TECHNICKÁ UNIVERZITA V KOŠICIACH
STROJNÍCKA FAKULTA

ZÁKLADNÉ ŠTATISTICKÉ METÓDY

Dušan Knežo, Miriam Andrejiová, Gabriela Ižaríková

2011

RECENZOVALI: prof. RNDr. Martin Bača, CSc.
prof. RNDr. Jozef Džurina, CSc.

Predhovor

Tento učebný text je určený poslucháčom prvého ročníka bakalárskeho štúdia Strojníckej fakulty Technickej univerzity v Košiciach a je zameraný na potreby vyučovania predmetov orientovaných na základy matematickej štatistiky. Rovnako dobre však môže poslúžiť každému čitateľovi, ktorý chce získať základné vedomosti a zručnosti zo základov matematickej štatistiky.

V učebnom texte sú uvedené podstatné teoretické poznatky potrebné ku riešeniu úloh, riešené príklady a neriešené úlohy týkajúce sa popisnej štatistiky, teórie pravdepodobnosti, náhodnej veličiny, teórie odhadu, testovania hypotéz a regresnej a korelačnej analýzy. V závere textu sú uvedené základné tabuľky potrebné pre riešenie úloh, ktoré text obsahuje. Obsah je dostatočným základom pre štúdium a úspešné absolvovanie spomínaných predmetov.

Obom recenzentom prof. RNDr. Martinovi Bačovi, CSc. a prof. RNDr. Jozefovi Džurinovi, CSc.ďakujeme za dôsledné posúdenie tejto učebnej pomôcky. Ich cenné pripomienky, rady a odporúčania prispeli ku zvýšeniu kvality tejto publikácie.

V Košiciach 15. 11. 2011 Autori

Obsah

Kapitola 1

Popisná štatistika

1.1 Štatistické spracovanie údajov

1.1.1 Základné štatistické pojmy

Termín štatistika pochádza z latinského slova „status“, ktoré znamená stav alebo štát a vyjadroval sa ním súbor poznatkov o štáte.

- *Popisná štatistika*, úradnícka štatistika. K úradným zisťovaniam dochádzalo už niekoľko tisíc rokov pred naším letopočtom v starom Egypte, resp. v Číne, kedy vtedajší vládcovia potrebovali poznať čo najpresnejšie údaje pre vojenské a finančné účely.
- *Politická aritmetika* sa využívala na úplný popis obyvateľstva: natalita, mortalita, vývoj obyvateľstva. Skúmali sa teda hromadné javy, ktorými bolo možné po ich preštudovaní ovplyvňovať mocensky štát (politicky).
- *Induktívna štatistika* (moderná, analytická) vznikla na začiatku 20. storočia, rozvíjala metódy umožňujúce robiť závery o celku na základe výberov a čiastkových zisťovaní.

Štatistiku ako vednú disciplínu možno rozdeliť na dve časti:

- **Popisná (deskriptívna) štatistika**, ktorá sa zaoberá metódami zberu, spracovania, prezentácie a analýzy dát.
- **Matematická štatistika**, ktorá nachádza uplatnenie vtedy, ak zozbieranie údajov nie je z akýchkoľvek dôvodov možné a robíme záver o skúmanej udalosti na základe čiastkových údajov.

Štatistika je jazyk pre zhromažďovanie údajov, manipuláciu s nimi a ich interpretáciu. Úlohou štatistiky pri hodnotení informácií je spracovanie hromadných pozorovaní, ich interpretácia a analýza. Predmetom štatistiky ako vednej disciplíny sú *hromadné javy*, definujeme ju ako vedu o metódach kvantitatívneho hodnotenia vlastností hromadných javov. Nevyhnutným predpokladom každého štatistického skúmania je teda hromadnosť pozorovania. Hromadné javy sa za presne definovaných podmienok vecných, časových a priestorových viackrát vyskytujú.

Hromadný náhodný jav je ľubovoľný jav, ktorý sa opakuje a skladá sa z veľkého počtu prvkov. Hromadný jav sa skladá z mnohých individuálnych javov. Nositelia týchto javov sa nazývajú štatistické jednotky.

Štatistická jednotka je základný prvok, na ktorom pozorujeme konkrétny prejav určitej hromadnej udalosti. Teda štatistickou jednotkou je prvok, ktorý je predmetom skúmania a zbierame údaje o ňom s cieľom získať informácie o celku, do ktorého uvedený prvok patrí. Môže to byť osoba, meraná veličina, udalosť a pod.

Štatistické jednotky môžu byť vymedzené z priestorového hľadiska (vymedzenie priestoru, napríklad mesto, kraj), z časového hľadiska (vymedzenie obdobia, resp. okamihu, napríklad mesiac, kalendárny rok) a z vecného hľadiska (obsahové vymedzenie, napríklad študenti 1. ročníka SjF, obyvatelia Košíc).

Štatistický znak je vonkajší postihnuteľný, merateľný prejav skúmanej premenlivej vlastnosti štatistických jednotiek. Je to teda vlastnosť štatistických jednotiek. Podľa spôsobu vyjadrenia ich možno rozdeliť na:

- **Kvantitatívne** (číselné, merateľné) vyjadrujúce merateľné vlastnosti štatistických jednotiek číslami, delíme ich na spojité a diskrétne. *Spojité* nadobúdajú ľubovoľné hodnoty z ohraničeného alebo neohraničeného intervalu, napríklad teplota vody v bázene. *Diskrétne* nadobúdajú izolované, väčšinou celočíselné hodnoty, napríklad počet áut na parkovisku.

- **Kvalitatívne** (slovné), ktoré slovne vyjadrujú vlastnosti štatistických jednotiek a delíme ich na množné a alternatívne. *Množné*, ak slovný znak nadobúda viacero variantov, napríklad stupeň ukončeného vzdelania. *Alternatívne*, ak slovný znak nadobúda iba dve obmeny, napríklad pohlavie.

Štatistický súbor je množina štatistických jednotiek, ktoré majú požadované spoločné vlastnosti a ktoré vymedzujú štatistický súbor z hľadiska časového, priestorového a vecného. Každý štatistický súbor má svoj rozsah a obsah. *Rozsah súboru* je určený počtom štatistických jednotiek v štatistickom súbore, označujeme n. *Obsah súboru* je vymedzený štatistickými znakmi. **Základný súbor** je štatistický súbor vytvorený zo všetkých štatistických jednotiek, ktoré doň patria a na ktorých sledujeme hodnoty štatistického znaku. **Výberový súbor** tvoria vybrané štatistické jednotky, ktoré predstavujú reprezentatívnu podmnožinu základného súboru.

1.1.2 Štatistické triedenie

Triedenie štatistického súboru je prvým krokom pri spracovaní údajov. Pri vyšetrovaní kvantitatívneho znaku štatistického súboru zapisujeme číselné údaje v takom poradí, v akom sme ich získali, dostaneme *prvotnú tabuľku.*

Ak usporiadame hodnôty štatistického znaku podľa veľkosti, od najmenšej hodnoty po najväčšiu, pričom rovnaké hodnoty zapíšeme toľkokrát, koľkokrát sa vyskytujú v prvotnej tabuľke, dostaneme *variačný rad.* Označme si $x_{(i)}$ hodnoty štatistického znaku usporiadané do variačného radu z prvotnej tabuľky, kde i je priebežný index, za ktorý dosadzujeme poradové číslo hodnoty znaku usporiadaného podľa veľkosti jeho hodnôt. Variačný rad môžeme teda napísať v tvare:

$$x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq \ldots \leq x_{(n-1)} \leq x_{(n)}.$$

Celkový počet rôznych hodnôt znaku je l, $l \leq n$, rovnosť nastáva len v prípade, ak každá jednotka súboru nadobúda rôzne hodnoty sledovaného znaku.

