

## **Тақырыбы: Таңдама тәсіл**

### **Бас жиынтық және таңдама**

Біртекті объектілердің жиынтығын оларды сипаттайтын сапалық және сандық белгісі бойынша зерттеу керек болсын. Мысалы, бір партия құралдар бар болса, олардың сапалық белгісі ретінде стандарттылығы сандық белгісі ретінде-құралдың өлшемі қарастырылады.

Кей кезде жиынның әр объектісін қажет белгісіне қатысты зерттеп, жаппай зерттеу жүргізеді. Алайда, практикада мұндай жаппай зерттеу салыстырмалы түрде сирек қолданылады.

Таңдама жиынтық немесе жай таңдама деп кездейсоқ таңдап алынған объектілер жиынын айтады.

Бас жиынтық деп ішінен таңдама жүргізілетін объектілер жиынын айтады. Жиынтық көлемі деп осы жиынтықтың объектілер санын айтады.

Практикада таңдаманың түрлі әдістері қолданылады. Бұл әдістерді 2 түрге топтауға болады.

I. Бас жиынтықты бөлшектеп бөлуді қажет етпейтін таңдама, оған:

а) жай кездейсоқ қайталанымсыз таңдама;

б) жай кездейсоқ қайталанбалы таңдама;

II. Бас жиынтық бөлшектеп бөлінетін таңдама, оған:

а) типтік таңдама;

ә) механикалық таңдама;

б) сериялық таңдама

жатады.

Жай кездейсоқ таңдама деп барлық бас жиынтықтан объектілерді бір-бірден алатын таңдаманы атайды. Егер алынған карточкаларды бумаға қайтармаса, онда таңдама жай кездейсоқ қайталанымсыз болады.

Типтік таңдама деп, объектілер бас жиынтықтың барлығынан емес, оның әрбір «типтік бөлігінен алынатын таңдаманы атайды.

Механикалық таңдама деп бас жиынтық таңдамаға қанша объект қажет болса, сонша топқа бөлінетін таңдаманы атайды, әрбір топтан бір объект алынады.

Сериялық таңдама деп бас жиынтықтан объектілерді бір-бірден емес, жаппай зерттеуге ұшырайтын объектілер «сериялармен» таңдап алатын таңдаманы атайды.

### **Таңдаманың статистикалық таралуы.**

Алынған таңдамалық зерттеулерді жүйелендіруде таралудың статистикалық дискретті және интервалды қатарлар қолданылады.

**Вариациялық қатар** — сәйкес жиіліктерімен бірге ранжирленген ретпен орналасқан белгінің сандық мәндері.

Вариациялық қатардың негізгі белгіленулері:

$x_i$  - варианта, зерттеліп отырған белгінің сандық өрнектелуі;

$n_i$  - вариантаның жиілігі (“салмақ”), оның вариациялық қатарда қайталану саны;

$n$  — бақылаудың жалпы саны, таңдама көлемі (яғни, барлық жиіліктердің қосындысы,  $n = \sum n_i$ );

$x_{max}, x_{min}$  - вариациялық қатарды шектейтін шеткі варианттар, (қатардың лимиттері);

$R$  - қатардың амплитудасы (яғни, ең жоғарғы және ең төменгі варианттардың айырымы  $R = x_{max} - x_{min}$ ).

Вариациялық қатардың қолданылуы:

Вариациялық қатар орта шаманы ( $\bar{x}$ ) және әртүрлілік критерийлерін ( $\sigma, C$ ) анықтау үшін қолданылады.

Орта шама – зерттеліп отырған белгі өлшемінің жалпылама сипаттамасы. Ол бір санмен сапалы біртекті жиынтықты сандық сипаттауға мүмкіндік береді.

Бас жиынтықтан таңдама алынсын, және  $x_1-n_1$  рет,  $x_2-n_2$  рет, ...,  $x_k-n_k$  рет қайталанады, ал  $\sum n_i = n$  -таңдама көлемі.  $x_1$  мәндерін варианттар деп, ал өсу ретімен жазылған варианттар тізбегін вариациялық қатар деп атайды. Қарастырылатын мәндер санын жиіліктер, ал олардың таңдама көлеміне қатынасын салыстырмалы жиіліктер  $\frac{n_i}{n} = W_i$  - деп айтады.

Таңдаманың статистикалық таралуы деп варианттар мен оларға сәйкес жиіліктер немесе салыстырмалы жиіліктердің тізімі аталады.

1. Статистикалық таралуды интервалдар тізбегі және оларға сәйкес жиіліктер (интервалға сәйкес жиілік ретінде осы интервалға түскен жиіліктер қосындысын қабылдайды) тізбегі түрінде беруге болады. Мода (*Mo*) – кездейсоқ шаманың ең жиі кездесетін мәні

2. Медиана (*Me*) – таңдаманы қақ ортасынан бөлетін кездейсоқ шаманың мәні. Барлық мәндерді ранжрленген қатарға орналастыру керек.

Таңдаманың 50% жоғары не төмен орналасқан мәнді көрсетеді.

3. Минимал мәні  $x_{\min}$

4. Максимал мәні  $x_{\max}$

5. Қадам  $R = x_{\max} - x_{\min}$

6. Орташа қатесі (стандартты) – таңдама көрсеткіштің (статистика) оның генеральды

параметрінен ауытқу шамасы:  $S = \frac{\sigma}{\sqrt{n}}$

### Полигон және гистограмма.

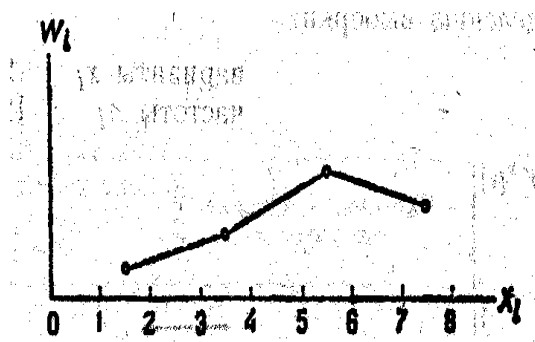
Көрнекілік үшін статистикалық таралудың түрлі графиктер салынады, соның ішінде полигон мен гистограмма тұрғызылады.

