Statistika I (KMI/PSTAT)

Cvičení první

aneb

Sumační symbolika, úvod do popisné statistiky

Obsah hodiny

Po dnešní hodině byste měli být schopni:

- správně používat sumační symboliku
- správně používat pojmy statistický soubor, statistická jednotka, znak statistické jednotky, hodnota znaku, statistická proměnná,
- rozlišovat jednotlivé typy statistických proměnných podle jejich nejrůznějších vlastností (kategoriální $\times$ nekategoriální, kvalitativní $\times$ kvantitativní (diskrétní $\times$ spojité), nominální $\times$ ordinální, alternativní $\times$ množné),
- vypočítat četnosti hodnot statistické proměnné (prostá absolutní četnost, kumulovaná absolutní četnost, prostá relativní četnost, kumulovaná relativní četnost) a sestavit tabulku těchto četností,
- nakreslit podle tabulky četností vhodný graf statistické proměnné (histogram, koláčový graf, polygon četností atd.),
- z hodnot dvou statistických proměnných sestavit kontingenční tabulku a umět číst údaje z těchto kontingenčních tabulek.

Základní pojmy

- sumační symbolika
- statistický soubor
- statistická jednotka
- rozsah souboru
- statistický znak
- hodnota znaku
- statistická proměnná
- četnosti
- tabulky četností
- intervalové rozdělení četností
- kontingenční tabulka

Sumační symbolika

Sumační symbolika I

V následujících příkladech rozepište výrazy:

- $\sum_{n=1}^{10} n$

Sumační symbolika

Sumační symbolika I

V následujících příkladech rozepište výrazy:

- $$\sum_{n=1}^{10} n = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10$$

Sumační symbolika I

V následujících příkladech rozepište výrazy:

- $\sum_{n=1}^{10} n = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10$
- $\sum_{n=1}^{4} (3n + 5)$

Sumační symbolika I

V následujících příkladech rozepište výrazy:

- $\sum_{n=1}^{10} n = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10$
- $\sum_{n=1}^{4} (3n + 5) = (3 \cdot 1 + 5) + (3 \cdot 2 + 5) + (3 \cdot 3 + 5) + (3 \cdot 4 + 5)$

Sumační symbolika I

V následujících příkladech rozepište výrazy:

- $\sum_{n=1}^{10} n = 1+2+3+4+5+6+7+8+9+10$
- $\sum_{n=1}^{4} (3n+5) = (3\cdot 1+5)+(3\cdot 2+5)+(3\cdot 3+5)+(3\cdot 4+5)$
- $\sum_{n=1}^{3} 2^n$
- $\sum_{n=1}^{4} x^n$
- $\sum_{x=1}^{4} x^n$
- $\sum_{i=1}^{5} x_i^2$

Sumační symbolika II

V následujících příkladech rozepište výrazy:

- $\sum_{i=1}^{6} n$
- $\sum_{n=1}^{6} 1$

Sumační symbolika II

V následujících příkladech rozepište výrazy:

- $\sum_{i=1}^{6} n$
- $\sum_{n=1}^{6} 1$
- $\sum_{i=1}^{3}\sum_{j=1}^{2} a_{ij}$
- $\sum_{i=1}^{3}\sum_{j=1}^{3}(x_i - x_j)$

Sumační symbolika II

V následujících příkladech rozepište výrazy:

- $\sum_{i=1}^{6} n$
- $\sum_{n=1}^{6} 1$
- $\sum_{i=1}^{3} \sum_{j=1}^{2} a_{ij}$
- $\sum_{i=1}^{3} \sum_{j=1}^{3} (x_i - x_j)$

Sumační symbolika III

V následujících příkladech zapište výrazy pomocí sumační symboliky:

- $2 + 4 + 6 + 8 + 10 + 12$
- $11 + 13 + 15 + 17 + 19 + 21 + 23 + 25$
- $n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7$
- $n_{11} + n_{12} + n_{13} + n_{14} + n_{21} + n_{22} + n_{23} + n_{24} + n_{31} + n_{32} + n_{33} + n_{34}$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{i=1}^{n} (x_i + y_i)$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{i=1}^{n} (x_i + y_i)$

