

ДӘРІС

Вариациялық қатардың сандық
сипаттамалары

Жоспар:

- Вариациялық қатар дегеніміз не?
- Вариациялық қатарды сипаттайтын белгілер?
- Қандай жағдайда жай вариациялық қатарды құрады?
- Топтастырылған вариациялық қатар қандай жағдайда құрады?
- Топтастырылған вариациялық қатар құру кезеңдері?
- Топтастырылған вариациялық қатарда топ саны қалай анықталады?
- Вариантаның топ аралығы қалай табылады?
- Статистикалық өлшемдер

Статистика кең мағынада, табиғат пен қоғамның көптеген құбылыстарының сапалық ерекшеліктерін айқындау үшін сол құбылыстарға жүргізілетін сандық талдау туралы ғылым. Статистика жекелеген бірліктерді емес, сол жеке бірліктердің жиынтығын зерттеу үшін пайдаланылады. Биометрияның негізін 1889 жылы ағылшын ғалымы Ф. Гальтон (Дарвиннің немере інісі) қалаған. «Варияциалық статистика» термині «биометрия» терминінің синонимі болып табылады. Бірінші рет бұл терминді ғылымға 1899 Дункер енгізген. Оны тұқымқулаушылық пен өзгергіштіктің сұрақтары қатты қызықтырған. Ол адамға зерттеулер жүргізді. Көбінесе адамның дене құрылысындағы сандық белгілерді зерттеді. Оны ғылыми криминалистиканың негізін қалаушы деп санаған. Өйткені, ол қылмыскерлердің ұқсастықтарына зерттеу жүргізген. Белгілердің сандық көрсеткіші варианты немесе дата деп аталады және v немесе x – пен белгіленеді.

Статистикалық жиынтық – дегеніміз белгілі бір кеңістікте және уақытта алынған салыстырмалы біртекті элементтерден тұратын топ. Статистикалық жиынтық – 2 бөлінеді: жалпы және таңдамалы.

Әр бір бақылау бірлігі көптеген сипаттардан тұрады, бірақ соның ішінде біздің мақсатымызға қажетті белгілер алынады. Белгілер сипатына қарай—атрибутивті және сандық болып бөлінеді. Атрибутивті – адамның жынысы, мамандығы, тұратын жері туралы мәлімет.

Сандық – адамның салмағы, жасы, бойы, туралы мәлімет.

Белгілі бір іріктеуде әрбір вариантының неше рет кездесетіндігін көрсететін сан жиілік деп аталады және f әрпімен белгіленеді. Басты жиынтықтың көлемі N әріппен белгіленеді, таңдамалы жиынтық көлемі — n . Зерттелінетін жиынтықтың элементтері — белгі немесе варианта деп аталады, Белгіленуі — X_i .

- **Вариация** — жиынтықтың жеке бірліктерінің белгісі мәнінің аутқуы, әртүрлілігі, өзгерілуі.
- **Вариация** деп зерттелетін жиынтықтың кез-келген бірлігі белгісі мәнінің бір кезеңдегі және мезеттегі құбылмалылығын айтады.
- **Вариация мәні:**
- Оның өзгеруі басқа өзгермелі белгілердің сол белгіге әсер ету деңгейін бағалауға мүмкіндік береді.
- Статистикалық моделдерді құруда қолданылады.

Вариация өлшемі деп бегінің ауытқушылығын көрсететін абсолюттік және қатысты көрсеткіштерді айтады. Вариацияның абсолюттік көрсеткіштеріне : вариация ауқымы, оташа ауытқу, дисперсия, орта квадраттық ауытқу жатады. Вариацияның қатысты көрсеткіштеріне: осцилляция коэффициенті, вариацияның сызықты коэффициенті, вариация коэффициенті жатады.

Жиынтықтағы варианттарды зерттеу үшін, оларды вариациялық қатарлар ретінде жазады. Вариациялық қатардағы ең көп кездесетін вариантаны мода деп атайды. Вариациялық қатардың ортасында орналасқан вариантаны медиана деп аталады. Мода M_o , ал медиана M_e белгілерімен белгіленеді.

Вариациялық қатар — зерттелетін белгінің сандық мелшерлерін жоғарылату немесе темендету ретімен орналастыру. Вариациялық қатар статистикалық жиынтықтың белгілерінің сан түріндегі мәнін көрсетеді және орта шаманы есептегенде пайдаланылады.

Вариациялық қатардың құрылымдық сипаттамасына: мода, медиана, децили, квартили, перцентили жатады.