1. X дискретті белгісінің таралуы.

Жиіліктер полигоны деп  $(x_1, n_1), (x_2, n_2), \dots, (x_k, n_k)$  нүктелерін қосатын сынық сызықты айтады жиіліктер полигонын тұрғызу үшін абциссалар осінде  $x_i$  варианттарын, ал ординаталр осінде оларға сәйкес  $n_i$  жиіліктерді орналастырады.  $(x_i, n_i)$  нүктелерін түзудің кесінділерімен қосып, жиіліктер полигонын салады.

Салыстырмалы жиіліктер полигоны деп кесінділері  $(x_1, W_1), (x_2, W_2), \dots, (x_k, W_k)$ , нүктелерін қосатын сынық сызықты айтады. Салыстырмалы жиіліктер полигонын тұрғызу үшін абциссалар осінде  $x_i$  варианттарын, ал ординаталар осінде оларға сәйкес  $W_i$  салыстырмалы жиіліктер полигонын тұрғызады.

1-суретте келесі салыстырмалы жиіліктер таралуының полигоны бейнеленген.



1-сурет

2. X үзіліссіз белгісінің таралуы.

Статистикалық үзіліссіз қатарды (немесе таралудың эмпирикалық функциясын) әдетте таңдамада бір-бірінен өзгеше варианттар көп болмаған жағдайда немесе осы себеппе зерттеуші үшін үзіліссіздік бар болған жағдайда қолданады. Егер бізді қызықтыратын бас жиынтықтың X белгісі үзіліссіз болса, онда варианттар интервалдарға топтастырылады.

Көлемі n таңдамадағы берілгендерді топтастырғанда барлық варианттарды қамтитын  $[a, b]$  кесіндісін k интервалдарға  $(k_{i-1}, k_i, i = 1, \dots, k : a = h_0 < h_1 < \dots < h_k = b)$  бөледі. (әдетте  $a \approx x_1, b \approx x_n = X_n$ ).

Осыдан соң, әрбір интервалдың жиілігін, яғни  $i$ -ші интервалға түсетін  $n_i$  бақылаулар санын анықтау керек. Айқындылық үшін  $n_i - [h_{i-1}, h_i)$  жарты интервалында жататын варианттар санына тең деп ұйғарамыз. Барлық жағдайда аралық  $h_i$  - ге тең болады.  $k$  – интервалдар саны.

Ескерту. Көбіне барлық  $i$  үшін  $h_i - h_{i-1} = h$ , яғни топтастыру тең  $h$  қадаммен алынады. Бұл жағдайда  $a, k, h_i$  табу үшін келесі ұсынысты жетекшілікке алуға болады.

1.  $R = X_{\max} - X_{\min}$  табамыз, мұндағы  $R$  – ең үлкен және ең кіші варианттардың айырымы.

2.  $h = \frac{R}{k}$   $k$ -топтар саны,  $h$ -қадам.

3.  $k \geq 1 + 3,321 \lg n$  (Стерджес формуласы)

4.  $a = X_{\min}, b = X_{\max}$

5.  $h_i = a + ih, i = 0, 1, \dots, k$

Алынған топтастыруды жиілік кестесі түрінде ұсыну қолайлы. Бұл кестені таралудың статистикалық интервалдық қатары деп атайды.

Топтастырудың интервалы	$[h_0, h_1)$	$[h_1, h_2)$	...	$[h_{k-2}, h_{k-1})$	$[h_{k-1}, h_k]$
Жиіліктер	$n_1$	$n_2$	...	$n_{k-1}$	$n_k$

$$\sum n_k = n.$$

Осы кестені  $n_i$  жиіліктерді салыстырмалы жиіліктермен алмастырып мынадай түрде жазуға болады:

Топтастыру интервалы	$[h_0, h_1)$	$[h_1, h_2)$	...	$[h_{k-2}, h_{k-1})$	$[h_{k-1}, h_k]$
Салыстырмалы жиіліктер	$w_1$	$w_2$	...	$w_{k-1}$	$w_k$

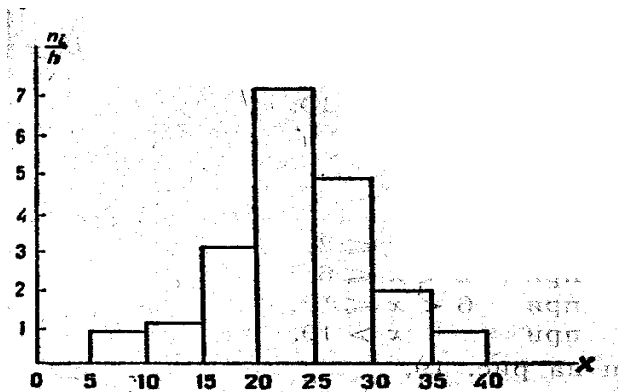
$$\sum w_i = 1.$$

Жиіліктердің графиктік түрі - жиіліктер гистограммасы деп аталатын арнайы график болып табылады.

Жиіліктер гистограммасы деп табандары  $h$ -қа, биіктіктері  $\frac{n_i}{h}$  (жиілік тығыздығы) қатынасына тең тіктөртбұрыштардан тұратын баспалдақты фигураны айтады.

Үзіліссіздік белгісі жағдайында гистограмма салған жөн, ол үшін белгінің барлық бақыланатын мәндер жататын интервалды ұзындығы  $h$ -қа тең бірнеше дербес (жеке) интервалдарға бөліп, әрбір дербес  $n_i$  интервал үшін  $i$ -ші интервалға түскен варианттар жиіліктерінің қосындысын табады.  $i$ -ші

дербес төртбұрыштың ауданы  $h \cdot \frac{n_i}{h} = n_i$  -  $i$ -ші интервалға түскен варианттар жиіліктерінің қосындысына тең, сондықтан жиіліктер гистограммасының ауданы барлық жиіліктердің қосындысына тең, яғни таңдаманың көлеміне тең.



Салыстырмалы жиіліктер гистограммасы деп табандары  $h$ -қа, биіктіктері  $\frac{w_i}{h}$  (салыстырмалы жиілік тығыздығы) қатынасына тең тіктөртбұрыштардан тұратын баспалдақты фигураны айтады.

Салыстырмалы жиіліктер гистограммасын тұрғызу үшін абциссалар осінде дербес интервалдар салып, олардың үстінен абцисса осіне параллель  $\frac{w_i}{h}$  қашықтықта кесінділер жүргіземіз.  $i$ -ші дербес төртбұрыштың ауданы  $h = \frac{w_i}{h} = w_i$   $i$  - ші интервалға түскен варианттардың салыстырмалы жиіліктеріне тең. Сонымен, салыстырмалы жиіліктер гистограммасының ауданы барлық салыстырмалы жиіліктер қосындысына, яғни 1-ге тең.