Sumační symbolika - suma součtu, rozdílu

$$\sum_{i=1}^{n} [f(i) + g(i)] = \sum_{n=1}^{n} [f(i)] + \sum_{n=1}^{n} [g(i)]$$

$$\sum_{i=1}^{n} [f(i) - g(i)] = \sum_{n=1}^{n} [f(i)] - \sum_{n=1}^{n} [g(i)]$$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{i=1}^{n}(x_i + y_i)$

Sumační symbolika - suma součtu, rozdílu

$$\sum_{i=1}^{n}[f(i)+g(i)] = \sum_{n=1}^{n}[f(i)] + \sum_{n=1}^{n}[g(i)]$$

$$\sum_{i=1}^{n}[f(i)-g(i)] = \sum_{n=1}^{n}[f(i)] - \sum_{n=1}^{n}[g(i)]$$

$$\sum_{n=1}^{10}(5n^3 - 3n^2 + 2n - 11) = \sum_{n=1}^{10}(5n^3) - \sum_{n=1}^{10}(3n^2) + \sum_{n=1}^{10}(2n) - \sum_{n=1}^{10}(11)$$

Sumační symbolika - pravidla

Rozepište výraz $\sum_{n=1}^{6} (5n^2)$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{n=1}^{6} (5n^2)$

Sumační symbolika - suma součinu s konstantou

$$\sum_{i=1}^{n} [c \cdot f(i)] = c \cdot \sum_{i=1}^{n} [f(i)]$$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{n=1}^{6} (5n^2)$

Sumační symbolika - suma součinu s konstantou

$$\sum_{i=1}^{n} [c \cdot f(i)] = c \cdot \sum_{i=1}^{n} [f(i)]$$

$$\sum_{n=1}^{10} (5n^3 - 3n^2 + 2n - 11) = 5 \cdot \sum_{n=1}^{10} (n^3) - 3 \cdot \sum_{n=1}^{10} (n^2) + 2 \sum_{n=1}^{10} (n) - 11 \sum_{n=1}^{10} (1)$$

Sumační symbolika

Sumační symbolika - pravidla

Rozepište výraz $\sum_{n=1}^{6} (5n^2)$

Sumační symbolika - suma součinu s konstantou

$$\sum_{i=1}^{n} [c \cdot f(i)] = c \cdot \sum_{i=1}^{n} [f(i)]$$

$$\sum_{n=1}^{10} (5n^3 - 3n^2 + 2n - 11) = 5 \cdot \sum_{n=1}^{10} (n^3) - 3 \cdot \sum_{n=1}^{10} (n^2) + 2 \sum_{n=1}^{10} (n) - 11 \sum_{n=1}^{10} (1)$$

$$\sum_{k=1}^{n} (2k - 1) = ???$$

Základní pojmy

statistická jednotka	statistický znak	hodnota znaku
Jan Novák	výška	184 cm
	hmotnost	92 cm
	barva vlasů	černá
	počet sourozenců	1
	pohlaví	muž
Jiří Novotný	výška	176 cm
	hmotnost	72 kg
	barva vlasů	hnědá
	počet sourozenců	2
	pohlaví	muž
Jana Rychtářová	výška	171 cm
	hmotnost	65 kg
	barva vlasů	hnědá
	počet sourozenců	1
	pohlaví	žena
Jitka Kovářová	výška	166 cm
	hmotnost	64 kg
	barva vlasů	blond
	počet sourozenců	0
	pohlaví	žena

Typy veličin

- **kvalitativní**
  - pohlaví (muž, žena, ...)
  - výsledek přijímacího řízení (uspěl, neuspěl)
  - barva očí (modrá, hnědá, zelená, šedivá, ...)
  - nejvyšší dosažené vzdělání (ZŠ, SŠ, VŠ)
- **kvantitativní**
  - hmotnost (64 kg)
  - pořadí v závodě (1. místo)
  - cena akcie (542 Kč)
  - teplota (15°C)

Číselné (kvantitativní) proměnné dále dělíme na

- **nespojité** (diskrétní): počet sourozenců, počet vypůjčených knih, počet vlastněných mobilů atd.
- **spojité** (kontinuální): hmotnost, výška, čas atd.

