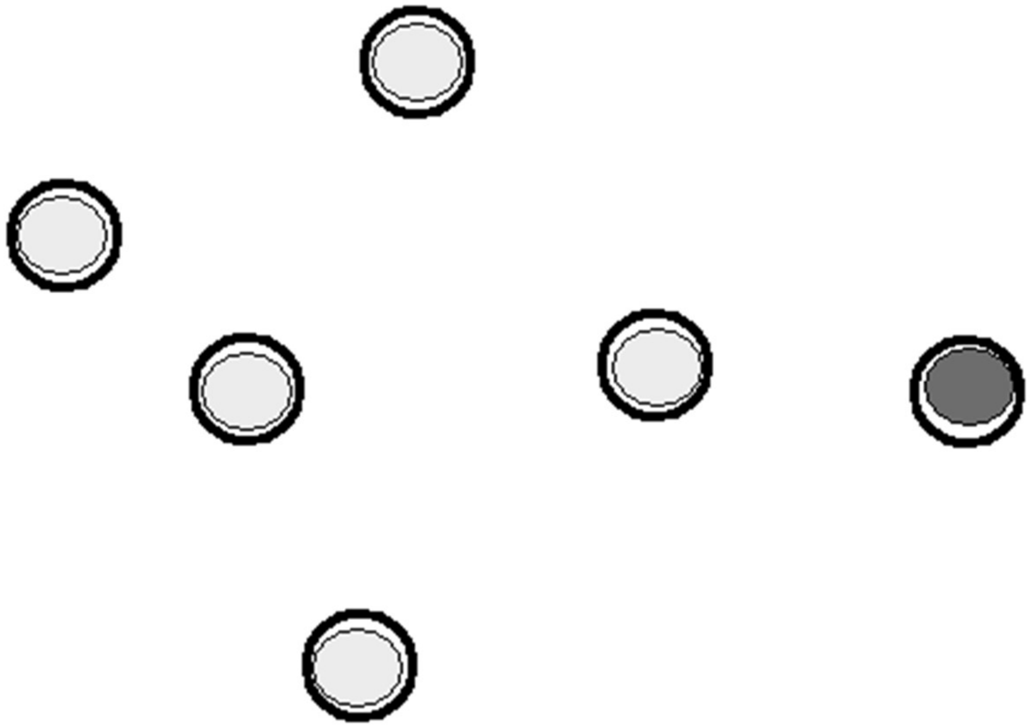
VÝBĚROVÉ ŠETŘENÍ

NE/REPREZENTATIVNOST ?



VÝBĚROVÉ ŠETŘENÍ

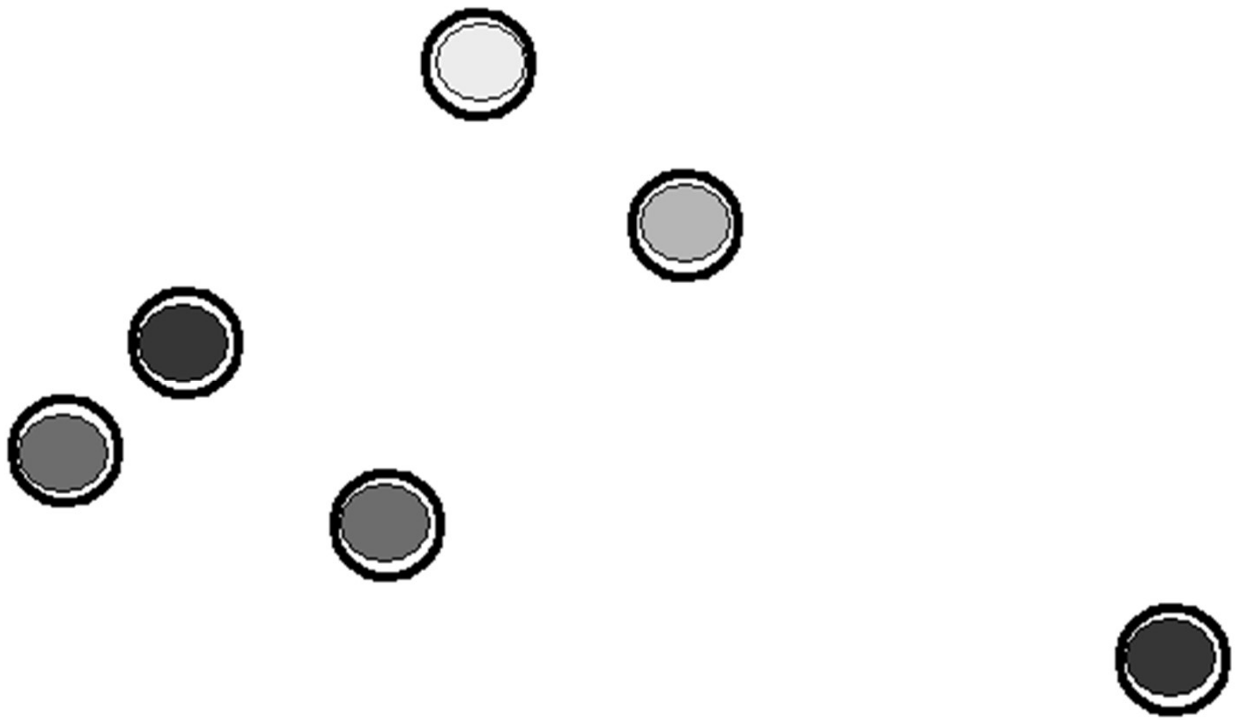
!!! NEREPREZENTATIVNOST !!!



VÝBĚROVÉ ŠETŘENÍ

!!! REPREZENTATIVNOST !!!