Počet štatistických jednotiek s rovnakou hodnotou x_i, $i = 1, 2, \ldots, l$ nazývame *absolútnou početnosťou* hodnoty x_i. Označujeme ju n_i.

Pomocou absolútnych početností n_i môžeme definovať *absolútnu kumulatívnu početnosť* vzťahom $N_r = \sum\limits_{i=1}^{r} n_i$.

Relatívnu početnosť označujeme f_i a definujeme vzťahom $f_i = \dfrac{n_i}{n}$, pre $i = 1, 2, \ldots, l$.

Pomocou relatívnych početností f_i môžeme definovať *relatívnu kumulatívnu početnosť*, ktorú označujeme F_i a definujeme vzťahom $F_r = \sum_{i=1}^{r} f_i$. Relatívne početností sa často udávajú v percentách.

Ak štatistický súbor ma l rôznych hodnôt, potom platí:

$$\sum_{i=1}^{l} n_i = N_l = n, \qquad \sum_{i=1}^{l} f_i = F_l = 1.$$

Z variačného radu si môžeme pomocou čiarkovacej metódy utvoriť *variačnú tabuľku* (tabuľku početností). Jej podstata spočíva v tom, že vedľa hodnoty $x_{(1)}$, do druhého stĺpca, urobíme toľko čiarok, koľkokrát sa hodnota $x_{(1)}$ vo variačnom rade vyskytuje, pre rýchlejšie spočítavanie piatu čiarku zaznačíme vodorovne, cez štyri zvislé čiarky. Tento postup zopakujeme pre ďalšie hodnoty $x_{(i)}$.

i	x_i	čiark. metóda	n_i	f_i	N_i	F_i
1	$x_{(1)}$	$\lvert\ldots\rvert$	n_1	$f_1 = \frac{n_1}{n}$	$N_1 = n_1$	$F_1 = \frac{N_1}{n}$
2	$x_{(2)}$	$\lvert\ldots\rvert$	n_2	$f_2 = \frac{n_2}{n}$	$N_2 = n_1 + n_2$	$F_2 = \frac{N_2}{n}$
3	$x_{(3)}$	$\lvert\ldots\rvert$	n_3	$f_3 = \frac{n_3}{n}$	$N_3 = n_1 + n_2 + n_3$	$F_3 = \frac{N_3}{n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
l	$x_{(l)}$	$\lvert\ldots\rvert$	n_l	$f_l = \frac{n_l}{n}$	$N_l = n$	$F_l = 1$

Triedenie do intervalov

Ak ide o štatistické súbory väčšieho rozsahu ($n \geq 50$), alebo aj v prípade menšieho počtu rôznych hodnôt, skúmany štatistický súbor triedime do skupín (intervalov). Ide o variačné triedenie, pri ktorom sa celkový počet štatistických jednotiek v štatistickom súbore rozsahu n rozdelí do k tried.

Určenie počtu tried

V literarúre je uvedených veľa postupov a vzorcov pre odhad počtu tried. Počet tried nemá byť príliš malý, lebo by sa stratila podstatná časť informácie, pretože v tomto prípade je dĺžka triednych intervalov väčšia a výpočty sú menej presné. Naopak zvyšovanie počtu tried môže znížiť prehľadnosť a zvyšuje pracnosť výpočtov. Počet tried je ovplyvnený aj rozsahom štatistického súboru. Uvedieme niekoľko vzorcov na určenie počtu tried:

$$k \sim \sqrt{n}; \quad 0{,}55 \cdot n^{0{,}4} \leq k \leq 1{,}25 \cdot n^{0{,}4}; \quad k \leq 5 \cdot \log n; \quad k \sim 1 + 3{,}322 \cdot \log n.$$

Základné kritériá pri triedení

Pri triedení štatistických jednotiek do tried musia byť splnené dve zásady: zásada úplnosti - triedy musia byť vytvorené tak, aby každá jednotka mala šancu byť do niektorej z tried zatriedená a zásada jednoznačnosti - triedy musia byť vytvorené tak, aby o každej jednotke bolo jednoznačne rozhodnuté, do ktorej z tried má byť zaradená.

Nech j znamená poradové číslo triedy, t.j. $1 \leq j \leq k$. Ako I_j označujeme j-tý triedny interval, rozumieme ním $I_j = (t_j, t_{j+1}\rangle$ alebo $I_j = \langle t_j, t_{j+1})$, kde t_j je dolná hranica a t_{j+1} horná hranica.

Triedny znak z_j je stredom j-tého intervalu a vypočítame ho podľa vzťahu: $z_j = \dfrac{t_j + t_{j+1}}{2}$.

Dĺžka intervalu závisí od *variačného rozpätia* štatistického súboru. Pre variačné rozpätie platí:

$$R_v = x_{max} - x_{min}.$$

Rozdiel medzi dolnou a hornou hranicou intervalu je *dĺžka triedneho intervalu*, ktorú označujeme h a určíme podľa vzťahu: $h \doteq \frac{R_v}{k}$, upravíme ho zaokruhlením nahor.

Prvý riadok variačnej tabuľky má tvar:

j	$\langle t_j;\ t_{j+1}\rangle$	ČM	z_j	n_j	f_j	N_j	F_j

Triedenie štatististických jednotiek do tried si ukážeme konkrétne v príklade 2.

1.1.3 Grafické zobrazenie štatistického súboru

Grafy, vedľa tabuliek, sú druhým hlavným vyjadrovacím prostriedkom štatistiky. Pomocou grafov názorne prezentujeme výsledky štatistického spracovania údajov. *Bodové diagramy* sú grafy vytvorené bodmi $[x_j; w_j]$, kde za w_j môžeme dosadiť n_j, N_j, f_j, F_j, $j = 1, 2, \dots k$. *Spojnicové diagramy* sú v podstate bodové diagramy, v ktorých sú jednotlivé body spojené lomenou čiarou. *Polygón* je spojnicový graf, v ktorom sú jednotlivé susedné body spojené úsečkami. *Histogram početnosti* je stĺpcový graf, ktorý zostrojíme, tak, že nad úsekom na osi x, rovným dĺžke triedneho intervalu, nakreslíme obdĺžnik, ktorého výška je úmerná w_j. *Kruhové, koláčové grafy* tvorí kruh rozdelený na výseky.

Príklad 1. Na zistenie priemernej dennej teploty sme urobili 10 meraní: 27; 26; 25; 28; 29; 25; 28; 27; 25; 26. Zostavme tabuľku početností pre tento kvantitatívny znak a hodnoty znázorníme pomocou bodového grafu a polygónu.

Riešenie.

Najprv zoradíme hodnoty do variačného radu:

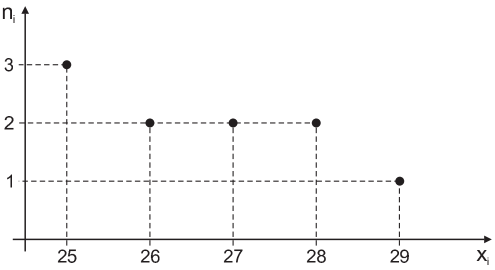
$$25 \leq 25 \leq 25 < 26 \leq 26 < 27 \leq 27 < 28 \leq 28 < 29.$$

Rozsah súboru je $n = 10$. Tabuľka početností má päť riadkov $l = 1, 2, \dots, 5$, keďže sa nám opakuje päť hodnôt.