Мода деп – зерттелетін жиынтықта басқаларына қарағанда жиі кездесетін **вариантаны** айтады.

Дискреттік қатарда мода деп – көп рет кездесетін белгіні айтады. Мода мысалы, сатып алушыларда үлкен сұранысқа ие киім мен аяқ киімінің размерін анықтау үшін жиі қолданылады.

Интервалды вариацияның қатарда моданы есептеу үшін алдымен мода орналасқан модальді интервалды анықтау керек, ал одан кейін белгінің модальді шамасының мәнін анықтау керек.

Медиана (Me) - белгілі бір тәртіппен орналасқан, өсуі бойынша немесе азаюы бойынша реттелген қатардың варианттарының бірінің ортасасы. Ол мұндай қатарды ортасынан бөледі. Медиананы табу үшін реттелген қатардың ортасында орналасқан белгінің мәнін табу керек. Тақ қатардағы реттелген қатардың медианасы деп ортасында тұрған белгінің шамасын айтады. Тақ қатардағы реттелген қатардың медианасы **номері** келесі формуламен есептеледі:

$$N_{Me} = \frac{(n+1)}{2}, \quad \text{мұнда } n - \text{қатардың мүшесі саны;}$$

Жұп сандағы реттелген қатардың медианасы деп қатардың ортасында орналасқан екі варианттың бірінің орташа арифметикалығын айтады.

Дискретті вариациялық қатардың медианасы жиынтық көлемінің ортасынан асатын жинақталған жиілігін сомасымен табылады.

Вариациялық қатардағы моданы және медиананы тапқаннан кейін қатарды 4 бірдей, 10 және 100 бөлікке бөлетін бірліктердің белгісі мәнің табуға болады. Бұл шамаларды «децили», «квартили», «перцентили» деп атайды.

Вариация ауқымы – белгі мәнінің үлкені және кішісі арасындағы әртүрлілігі.

Вариация коэффициенті – орташа квадраттық ауытқудың белгінің орташа шамасының пайыздық қатынасына тең.

Мода – жиілігі көп қайталанатын белгінің мәні.

Медиана - реттелген жиынтықтың ортасына келетін белгінің мәні.

Квартили – реттелген жиынтықты 4 бірдей бөлікке бөлетін белгінің мәні.

Децили – реттелген қатарды 10 бірдей бөлікке бөлетін белгінің мәні.

Перцентили – қатарды 100 бөлікке бөлетін белгінің мәні.

Вариациялық қатар – берілген варианттардың және оларға сәйкес келетін жиіліктердің өсу немесе кему ретімен орналасуы. Дискреттік вариациялық қатарда варианттар дискретті белгілер ретінде анықталады, ал интервалдық вариациялық қатарда варианттар үздіксіз белгілер ретінде анықталады.

Дискреттік вариациялық қатар анықтама бойынша құрылады.

Интервалдық вариациялық қатарды құру үшін Стерджесс формуласын

қолданады:
$$h = \frac{x_{\max} - x_{\min}}{1 + 3,322 \lg n},$$
 мұндағы x_{\max} и x_{\min} - сәйкесінше қарастырылып жатқан жиынтықтың максималды және минималды варианттары; h - интервал ұзындығы, $1 + 3,322 \lg n$ интервалдың тиімді саны.

Вариациялық қатарды графика арқылы бейнелеу вариациялық қисық деп аталады.

Вариациялық қатарларды графикалық суреттеу үшін келесі бейнелер пайдаланады:

- **ПОЛИГОН** — дискреттік және интервалдық қатарларды бейнелеуге арналған график. Координаталар жүйесінде X бойымен варианттар, U бойымен жиіліктер алынады.
- **ГИСТОГРАММА** — тек қана интервалдық қатарды бейнелейді. Координаталар жүйесінде X бойымен интервалдар, U бойымен жиіліктер алынады.
- **КУМУЛЯТА** — жинақталған жиіліктер қисығы. Координаталар жүйесінде X бойымен варианттар, U бойымен жинақталған жиіліктер алынады.
- **ОГИВА** — кумулятаға керісінше, X бойымен жинақталған жиіліктер, U бойымен варианттар алынады.

Есептің қойылуы: Қатардың таңдамалы көрсеткіштерін Microsoft Excel бағдарламасында есептеу.