(zajištěna např. NÁHODNÝM výběrem)



VÝBĚROVÉ ŠETŘENÍ - výběr

- zcela náhodný
- systematický
- stratifikovaný …

VÝBĚROVÉ ŠETŘENÍ - průběh

- vlastní měření … s.j. je objekt
- anketa … s.j. je subjekt (i dále):
- řízený rozhovor
- vyplnění dotazníku

Jak jinak získat data (např. k BP)?

JIŽ „HOTOVÁ“ (tj. sebraná), NAPŘ OD:

- ČSÚ (www stránky)
- Eurostatu (www stránky)
- úřadů státní (samo)správy
- firem (?) …

Zpracování dat

a) „ručně“

b) pomocí SW

- MS Excel
- STATISTICA, SPSS, …
- freeware (R-project)

DOTAZNÍK pro zaměstnance firmy

JMÉNO

DOBA

ÚKOL

POHLAVÍ

VZDĚLÁNÍ

VĚK

POBOČKA

POČET DĚTÍ

DOTAZNÍK pro zaměstnance firmy

JMÉNO	***identifikátor***
DOBA	***veličina (značena např. X)***
ÚKOL	***veličina (značena např. Q)***
POHLAVÍ	***…***
VZDĚLÁNÍ	***…***
VĚK	***…***
POBOČKA	***…***
POČET DĚTÍ	***veličina (značena např. Y)***

DOTAZNÍK pro zaměstnance firmy

	možné hodnoty :
JMÉNO	**textový řetězec**
DOBA	**1,2,…(počet dní proškolení)**
ÚKOL	**ano / ne (splněn nový úkol?)**
POHLAVÍ	**m / z**
VZDĚLÁNÍ	**z, s, v (nejvyšší dosažené)**
VĚK	**v rocích**
POBOČKA	**a,b,c (1 ze 3 poboček firmy)**
POČET DĚTÍ	**0,1,2,…**

DOTAZNÍK - příklad vyplnění (1. statistická jednotka)

JMÉNO	Frank
DOBA	14
ÚKOL	ne
POHLAVÍ	m
VZDĚLÁNÍ	z
VĚK	30
POBOČKA	a
POČET DĚTÍ	1

DOTAZNÍK data (začátek)

Poradi	JMENO	DOBA	UKOL	POHL	VZDEL	VEK	POBOCKA	DETI
1	Frank	14	ne	m	z	30	a	1
2	Henry	29	ne	m	s	41	a	3
3	Tom	6	ne	m	v	32	b	0
4	Beth	25	ano	z	z	38	b	2
5	Susan	18	ano	z	s	44	b	2
6	Harry	4	ne	m	z	44	c	4
7	Paul	18	ne	m	s	39	c	4
8	Pete	12	ne	m	v	46	a	3
9	Diana	22	ano	z	z	45	a	2
10	Louise	6	ne	z	s	33	c	0
11	Fred	30	ano	m	z	32	a	1
12	Hank	11	ne	m	s	48	b	5
13	Steven	30	ano	m	v	42	b	3

DOTAZNÍK data (dokončení)

Poradi	JMENO	DOBA	UKOL	POHL	VZDEL	VEK	POBOCKA	DETI
14	Tod	5	ne	m	z	23	b	0
15	Take	20	ano	z	s	44	c	1
16	Sam	13	ne	m	z	34	c	2
17	Gail	9	ne	z	s	54	a	3
18	Thomas	32	ano	m	v	39	c	2
19	Theodore	24	ne	m	z	30	a	1
20	Charles	13	ano	m	s	39	a	3
21	Ellein	19	ne	z	z	51	b	0
22	Lori	4	ne	z	s	31	b	0
23	Ann	28	ano	z	v	38	c	1
24	Valerie	22	ano	z	z	35	c	2
25	Anke	8	ano	z	s	26	b	0

Značení dat (pozorování)

Např. veličina X – DOBA:

1. pozorování: x_1=14

2. pozorování: x_2=29

…

n. (poslední) pozorování: $x_n = x_{25} = 8$

n značí počet pozorování (rozsah souboru), zde n=25

DOTAZNÍKY - značení a data

JMENO	DOBA (X)	UKOL (Q)	...	DETI (Y)
Frank	$x_1 = 14$	$q_1 =$ ne	...	$y_1 = 1$
Henry	$x_2 = 29$	$q_2 =$ ne	...	$y_2 = 3$
...	...	...	...	...
Anke	$x_{25} = 8$	$q_{25} =$ ano	...	$y_{25} = 0$

TYPY VELIČIN (*X*, *Y*,…)
(proměnných; znaků; angl.VARIABLE)

KATEGORIÁLNÍ

- alternativní (2 možnosti: 0-1)
  např. POHL (m/z), UKOL (splněn: ano/ne)
- slovní – neuspořádané (nominální)
  např. POBOCKA (a, b, c)
- slovní – uspořádané (ordinální)
  např. VZDEL (z < s < v)
- číselné (diskrétní)
  např. DETI (počet dětí)

TYPY VELIČIN (X, Y,…) (proměnných; znaků)

NEKATEGORIÁLNÍ

- číselné (<u>spojité</u>)

např. DOBA (počet dní výcviku), VĚK (v rocích), HMOTNOST (kg), PLAT (tis.Kč)…

v příkladech je údaj vlastně zaokrouhlen, záleží na zvolené přesnosti;

lze převést na kategoriální typ (jak, jaký?)

DOTAZNÍKY – příklad zpracování (*Y*)

POČET DĚTÍ (*Y*)			kumul.	kumul.
y_i	n_i	p_i	n_i*	p_i*
0	6	0,24	6	0,24
1	5	0,20	11	0,44
2	6	0,24	17	0,68
3	5	0,20	22	0,88
4	2	0,08	24	0,96
5	1	0,04	25	1,00
suma	25	1,00	xxx	xxx

DOTAZNÍKY – příklad zpracování (Y)

POZOR NA PODOBNÉ ZNAČENÍ:

a) pro jednotlivá pozorování veličiny Y bylo

$y_1 = 1$, $y_2 = 3$, …, $y_{25} = 0$ ($n = 25$)

b) pro kategorie veličiny Y bylo

$y_1 = 0$, $y_2 = 1$, …, $y_6 = 5$ (K=6)

V praxi je rozdíl v použití jasný z kontextu.