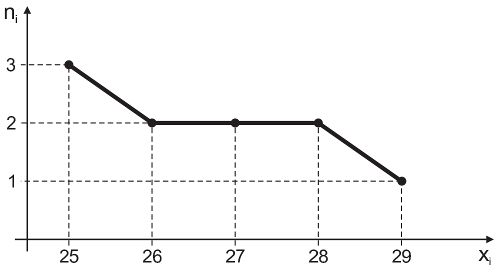
Tabuľka početností

i	x_i	ČM	n_i	f_i	N_i	F_i
1	25	|||	3	$\frac{3}{10} = 0{,}3$	3	$\frac{3}{10} = 0{,}3$
2	26	||	2	$\frac{2}{10} = 0{,}2$	3+2=5	$\frac{5}{10} = 0{,}5$
3	27	||	2	$\frac{2}{10} = 0{,}2$	5+2=7	$\frac{7}{10} = 0{,}7$
4	28	||	2	$\frac{2}{10} = 0{,}2$	7+2=9	$\frac{9}{10} = 0{,}9$
5	29	|	1	$\frac{1}{10} = 0{,}1$	9+1=10	$\frac{10}{10} = 1{,}0$

Príklad 2. V ankete sa medzi 20 majiteľmi rodinných domov sledovalo 5 znakov: súbor 1 - počet členov domácnosti, súbor 2 - počet bodov, ktoré získali po vyplnení dotazníka o bývaní (od 20 do 50), súbor 3 - ročná spotreba studenej vody (m^3), súbor 4 - prevažný spôsob vykurovania (drevo, elektrina, plyn), súbor 5 - spokojnosť so separáciou odpadu v tejto mestskej časti (hodnotená štvorbodovou škálou: spokojný, menej spokojný, nespokojný, vôbec nespokojný). Údaje sú v tabuľke. Zostavme tabuľky početností pre jednotlivé znaky. Súbor 1 a 3 znázorníme



Bodový graf



Polygón

pomocou histogramu a polygónu. Súbor 4 znázorníme pomocou kruhového diagramu.

Majiteľ	Súbor 1	Súbor 2	Súbor 3	Súbor 4	Súbor 5
1	4	26	105	plyn	vôbec nespokojný
2	2	28	48	elekrina	vôbec nespokojný
3	3	36	65	drevo	nespokojný
4	5	45	158	plyn	spokojný
5	4	34	143	plyn	spokojný
6	3	32	79	elekrina	vôbec nespokojný
7	2	27	86	elekrina	menej spokojný
8	2	29	54	plyn	spokojný
9	5	37	175	plyn	nespokojný
10	6	43	185	drevo	menej spokojný
11	5	48	120	elekrina	spokojný
12	4	26	163	drevo	menej spokojný
13	5	37	124	elekrina	vôbec nespokojný
14	2	26	60	elekrina	spokojný
15	4	32	95	plyn	menej spokojný
16	3	40	86	elekrina	spokojný
17	5	48	148	plyn	menej spokojný
18	6	48	196	drevo	menej spokojný
19	4	45	102	plyn	spokojný
20	3	46	87	plyn	nespokojný

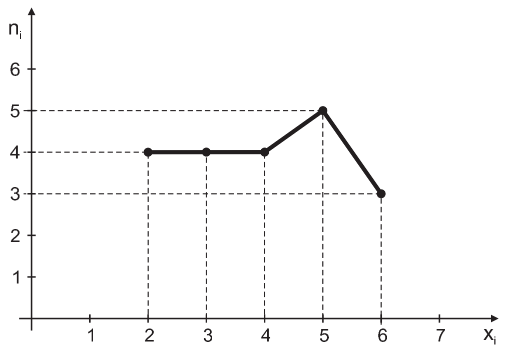
Riešenie.

Znak *Počet členov domácnosti - súbor 1* je kvantitatívny diskrétny znak s malým počtom obmien, čiže do tabuľky početností vypíšeme všetky možné obmeny hodnôt. Medzi údajmi sa opakuje 5 rôznych hodnôt, to znamená, že tabuľka bude mať päť riadkov. Čiarkovacou metódou zistíme počet opakovaní.

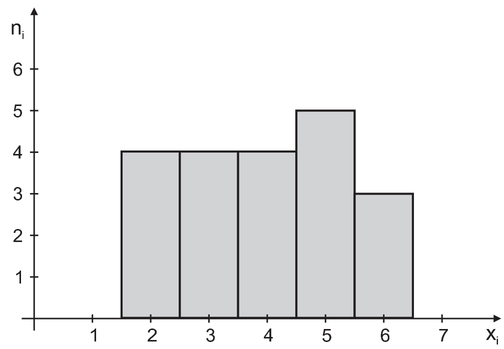
Tabuľka početností

<table>
<tr><th>i</th><th>x_i</th><th>ČM</th><th>n_i</th><th>f_i</th><th>N_i</th><th>F_i</th></tr>
<tr><td>1</td><td>2</td><td>||||</td><td>4</td><td>0,20</td><td>4</td><td>0,20</td></tr>
<tr><td>2</td><td>3</td><td>||||</td><td>4</td><td>0,20</td><td>8</td><td>0,40</td></tr>
<tr><td>3</td><td>4</td><td>||||</td><td>4</td><td>0,20</td><td>12</td><td>0,60</td></tr>
<tr><td>4</td><td>5</td><td>𝍸</td><td>5</td><td>0,25</td><td>17</td><td>0,85</td></tr>
<tr><td>5</td><td>6</td><td>|||</td><td>3</td><td>0,15</td><td>20</td><td>1</td></tr>
</table>

Z tabuľky napríklad vyplýva, že medzi obyvateľmi rodinných domov sú tri šesťčlenné rodiny, 25% rodín je päťčlenných a 60% rodín má najviac štyroch členov.



Polygón



Histogram

Znak *Počet bodov v dotazníku - súbor 2* je kvantitatívny diskrétny znak s veľkým počtom obmien, preto použijeme intervalové rozdelenie početnosti. Dáta rozdelíme do intervalov tak, aby bola splnená podmienka jednoznačnosti a úplnosti. Na určenie počtu tried použijeme vzťah

$$0{,}55 \cdot n^{0,4} \leq k \leq 1{,}25 \cdot n^{0,4},$$

kde $n = 20$, $x_{min} = 26$, $x_{max} = 48$.

- počet tried: $0{,}55 \cdot 20^{0,4} \leq k \leq 1{,}25 \cdot 20^{0,4}$; $1{,}82 \leq k \leq 4{,}13$; zvolíme napr. $k = 3$,
- variačné rozpätie: $R_v = x_{max} - x_{min} = 48 - 26 = 22$,
- dĺžka triedy: $h = \frac{R_v}{k} = \frac{22}{3} = 7{,}33 \sim 8$, $h = 8$.

Tabuľka početností

<table>
<tr><th>j</th><th>$(t_j,\, t_{j+1}\rangle$</th><th>z_j</th><th>ČM</th><th>n_j</th><th>f_j</th><th>N_j</th><th>F_j</th></tr>
<tr><td>1</td><td>$(25,\, 33\rangle$</td><td>29</td><td>𝍸 ||</td><td>7</td><td>0,35</td><td>7</td><td>0,35</td></tr>
<tr><td>2</td><td>$(33,\, 41\rangle$</td><td>37</td><td>𝍸</td><td>5</td><td>0,25</td><td>12</td><td>0,60</td></tr>
<tr><td>3</td><td>$(41,\, 49\rangle$</td><td>45</td><td>𝍸 |||</td><td>8</td><td>0,40</td><td>20</td><td>1,00</td></tr>
</table>

Z tabuľky je zrejmé, že medzi majiteľmi rodinných domov sú siedmi, ktorí v dotazníku mali 25 až 33 bodov, 25% je 33 až 41 bodových a 60% majiteľov dosiahlo nanajvýš 41 bodov.