**1-ден 10 дейінгі сандардың ондық логарифмдері**

$n$	1	2	3	4	5	6	7	8	9	10
$\lg n \approx$	0	0,30	0,48	0,60	0,70	0,78	0,85	0,90	0,95	1

### Бас орта

$x_B$  бас орта деп бас жиынтық белгісінің орта арифметикалық мәнін айтады.

Егер  $N$  көлемді бас жиынтық белгісінің барлық  $x_1, \dots, x_N$  мәндері әртүрлі болса, онда

$$x_B = \frac{x_1 + x_2 + \dots + x_N}{N}$$

Егер  $x_1, x_2, \dots, x_k$  белгісінің мәндері сәйкесінше  $N_1, N_2, \dots, N_k$  жиіліктерге ие болса, мұнда  $N_1 + N_2 + \dots + N_k = N$

$$x_B = \frac{x_1 N_1 + x_2 N_2 + \dots + x_k N_k}{N};$$

### Таңдама орта

$X$  сандық белгісіне қатысты бас жиынтықты зерттеу үшін  $n$  көлемді таңдама алынсын.

$X_T$  таңдама орта деп таңдама жиынтық белгісінің орта арифметикалық мәнін айтады.

Егер  $n$  көлемді таңдаманың барлық  $x_1, x_2, \dots, x_n$  мәндері әр түрлі болса, онда

$$X_T = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Егер таңдаманың  $x_1, x_2, \dots, x_k$  мәндерінің сәйкесінше жиіліктері  $n_1, n_2, \dots, n_k$  болса, және  $n_1 + n_2 + \dots + n_k = n$  онда

$$X_T = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n}$$

немесе 
$$\bar{X}_T = \frac{\left( \sum_{i=1}^k n_i x_i \right)}{n}$$

Көлемі  $n$  - ге тең  $X$  сандық белгінің мәндер жиынтығын қарастырамыз.

Белгінің мәні	$x_1$	$x_2$	...	$x_k$
Жиілік	$n_1$	$n_2$	...	$n_k$

мұнда  $\sum_{i=1}^k n_i = n$ .

Жазуға қолайлы болу үшін  $\sum_{i=1}^k$  қосынды таңбасы  $\sum$  таңбасымен ауысады.

### Бас дисперсия

Бас жиынтықтың  $X$  сандық белгісі мәндерінің өз орта мәнінің маңайында шашырауын сипаттау үшін бас дисперсия сипаттамасы енгізіледі.

$D_A$  бас дисперсиясы деп бас жиынтық белгісі мәндерінің  $x_B$  орта мәнінен ауытқуының квадратының орта арифметикалық мәнін айтады.

Егер  $N$  көлемді бас жиынтық белгісінің барлық  $x_1, x_2, \dots, x_N$  мәндері әртүрлі болса, онда

$$D_B = \frac{\sum_{i=1}^N (x_i - \bar{x}_B)^2}{N}$$

Егер белгінің барлық  $x_1, x_2, \dots, x_k$  мәндерінің сәйкес жиіліктері  $N_1, N_2, \dots, N_k$  бар болса, және  $N_1 + N_2 + \dots + N_k = N$ , онда

$$D_B = \frac{\sum_{i=1}^k N_i (x_i - \bar{x}_B)^2}{N}$$

Бас жиынтықтың сандық белгісі мәндерінің өз орта мәнінің маңайында шашырауын сипаттау үшін дисперсиядан басқа орта квадраттық ауытқуды пайдаланады.

Бас орташа квадраттық ауытқу деп бас дисперсиядан алынған квадрат түбірді айтады:  $\sigma_B = \sqrt{D_B}$ .

Таңдама дисперсия  $D_T$  деп белгінің бақыланатын мәндерінің  $x_T$  орта мәнінен ауытқу квадраттарының орта арифметикалық мәнін айтады.

Егер  $n$  көлемді таңдаманың барлық  $x_1, x_2, \dots, x_n$  белгілерінің мәндері әр түрлі болса, онда

$$D_T = \frac{\sum_{i=1}^n (x_i - x_T)^2}{n}$$

Егер  $x_1, x_2, \dots, x_n$  мәндерінің жиіліктері бар және сәйкесінше  $n_1, n_2, \dots, n_k$  болса, мұндағы

$$n_1 + n_2 + \dots + n_k = n, \text{ онда } D_T = \frac{\sum_{i=1}^n (x_i - x_T)^2}{n}.$$

Теорема: Дисперсия таңдама мәндерінің квадраттарының орта мәні мен орта мәнінің квадратының айырымына тең:  $D = \bar{x}^2 - [\bar{x}]^2$ .

Орта квадраттық ауытқу деп дисперсиядан алынған квадрат түбірді айтады:  $\sigma_B = \sqrt{D_B}$ .

Орта квадраттық ауытқудың қолданылуы:

а) вариациялық қатардың өзгергіштігі жөнінде пайымдау және орта шамалардың типтілігін салыстырмалы түрде бағалау. Бұл белгілердің тұрақтылығын анықтау барысында дифференциальды диагностикада аса қажет;

б) вариациялық қатарды қайта құру, яғни оның жиілік сипаттамасын «үш сигма» ережесі негізінде қалпына келтіру үшін.  $M \pm 3\sigma$  аралығында барлық варианттардың 99,7%,  $M \pm 2\sigma$  аралығында — 95,5% және  $M \pm 1\sigma$  аралығында — 68,3%;

в) «қырғып шығушы» варианттарды айқындау үшін (нақты және қайта құрылған вариациялық қатарларды салыстырғанда);

г) сигмальдық бағалар көмегімен қалыпты жағдай мен патологиялық жағдайдағы параметрлерді анықтау үшін;

д) вариация коэффициентін есептеу үшін;

е) орта арифметикалық шаманың қатесін есептеу үшін;

Вариация коэффициенті ( $C$ ) - орта квадраттық ауытқудың орта арифметикалық шамаға проценттік қатынасы:

$$C = \frac{\sigma}{\bar{x}} \cdot 100\%$$

Вариация коэффициенті – бұл вариациялық қатардың өзгергіштігінің салыстырмалы өлшемі.