ČETNOSTI

ABSOLUTNÍ

n_i … počet výskytů i-té kategorie, i=1…K

$$\Sigma\, n_i = n$$

RELATIVNÍ

p_i … rel. výskyt i-té kategorie, i=1…K

$$p_i = n_i/\mathrm{n} \qquad \Sigma\, p_i = 1$$

$$p_i = (n_i/\mathrm{n})\cdot 100\% \qquad \Sigma\, p_i = 100\ (\%)$$

Oba typy lze určit u každé kategoriální veličiny (K=počet kategorií).

ČETNOSTI

Příklad 1.

Y - známky žáka. Popořadě:

3, 4, 2, 3, 2, 3, 3, 3.

Tabulka četností:

i	1	2	3	suma
y_i	2	3	4	xxx
n_i	2	5	1	8
p_i	0,250	0,625	0,125	1,000

ČETNOSTI

KUMULOVANÉ ABSOLUTNÍ

n_i^* … počet výskytů do i-té kategorie včetně,

$$n_i^* = n_1 + \ldots + n_i$$

KUMULOVANÉ RELATIVNÍ

p_i^* … rel. výskyt do i-té kategorie včetně,

$$p_i^* = p_1 + \ldots + p_i$$

$$p_i^* = n_i^*/n$$

Oba typy mají smysl jen u veličin ordinálních či diskrétních.

ČETNOSTI

Příklad 1 - pokračování.

i	1	2	3	suma
y_i	2	3	4	xxx
n_i	2	5	1	8
p_i	0,250	0,625	0,125	1,000
n_i^*	2	7	8	xxx
p_i^*	0,250	0,875	1,000	xxx

ČETNOSTI

MODUS (skloňujeme: bez modu,…,s modem)

Je (jsou) kategorie s největší četností (samozřejmě nikoli kumulovanou).

Značen $\hat{y}$ (se stříškou).

Lze určit u každé kategoriální veličiny.

Příklad 1 – pokračování:

$\hat{y}$=3

(druhá kategorie se vyskytla nejčastěji, a to pětkrát; nejčastější známkou byla trojka)

Obecné momenty – prosté tvary

První obecný moment: $(\Sigma y_i)/n$, $i=1\ldots n$

aritmetický průměr, „těžiště“ dat

$\bar{y}$

Druhý obecný moment: $(\Sigma y_i^2)/n$, $i=1\ldots n$

$\overline{y^2}$

Obecné momenty – prosté tvary

Příklad 1 – pokračování:

$\bar{y} = (3+4+2+3+2+3+3+3)/8 = 23/8 = 2,875$

$\overline{y^2} = (3^2+4^2+2^2+3^2+2^2+3^2+3^2+3^2)/8 =$

$= (9+16+4+9+4+9+9+9)/8 = 69/8 = 8,625$

Obecné momenty – prosté tvary

Příklad 1 – Poznámka:

Hodnota SOUČTU všech pozorování
(3+4+2+3+2+3+3+3) = 23
je tzv. **<u>ÚHRN</u>** (úhrnná hodnota).

Platí: průměr = úhrn / počet

aneb: úhrn = průměr · počet

aneb: počet = úhrn / průměr

Obecné momenty – vážené tvary

Při výpočtu lze využít četností absolutních:

$\overline{y} = (\Sigma y_i \, n_i)/n$, $i=1\ldots K$

$\overline{y^2} = (\Sigma y_i^2 \, n_i)/n$, $i=1\ldots K$

Obecné momenty – vážené tvary

Příklad 1 - pokračování:

$\bar{y} = (2 \cdot 2 + 3 \cdot 5 + 4 \cdot 1)/8 = 23/8 = 2,875$

$\overline{y^2} = (2^2 \cdot 2 + 3^2 \cdot 5 + 4^2 \cdot 1)/8 = (4 \cdot 2 + 9 \cdot 5 + 16 \cdot 1)/8 =$

$= 69/8 = 8,625$

Obecné momenty – vážené tvary

Při výpočtu lze využít také četností relativních:

$$\overline{y} = \Sigma y_i \, p_i \,, \qquad i=1 \ldots K$$

$$\overline{y^2} = \Sigma y_i^2 \, p_i, \qquad i=1 \ldots K$$

Obecné momenty – vážené tvary

Příklad 1 - pokračování:

$$\bar{y} = 2 \cdot 0{,}250 + 3 \cdot 0{,}625 + 4 \cdot 0{,}125 = 2{,}875$$

$$\overline{y^2} = 2^2 \cdot 0{,}250 + 3^2 \cdot 0{,}625 + 4^2 \cdot 0{,}125 =$$

$$= 4 \cdot 0{,}250 + 9 \cdot 0{,}625 + 16 \cdot 0{,}125 = 8{,}625$$

Centrované momenty prosté tvary

Druhý centrovaný moment:

$$M_2(y) = \Sigma(y_i - \bar{y})^2 / n, \qquad i=1\ldots n$$

= <u>ROZPTYL</u> aneb „průměrná čtvercová odchylka od aritmetického průměru“

$$\sqrt{M_2(y)}$$

= <u>směrodatná odchylka</u>

Centrované momenty prosté tvary

Příklad 1 - pokračování:

$$M_2(y)=[(3-2{,}875)^2+(4-2{,}875)^2+(2-2{,}875)^2+$$
$$+(3-2{,}875)^2+(2-2{,}875)^2+(3-2{,}875)^2+$$
$$+(3-2{,}875)^2+(3-2{,}875)^2\,]\,/\,8=0{,}359$$

$$\sqrt{M_2(y)}=\sqrt{0{,}359}=0{,}599$$

Známky byly zhruba v rozmezí 2,875±0,599 aneb v rozmezí 2,276 až 3,474.