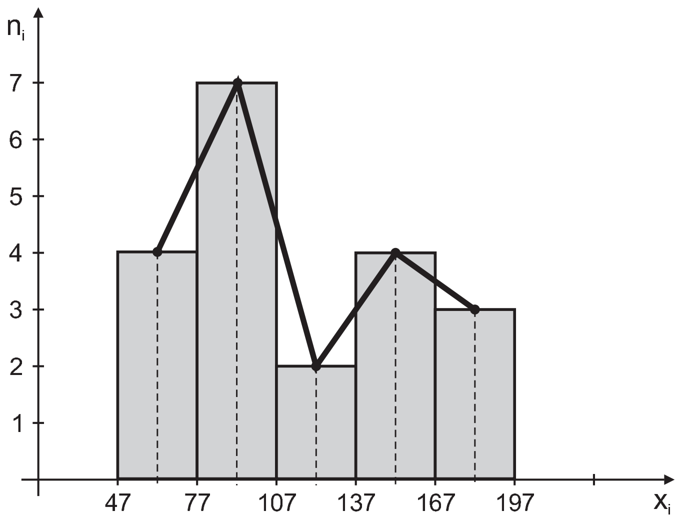
Znak *Ročná spotreba studenej vody - súbor 3* je kvantitatívny spojitý znak, preto použijeme rozdelenie hodnôt do tried. Dáta rozdelíme do intervalov tak, aby sa každá hodnota nachádzala práve v jednom intervale. Na určenie počtu tried môžeme použiť vzťah $k \sim \sqrt{n}$, kde $n = 20$, $x_{min} = 48$, $x_{max} = 196$.

- počet tried: $k \sim \sqrt{n} = \sqrt{20} = 4{,}47$; volíme $k = 5$,
- variačné rozpätie: $R_v = x_{max} - x_{min} = 196 - 48 = 148$,
- dĺžka triedy: $h = \frac{R_v}{k} = \frac{148}{5} = 29{,}6 \sim 30$, $h = 30$.

Tabuľka početností

j	$(t_j, t_{j+1}\rangle$	z_j	ČM	n_j	f_j	N_j	F_j
1	$(47, 77\rangle$	$\frac{47+77}{2} = 62$	||||	4	0,20	4	0,2
2	$(77, 107\rangle$	92	~~||||~~ ||	7	0,35	11	0,55
3	$(107, 137\rangle$	122	||	2	0,10	13	0,65
4	$(137\ 167\rangle$	152	||||	4	0,20	17	0,85
5	$(167, 197\rangle$	182	|||	3	0,15	20	1,00

Z tabuľky vyplýva, že v siedmich rodinných domoch je ročná spotreba studenej vody v intervale $(77,\ 107\rangle$, 10% zo všetkých domácnosti má spotrebu v intervale $(107,\ 137\rangle$ a 85% má spotrebu maximálne 167 m^3.



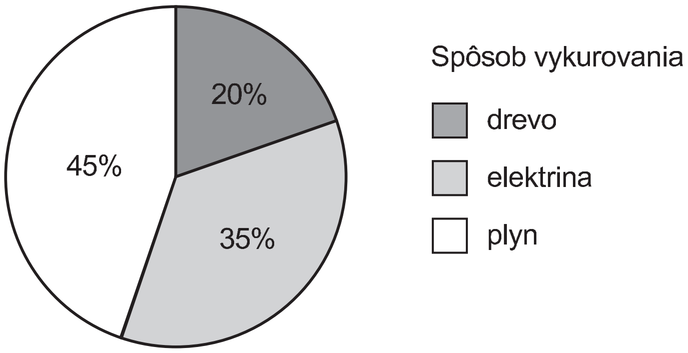
Histogram a polygón

Znak *Prevažný spôsob vykurovania - súbor 4* je slovný znak s tromi obmenami - drevo, elektrina, plyn.

Tabuľka početností

Vykurovanie	Obmena znaku	Početnosť	
	a_i	absolútna	relatívna
drevo	a_1	4	0,20
elektrina	a_2	7	0,35
plyn	a_3	9	0,45

Z tabuľky vyplýva, že v siedmich rodinných domoch na vykurovanie využívajú prevažne elektrinu, 45% zo všetkých domácnosti využíva na vykurovanie prevažne plyn.



Kruhový graf

Znak *Spokojnosť so separáciou odpadu - súbor 5* je kvalitatívný znak so štyrmi obmenami. Slovným znakom priradíme kvantitatívne dáta: 1 - spokojný, 2 - menej spokojný, 3 - nespokojný, 4 - vôbec nespokojný.
Tabuľka početností

Spokojnosť so separáciou	Obmena znaku	Početnosť	
	a_i	absolútna	relatívna
spokojný	1	7	0,35
menej spokojný	2	6	0,30
nespokojný	3	3	0,15
vôbec nespokojný	4	4	0,20

Z tabuľky vieme zistiť, že traja majitelia rodinných domoch sú nespokojní a 35% zo všetkých majiteľov je spokojných so separáciou odpadu.

Úlohy

V nasledujúcich úlohách nájdite rozdelenie početností a stanovte kompletnú tabuľku početností. Štatistický súbor znázornite graficky pomocou histogramu a polygónu.

1.1 Pri kontrole automatickej linky boli zistené údaje o váhe automaticky balených výrobkov. Výsledky (v g) sú: 40; 43; 45; 42; 44; 43; 47; 41; 43; 44; 41; 43; 42; 42; 44; 43; 42; 42; 43; 47.

1.2 Energetické náklady na prevádzku závodu počas 46 týždňov (v tis.€) sú: 38,1; 36,7; 37,5; 38,7; 38,1; 37,5; 36,7; 38,1; 36,7; 37,5; 38,7; 38,1; 40,1; 38,1; 36,7; 36,1; 37,5; 39,5; 35,5; 38,7; 38,1; 37,5; 36,7; 38,1; 37,5; 38,7; 36,7; 37,5; 38,1; 38,7; 37,5; 36,1; 38,1; 37,5; 38,1; 36,7; 38,1; 38,7; 37,5; 36,7; 38,7; 38,1; 36,7; 38,1; 39,5; 37,5.

1.3 Súčasťou štatistického prieskumu bolo získanie údajov o priemernej hodinovej mzde. Zistené údaje (v €) sú : 8,1; 10,5; 22; 15,7; 16,2; 10,5; 12,6; 16,3; 10,5; 18; 18,5; 16,4; 25,7; 22,3;

20,1; 18,3; 17,2; 18,9; 21,5; 23,7; 9,5; 11,4; 19,6; 16,6.

1.4 V rovnakých časových intervaloch sa meralo elektrické napätie (V) a boli namerané tieto hodnoty:
227; 220; 226; 223; 212; 226; 222; 215; 221; 209; 222; 219; 220; 227; 232; 212; 215; 219; 219; 220; 221; 220; 217; 220; 221; 230; 217; 219; 223; 220; 218; 219; 222; 221; 217; 216; 221; 225; 224; 218; 216; 222; 218; 214; 213; 225; 230; 218; 231; 232; 226; 217; 230; 231; 215; 219; 221; 227; 228; 231.

1.5 Medzi 15 respondentmi ankety sa okrem odpovedí na otázky sledovalo aj pohlavie a vzdelanie. Dáta sú v tabuľke. Zostavte tabuľku početností pre alternatívny znak pohlavie (žena, muž) a množný znak vzdelanie (základne - ZŠ, stredoškolské - SŠ, vysokoškolské - VŠ). Roztrieďte tieto údaje do tabuľky podľa obidvoch znakov.
a) Koľko žien sa zúčastnilo ankety?
b) Koľko percent respondentov ankety bolo vysokoškolsky vzdelaných?
c) Koľko mužov so stredoškolským vzdelaním sa zúčastnilo ankety?

