

Základy statistické indukce

■ Zákony velkých čísel

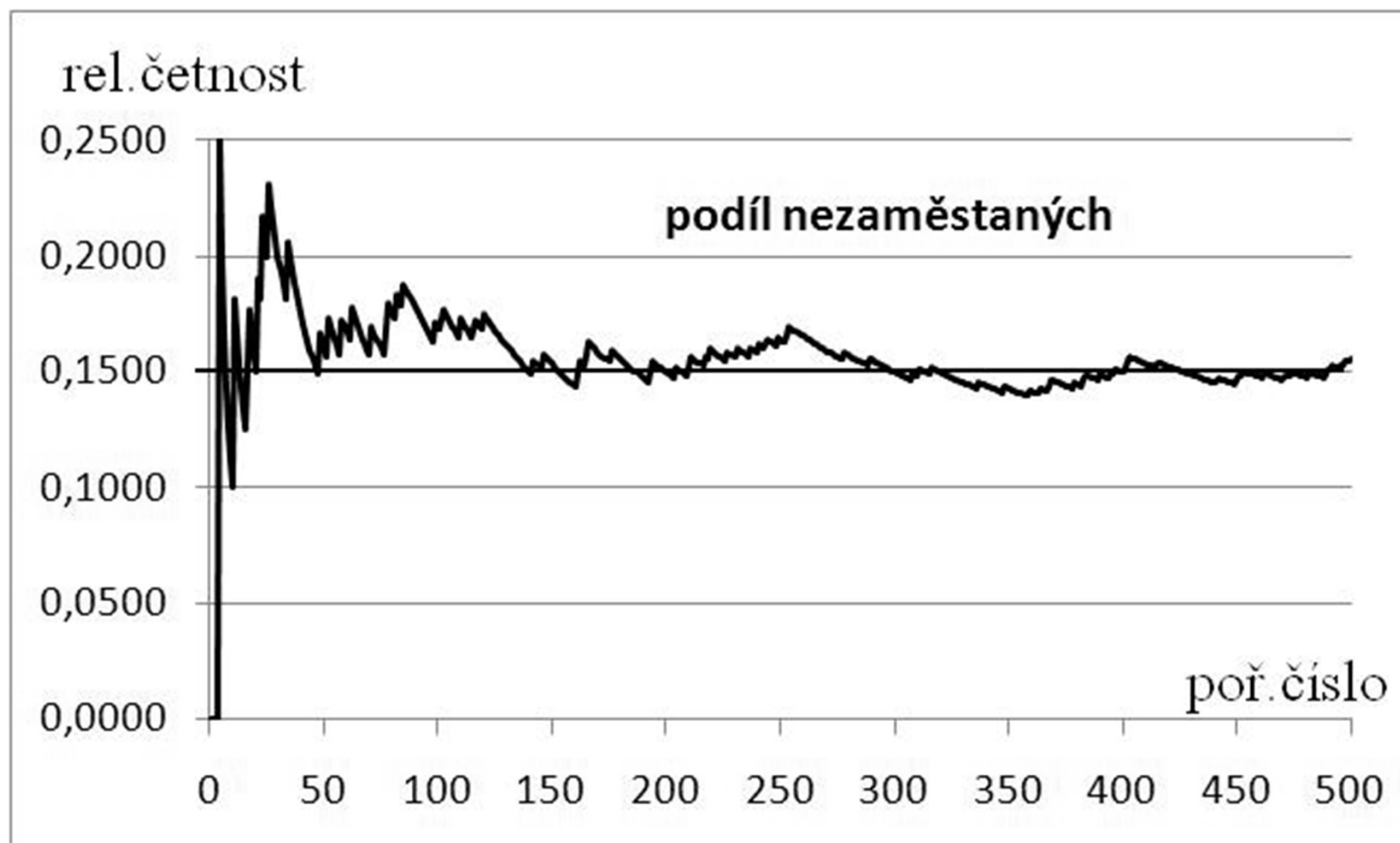
(chování bodových odhadů):