Centrované momenty
vážené tvary

K výpočtu lze užít četností absolutních:

$$M_2(y) = \Sigma(y_i - \bar{y})^2 \, n_i / n, \qquad i=1 \ldots K$$

nebo relativních:

$$M_2(y) = \Sigma(y_i - \bar{y})^2 \, p_i \,, \qquad i=1 \ldots K$$

Centrované momenty vážené tvary

Příklad 1 - pokračování:

$$M_2(y) = [(2-2{,}875)^2 \cdot 2 + (3-2{,}875)^2 \cdot 5 + (4-2{,}875)^2 \cdot 1] \,/\, 8 = 0{,}359$$

$$M_2(y) = (2-2{,}875)^2 \cdot 0{,}250 + (3-2{,}875)^2 \cdot 0{,}625 + (4-2{,}875)^2 \cdot 0{,}125 = 0{,}359$$

MOMENTY

Mají smysl jen u číselných veličin
(diskrétních či spojitých)

Vážené tvary má smysl používat jen tehdy, jsou-li k dispozici četnosti (tj. u diskrétních veličin)

Použití prostého i váženého tvaru musí dát stejný výsledek (jde o totéž, jen jinak počítáno)

MOMENTY

Chování při aditivní lineární transformaci:

Známe průměr a rozptyl.

Každé pozorování změníme o stejnou konstantu c.

Průměr se tak změní o tutéž konstantu c, rozptyl se nezmění.

MOMENTY

Příklad 1 - pokračování:

Pokud bychom žákovi všechny udělené známky o jeden stupeň zlepšili (c= -1), průměrná známka se změní z hodnoty 2,875 na hodnotu 1,875, ale rozptyl zůstane 0,359 (ověřte výpočtem).

MOMENTY

Chování při <u>multiplikativní lineární transformaci</u>:

Známe průměr a rozptyl.

Každé pozorování vynásobíme stejnou konstantou c.

Průměr se také musí vynásobit konstantou c, rozptyl se musí vynásobit její druhou mocninou c^2.

MOMENTY

Příklad 2:

Byly měřeny délky vyrobených trubek v centimetrech. Průměrná délka činí 49 cm a rozptyl 144 (cm^2). Nyní je nutno všechna měření převést na decimetry. Jak se změní hodnoty momentů?

Jelikož c=0,1, průměr bude činit 4,90 dm a rozptyl 1,44 (dm^2).

MOMENTY

ROZPTYL – rychlý výpočetní tvar

$$M_2(y) = \overline{y^2} - (\overline{y})^2$$

Příklad 1 - pokračování:

$M_2(y) = 8{,}625 - 2{,}875^2 = 0{,}359$

Poznámka: Vždy musí vyjít $M_2(y) \geq 0$!

MOMENTY

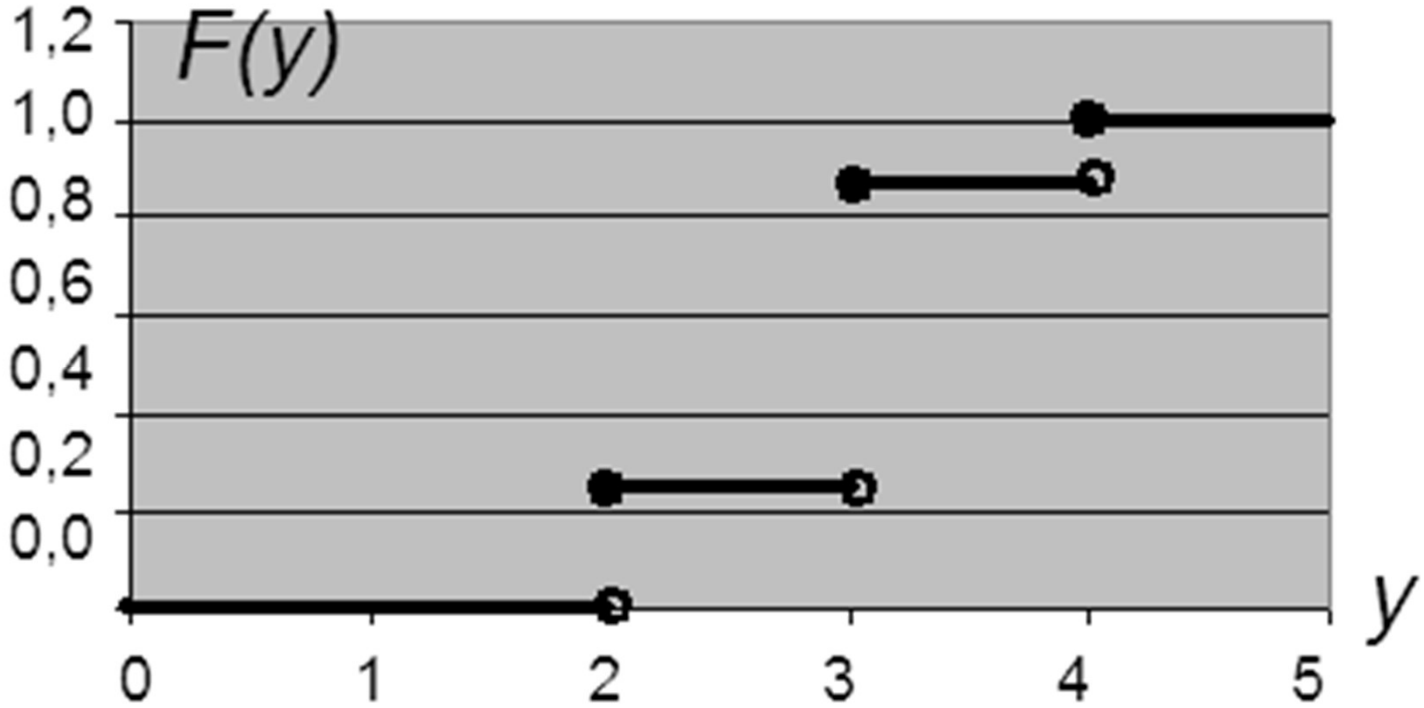
- Rozptyl – čtvercová míra variability (proměnlivosti) dat
- Relativní mírou variability je variační koeficient

$$V_y = \sqrt{M_2(y)} / \bar{y}$$

může porovnávat variabilitu souborů, v nichž je veličina zaznamenána v různých měrných jednotkách – např. platy u nás v Kč versus platy v Německu v Euro), či je na jiné úrovni (poloze)