**Вариация коэффициентінің қолданылуы.**

а) әрбір нақты вариациялық қатардың әртүрлілігін бағалау және, сәйкес, әр орта мәннің типтілігін пайымдау үшін.  $C < 10\%$  болғанда қатардың әртүрлілігі әлсіз,  $10\% < C < 20\%$  - орта, ал  $C > 20\%$  - күшті болып саналады. Қатардың әртүрлілігінің күшті болуы сәйкес орта шаманың типтілігінің аз екендігін көрсетеді, сондықтан оны практикалық мақсатта қолдану маңызды емес.

#### ҮЛГІ-ЕСЕП

*Есептің шарты.* N қаласында 2000 ж. 7 жасар ұл балалардың дене салмақтарын өлшеу жүргізілді. N қаласында 1990 ж. жүргізілген осыған ұқсас зерттеудің деректері бойынша 7-жасар ұл балалардың орташа дене салмақтары 23,8 кг,  $\sigma = \pm 3,6$  кг. болған.

**Тапсырма.** 1. Орта арифметикалық шаманы ( $\bar{x}$ ) және вариациялық қатардың әртүрлілік критерийлерін ( $\sigma$ , C) есептендер.

2. Алынған нәтижелерді бағаландар, оларды алдыңғы зерттеудің деректерімен салыстырындар, сәйкес қорытынды жасандар.

#### ЕСЕПТІҢ ШЕШУІ

**N қаласында 2000 ж. 7 жасар ұл балалардың дене салмақтарын өлшеу нәтижелері.**

Дене салмағы (кг) $x_i$	Интервалдың ортасы (ортадағы варианта) $x$	Ұл балалар саны $n_i$	$x \cdot n_i$	$d = x - \bar{x}$	$d^2$	$d^2 \cdot n_i$
15-18,9	17	16	272	-7	49	784
19-22,9	21	27	567	-3	9	243
23-26,9	25	32	800	+1	1	32
27-30,9	29	16	464	+5	25	400
31-34,9	33	9	297	+9	81	729
		$n = \sum n_i = 100$	$\sum x n_i = 2400$			$\sum d^2 \cdot n_i = 2188$

Топталған вариациялық қатарда ортадағы варианта көрші интервалдардың бастапқы варианттарының жарты қосындысы ретінде есептеледі.

$$\bar{x} = \frac{\sum x \cdot n_i}{n} = \frac{2400}{100} = 24 \text{ кг}$$

$$\sigma = \sqrt{\frac{\sum d^2 \cdot n_i}{n}} = \sqrt{\frac{2188}{100}} = \pm 4,68 \text{ кг}$$

$$C = \frac{\sigma}{\bar{x}} = \frac{4,68}{24} \cdot 100\% = 19,5\%$$

*Қорытынды:*

1 N қаласында 2000 ж. 7 жасар ұл балалардың дене салмақтарының орташа мәні 24,0 кг.

2.  $\sigma = \pm 4,68$  кг.

3. 19,5% -ке тең вариация коэффициентінің мәні белгінің әртүрлілігінің күштіге жақын орта екендігін көрсетеді.

Сонымен, дене салмағының алынған орташа мәні зерттеліп отырған жиынтық үшін жеткілікті типті деп санауға болады. Салыстыру нәтижесінде 1990 ж. қарағанда 2000 ж. 7 жасар ұл балалардың дене салмақтарында әлдеқайда үлкен вариабельділік бар екендігі белгілі болды. (4,68 кг қарсы 3,6 кг). Осыған ұқсас қорытынды вариация коэффициенттерін салыстырғаннан да келіп шығады (1990 ж. C тең  $(3,6 \cdot 100) / 23,8 = 15,1\%$ ).

#### СЕНІМДІЛІК ИНТЕРВАЛЫ

Белгілі нүктелік таңдама сипаттамалар бойынша интервалды баға немесе сенімділік интервалын құруға болады, онда қандай да бір ықтималдықпен генеральды паратер орналасады. Белгілі таңдама

көрсеткіштер негізіндегі генеральды параметрлер туралы сенімді түрде айтуға келетін болып есептелген ықтималдықтар сенімді деп аталады. Әдетте медициналық-биологиялық зерттеулерде  $P=0,95$  (95%) сенімділік ықтималдығының мәні қолданылады. Және де параметрдің нақты мәнінің осы шектерден шығу ықтималдығы  $1-0,95=0,05$  (5%)-тен аспайды. Сенімділік ықтималдықты толықтыратын шаманы әдетте  $\alpha$  деп белгілейді.

Орталық шекті теоремадан білетініміздей, таңдамалар алынған бастапқы жиынтықтың тарамдалуына тәуелсіз таңдама орташалар жуықтап алғанда қалыпты тарамдалуға ие. Осылайша, таңдама орташа мән үшін сенімділік интервалы  $\bar{x} - t_{\alpha} S_x$  және  $\bar{x} + t_{\alpha} S_x$  мәндерінің арасында орналасқан, мұндағы  $S_x$  орташаның стандартты қатесі,  $t_{\alpha}$  – Стьюдент коэффициенті,  $n$  таңдама көлеміне тәуелді (немесе сәйкес  $df=n-1$  бостандық дәрежелерінің саны) шама және сенімділік ықтималдығының таңдалған деңгейіне тәуелді шама Стьюдент тарамдалу кестелері бойынша анықталады.  $t_{\alpha}$  коэффициент шамасы сенімділік ықтималдығын 1-ге дейін толықтыратын  $\alpha$  деңгейінде кесте бойынша анықталады, яғни сенімділік ықтималдығы 95% жағдайда интервал симметриясы ескерілгенде  $(1-0,95)=0,05$  деңгейінде.

### 1. Шынайылықты репрезентативтілік қатесін анықтау арқылы бағалау тәсілі.

Орта шаманың орташа қатесінің формуласы:

$$S_x = \pm \frac{\sigma}{\sqrt{n}}, \text{ мұндағы } \sigma - \text{ ортаквадраттық ауытқу; } n - \text{ бақылау саны.}$$

Салыстырмалы көрсеткіштің қатесінің формуласы:

$$S_x = \pm \sqrt{\frac{P \cdot q}{n}}, \text{ мұндағы } p - \% , \%o, \%oo \text{ өрнектелген көрсеткіш.}$$

$q = (100 - p)$  егер  $p - \%$  өрнектелсе; немесе  $(1000 - p)$  егер  $p - \%o$  өрнектелсе;

$(10\,000 - p)$  егер  $p - \%oo$  өрнектелсе және с.с.