Typy veličin

- **nominální** (názvové) - nelze objektivně stanovit pořadí hodnot
  - barva vlasů (*světlé, zrzavé, černé, fialové, zelené, ...*)
  - náboženské vyznání (*katolíci, protestanté, hinduisté, ...*)
  - oblíbený sport (*běh, fotbal, hokej, ...*)
- **ordinální** (pořadové) lze jednoznačně objektivně seřadit od nejnižší obměny k obměně nejvyšší z hlediska stupně sledované vlastnosti
  - nejvyšší dosažené vzdělání (*ZŠ, SŠ, VŠ*)
  - známka při zkoušení (*výborně, velmi dobře, dobře, nevyhověl*)
  - cena zboží (*25 Kč, 27 Kč, 28 Kč, 31 Kč*)

- **alternativní** - hodnoty mohou nabýt pouze dvě obměny
  - pohlaví (*muž, žena*)
  - výsledek zápočtu (*započteno, nezapočteno*)
- **množné** - více než dvě možné obměny hodnot znaku
  - oblíbený nápoj (*pivo, limo, káva, víno, voda*)
  - nejvyšší dosažené vzdělání (*ZŠ, SŠ, VŠ*)

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

b) Kolik respondentů mělo vysokoškolské vzdělání?

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

b) Kolik respondentů mělo vysokoškolské vzdělání?

c) Jaký byl podíl lidí s VŠ vzděláním?

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

b) Kolik respondentů mělo vysokoškolské vzdělání?

c) Jaký byl podíl lidí s VŠ vzděláním?

d) Kolik respondentů mělo nejvýše středoškolské vzdělání?

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

b) Kolik respondentů mělo vysokoškolské vzdělání?

c) Jaký byl podíl lidí s VŠ vzděláním?

d) Kolik respondentů mělo nejvýše středoškolské vzdělání?

e) Kolik respondentů bylo ve věku 35 let?

Četnosti

V rámci dotazníkového šetření byla shromážděna data od 30 respondentů. Tito byli dotázáni na svůj věk, nejvyšší dosažené vzdělání a pohlaví. Zjištěné výsledky jsou uvedeny v tabulce.

respondent č.	1	2	3	4	5	6	7	8	9	10
vzdělání	ZŠ	ZŠ	VŠ	SŠ	VŠ	SŠ	SŠ	SŠ	VŠ	SŠ
pohlaví	M	M	M	Ž	M	Ž	M	M	Ž	Ž
věk	23	15	35	27	28	34	65	21	43	25

respondent č.	11	12	13	14	15	16	17	18	19	20
vzdělání	SŠ	SŠ	VŠ	VŠ	SŠ	VŠ	ZŠ	VŠ	SŠ	VŠ
pohlaví	M	Ž	M	Ž	Ž	Ž	Ž	Ž	Ž	Ž
věk	35	24	27	28	36	32	41	56	49	37

respondent č.	21	22	23	24	25	26	27	28	29	30
vzdělání	VŠ	VŠ	ZŠ	SŠ	SŠ	SŠ	SŠ	VŠ	SŠ	VŠ
pohlaví	M	M	Ž	M	Ž	M	Ž	M	M	Ž
věk	47	28	23	34	48	52	18	24	31	42

a) Kolik bylo mezi respondenty mužů a kolik žen?

b) Kolik respondentů mělo vysokoškolské vzdělání?

c) Jaký byl podíl lidí s VŠ vzděláním?

d) Kolik respondentů mělo nejvýše středoškolské vzdělání?

e) Kolik respondentů bylo ve věku 35 let?

f) Kolik respondentů bylo ve věku nejvýše 35 let?