DISTRIBUČNÍ FUNKCE

- Slouží k popisu rozdělení (distribuce) číselných dat; jakési „zobecnění“ kumul.relat.četností:
- *F*(*y*) … udává podíl pozorování s hodnotou nejvýše *y*, viz graf pro data z příkladu 1:



KVANTILY

- Udávají pro číselnou veličinu *Y* hodnotu *y*, pod níž leží požadovaný podíl pozorování
- Značíme $\tilde{y}_{100p}$, kde *p* je onen podíl (údaj mezi 0 – 1)
- *p*=0,5 … 50% kvantil – <u>medián</u>

(odděluje polovinu nižších od zbytku vyšších pozorování; značen obvykle jen $\tilde{y}$)

- *p*=0,25 … 25% kvantil – <u>dolní kvartil</u> $\tilde{y}_{25}$
- *p*=0,75 … 75% kvantil – <u>horní kvartil</u> $\tilde{y}_{75}$

KVANTILY

1. Určení pomocí distribuční funkce

Najdeme bod *y,* v němž poprvé *F(y)* dosáhne úroveň *p.*

Příklad 1 – pokračování:

Nalezněte medián známek žáka.

Medián…tj. p=0,5.

Kde poprvé příslušná *F(y)* dosáhla úroveň 0,5?

Podle grafu je mediánem hodnota $\tilde{y}$ =3.

Interpretace: Polovina známek měla hodnotu nejvýše 3 (tj. 3 nebo nižší)

KVANTILY

2. Určení přímo pomocí dat

a) Data uspořádáme vzestupně dle velikosti

b) Nalezneme celočíselné z vyhovující nerovnicím:

$$n \cdot p < z < n \cdot p + 1$$

c) Hledaným kvantilem je hodnota s pořadovým číslem z.

Příklad 1 – pokračování:

Nalezněte 60% kvantil známek žáka.

Známky seřadíme: 2, 2, 3, 3, 3, 3, 3, 4.

Jelikož n=8 a p=0,6, vyjde $n \cdot p$=4,8, takže hledáme z :

$$4{,}8 < z < 5{,}8$$

Tudíž z=5, aneb hledaným kvantilem je 5. známka, čili trojka: $\tilde{y}_{60}=3$.

KVANTILY

Příklad 1 – pokračování:

Interpretujte nalezený výsledek: $\tilde{y}_{60}$=3.

Znamená to, že 60 % zjištěných známek mělo hodnotu nejvýše 3 (tj. 3 nebo nižší).

Podobně interpretujte to, kdyby pro veličinu plat měl např. horní kvartil hodnotu 32 tis. Kč:

Znamenalo by to, že 75 % zjištěných platů mělo hodnotu nejvýše 32 tis. Kč (*tj. zbylých 25 % platů bylo jakých*?)

KVANTILY

2. Určení přímo pomocí dat – možný problém

a) Data uspořádáme vzestupně dle velikosti

b) Nenalezneme celočíselné z vyhovující nerovnicím:

$$n \cdot p < z < n \cdot p + 1$$

Stane se to tehdy, když obě hodnoty ($n \cdot p$; $n \cdot p+1$) vyjdou přesně celočíselně.

c) Nalezneme hodnotu s pořadovým číslem $n \cdot p$ a s následujícím pořadovým číslem $n \cdot p+1$.

d) Za hledaný kvantil prohlásíme průměr obou nalezených hodnot.

KVANTILY

<u>2. Určení přímo pomocí dat – možný problém</u>

Příklad 1 – pokračování:

Nalezněte medián známek žáka.

Známky seřadíme: 2, 2, 3, 3, 3, 3, 3, 4.

Jelikož n=8 a p=0,5, vyjde $n \cdot p$ =4, takže hledáme z :

$$4 < z < 5 \qquad \text{(neexistuje)}$$

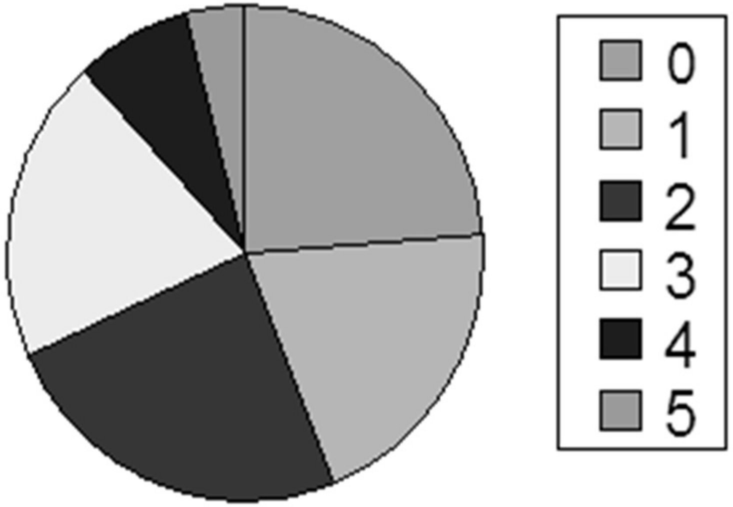
Najdeme tedy jak 4., tak 5. známku.