Егер бақылау саны 30-дан кіші болса, онда репрезентативтілік қатесінің сәйкес формулалары:

$$S_x = \frac{\sigma}{\sqrt{n-1}} \text{ и } S_x = \sqrt{\frac{P \cdot q}{n-1}}$$

### 2. Орта және салыстырмалы шамалардың сенім шекараларын анықтау.

Сенім шекараларын анықтау формулалары:

Орта шама үшін  $(\bar{x})$ :  $\bar{x} = \bar{x} \pm t S_x$   
 бас таңдама

Салыстырмалы шамалар үшін (P):  $P_{\text{бас}} = P_{\text{таңдама}} \pm t S_x$ , мұндағы  $\bar{x}$  және  $P_{\text{бас}}$  — сәйкес бас таңдама

жиынтықтың орта шамасы мен салыстырмалы көрсеткіштерінің мәндері;

$\bar{x}$  және  $P_{\text{таңдама}}$  - сәйкес таңдама жиынтықтың орта шамасы мен салыстырмалы таңдама

көрсеткіштерінің мәндері;

$S_x$  - репрезентативтілік қатесі;  $t$  – шынайылық критерийі (сенім коэффициенті).

Бұл тәсіл таңдама жиынтық нәтижелері бойынша бас жиынтықтағы зерттеліп отырған құбылыстың (немесе белгінің) өлшемдері жөнінде пайымдау қажет болған жағдайда қолданылады.

Әдісті қолданудың міндетті шарты таңдама жиынтықтың репрезентативті болуы. Таңдама тәсілдің нәтижелерін бас жиынтыққа аудару үшін сенім ықтималдығының (қатесіз болжау ықтималдығы P) дәрежесі қажет, ол зерттеліп отырған белгілер бойынша таңдама тәсіл нәтижелері жағдайлардың қанша пайызында бас жиынтықта орын алатындығын көрсетеді.

Орта шаманың немесе салыстырмалы көрсеткіштің сенім шекараларын анықтау барысында сенім ықтималдығының қажетті дәрежесін зерттеуші өзі тағайындайды.

Медициналық-биологиялық зерттеулердің көбісінде сенім ықтималдығының дәрежесі  $P=95,5\%$ , жеткілікті деп саналады, яғни бас жиынтықта таңдама тәсіл бойынша анықталған заңдылықтан ауытқу байқалуы мүмкін жағдайлар саны 5%-тен аспайды. Кейбір зерттеулерде, мысалы, жоғарытоксикологиялық заттарды, вакциналарды, оперативті емдеу тәсілдерін және т.б. қолданумен байланысты болатын, нәтижесінде сырқаттың ауыр түрі, асқыну, летальды нәтижелер тууы мүмкін жағдайларда сенім ықтималдығының  $P=99,7\%$ , дәрежесі қолданылады, яғни

бас жиынтықта 1%-тен аспайтын жағдайлар санында таңдама тәсіл бойынша анықталған заңдылықтан ауытқу байқалуы мүмкін.

P сенім ықтималдығының берілген дәрежесіне t критерийінің белгілі бір анықталған мәні сәйкес келеді.

**n>30 болғанда сенім ықтималдығының P=99,7% дәрежесіне t=3 мәні, ал P=95,5% дәрежесіне t=2 мәні сәйкес келеді. n<30 болғанда t шамасы сенім ықтималдығының дәрежесіне сәйкес арнайы кесте бойынша анықталады. (Н.А.Плохинского).**

ҮЛГІ-ЕСЕП

Бақылау саны 30-дан үлкен болған жағдайда репрезентативтілік қатені ( $S_x$ ) және бас жиынтықтың орта шамасының ( $\bar{x}$ ) сенім шекараларын анықтауға арналған.

Есептің шарты: адам ағзасына шу мен төмен жиілікті дірілдің құрамдасқан әсерін зерттеу барысында тексерілген ауылшаруашылығы көліктерінің 36 жүргізушісінің 1 сағ. жұмыс істегеннен кейінгі тамыр соғу жиіліктерінің орташа мәні минутына 80 соққы,  $\sigma = \pm 6$  мин/соққы болғаны тағайындалды.

Тапсырма: репрезентативтілік қатені ( $S_x$ ) және бас жиынтықтың орта шамасының ( $\bar{x}$ ) сенім шекараларын анықтау керек.

ШЕШУІ

1. Орта шаманың орта қатесін есептеу ( $S_x$ ):

$$S_x = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{36}} = \pm 1 \text{ мин./соққы.}$$

2. Бас жиынтықтың орта шамасының ( $\bar{x}$ ) сенім шекараларын есептеу.

Ол үшін:

а) сенім ықтималдығының дәрежесін беру: P=95,5%;

б) t критерийінің шамасын анықтау: t=2.

Онда  $\bar{x} \pm t S_x = 80 \pm 2 * 1 = 80 \pm 2$  мин./соққы.

Қорытынды: 95,5% сенім ықтималдығымен тағайындалды: бас жиынтықтағы, яғни ауылшаруашылығы көліктерінің барлық жүргізушілерінің 1 сағ. жұмыс істегеннен кейінгі тамыр соғу жиіліктерінің орташа мәні 78 минутына соққыдан 82 минутына соққыға дейінгі аралықта болады, яғни тамыр соғу жиіліктерінің орташа мәні 78 минутына соққыдан кем, 82 минутына соққыдан жоғары болу мүмкіндігі кездесетін жағдайлар саны бас жиынтықта 5%-тен аспайды.

ҮЛГІ-ЕСЕП

репрезентативтілік қатені (m) және бас жиынтықтың салыстырмалы көрсеткішінің (Pбас) сенім шекараларын анықтауға арналған.

Есептің шарты: Н қаласының аудандарының бірінде тұратын 3 жастағы 164 баланы медициналық тексеруден өткізгенде 18% жағдайда мүсіннің функционалдық сипатта бұзылуы табылды.

Тапсырма: репрезентативтілік қатені (mP) және бас жиынтықтың салыстырмалы көрсеткішінің (Pбас) сенім шекараларын анықтау керек.

ШЕШУІ

1. Салыстырмалы көрсеткіштің репрезентативтілік қатесін есептеу:

$$S_x = \sqrt{\frac{P \cdot q}{n - 1}} = \sqrt{\frac{18 * (100 - 18)}{164}} = \pm 3\%$$

2. Бас жиынтықтың салыстырмалы көрсеткіштерінің сенім шекараларын есептеу:

а) сенім ықтималдығының дәрежесін беру қажет (P =95%);

б) берілген ықтималдық дәрежесі және 30- дан үлкен бақылаулар санына сәйкес t критерийінің шамасы 2-ге тең (t = 2).