Např. rel.čet. $\rightarrow$ pravděpodobnost (zde $\pi=0,15$)

Dotázaný č.	1	2	3	4	5	...	498	499	500
nezam.? (1-ano)	0	0	0	1	0	...	1	0	1
rel.čet. nezam.	0	0	0	0,25	0,20	...	0,155	0,154	0,156

Základy statistické indukce

- Zákony velkých čísel (pokračování ilustrace):



Základy statistické indukce

■ Zákony velkých čísel

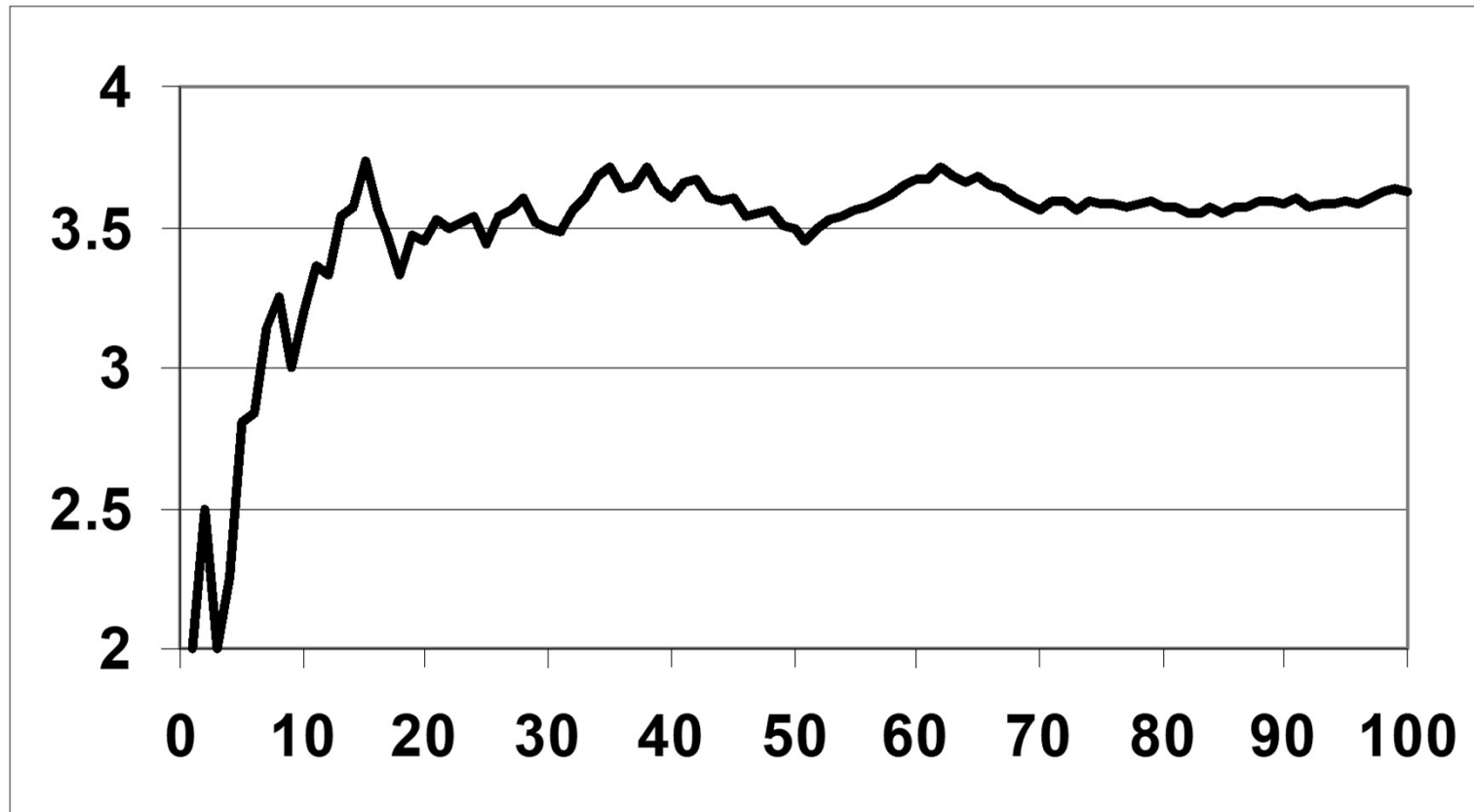
(chování bodových odhadů):