Jelikož jsou obě trojky, jejich průměrem je také trojka a to je hledaný medián: $\tilde{y}$=3.

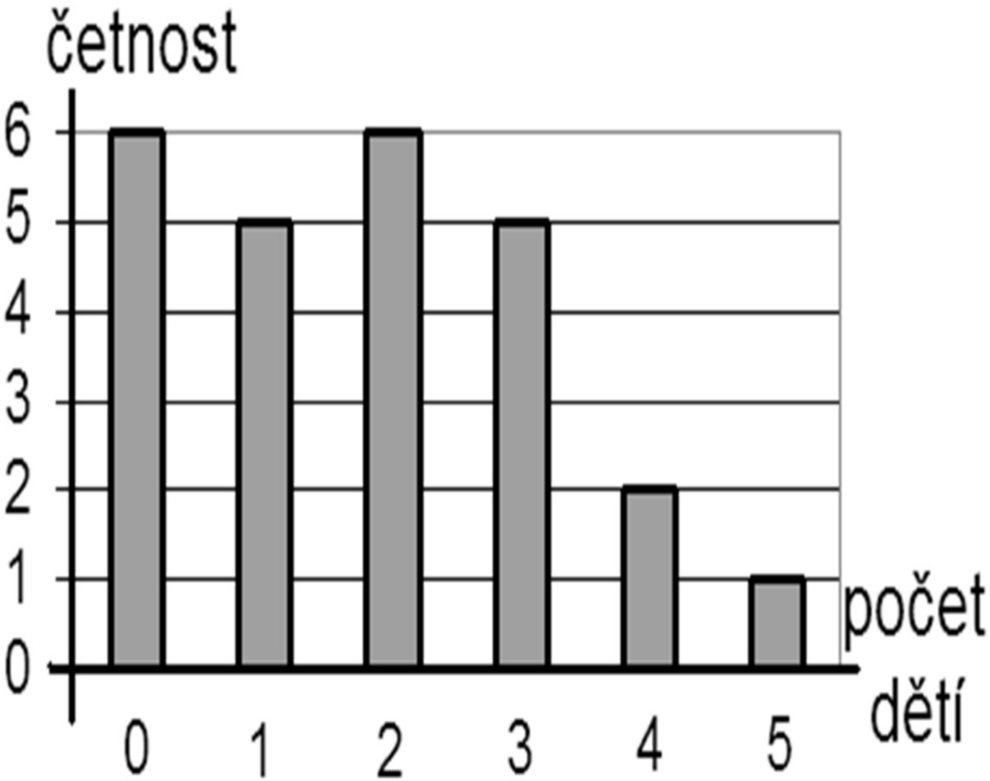
GRAFICKÉ ZNÁZORNĚNÍ DAT

- Číselné veličiny (i spojité!) - graf *F(y)*
- Kategoriální veličiny - graf četností (vel. DETI):