Онда Pбас=Pt  $\pm$  t S<sub>x</sub> = 18%  $\pm$  2\*3 = 18%  $\pm$  6%.



**Қорытынды:** P=95% сенім ықтималдығымен Н. қаласында тұратын 3 жасар балалардың мүсіннің функционалдық сипатта бұзылуының жиілігі 100 балаға шаққанда 12%-тен 24% аралығында болатындығы тағайындалды.

**Қалыпты таралудың негізгі сипаттамалары:**

- Сандық сипаттамалардың теңдігі (орта мән, мода және медиана өз ара тең);
- орта мәннен ауытқудың симметриялылығы;
- қисық астындағы жалпы аудан 1 ге тең;
- қисықтың ұштары екі бағытта да абцисса осіне үздіксіз жақындай отырып, алайда ешқашан онымен жанаспай шексіздікке ұмтылады.
- қисықтың түрі бас жиынтықтың орта квадраттық ауытқуымен анықталады;
- орта квадраттық ауытқуы аз таралуға жіңішке, жоғары созылған қисықтар, ал орта квадраттық ауытқуы үлкен таралуға жазыңқы қисықтар сәйкес келеді.

**3 сигма ережесі.**

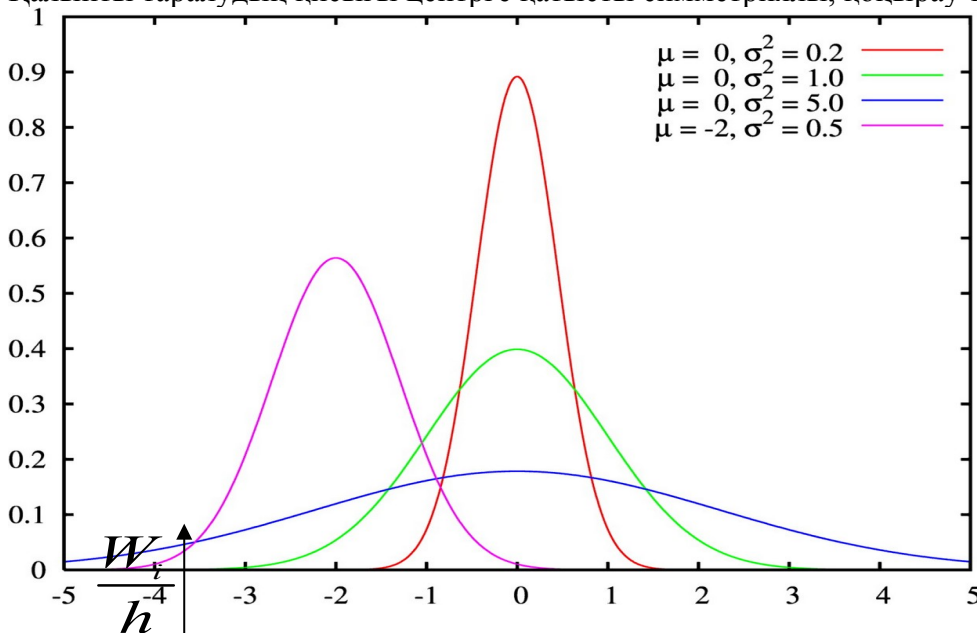
барлық мәндердің 68,26% -і  $\bar{x} \pm \sigma$  аралығында жатады (орта мәннен  $\pm 1$  орта квадраттық ауытқу);  
 барлық мәндердің 95,44% -і  $\bar{x} \pm 2\sigma$  аралығында жатады (орта мәннен  $\pm 2$  орта квадраттық ауытқулар);  
 барлық мәндердің 99,73% -і  $\bar{x} \pm 3\sigma$  аралығында жатады (орта мәннен  $\pm 3$  орта квадраттық ауытқулар).

**Гаусс қисығы**

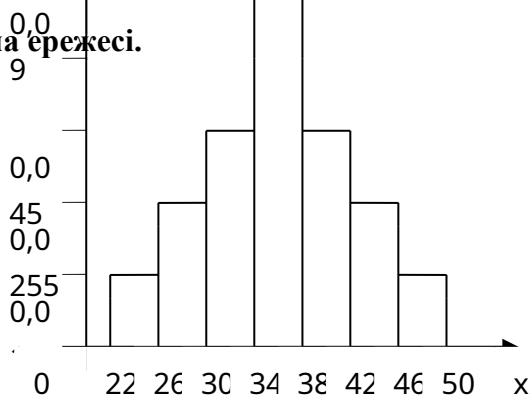
Қалыпты таралу тығыздығының графигін қалыпты қисық немесе Гаусс қисығы деп атайды.

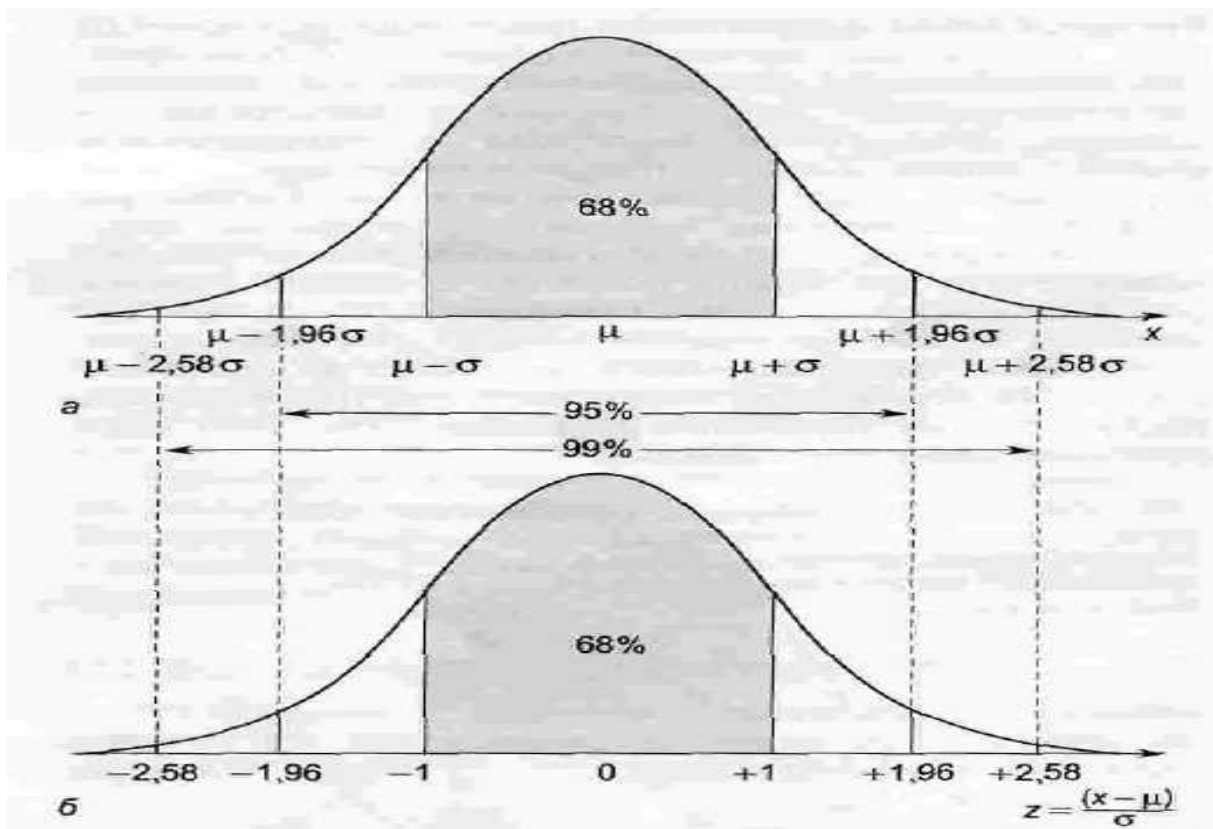
$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Қалыпты таралудың қисығы центрге қатысты симметриялы, қоңырау тәрізді түрі бар