Шешімі:

1. Microsoft Excel бағдарламасын іске қосу.
2. Бастапқы деректерді 1 суреттегідей енгізу

	A
1	11300
2	12100
3	12100
4	12500
5	12700
6	12700
7	12700
8	12900
9	13200
10	14600

Сурет 1 - Бастапқы деректер

- Деректерді дискретті қатар ретінде топтастырамыз және кестеге толтырамыз

C	D
Признак, X_i	Локальная частота, n_i
11300	
12100	
12500	
12700	
12900	
13200	
14600	
	0

Сурет 1.1 - Дискретті үлестірім қатар

Жиілікті (ni) есептеу үшін СТАТИСТИЧЕСКИЕ категориясынан - СЧЕТЕСЛИ функциясын пайдаланамыз.

- D2 ұяшығын белгілеу – Вставка-Функция командасын орындау; Статистические категориясынан СЧЕТЕСЛИ функциясын таңдау;
- «Аргументы функции» терезесінде «Диапазон» өрісін A1:A10 ұяшықтар диапазонымен толтыру, «Критерий» өрісінде – 11300 (C2 ұяшығы);
- D3:D8 ұяшықтарына формуланы көшіру
- барлық жиіліктерінің сомасын есептеу (1.2 сурет)

С	Д
Признак, X_i	Локальная частота, n_i
11300	1
12100	2
12500	1
12700	3
12900	1
13200	1
14600	1
	10

Сурет 1.2 - Дискретті үлестірім қатар

Қатардың көрсеткіштерін есептеу үшін
СТАТИСТИЧЕСКИЕ категориясынан келесі
функцияларды пайдаланамыз.

Арифметикалық орташа – СРЗНАЧ функциясы

Дисперсия – ДИСПР функциясы

Орташа квадраттық ауытқу – СТАНДОТКЛОНП
функциясы

Мода – МОДА

Медиана – МЕДИАНА

Ассиметрия коэффициенті - СКОС

Эксцесс көрсеткіші – ЭКСЦЕСС

Көрсеткіштер мәндерін қосымша кестеге толтыру
(1.3 сурет)

12		Характеристика	Значение
13		Среднее значение признака	=СРЗНАЧ(А1:А10)
14		Дисперсия	=ДИСПР(А1:А10)
15		СКО	=СТАНДОТКЛОНП(А1:А10)
16		Мода	=МОДА(А1:А10)
17		Медиана	=МЕДИАНА(А1:А10)
18		Ассиметрия	=СКОС(А1:А10)
19		Эксцесс	=ЭКСЦЕСС(А1:А10)

Сурет 1.7 - Қатар көрсеткіштерін есептеу

Нәтижесінде:

12		Характеристика	Значение
13		Среднее значение признака	12680
14		Дисперсия	661600
15		СКО	813,39
16		Мода	12700
17		Медиана	12700
18		Ассиметрия	0,93
19		Эксцесс	2,74

Сурет 1.4 - Көрсеткіштер мәндері

- Статистикалық талдаудың бірінші кезеңінде зерттелініп жатқан бірліктер бойынша мәліметтер жиналады.
- Статистикалық талдаудың екінші кезеңінде сол деректерді топтастырып, жинақтауға жөн.
- Жинақтау нәтижелері екі түрлі вариациялық қатарлар ретінде анықталуы мүмкін.
- Вариациялық қатарлардың екі түрлерін ажыратады — дискретті және интервалдық.

- Есептің қойылуы:
- Қостанай облысы бойынша негізгі тамақ өнімдеріне орташа бағалар белгілі (2006 ж 16.01. сәйкес)

Аудан	Сиыр еті (килограмына теңгемен)
1. Аманкелді	370
2. Әуликөл	345
3. Денисов	340
4. Жанкелдин	345
5. Жітіқара	340
6. Қамысты	330
7. Қарасу	330
8. Мендіқара	330
9. Наурзым	336
10. Сарыкөл	330

- Тапсырма: 1.1 кестедегі мәліметтер бойынша дискреттік вариациялық қатар құру, қатардың негізгі көрсеткіштерін: арифметикалық орташа, дисперсия, орташа квадраттық ауытқу (ОКА), мода, медиана, вариация коэффициентің, вариация құбылуын есептеу. Нәтижесі бойынша қорытынды жасау.

- Шешімі:

- Дискреттік вариациялық қатарды құру үшін, кестедегі мәліметтерді өсу немесе кему ретімен орналастырамыз: 330, 330, 330, 330, 336, 340, 340, 345, 345, 370.