Např. průměr → střední hodnota (zde $EX=3,500$)

Pořadí hodu	1	2	3	4	5	...	98	99	100
Hozeno	2	3	1	3	5	...	5	5	3
Průměr	2,000	2,500	2,000	2,250	2,800	...	3,622	3,636	3,630

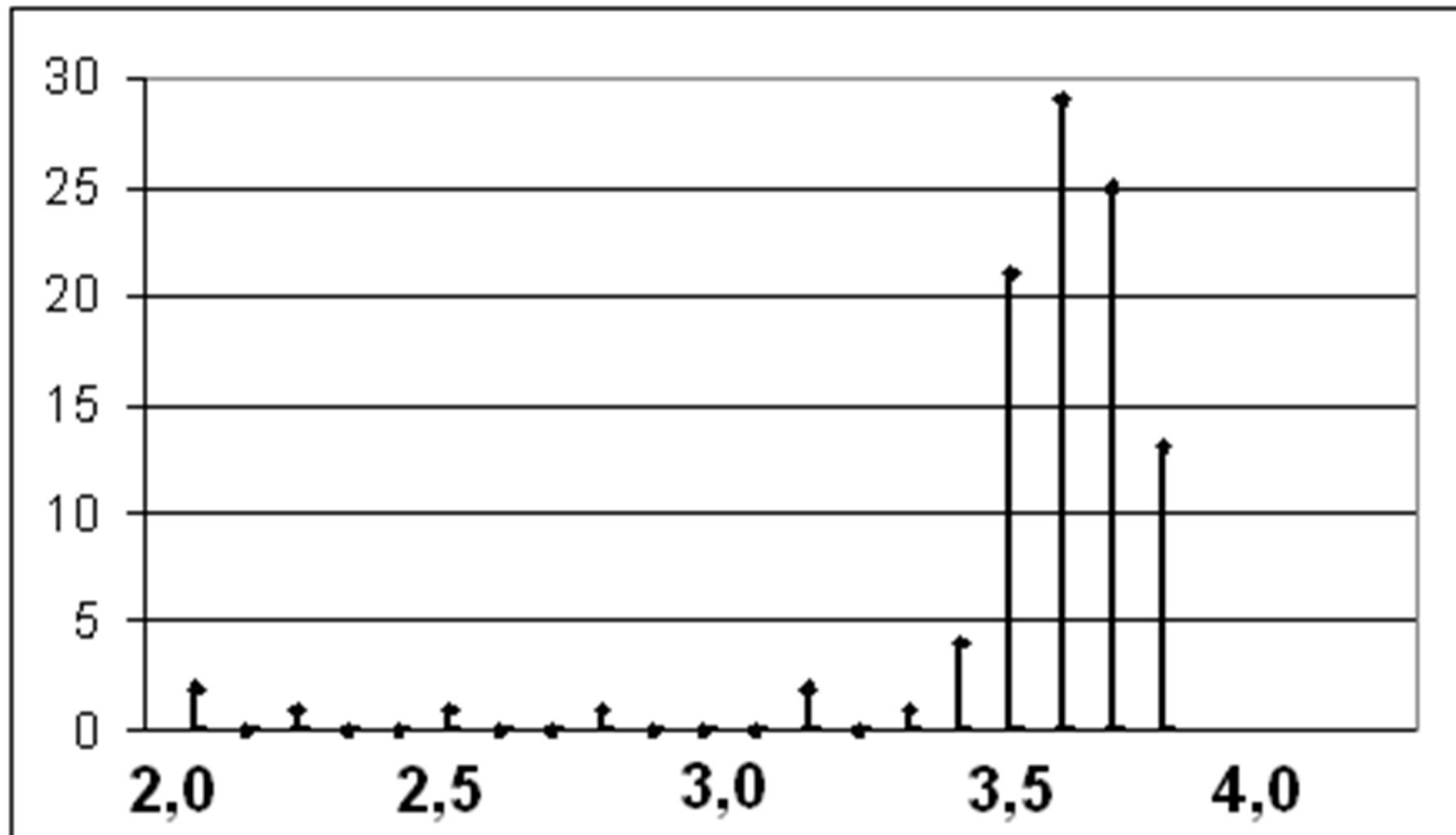
Základy statistické indukce

- Zákony velkých čísel (pokračování ilustrace):



Základy statistické indukce

- Zákony velkých čísel (pokr.-četnosti průměrů):



Základy statistické indukce

■ Centrální limitní věty (CLV)

Popisují asymptotické (tj. v limitě, v praxi pro „dostatečně velký“ počet stat. dat) chování testových charakteristik T_n jakožto náhodných veličin.

Např.

$$T_n = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})} \rightarrow U \sim N(0,1)$$

Intervaly spolehlivosti

= intervalové odhady neznámého parametru (odhad pro π , μ , $\sigma^2, \dots$),
odvozují se z příslušné CLV

spolehlivost = $1-\alpha$

= pravděpodobnost, že neznámá hodnota parametru je intervalem pokryta;

nejčastěji volba $1-\alpha = 0,95$ (95% I.S.)

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ při známém σ :

$$\bar{x} \pm u_{1-\alpha/2} \sigma/\sqrt{n}$$

Pro střední hodnotu μ při neznámém σ :

$$\bar{x} \pm t_{1-\alpha/2 (n-1)} s/\sqrt{n}$$

kde $n-1$ = počet stupňů volnosti (DF)

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ - vzorce:

STATISTIKA – přehled vzorců

$\left(p - \frac{1}{2n} - u_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}} ; p + \frac{1}{2n} + u_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}} \right)$	
$\left(\bar{x} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} ; \bar{x} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right),$	$\left(\bar{x} - t_{1-\alpha/2}(n-1) \cdot \frac{s}{\sqrt{n}} ; \bar{x} + t_{1-\alpha/2}(n-1) \cdot \frac{s}{\sqrt{n}} \right)$
$H_0: \pi = \pi_0 \quad T = \frac{(p - \pi_0) \cdot \sqrt{n}}{\sqrt{\pi_0(1 - \pi_0)}}$	$H_1: \pi \neq \pi_0 \quad \dots \quad W = (-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; \infty)$
	$H_1: \pi > \pi_0 \quad \dots \quad W = (u_{1-\alpha}; \infty)$
	$H_1: \pi < \pi_0 \quad \dots \quad W = (-\infty; -u_{1-\alpha})$
$H_0: \mu = \mu_0 \quad T = \frac{(\bar{x} - \mu_0) \cdot \sqrt{n}}{\sigma}$	$H_1: \mu \neq \mu_0 \quad \dots \quad W = (-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; \infty)$
	$H_1: \mu > \mu_0 \quad \dots \quad W = (u_{1-\alpha}; \infty)$

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ („ručně“):

Př. Dle věku osmi náhodně vybraných čtenářů dětského časopisu odhadněte střední věk čtenářů tohoto časopisu.

Věky popořadě:

12, 14, 15, 12, 15, 14, 12, 15.