**3 сигма ережесі.**





**Мысал.** Өте үлкен ампулалар партиясынан көлемі 50 болатын кездейсоқ таңдама алынды.  $X$  белгісі – 1 см-ге дейінгі дәлдікпен өлшенген ампулалардың ұзындығы келесі вариациялық қатар түрінде берілген: 22, 24, 26, 26, 27, 28, 28, 31, 31, 31, 32, 32, 33, 33, 33, 33, 34, 34, 34, 34, 34, 35, 35, 36, 36, 36, 36, 37, 37, 37, 37, 37, 37, 38, 38, 40, 40, 40, 40, 40, 41, 41, 43, 44, 44, 45, 45, 47, 50. Таралудың статистикалық дискретті, интервалдық қатарларын, мода, медиана, таңдама ортасын, дисперсияны, орта квадраттық ауытқу, сенімділік интервалын табу керек және тең қадаммен жиілік және салыстырмалы жиілік гистограммасын тұрғызу керек. Таңдаманың қалыпты таралу заңына бағынатындығын немесе бағынбайтындығын тексеру керек (3 сигма ережесінің орындалуы).

**Шешімі.** Таңдаманың дискретті статистикалық таралуы

$x_i$	22	24	26	27	28	31	32	33	34	35	36	37	38	40	41	43	44	45	47	50
$n_i$	1	1	2	1	2	3	2	4	5	2	5	6	2	5	2	1	2	2	1	1
$w_i$	0,02	0,02	0,04	0,02	0,04	0,06	0,04	0,08	0,1	0,04	0,1	0,12	0,04	0,1	0,04	0,02	0,04	0,04	0,02	0,02



Мода  $M_0=37$ .

$$\text{Медиана } M_i = \left| \frac{n}{2} = \frac{50}{2} = 25 \right| = \frac{x_{25} + x_{26}}{2} = \frac{36 + 36}{2} = 36$$

$$\text{Таңдама орта } \bar{X}_T = \frac{\sum_{i=1}^k n_i x_i}{n}$$

$$\bar{x}_T = \frac{1}{50} (22 * 1 + 24 * 1 + 26 * 2 + 27 * 1 + 28 * 2 + 31 * 3 + 32 * 2 + 33 * 4 + 34 * 5 + 35 * 2 + 36 * 5 + 37 * 6 + 38 * 2 + 40 * 5 + 41 * 2 + 43 * 1 + 44 * 2 + 45 * 2 + 47 * 1 + 50 * 1) = 35,76$$

Таңдама дисперсия

$$D_T = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_T)^2 n_i = \frac{1}{50} [(22 - 35,76)^2 * 1 + (24 - 35,76)^2 * 1 + (26 - 35,76)^2 * 2 + (27 - 35,76)^2 * 1 + (28 - 35,76)^2 * 2 + (31 - 35,76)^2 * 3 + (32 - 35,76)^2 * 2 + (33 - 35,76)^2 * 4 + (34 - 35,76)^2 * 5 + (35 - 35,76)^2 * 2 + (36 - 35,76)^2 * 5 + (37 - 35,76)^2 * 6 + (38 - 35,76)^2 * 2 + (40 - 35,76)^2 * 5 + (41 - 35,76)^2 * 2 + (43 - 35,76)^2 * 1 + (44 - 35,76)^2 * 2 + (45 - 35,76)^2 * 2 + (47 - 35,76)^2 * 1 + (50 - 35,76)^2 * 1] = 34,2224$$

Таңдама орта квадраттық ауытқу деп таңдама дисперсиядан алынған квадрат түбірді айтады:  
 $\sigma_T = \sqrt{D_T} = \sqrt{34,2224} = 5,85$ .

Таңдама органың стандартты қатесі:

$$S = \frac{\sigma}{\sqrt{n}} = \frac{5,85}{\sqrt{50}} = 0,83$$

Бас орта  $\bar{x} - t_{\alpha} S_x$  және  $\bar{x} + t_{\alpha} S_x$  сенім интервалының арасында жатады.  
 $t_{\alpha} = |\alpha = 0,05, df = 50 - 1 = 49| = 2,009$

$$(\bar{x} - t_{\alpha} S_x; \bar{x} + t_{\alpha} S_x) \Rightarrow (35,76 - 2,009 * 0,83; 35,76 + 2,009 * 0,83) \Rightarrow (35,59; 37,43)$$

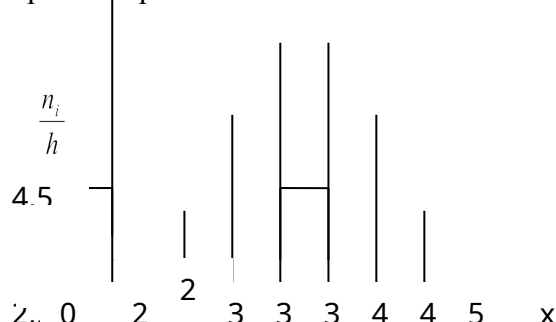
### Таңдаманың интервалдық статистикалық таралуы