Četnosti hodnot znaku

- **absolutní četnost** n_i **hodnoty** x_i: počet znaků s hodnotou x_i

---

- **relativní četnost** p_i **hodnoty** x_i: podíl znaků s hodnotou x_i

$$p_i = \frac{n_i}{n}$$

- **kumulovaná absolutní četnost** n_i^* **hodnoty** x_i: počet znaků s hodnotou nejvýše x_i

$$n_i^* = \sum_{k=1}^{i} n_k = n_1 + n_2 + \ldots + n_i$$

- **kumulovaná relativní četnost** p_i^* **hodnoty** x_i: podíl znaků s hodnotou nejvýše x_i

$$p_i^* = \sum_{k=1}^{i} p_k = p_1 + p_2 + \ldots + p_i = \frac{n_1}{n} + \frac{n_2}{n} + \ldots + \frac{n_i}{n} = \frac{n_i^*}{n}$$

!!! Kumulované četnosti má smysl uvádět pouze pro ordinální veličiny!!!

Četnosti

i	A	B	C	D	E
1	3	33	ov	15 214	ano
2	2	27	dr	21 463	ano
3	3	35	ov	22 274	ano
4	4	47	ov	32 236	ne
5	5	52	ov	27 148	ano
6	4	43	nv	31 884	ne
7	1	67	ov	15 224	ne
8	4	50	dr	26 398	ano
9	3	32	dr	23 594	ne
10	2	25	ov	45 395	ano
11	1	30	nv	18 964	ano
12	4	35	ov	25 982	ano
13	5	43	ov	27 134	ne
14	2	25	dr	22 546	ne
15	3	29	dr	23 641	ano
16	2	29	dr	25 362	ne
17	2	27	dr	13 547	ano
18	3	25	ov	42 378	ne
19	4	43	dr	38 221	ano
20	4	42	ov	24 635	ne

Zadání příkladů

i ... pořadové číslo domácnosti
A ... počet členů domácnosti
B ... věk nejstaršího člena domácnosti
C ... druh vlastnictví bytu (osobní, družstevní, nájemní)
D ... průměrné měsíční výdaje domácnosti v Kč
E ... vlastnictví PC.

Četnosti

i	A	B	C	D	E
1	3	33	ov	15 214	ano
2	2	27	dr	21 463	ano
3	3	35	ov	22 274	ano
4	4	47	ov	32 236	ne
5	5	52	ov	27 148	ano
6	4	43	nv	31 884	ne
7	1	67	ov	15 224	ne
8	4	50	dr	26 398	ano
9	3	32	dr	23 594	ne
10	2	25	ov	45 395	ano
11	1	30	nv	18 964	ano
12	4	35	ov	25 982	ano
13	5	43	ov	27 134	ne
14	2	25	dr	22 546	ne
15	3	29	dr	23 641	ano
16	2	29	dr	25 362	ne
17	2	27	dr	13 547	ano
18	3	25	ov	42 378	ne
19	4	43	dr	38 221	ano
20	4	42	ov	24 635	ne

Zadání příkladů

i . . . pořadové číslo domácnosti
A . . . počet členů domácnosti
B . . . věk nejstaršího člena domácnosti
C . . . druh vlastnictví bytu (osobní, družstevní, nájemní)
D . . . průměrné měsíční výdaje domácnosti v Kč
E . . . vlastnictví PC.

1. Sestavte **tabulku rozdělení četností** veličiny *počet členů domácnosti*.
2. Sestavte **tabulku rozdělení četností** veličiny *druh vlastnictví bytu*.
3. Sestavte **kontingenční tabulku** pro veličiny *druh vlastnictví bytu* a *vlastnictví PC*.

Intervalové rozdělení četností

Pro spojitou (metrickou) veličinu sestavujeme **intervalové rozdělení četností**, tj. neurčujeme četnosti pro konkrétní hodnoty, ale četnosti hodnot, které se nacházejí v jistém rozmezí hodnot (tj. v jistém intervalu reálných čísel).

Obecně: pokud by při diskrétním rozdělení byly četnosti jednotlivých znaků menší než celkový počet znaků, použijeme intervalové rozdělení četností.

Intervalové rozdělení četností

Pro spojitou (metrickou) veličinu sestavujeme **intervalové rozdělení četností**, tj. neurčujeme četnosti pro konkrétní hodnoty, ale četnosti hodnot, které se nacházejí v jistém rozmezí hodnot (tj. v jistém intervalu reálných čísel).