Дискреттік қатарды кесте ретінде анықтаймыз: берілген сиыр етінің бағасы – варианта, оларды X_i белгілеп, өсу ретімен кестеге толтырамыз, бірнеше рет кездесетін белгілерді бір рет жазамыз. Кестенің екінші бағанында – жиіліктер. Бұл көрсеткіш вариантының қарастырылып жатқан жиынтықта неше рет кездесетінін анықтайды. Мысалы, 330 сиыр етінің бағасы 4 ауданда кездеседі, 336 сиыр етінің бағасы – 1 ауданда.... Кестенің 3 бағанында жинақталған жиіліктер, бұл көрсеткіш варианттардың берілген варианттан кем немесе тең варианттар санын анықтайды. Мысалы, 330 сиыр етінің бағасынан төмен бағалар кездескен емес, ал оған тең – 4 ауданда, қосындысы -4; 336 сиыр етінің бағасынан төмен бағалар – 4 ауданда, ал оған тең – 1 ауданда, қосындысы -5....

Кесте 1.2

Варианталар, X_i	Жиіліктер, n_i	Жинақталған жиіліктер, n_i жин
330	4	4
336	1	5
340	2	7
345	2	9
370	1	10
Барлығы	10	

Қатардың көрсеткіштерін есептейміз:

Вариациялық қатарларды талдау көрсеткіштері 4 топқа бөлінеді:

1) Орташа шамалар – арифметикалық орташа, квадраттық орташа

2) Құрылымдық орташа шамалар – мода, медиана

3) Вариация көрсеткіштері – дисперсия, орташа квадраттық ауытқу, вариация коэффициенті, вариация өлшемі.

4) Асимметрия және эксцесс.

Арифметикалық орташа:

$$\bar{X} = \frac{\sum x_i * n_i}{\sum n_i} = \frac{330 * 4 + 336 * 1 + 340 * 2 + 345 * 2 + 370 * 1}{10} = \frac{3396}{10} = 339.6$$

Дисперсия:

$$\sigma^2 = \frac{(330-339,6)^2 * 4 + (336-339,6)^2 * 1 + (340-339,6)^2 * 2 + (345-339,6)^2 * 2 + (370-339,6)^2 * 1}{10} = 136,44$$

Орташа квадраттық ауытқу

$$\sigma = \sqrt{\sigma^2} = \sqrt{136,44} = 11,68$$

Мода, Медиана:

Дискреттік қатар үшін МОДА және МЕДИАНА анықтама бойынша табылады.

$$M_0=330$$

$$Me=340$$

Вариация өлшемі:

$$R = x_{\max} - x_{\min} = 370 - 330 = 40 \text{ мг}$$

Вариация коэффициенті:

$$V = \frac{\sigma}{x} * 100\% = \frac{11.68}{339.6} * 100 = 3.4\%$$

Асимметрия:

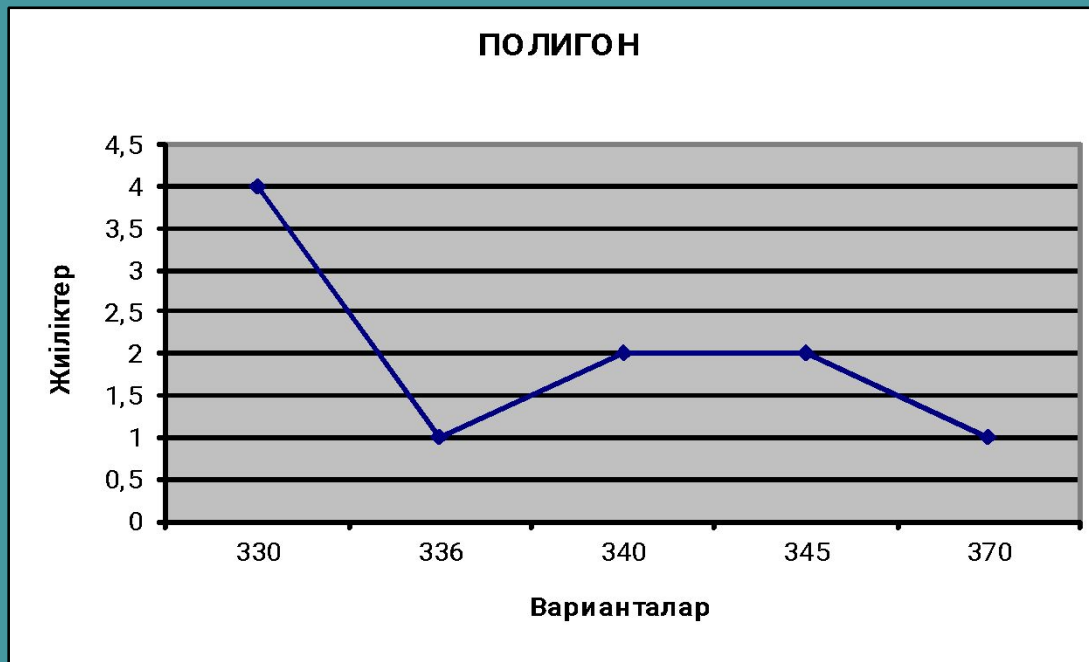
$$A = \frac{\sum (x_i - \bar{x})^3 * n_i}{n * \sigma^3} = \frac{(330 - 3396)^3 * 4 + (336 - 3396)^3 * 1 + (340 - 3396)^3 * 2 + (345 - 3396)^3 * 2 + (370 - 3396)^3 * 1}{10 * 11.68^3} = 1,5$$