Respondent	Pohlavie	Vzdelanie
1	muž	SŠ
2	muž	SŠ
3	muž	SŠ
4	muž	VŠ
5	žena	VŠ
6	muž	SŠ
7	muž	SŠ
8	muž	SŠ
9	žena	VŠ
10	žena	VŠ
11	žena	SŠ
12	muž	SŠ
13	muž	VŠ
14	muž	SŠ
15	žena	VŠ

1.2 Číselné charakteristiky

Rozdelenia početností, prezentované formou tabuľky alebo grafom poskytujú len základnú predstavu o štatistickom súbore. Samotné roztriedenie je podkladom pre popis a vzájomné porovnanie viacerých súborov, ale nemôže ho nahradiť. Pre jednoznačné vzájomné porovnávanie dvoch alebo niekoľkých štatistických súborov potrebujeme také veličiny, ktoré číselne určujú základné vlastnosti rozdelenia početností. Takéto veličiny nazývame **číselnými charakteristikami**.

Patria sem charakteristiky polohy, charakteristiky variability, miery šikmosti a špicatosti.

1.2.1 Charakteristiky polohy

Charakteristiky polohy vyjadrujú určitú úroveň (polohu) znaku, okolo ktorej sú ostatné hodnoty viac či menej koncentrované. Poloha sa meria pomocou rôznych druhov stredných hodnôt.

Stredné hodnoty rozdeľujeme na:

Priemery, sú to hodnoty, ktoré závisia od veľkostí hodnôt všetkých jednotiek štatistického súboru. Patria sem aritmetický, harmonický a geometricky priemer. Každý z nich môže byť jednoduchý (z netriedených údajov) alebo vážený (z údajov usporiadaných do radu rozdelenia alebo z intervalového rozdelenia).

Ostatné stredné hodnoty, ktoré sú založené len na niektorých vybraných hodnotách súboru, napríklad medián a modus.

Kvantily, ktoré rozdeľujeme na kvartily, decily a percentily.

Poznámka: vzorce pre výpočet číselných charakteristík sú označené nasledovne; (1) pre súbor z netriedených údajov, (2) pre súbor údajov usporiadaných do variačného radu, (3) pre súbor údajov triedených do intervalov.

Priemery sú základnými charakteristikami polohy a platia pre ne nasledujúce vzťahy:

Aritmetický priemer

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (1), \qquad \overline{x} = \frac{1}{n}\sum_{i=1}^{l} x_i \cdot n_i \quad (2), \qquad \overline{x} = \frac{1}{n}\sum_{j=1}^{k} z_j \cdot n_j \, (3).$$

Harmonický priemer

$$x_H = \frac{n}{\displaystyle\sum_{i=1}^{n}\frac{1}{x_i}} \quad (1), \qquad x_H = \frac{n}{\displaystyle\sum_{i=1}^{l}\frac{n_i}{x_i}} \quad (2), \qquad x_H = \frac{n}{\displaystyle\sum_{j=1}^{k}\frac{n_j}{z_j}} \quad (3).$$

Geometrický priemer

$$x_G = \sqrt[n]{\prod_{i=1}^{n} x_i} \quad (1), \qquad x_G = \sqrt[n]{\prod_{i=1}^{l} {x_i}^{n_i}} \quad (2), \qquad x_G = \sqrt[n]{\prod_{j=1}^{k} {z_j}^{n_j}} \quad (3).$$

Platí: $x_H \leq x_G \leq \overline{x}$. Rovnosť nastáva len v prípade, keď sú všetky hodnoty znaku v súbore rovnaké.

Ostatné stredné hodnoty

Modus chápeme ako najpočetnejšiu hodnotu štatistického znaku, označujeme $Mo = \hat{x}$. Modálnu hodnotu z radu rozdelenia početností určíme ako hodnotu znaku s najväčšou početnosťou. Ak táto hodnota je jediná, hovoríme, že rozdelenie početností je unimodálne, v opačnom prípade je polymodálne. V prípade intervalového rozdelenia početností modus vypočítame podľa vzťahu:

$$Mo = a_o + h\frac{d_1}{d_1 + d_2},$$

a_o je začiatok modálneho intervalu,
h je dĺžka intervalu,
d_1 je rozdiel medzi absolútnou početnosťou modálneho a predchádzajúceho intervalu,
d_2 je rozdiel medzi absolútnou početnosťou modálneho a nasledujúceho intervalu.

Medián chápeme ako prostrednú hodnotu štatistického súboru. Medián je číslo, ktoré leží uprostred variačného radu, označujeme $Me = \widetilde{x}$. Ak n je rozsah štatistického súboru, potom platí:

pre n párne: $\widetilde{x} = \dfrac{x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}}{2}$ a pre n nepárne: $\widetilde{x} = x_{(\frac{n+1}{2})}$.
V prípade intervalového rozdelenia početností medián vypočítame podľa vzťahu:

$$Me = a_e + h\frac{\frac{n+1}{2} - N_{j-1}}{n_j},$$

a_o je začiatok mediánového intervalu,
h je dĺžka intervalu,
N_{j-1} je absolútna kumulatívna početnosť intervalu pred mediánovým,
n_j je absolútna početnosť mediánového intervalu.

Vzájomná poloha modusu, mediánu a aritmetického priemeru toho istého štatistického súboru charakterizuje tvar určitého rozdelenia početností. Ak $\hat{x} = \widetilde{x} = \overline{x}$, ide o symetrické rozdelenie početnosti. Pre prípad nesymetrického rozdelenia rozlišujme: kladné zošikmenie, ak $\widetilde{x} < \hat{x} < \overline{x}$ a záporne zošikmenie, ak $\overline{x} < \widetilde{x} < \hat{x}$.

Kvantily
Kvantily sú také reálne číselné hodnoty, ktoré rozdeľujú rad vzostupne usporiadaných hodnôt štatistického znaku $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n)}$ na r rovnako početných častí. Najpoužívanejšie kvantily sú: medián $(r = 2)$, kvartily $(r = 4)$, decily $(r = 10)$ a percentily $(r = 100)$.

1.2.2 Charakteristiky variability

Uvažujme tri trojice čísel 9, 10, 11, 5, 10, 15, 1, 10, 19. Po jednoduchom výpočte zistíme, že všetky trojice majú aritmetický priemer 10, mediány sú tiež rovnaké a rovné 10 a predsa nie sú tieto súbory identické. Odlišnosť hodnôt štatistického znaku nazývame variabilitou. Čím väčšia je variabilita, tým menej je reprezentatívna charakteristika polohy. Variabilitu, rozptyl hodnôt znaku v štatistickom súbore vyjadrujú **charakteristiky variability**.

Z charakteristík variability sa najčastejšie používajú:
Priemerná odchýlka štatistického súboru udáva, o koľko sa zistené hodnoty v štatistickom súbore v priemere odchyľujú od ich aritmetického priemeru. Pretože platí: $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$, nemôžeme použiť súčet odchýliek ako charakteristiku variability, preto zavádzame priemernú odchýlku, kde je nutné použiť absolútne hodnoty odchýliek.

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{x}| \quad (1), \quad \overline{d} = \frac{1}{n}\sum_{i=1}^{l}|x_i - \overline{x}| \cdot n_i \quad (2), \quad \overline{d} = \frac{1}{n}\sum_{j=1}^{k}|z_j - \overline{x}| \cdot n_j \quad (3).$$

Rozptyl je najpoužívanejšou charakteristikou variability. Rozptyl je aritmetický priemer štvorcov odchýliek hodnôt štatistického znaku od ich aritmetického priemeru. Rozptyl označujeme s^2 a platí:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \quad (1), \quad s^2 = \frac{1}{n}\sum_{i=1}^{l}(x_i - \overline{x})^2 \cdot n_i \quad (2), \quad s^2 = \frac{1}{n}\sum_{j=1}^{k}(z_j - \overline{x})^2 \cdot n_j \quad (3).$$

Smerodajná odchýlka je kladná odmocnina z rozptylu, označujeme ju s, $s = \sqrt{s^2}$. Čím väčšia je variabilita hodnôt x_i, tým väčšie sú odchýlky $|x_i - \overline{x}|$ a tým väčšia je hodnota s^2 a aj s.