Obecně: pokud by při diskrétním rozdělení byly četnosti jednotlivých znaků menší než celkový počet znaků, použijeme intervalové rozdělení četností.

Počet kategorií (tj. intervalů) lze určit např. pomocí tzv. **Sturgesova pravidla**.

Sturgesovo pravidlo

Počet kategorií: $k = 1 + 3,3 \cdot \log n$,

k ... počet kategorií (intervalů),
n ... počet hodnot (pozorování).

Intervalové rozdělení četností

Pro spojitou (metrickou) veličinu sestavujeme **intervalové rozdělení četností**, tj. neurčujeme četnosti pro konkrétní hodnoty, ale četnosti hodnot, které se nacházejí v jistém rozmezí hodnot (tj. v jistém intervalu reálných čísel).

Obecně: pokud by při diskrétním rozdělení byly četnosti jednotlivých znaků menší než celkový počet znaků, použijeme intervalové rozdělení četností.

Počet kategorií (tj. intervalů) lze určit např. pomocí tzv. **Sturgesova pravidla**.

Sturgesovo pravidlo

Počet kategorií: $k = 1 + 3,3 \cdot \log n$,

k ... počet kategorií (intervalů),
n ... počet hodnot (pozorování).

Počet kategorií

Při dotazníkovém šetření jsme zjišťovali údaje od 125 respondentů. Zjištěné hodnoty (spojité veličiny) se pohybovaly v rozmezí od 20 do 60. Navrhněte počet intervalů a jejich hranice, ve kterých budeme měřit četnosti hodnot.

Intervalové rozdělení četností

Pro spojitou (metrickou) veličinu sestavujeme **intervalové rozdělení četností**, tj. neurčujeme četnosti pro konkrétní hodnoty, ale četnosti hodnot, které se nacházejí v jistém rozmezí hodnot (tj. v jistém intervalu reálných čísel).

Obecně: pokud by při diskrétním rozdělení byly četnosti jednotlivých znaků menší než celkový počet znaků, použijeme intervalové rozdělení četností.

Počet kategorií (tj. intervalů) lze určit např. pomocí tzv. **Sturgesova pravidla**.

Sturgesovo pravidlo

Počet kategorií: $k = 1 + 3,3 \cdot \log n$,

k ... počet kategorií (intervalů),
n ... počet hodnot (pozorování).

Počet kategorií

Při dotazníkovém šetření jsme zjišťovali údaje od 125 respondentů. Zjištěné hodnoty (spojité veličiny) se pohybovaly v rozmezí od 20 do 60. Navrhněte počet intervalů a jejich hranice, ve kterých budeme měřit četnosti hodnot.

počet kategorií: $k = 1 + 3,3 \cdot \log 125 \doteq 1 + 3,3 \cdot 2,097 \doteq 7,92 \doteq 8$
šířka intervalu: $(60 - 20)/8 = 5$

Kontingenční tabulka

		velikost bytu					
		1+0	1+1	2+1	3+1	4+1	**celkem**
počet dětí	0	54	46	84	62	24	270
	1	23	59	112	84	69	347
	2	6	41	76	76	45	244
	3	1	15	35	34	33	118
	4	0	3	6	7	5	21
celkem		84	164	313	263	176	1000

1. Kolik rodin s 2 dětmi bydlí v bytech o velikosti 3+1?
2. Kolik rodin bydlí v bytech 1+1?
3. Kolik rodin má právě 3 děti?
4. Kolik rodin bydlí v bytě s nejvýše 2 místnostmi?
5. Kolik rodin má více než 2 děti?
6. Jaký je průměrný počet dětí v rodinách, které bydlí v bytech 1+0?
7. Jaký je průměrný počet dětí v rodinách v bytech s nejvýše 3 místnostmi?
8. Jaký je průměrný počet místností v bytech, kde bydlí bezdětné rodiny?
9. Kolik dětí dohromady bydlí v bytech o velikosti 2+1?
10. V jaké velikosti bytů bydlí celkem nejvíce dětí? Kolik je těchto dětí?