Эксцесс:

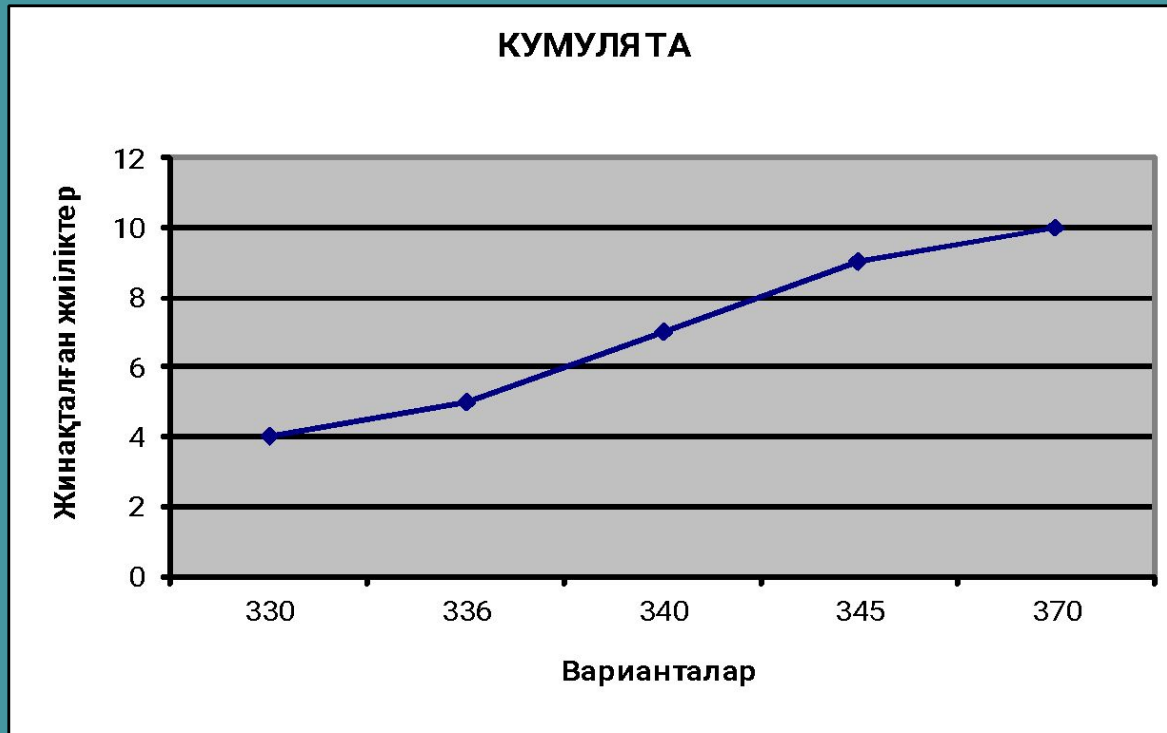
$$E = \frac{\sum (x_i - \bar{x})^4 * n_i}{n * \sigma^4} - 3 =$$

$$\frac{(330 - 339.6)^4 * 4 + (336 - 339.6)^4 * 1 + (340 - 339.6)^4 * 2 + (345 - 339.6)^4 * 2 + (370 - 339.6)^4 * 1}{10 * 11.68^4} - 3 = 1,7$$

Қатарды график түрінде бейнелейтін болсақ, ПОЛИГОН, КУМУЛЯТА, ОГИВА қисықтарын сызамыз.