1.2.3 Miery šikmosti a špicatosti

K ďalším číselným charakteristikám štatistických súborov patria momenty, pomocou ktorých zavedieme momentové charakteristiky: **koeficient šikmosti**, **koeficient špicatosti**.

Všeobecný moment r-tého radu daného štatistického súboru označujeme v_r a definujeme nasledovne:

$$v_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r \quad (1), \quad v_r = \frac{1}{n}\sum_{i=1}^{l} x_i^r \cdot n_i \quad (2), \quad v_r = \frac{1}{n}\sum_{j=1}^{k} z_j^r \cdot n_j \quad (3),$$

kde r je prirodzené číslo.
Centrálny moment r-tého radu daného štatistického súboru označujeme μ_r a definujeme nasledovne:

$$\mu_r = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^r \quad (1), \quad \mu_r = \frac{1}{n}\sum_{i=1}^{l} (x_i - \overline{x})^r \cdot n_i \quad (2), \quad \mu_r = \frac{1}{n}\sum_{j=1}^{k} (z_j - \overline{x})^r \cdot n_j \quad (3),$$

kde r je prirodzené číslo a $\overline{x}$ je aritmetický priemer daného štatistického súboru.
Koeficient šikmosti (asymetrie) je momentová miera šikmosti. Koeficient šikmosti γ_3 definujeme nasledovne:

$$\gamma_3 = \frac{\mu_3}{s^3},$$

kde μ_3 je tretí centrálny moment, ktorý možno určiť podľa vzťahu

$$\mu_3 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^3 \quad (1), \quad \mu_3 = \frac{1}{n}\sum_{i=1}^{l} (x_i - \overline{x})^3 \cdot n_i \quad (2), \quad \mu_3 = \frac{1}{n}\sum_{j=1}^{k} (z_j - \overline{x})^3 \cdot n_j \quad (3).$$

Ak je $\gamma_3 < 0$, potom väčšina hodnôt štatistického znaku daného štatistického súboru leží naľavo od aritmetického priemeru (ľavostranná asymetria).
Ak $\gamma_3 = 0$, potom hodnoty štatistického znaku sú symetrický rozložené okolo aritmetického priemeru.
Ak $\gamma_3 > 0$, potom väčšina hodnôt štatistického znaku leží napravo od aritmetického priemeru (pravostranná asymetria).
Smer asymetrie vyjadruje znamienko koeficienta šikmosti a silu asymetrie vyjadruje jeho hodnota.
Koeficient špicatosti (excesu) meria stupeň koncentrácie hodnôt štatistického znaku okolo strednej hodnoty. Koeficient špicatosti γ_4 definujeme nasledovne:

$$\gamma_4 = \frac{\mu_4}{s^4} - 3,$$

kde μ_4 je štvrtý centrálny moment, ktorý možno určiť podľa vzťahu

$$\mu_4 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^4 \quad (1), \quad \mu_4 = \frac{1}{n}\sum_{i=1}^{l} (x_i - \overline{x})^4 \cdot n_i \quad (2), \quad \mu_4 = \frac{1}{n}\sum_{j=1}^{k} (z_j - \overline{x})^4 \cdot n_j \quad (3).$$

Ak $\gamma_4 < 0$, potom polygón relatívnych početností je plochejší ako krivka normálneho rozdelenia (pozri 3. kapitolu).
Ak $\gamma_4 = 0$, potom polygón relatívnych početností má rovnakú špicatosť ako krivka normálneho rozdelenia.
Akk $\gamma_4 > 0$, potom polygón relatívnych početností je špicatejší ako krivka normálneho rozdelenia.

Príklad 3. Pre kvantitatívne znaky (počet členov domácnosti a ročná spotreba studenej vody - súbor 1, 3) z predchádzajúceho príkladu vypočítajme číselne charakteristiky.

Riešenie.

Počet členov domácnosti - súbor 1
Charakteristiky polohy:

- Aritmetický priemer

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{5} x_i \cdot n_i,$$

$\overline{x} = \frac{1}{20}(2 \cdot 4 + 3 \cdot 4 + 4 \cdot 4 + 5 \cdot 5 + 6 \cdot 3) = \frac{79}{20} = 3{,}95.$

- Modus je najpočetnejšia hodnota, t.j

$$Mo = \widehat{x} = 5.$$

- Medián je prostredná hodnota a pre n -párne ($n = 20$) použijeme vzťah

$$Me = \widetilde{x} = \frac{x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}}{2}) = \frac{x_{(10)} + x_{(11)}}{2} = 4.$$

Charakteristiky variability:

- Variačné rozpätie

$$R_v = x_{max} - x_{min} = 6 - 2 = 4.$$

- Rozptyl

$$s^2 = \frac{1}{n}\sum_{i=1}^{5} (x_i - \overline{x})^2 \cdot n_i,$$

$$s^2 = \frac{1}{20}\left[(-1{,}95)^2 \cdot 4 + (-0{,}95)^2 \cdot 4 + 0{,}05^2 \cdot 4 + 1{,}05^2 \cdot 5 + 2{,}05^2 \cdot 3\right] = 1{,}8475.$$

- Smerodajná odchýlka

$$s = \sqrt{s^2} = 1{,}3592.$$

- Priemerná odchýlka

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{5} |x_i - \overline{x}| \cdot n_i,$$

$$\overline{d} = \frac{1}{20}(1{,}95 \cdot 4 + 0{,}95 \cdot 4 + 0{,}05 \cdot 4 + 1{,}05 \cdot 5 + 2{,}05 \cdot 3) = \frac{23{,}2}{20} = 1{,}16.$$

Koeficient šikmosti:

$$\gamma_3 = \frac{\mu_3}{s^3} = \frac{1}{n \cdot s^3}\sum_{i=1}^{5} (x_i - \overline{x})^3 \cdot n_i,$$

$$\gamma_3 = \frac{\left[(-1{,}95)^3 \cdot 4 + (-0{,}95)^3 \cdot 4 + 0{,}05^3 \cdot 4 + 1{,}05^3 \cdot 5 + 2{,}05^3 \cdot 3\right]}{20 \cdot 1{,}3592^3} = -0{,}029.$$

Koeficient špicatosti:

$$\gamma_4 = \frac{\mu_4}{s^4} - 3 = \frac{1}{n \cdot s^4} \sum_{i=1}^{5} (x_i - \overline{x})^4 \cdot n_i - 3,$$

$$\gamma_4 = \frac{\left[(-1{,}95)^4 \cdot 4 + (-0{,}95)^4 \cdot 4 + 0{,}05^4 \cdot 4 + 1{,}05^4 \cdot 5 + 2{,}05^4 \cdot 3\right]}{20 \cdot 1{,}3592^4} - 3 = -1{,}2397.$$

Ročná spotreba studenej vody - súbor 3
Charakteristiky polohy:

- Aritmetický priemer

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{5} z_j \cdot n_j,$$

$$\overline{x} = \frac{1}{20}(62 \cdot 4 + 92 \cdot 7 + 122 \cdot 2 + 152 \cdot 4 + 182 \cdot 3) = \frac{2290}{20} = 114{,}5.$$

- Modus

$$Mo = \widehat{x} = a_o + h\frac{d_1}{d_1 + d_2} = 77 + 30\frac{7-4}{(7-4)+(7-2)} = 88{,}25,$$

kde a_o je začiatok modálneho intervalu $\langle 77,\, 107\rangle$, d_1 je rozdiel absolútnej početnosti modálneho a predchádzajúceho intervalu a d_2 je rozdiel absolútnej početnosti modálneho a následujúceho intervalu.