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ („ručně“):

$$\text{průměrný věk} = 109/8 = 13,625$$

$$M_2 = 1499/8 - 13,625^2 = 1,734$$

$$s^2 = 1,734 \cdot 8/7 = 1,982$$

$$s = \sqrt{1,982} = 1,408$$

$$t_{0,975(7)} = 2,365$$

Kvantily rozdělení t

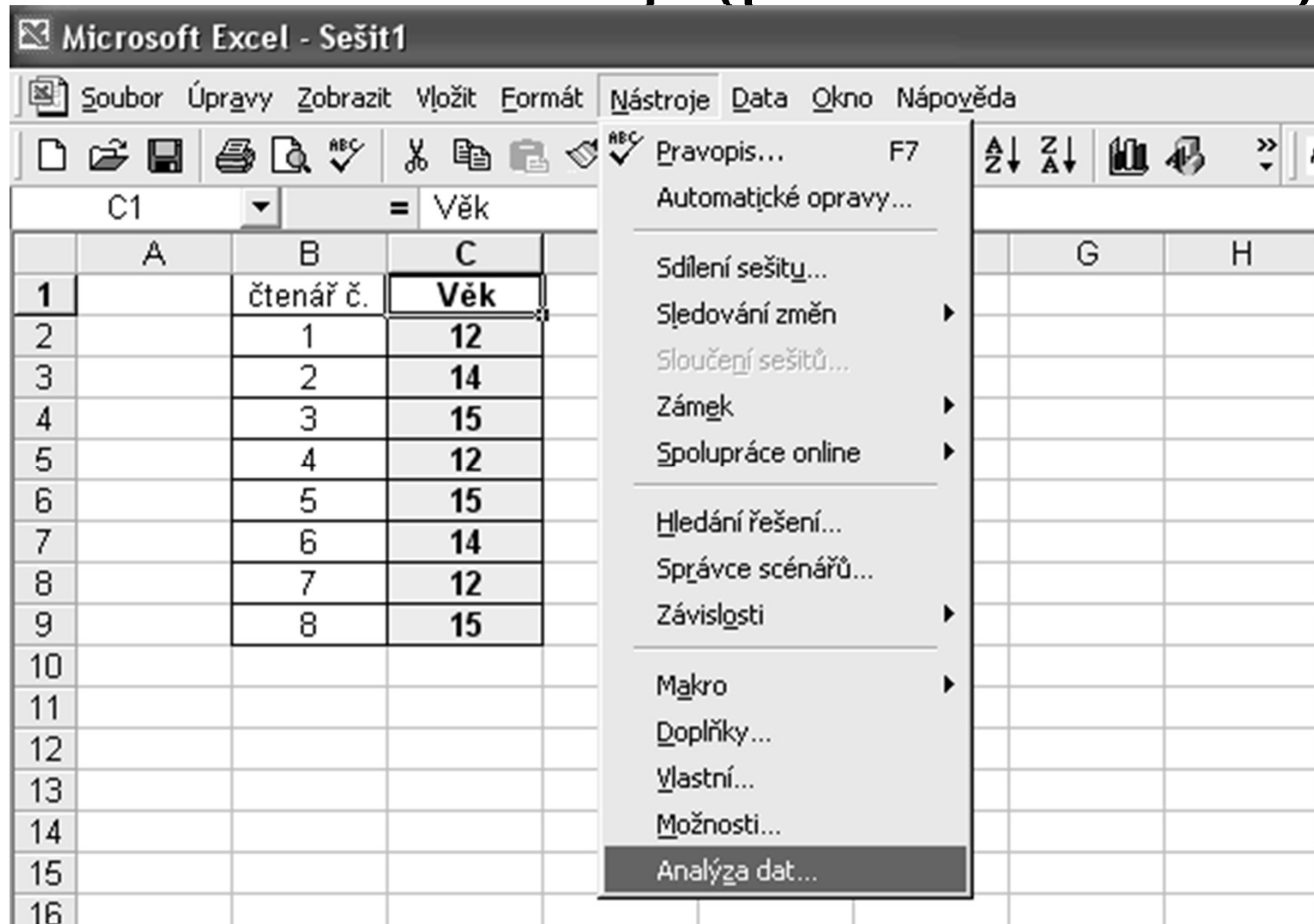
P	0,90	0,95	0,975	0,99
1	3,078	6,314	12,706	31,821
2	1,886	2,920	4,303	6,965
3	1,638	2,353	3,182	4,541
4	1,533	2,132	2,776	3,747
5	1,476	2,015	2,571	3,365
6	1,440	1,943	2,447	3,143
7	1,415	1,895	2,365	2,998
8	1,397	1,860	2,306	2,896
9	1,382	1,833	2,262	2,819