HISTOGRAM

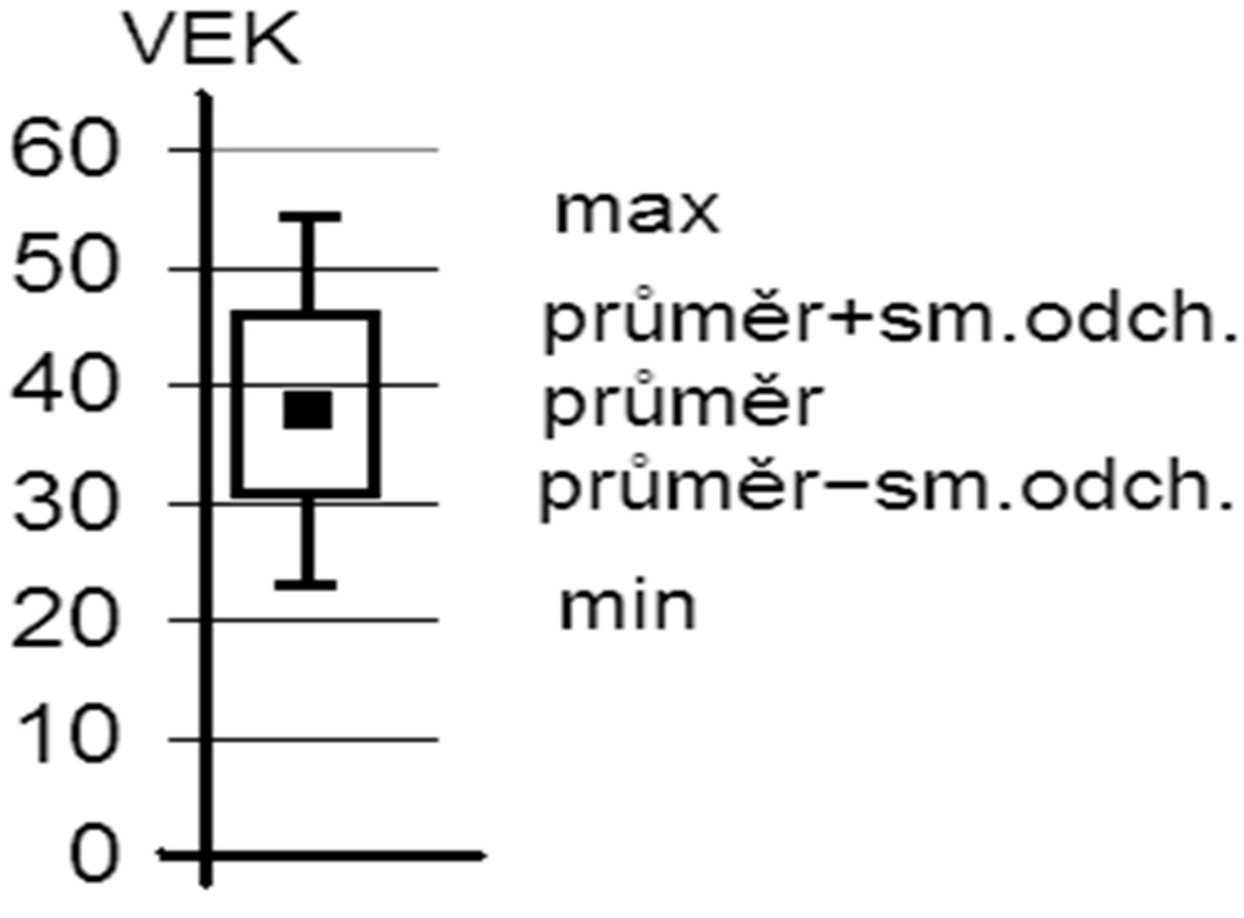


KRUHOVÝ (výsečový, „koláčový“) GRAF



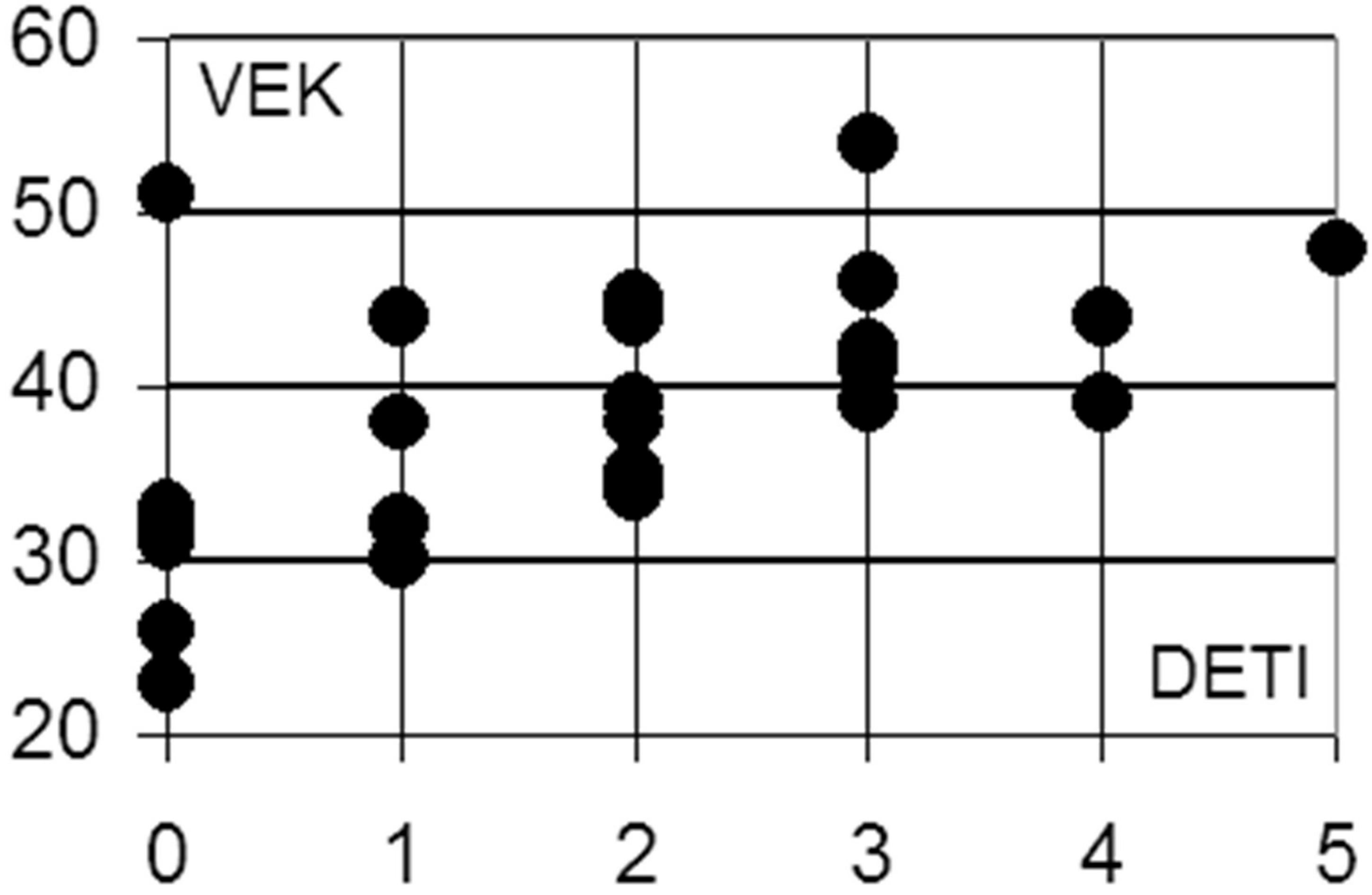
GRAFICKÉ ZNÁZORNĚNÍ DAT

- Spojité veličiny – krabičkový graf (box plot)



GRAFICKÉ ZNÁZORNĚNÍ DAT

- Dvojice číselných veličin – scatter plot (xy)



POPIS DAT („vlastnosti grafu“)

- PARAMETR POLOHY („těžiště“)

Průměr

Medián

Modus

- PARAMETR VARIABILITY („šířka“)

Rozptyl (směrodatná odchylka, var.koef.)