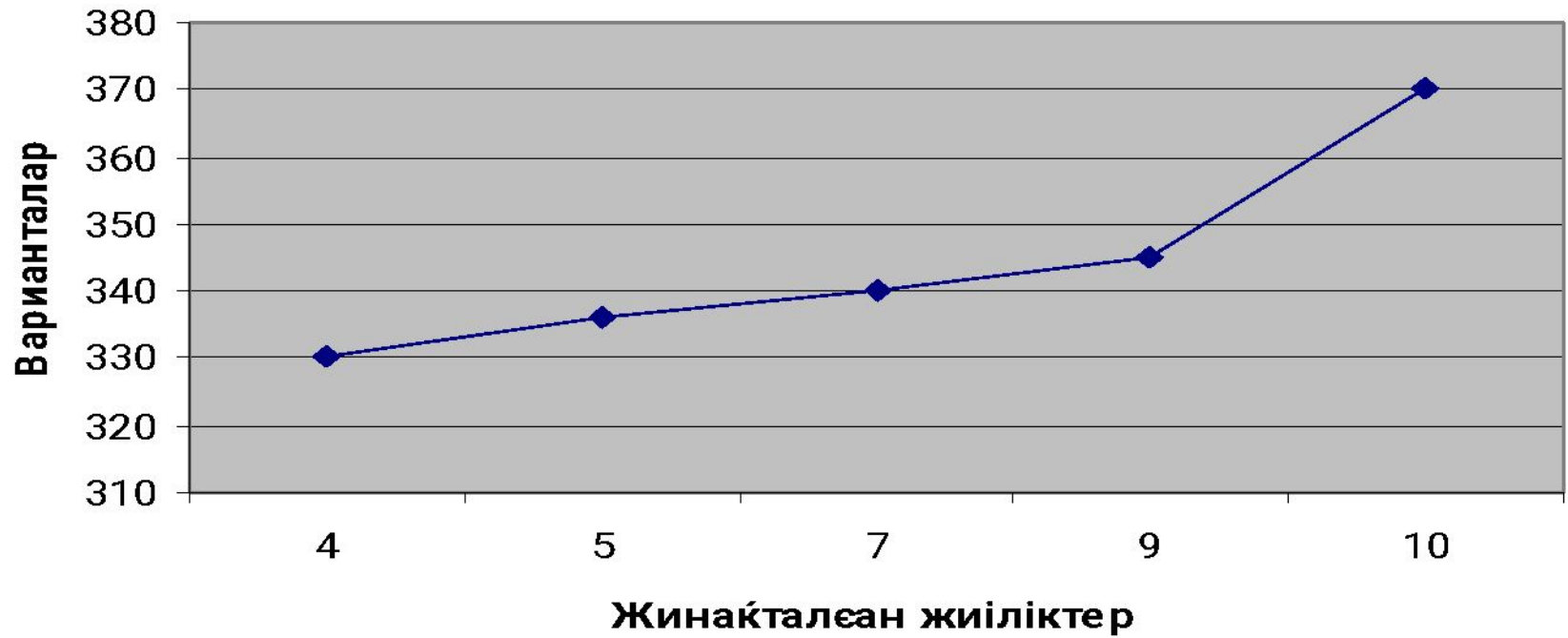


Сурет 1.1 Полигон



Сурет 1.2 Кумулята

ОГИВА



Сурет 1.3 Оги́ва

Қорытынды: қарастырылып жатқан жиынтықта сиыр етінің орташа бағасы 339,6 теңгеге тең, жиі кездесетін баға 330 теңге. Вариация коэффициенті 3,4 пайызға тең, сондықтан таңдамалы жиынтық көрсеткіштерін басты жиынтыққа тарата аламыз, яғни Қостанай облысына. Асимметрия коэффициенті $1,56 > 0$, қатар оң жақты асимметриялы. Экссесс $1,78 > 0$, үлестірім үшкір биікті.

- **Пайдаланған әдебиеттер:**

- 1. Ә.Н.Шыныбеков.Ықтималдықтар теориясы және математикалық статистика элементтері. Оқу құралы.-Алматы:Экономика, 2008.-236 б.
- 2. Лукьянова Е.А. Медицинская статистика.- М.: Изд. РУДН, 2002.
- 3. Медик В.А.,Токмачев М.С.,Фишман Б.Б. Статистика в медицине и биологии. М.: Медицина, 2000.
- 4. И.В. Павлушков и др. Основы высшей математики и математической статистики. (учебник для медицинских и фармацевтических вузов) М., «ГЭОТАР - МЕД»; 2008