$$\text{dolní mez} = 13,625 - 2,365 \cdot 1,408 / \sqrt{8} = \underline{\underline{12,448}}$$

$$\text{horní mez} = 13,625 + 2,365 \cdot 1,408 / \sqrt{8} = \underline{\underline{14,802}}$$

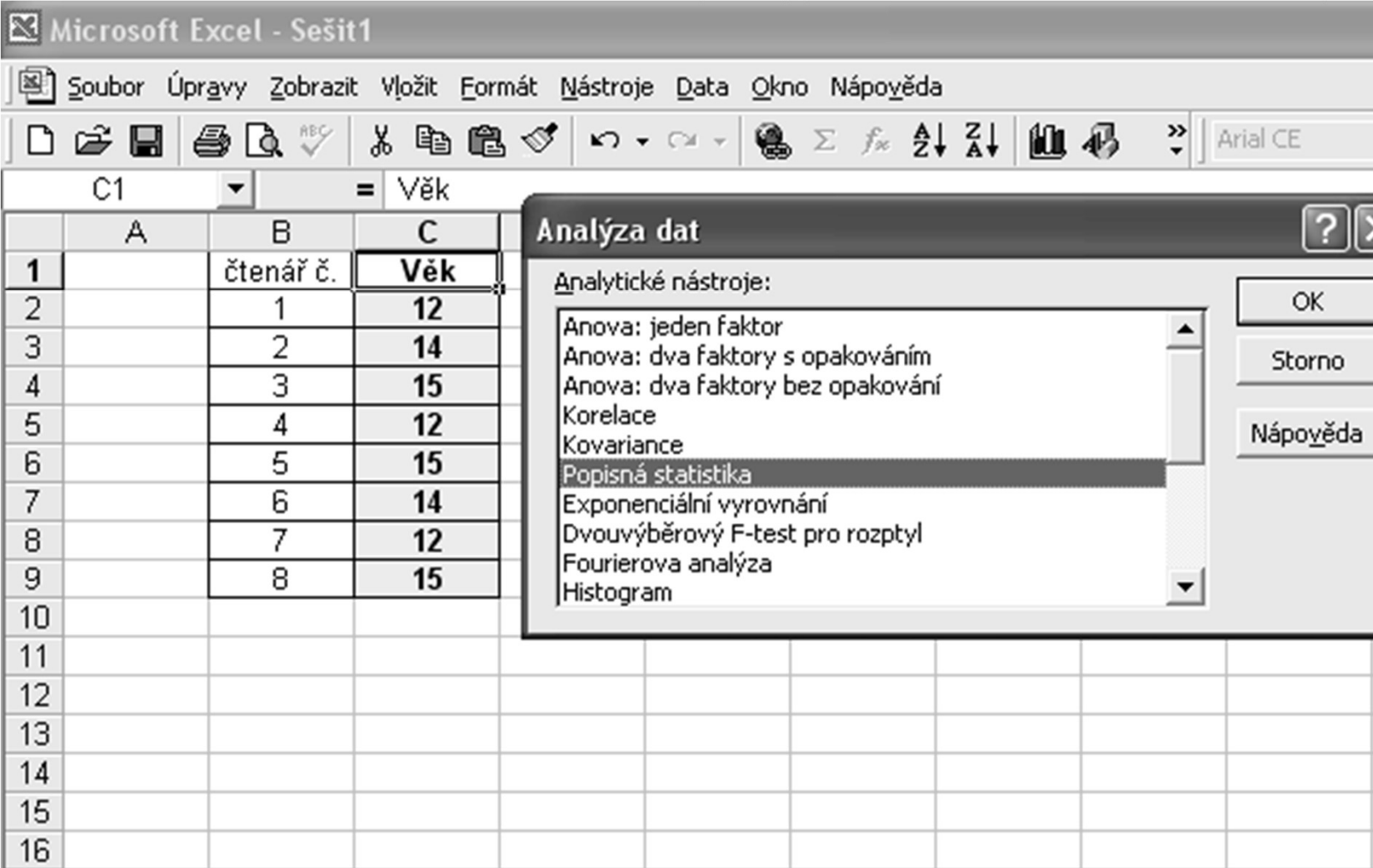
Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ (pomocí Excelu):



Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ (pomocí Excelu):



The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'Sešit1'. The spreadsheet contains data in columns A, B, and C. Column A is labeled 'čtenář č.' (reader number) and column B is labeled 'Věk' (age). The data is as follows:

	A	B	C
1		čtenář č.	Věk
2		1	12
3		2	14
4		3	15
5		4	12
6		5	15
7		6	14
8		7	12
9		8	15

The 'Analyza dat' (Data Analysis) task pane is open on the right side of the spreadsheet. It lists various statistical tools under the heading 'Analytické nástroje:' (Analytical tools:). The tools listed are:

- Anova: jeden faktor
- Anova: dva faktory s opakováním
- Anova: dva faktory bez opakování
- Korelace
- Kovariance
- Popisná statistika
- Exponenciální vyrovňání
- Dvouvýběrový F-test pro rozptyl
- Fourierova analýza
- Histogram

The 'Popisná statistika' (Descriptive Statistics) option is currently selected. The task pane also includes buttons for 'OK', 'Storno' (Cancel), and 'Nápověda' (Help).

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ (pomocí Excelu):

Microsoft Excel - Sešit1

Soubor Úpravy Zobrazit Vložit

C2 = Věk

	A	B	C
1		čtenář č.	Věk
2		1	12
3		2	14
4		3	15
5		4	12
6		5	15
7		6	14
8		7	12
9		8	15
10			
11			
12			
13			
14			
15			
16			

Popisná statistika

Vstup

Vstupní oblast: \$C\$2:\$C\$9

Sdružit:

☒ Sloupce

☐ Řádky

☐ Popisky v prvním řádku

Možnosti výstupu

☐ Výstupní oblast:

☒ Nový list:

☐ Nový sešit

☒ Celkový přehled

☒ Hladina spolehlivosti pro stř. hodnotu: 95 %

☐ K-té největší: 1

☐ K-té nejmenší: 1

OK Storno Nápověda

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ (pomocí Excelu):

	A	B
1	Sloupec1	
2		
3	Stř. hodnota	13,625
4	Chyba stř. hodnoty	0,498
5	Medián	14
6	Modus	12
7	Směr. odchylka	1,408
8	Rozptyl výběru	1,982
9	Špičatost	-2,135
10	Šikmost	-0,339
11	Rozdíl max-min	3
12	Minimum	12
13	Maximum	15
14	Součet	109
15	Počet	8
16	Hladina spolehlivosti (95,0%)	1,177

dolní mez:

$$13,625 - 1,177 = \underline{12,448};$$

horní mez:

$$13,625 + 1,177 = \underline{14,802}$$

Oboustranné intervaly spolehlivosti

Pro střední hodnotu μ (odpověď):

S 95% spolehlivostí je střední věk čtenářů daného časopisu z rozmezí 12,448 až 14,802 roku.

Zpřesnění odhadu (tj. zúžení IS)?

- a) zvýšit n (=změna dat);
- b) snížit spolehlivost (data stejná);
- c) snížit variabilitu (=změna populace).

Oboustranné intervaly spolehlivosti

Ilustrace vlivu zvýšení n (viz ZVČ):

Intervaly spolehlivosti pro střední hozené číslo na kostce