Mezikvartilové rozpětí

POPIS DAT aneb DESKRIPCE

- PARAMETR POLOHY – další možnosti

Geometrický průměr (je-li vhodný):

$$\sqrt[n]{(x_1 \cdot x_2 \cdot \ldots \cdot x_n)}$$

Harmonický průměr (je-li vhodný):

$$n \,/\, (1/x_1 + 1/x_2 + \ldots + 1/x_n)$$

PRAVDĚPODOBNOSTNÍ MODELY STAT. DAT

$\Omega=\{\omega_1,\ldots, \omega_k\}$ mna elementárních jevů (všech možných výsledků náhod.pokusu)

Náh.jevy A, B,…libovolné podmnožiny Ω

Příklad: Náh.pokus hod kostkou

$\Omega=\{1,2,\ldots,6\}$

$A=\{6\}$ hození šestky

$B=\{2,4,6\}$ hození sudého čísla

PRAVDĚPODOBNOSTNÍ MODELY STAT. DAT

PRAVDĚPODOBNOST je vhodně definovaná relativní míra výskytu náh.jevů

KLASICKÁ DEFINICE

||A|| značí počet elem.jevů tvořících A:

$$P(A) = ||A|| / ||\Omega||$$

Příklad – pokračování:

$P(A) = 1/6$ $P(B) = 3/6 = 1/2$

PRAVDĚPODOBNOSTNÍ MODELY STAT. DAT

Základní vlastnosti (příklady?):

a) $P(\{\ \}) = 0, \quad P(\Omega) = 1$

b) $0 \leq P(A) \leq 1$

c) A je podmnožina B => $P(A)<P(B)$

d) $P(A') = 1 - P(A)$ (*doplněk*)

e) $P(A \cap B) = P(A) \cdot P(B)$ (*A,B nezávislé*)

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

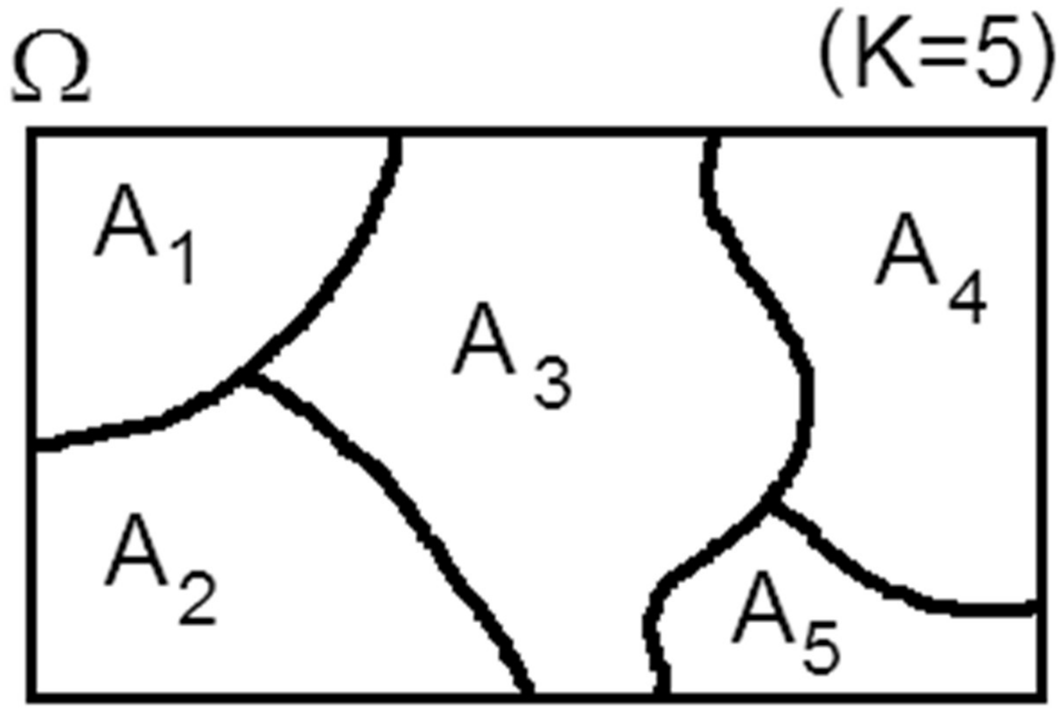
g) $P(A|B) = P(A \cap B) / P(B)$ (*podmínka*)

Rozklad Ω

Náhodné jevy (množiny) A_1 až A_K jsou rozklad Ω

$\Leftrightarrow$

$A_i \cap A_j = \{\}$ pro $i \neq j$, $A_1 \cup \ldots \cup A_K = \Omega$



Příklad 4:

Ω: výsledky 2 hodů kostkou (*X*=počet šestek)

A_1…nehozena ani jedna šestka (X=0)

A_2…hozena jedna šestka (X=1)

A_3…hozeny dvě šestky (X=2)

1,1	**1,2**	**1,3**	**1,4**	**1,5**	**1,6**
2,1	**2,2**	**2,3**	**2,4**	**2,5**	**2,6**
3,1	**3,2**	**3,3**	**3,4**	**3,5**	**3,6**
4,1	**4,2**	**4,3**	**4,4**	**4,5**	**4,6**
5,1	**5,2**	**5,3**	**5,4**	**5,5**	**5,6**
6,1	**6,2**	**6,3**	**6,4**	**6,5**	**6,6**

$P(A_1)=25/36$ aneb $P(X=0)=25/36$

$P(A_2)=10/36$ aneb $P(X=1)=10/36$

$P(A_3)=1/36$ aneb $P(X=2)=1/36$

Výpočty pomocí základních vlastností:

- $P(A_3)= P(A \cap A) = P(A) \cdot P(A) = 1/6 \cdot 1/6 = 1/36$

…jde o průnik výskytu A={6} při 1.hodu a při 2.hodu, tj.platí nezávislost

- $P(A_1)= P(A' \cap A') = P(A') \cdot P(A') = 5/6 \cdot 5/6 = 25/36$

…analog., ale v obou hodech má nastat doplněk

- $P(A_2)= P[(A \cap A') \cup (A' \cap A)] = P(A \cap A')+P(A' \cap A) =$

$= P(A) \cdot P(A') + P(A') \cdot P(A) = 1/6 \cdot 5/6 + 5/6 \cdot 1/6 =$

$= 5/36 + 5/36 = 10/36$

…buď hodíme šestku jen napoprvé, nebo jen napodruhé