$k \geq 1 + 3,32 \lg n = 1 + 3,32 * \lg 50 = 1 + 3,32 * 1,7 = 6,644 \Rightarrow k = 7$  екен,  $h$  табамыз

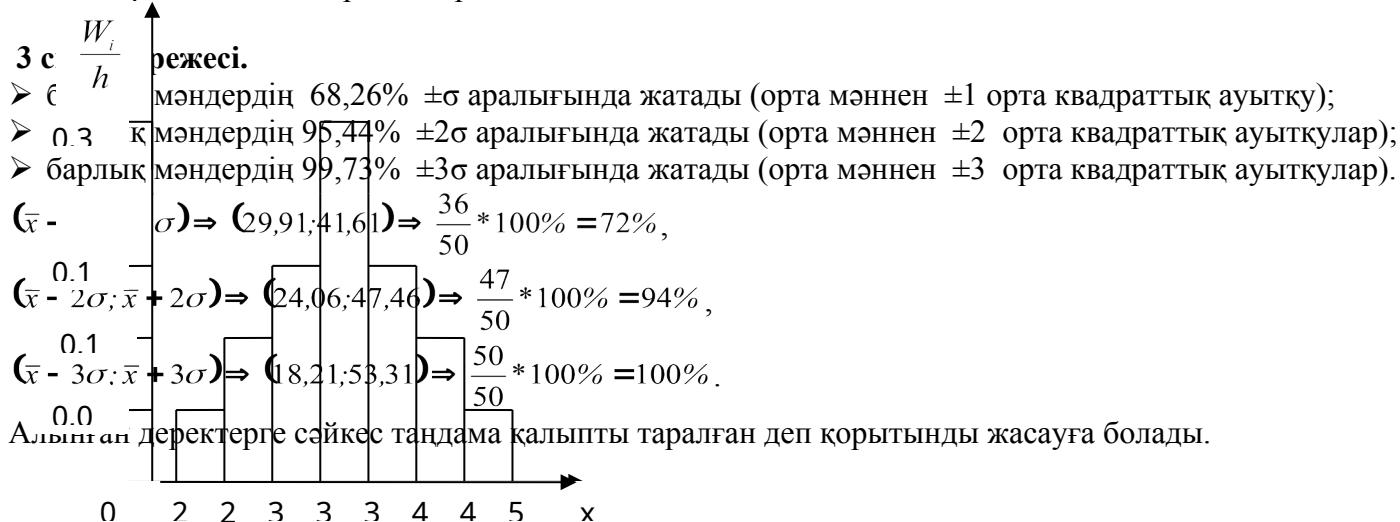
$$h = \frac{R}{k} = \frac{x_{max} - x_{min}}{k} = \frac{50 - 22}{7} = \frac{28}{7} = 4. \text{ Яғни,}$$

Топтастыру интервалы	[22;26)	[26;30)	[30;34)	[34;38)	[38;42)	[42;46)	[46;50]
$n_i$ жиіліктер	2	5	9	18	9	5	2
$w_i$ салыстырмалы жиіліктер	0,04	0,1	0,18	0,36	0,18	0,1	0,04

Жиіліктер гистограммасы



### Салыстырмалы жиіліктер гистограммасы



### Тапсырмалап

1. Екі жастағы 20 ұл балалардың бойының ұзындығы: 92, 91, 90, 93, 91, 93, 91, 92, 90, 91, 95, 94, 92, 92, 90, 95, 93, 94, 89 ( см ) тең. Таралудың статистикалық дискретті, интервалдық қатарларын, мода, медиана, таңдама ортасын, дисперсияны, орта квадраттық ауытқу, сенімділік интервалын табу керек және тең қадаммен салыстырмалы жиілік гистограммасын тұрғызу керек. Таңдаманың қалыпты таралу заңына бағынатындығын немесе бағынбайтындығын тексеру керек (3 сигма ережесінің орындалуы).

2. Ауылдық мекендер үшін 25 жас шамасындағы ер адамдардың бойлары қарастырылды. Көлемі 35 кездейсоқ таңдама бойынша: 175, 167, 168, 169, 168, 170, 174, 173, 177, 172, 174, 167, 173, 172, 171, 171, 170, 167, 174, 177, 171, 172, 173, 169, 171, 173, 173, 168, 173, 172, 166, 164, 168, 172, 174. Таралудың статистикалық дискретті, интервалдық қатарларын, мода, медиана, таңдама ортасын, дисперсияны, орта квадраттық ауытқу, сенімділік интервалын табу керек және тең қадаммен салыстырмалы жиілік гистограммасын тұрғызу керек. Таңдаманың қалыпты таралу заңына бағынатындығын немесе бағынбайтындығын тексеру керек (3 сигма ережесінің орындалуы).

3. Фармацевтикалық фабрикадан шығарылатын өнімдерден кездейсоқ жағдайда гомеопатикалық препараттардың 15 жәшігі таңдап алынды. Олардағы таблеткалардың саны сәйкесінше 50, 51, 48, 52, 50, 49, 50, 47, 50, 51, 49, 50, 52, 48, 49 – ге тең. Таралудың статистикалық дискретті, интервалдық қатарларын, мода, медиана, таңдама ортасын, дисперсияны, орта квадраттық ауытқу, сенімділік интервалын табу керек және тең қадаммен салыстырмалы жиілік гистограммасын тұрғызу керек. Таңдаманың қалыпты таралу заңына бағынатындығын немесе бағынбайтындығын тексеру керек (3 сигма ережесінің орындалуы).

4. Кездейсоқ жағдайда таңдалынған 30 наукастың артериалды қысымдарын өлшегенде келесі нәтижелер алынды: 151, 166, 133, 155, 179, 148, 143, 128, 138, 172, 168, 157, 158, 136, 169, 153, 142, 147, 134, 164, 167, 131, 152, 147, 176, 122, 149, 154, 161, 156. Таралудың статистикалық дискретті, интервалдық қатарларын, мода, медиана, таңдама ортасын, дисперсияны, орта квадраттық ауытқу,

сенімділік интервалын табу керек және тең кадаммен салыстырмалы жиілік гистограммасын тұрғызу керек. Таңдаманың қалыпты таралу заңына бағынатындығын немесе бағынбайтындығын тексеру керек (3 сигма ережесінің орындалуы).