- Medián

$$Me = \widetilde{x} = a_e + h\frac{\frac{n+1}{2} - N_{j-1}}{n_j} = 77 + 30\frac{\frac{21}{2} - 4}{7} = 104{,}8571,$$

kde a_e je začiatok mediánového intervalu $\langle 77,\, 107\rangle$.

Charakteristiky variability:

- Variačné rozpätie

$$R_v = x_{max} - x_{min} = 196 - 48 = 148.$$

- Rozptyl

$$s^2 = \frac{1}{n} \sum_{j=1}^{5} (z_j - \overline{x})^2 \cdot n_i,$$

$$s^2 = \frac{1}{20}\left[((-52{,}5)^2 \cdot 4 + (-22{,}5)^2 \cdot 7 + 7{,}5^2 \cdot 2 + 37{,}5^2 \cdot 4 + 67{,}5^2 \cdot 3\right] = 1698{,}75.$$

- Smerodajná odchýlka

$$s = \sqrt{s^2} = 41{,}2159.$$

- Priemerná odchýlka

$$\overline{d} = \frac{1}{n} \sum_{j=1}^{5} |z_j - \overline{x}| \cdot n_i,$$

$$\overline{d} = \frac{1}{20}(52{,}5 \cdot 4 + 22{,}5 \cdot 7 + 7{,}5 \cdot 2 + 37{,}5 \cdot 4 + 67{,}5 \cdot 3) = \frac{735}{20} = 36{,}75.$$

Koeficient šikmosti:

$$\gamma_3 = \frac{\mu_3}{s^3} = \frac{1}{n \cdot s^3} \sum_{j=1}^{5} (z_j - \overline{x})^3 \cdot n_i,$$

$$\gamma_3 = \frac{\left[(-52{,}5)^3 \cdot 4 + (-22{,}5)^3 \cdot 7 + 7{,}5^3 \cdot 2 + 37{,}5^3 \cdot 4 + 67{,}5^3 \cdot 3\right]}{20 \cdot 41{,}2159^3} = 0{,}3398.$$

Koeficient špicatosti:

$$\gamma_4 = \frac{\mu_4}{s^4} - 3 = \frac{1}{n \cdot s^4} \sum_{j=1}^{5} (z_j - \overline{x})^4 \cdot n_i - 3,$$

$$\gamma_4 = \frac{\left[((-52{,}5)^4 \cdot 4 + (-22{,}5)^4 \cdot 7 + 7{,}5^4 \cdot 2 + 37{,}5^4 \cdot 4 + 67{,}5^4 \cdot 3)\right]}{20 \cdot 41{,}2159^4} - 3 = -1{,}2262.$$

Úlohy

1.6 Vypočítajte aritmetický priemer, modus a medián známok z písomnej práce z matematiky: 3; 2; 1; 1; 2; 2; 3; 4; 4; 5.

1.7 Vypočítajte aritmetický priemer, modus a medián bodového hodnotenia 15 študentov: 158; 161; 168; 163; 174; 179; 176; 172; 178; 168; 179; 174; 173; 179; 177.

1.8 Vypočítajte a určte vzťah medzi aritmetickým priemerom, modusom a mediánom nasledujúcich hodnôt: 17; 21; 22; 20; 25; 15; 12; 23; 10; 8; 12; 23; 12; 5; 8; 10; 15; 23; 20; 11; 12; 7; 24; 5.

1.9 Vypočítajte a interpretujte modus krvných skupín: A; 0; 0; B; B; AB; A; A; 0; 0; 0; AB; B; 0; B; A; 0; AB; 0; 0; B; 0; A.

1.10 Vypočítajte a interpretujte medián z výsledkov skúšky:
a) C; E; B; D; A; A; B; FX; C; C; D (známky radíme v poradí A, B, C, D, E, FX).
b) 61; 49; 35; 74; 53; 82 (v %).

1.11 Vypočítajte priemer, rozptyl a smerodajnú odchýlku z údajov o hmotnosti jedenástich balení zemiakov (v kg): 43; 68; 65; 59; 48; 49; 52; 48; 57; 59; 48.

1.12 V ankete sa 50 vodičov áut pýtali, koľkokrát za posledný rok platili pokutu za prekročenie povolenej rýchlosti. Ich odpovede sú: 0; 2; 5; 5; 4; 2; 2; 0; 1; 5; 3; 2; 3; 1; 1; 2; 3; 2; 1; 6; 1; 2; 3; 4; 2; 3; 0; 4; 5; 0; 2; 1; 2; 2; 2; 0; 3; 4; 4; 4; 3; 3; 4; 6; 1; 2; 3; 4; 2; 1. Pre tento súbor vypočítajte modus, aritmetický priemer a smerodajnú odchýlku.

1.13 V nemocnici bolo v určitom období hospitalizovaných 150 osôb na chirurgickom oddelení s priemernou dĺžkou hospitalizácie 19 dní, 100 osôb na internom oddelení s priemernou dĺžkou hospitalizácie 7 dní a na detskom oddelení 90 detí s priemernou dĺžkou hospitalizácie 12 dní. Vypočítajte priemernú dĺžku hospitalizácie pre všetkých pacientov.

V nasledujúcich úlohách vypočítajte charakteristiky polohy, variability, koeficient šikmosti a špicatosti.

1.14 Vykonali sme 36 analýz na overenie koncentrácie chemickej látky v roztoku. Výsledky (v %) sú: 17; 12; 15; 16; 18; 17; 18; 13; 12; 15; 16; 11; 16; 17; 18; 12; 14; 20; 21; 17; 11; 14; 15; 20; 14; 13; 16; 15; 14; 15; 16; 15; 19; 15; 19; 14.

1.15 Čas čakania na trolejbus na zastávke MHD by nemal prekročiť 10 min. Náhodne sa urobilo 30 kontrol, pričom sa zaznamenával čas čakania (v min.): 3; 5; 2; 6; 4; 5; 7; 2; 6; 4; 5; 7; 6; 4; 6; 3; 5; 3; 6; 4; 5; 5; 3; 7; 4; 5; 5; 6; 4; 4.

1.16 V priebehu mesiaca odpracovalo 33 zamestnancov firmy tento počet pracovných dní: 22; 21; 23; 18; 20; 20; 23; 22; 22; 23; 17; 21; 23; 23; 23; 21; 23; 21; 23; 22; 24; 19; 23; 19; 21; 16; 24; 22; 20; 22; 22; 23; 21.

1.17 Pri kontrole kvality výrobku boli namerané odchýlky jedného rozmeru výrobku od stanovenej hodnoty v stotinách milimetra: 56; 60; 60; 61; 55; 59; 61; 60; 59; 61; 59; 61; 60; 62; 57; 61; 57; 59; 54; 62; 60; 58; 60; 60; 61; 59; 60; 59; 61; 57; 58.

1.18 30 zákazníkov si objednalo tovar cez internet. Od objednávky po doručenie tovaru uplynulo: 16; 18; 19; 16; 12; 14; 15; 18; 17; 14; 12; 13; 19; 18; 16; 15; 15; 17; 12; 16; 15; 14; 13; 15; 17; 18; 16; 15; 13; 14 dní.

1.19 Nasledujúce hodnoty udávajú priemer (v mm) 20 brúsnych kotúčov: 28,6; 28,5; 28,5; 28,2; 28,6; 28,3; 28,3; 28,8; 28,4; 28,5; 28,4; 28,4; 28,3; 28,4; 28,7; 28,4; 28,6; 28,3; 28,5; 28,6.

1.20 V rámci vyhodnotenia kvality obchodného reťazca bol počas určitého časového obdobia zaznamenávaný čas čakania pri pokladni. Zistené údaje (v min.) sú: 5; 6; 2; 3; 7; 5; 3; 5; 5; 3; 2; 3; 6; 5; 7; 7; 2; 7; 3; 4; 5; 5; 2; 5; 6; 3; 3; 7.

1.21 Výsledky meraní obsahu uhlíka v uhlí (v %) sú: 86; 87; 86; 81; 77; 85; 87; 86; 85; 87; 82; 84; 84; 83; 79; 83; 80; 85; 79; 78; 83; 77; 86; 81; 78; 82; 86; 83; 84; 84; 80; 86; 87.

1.22 Na začiatku semestra boli zistené údaje o veku študentov nastupujúcich do 1. ročníka externého štúdia na SjF. Vek poslucháčov je: 33; 30; 34; 26; 24; 27; 31; 32; 36; 34; 37; 26; 23; 28; 36; 30; 25; 29; 40; 33; 31; 35; 36; 38; 32; 36; 33; 39; 32; 33.

1.23 Meral sa percentuálny obsah balastných látok v tavbe. Získané údaje(v %) sú:
0,091; 0,097; 0,101; 0,099; 0,095; 0,099; 0,089; 0,097; 0,099; 0,093; 0,099; 0,105; 0,091; 0,103; 0,097; 0,099; 0,095; 0,101; 0,099; 0,101; 0,097; 0,093; 0,097; 0,101; 0,095; 0,099; 0,097; 0,091; 0,097; 0,101; 0,097; 0,099; 0,095; 0,101; 0,091; 0,099; 0,099; 0,089; 0,099; 0,097; 0,105; 0,097; 0,099; 0,103; 0,101; 0,093; 0,099; 0,095; 0,095; 0,101; 0,099; 0,097; 0,095; 0,101; 0,097; 0,099; 0,103; 0,097; 0,093.

1.24 Výsledky meraní obsahu cínu (v %) v 100 vzorkách rudy sú uvedené v tabuľke.

z_j	30	35	40	45	50	55	60	65	70	75
n_j	4	6	8	15	25	20	8	7	5	2

1.25 U 50 náhodne vybraných študentov sa zisťoval čas, ktorý doma venujú štúdiu. Hodnoty sú uvedené v tabuľke.

Čas štúdia (v min.)	15	45	75	105	135
Počet študentov	9	10	15	12	4

1.26 Pri prieskume trhu boli zistené výšky mesačných splátok medzi respondentmi (v €). Údaje sú uvedené v tabuľke.

z_j	145	155	165	175	185	195
n_j	3	8	10	15	8	6

1.27 Meral sa obsah Cu v istej zliatine (v %). Namerané údaje sú uvedené v tabuľke.

z_j	3	3,30	3,60	3,90	4,20	4,50	4,80
n_j	4	7	12	18	21	13	5

1.28 Pri prieskume cien istej komodity sa skúmalo predraženie jej ceny v percentách na 1 kg váhy. Údaje sú: 2,1; 1,3; 5,3; 4,9; 5,2; 9,4; 7,9; 8,7; 9,8; 3,2; 4,8; 6,0; 3,4; 7,8; 8,3; 6,7; 8,5; 10,0; 4,8; 3,6; 7,1; 4,1; 8,4; 11,1; 8,8; 8,6; 4,5; 8,1; 7,9; 5,3; 5,1; 9,4; 8,5; 7,3; 6,2; 5,1.

1.29 Počas testovania novej automatickej linky sa zaznamenávalo trvanie operácie (v s):
10,9; 9,4; 9,6; 9,7; 10,4; 9,5; 7,9; 10,2; 8,3; 7,3; 8,3; 11,4; 10,4; 9,3; 7,1; 9,6; 9,9; 8,2; 8,8; 11,9; 9,2; 10,4; 11,6; 9,8; 10,1; 11,8; 10,5; 8,1; 8,0; 12,1; 9,7; 9,8; 8,5; 10,2; 12,0; 7,0; 10,7; 9,0; 10,6; 10,0; 11,7; 9,6; 8,8; 8,6; 10,3; 8,9; 11,1; 10,0; 11,0; 9,0; 10,9; 9,4; 8,2; 8,5; 10,7; 8,6; 9,1; 9,5; 8,9; 9,9; 10,5; 9,8; 9,5; 7,5; 9,7; 9,3; 7,2; 11,2; 11,3; 7,0; 8,5; 10,0; 10,6; 7,8; 10,0; 10,2; 8,0; 11,2; 8,4; 7,8; 11,5; 9,2; 9,8; 8,3; 12,3; 9,0; 10,4; 9,3; 12,4; 8,7; 11,4; 10,2; 7,7; 11,0; 10,1; 8,4; 9,4; 10,5; 9,1; 7,4.

1.30 Zisťovali sme hmotnosť 60 vzoriek (v g). Výsledky experimentu sú:
4,20; 4,54; 4,37; 4,47; 4,29; 4,52; 4,38; 4,21; 4,53; 4,60; 4,56; 4,70; 4,59; 4,73; 4,61; 4,31; 4,42; 4,73; 4,37; 4,64; 4,46; 4,66; 4,50; 4,71; 4,32; 4,69; 4,39; 4,59; 4,34; 4,70; 4,70; 4,28; 4,41; 4,63; 4,80; 4,61; 4,44; 4,86; 4,76; 4,79; 4,56; 4,73; 4,31; 4,72; 4,68; 4,65; 4,33; 4,70; 4,33; 4,53; 4,21; 4,27; 4,42; 4,70; 4,42; 4,36; 4,65; 4,84; 4,78; 4,80.

1.31 Životnosť určitej súčiastky (v hod.) bola skúšaná na 70 vzorkách. Výsledky sú:
186; 153; 200; 166; 171; 158; 164; 174; 180; 189; 187; 163; 158; 149; 149; 204; 185; 178; 183; 176; 182; 184; 166; 151; 167; 197; 194; 210; 131; 159; 213; 159; 175; 186; 186; 215; 153; 128; 193; 176; 179; 216; 172; 198; 177; 203; 170; 142; 155; 181; 184; 163; 179; 171; 186; 169; 164; 185; 200; 145; 187; 183; 172; 160; 223; 183; 192; 138; 156; 115.

1.32 Vyšetrovalo sa 100 náhodne vybraných priadzí na pevnosť v ťahu (MPa). Výsledky merania sú uvedené v tabuľke.

I_j	$\langle 0{,}5; 0{,}7)$	$\langle 0{,}7; 0{,}9)$	$\langle 0{,}9; 1{,}1)$	$\langle 1, 1; 1{,}3)$	$\langle 1{,}3; 1{,}5)$	$\langle 1{,}5; 1{,}7)$	$\langle 1{,}7; 1{,}9)$
n_j	3	11	19	31	17	13	6

1.33 50 amatérskych pretekárov sa zúčastnilo jazdeckých pretekov. Počet prekonaných prekážok je uvedený v tabuľke.

I_j	$\langle 0;6)$	$\langle 6;12)$	$\langle 12;18)$	$\langle 18;24)$	$\langle 24;30)$	$\langle 30;36)$	$\langle 36;42\rangle$
n_j	5	7	12	10	6	3	7

1.34 Výsledky skúmania pevnosti na tlak (v MPa) pre 200 valcových brikiet sú uvedené v tabuľke.

I_j	$\langle 19;20)$	$\langle 20;21)$	$\langle 21;22)$	$\langle 22;23)$	$\langle 23;24)$	$\langle 24;25\rangle$
n_j	9	27	55	65	30	14