

OÑTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені		№ 35-11(Б)-2024 60 беттің 1 беті

ДӘРІС КЕШЕНІ

Пәннің атауы:	Биостатистика
Пәндер коды::	Biostat 2203
БББ:	6В10111 «Қоғамдық денсаулық»
Оқу сағаттары/кредиттерінің көлемі:	150/5
Оқытылатын курс пен семестр:	2/3
Дәріс көлемі:	10

Шымкент, 2024 жыл

ОҢТҮСТІК-ҚАЗАҚСТАН

**MEDISINA
AKADEMIASY**

«Оңтүстік Қазақстан медицина академиясы» АҚ



SOUTH KAZAKHSTAN

**MEDICAL
ACADEMY**

АО «Южно-Казакстанская медицинская академия»

Медициналық биофизика және ақпараттық технологиялар кафедрасы

«Биостатистика» пәні бойынша дәріс кешені

№ 35-11(Б)-2024

60 беттің 2 беті

Дәріс кешені бойынша МОБ сәйкес әзірленген «Денсаулық сақтаудағы ақпараттандыру және цифровизация» және кафедра мәжілісінде талқыланды.

Хаттама № 11 « 30 » 05 2024 ж.

Кафедра меңгерушісі: Иванова М.Б.

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 3 беті	

№ 1 Дәріс

1. Тақырыбы: Кіріспе. Биостатистика негіздері.

2. Мақсаты: Студенттерге «Биостатистика», оның пәні, даму тарихы және міндеті туралы негізгі түсінікті қалыптастыру. Студенттерді берілгендердің түрлерімен, оларды жинау әдістерімен және сызбалық ұсынуды, сондай-ақ өлшеу шкала түрлерімен, өлшеудің сенімділігі және ақиқаттылығымен таныстыру.

3. Дәріс тезистері:

Статистика - қоғамдық құбылыстар санынның сапалық жағымен тығыз байланыстылығын зерттейтін қоғамдық ғылым.

Статистикада бақылауға немесе өлшеуге болатын нысандардың немесе құбылыстардың қасиетін *белгі* деп атайды.

Биология, медицина, фармацевтика, гигиена және денсаулық сақтау саласына қатысты сұрақтарды зерттейтін статистиканы *биостатистика* деп атайды.

Тәжірибе жүзінде, жеке ғылыми жұмыста дәрігер, фармацевт, медбике үшін биостатистиканың маңызы зор.

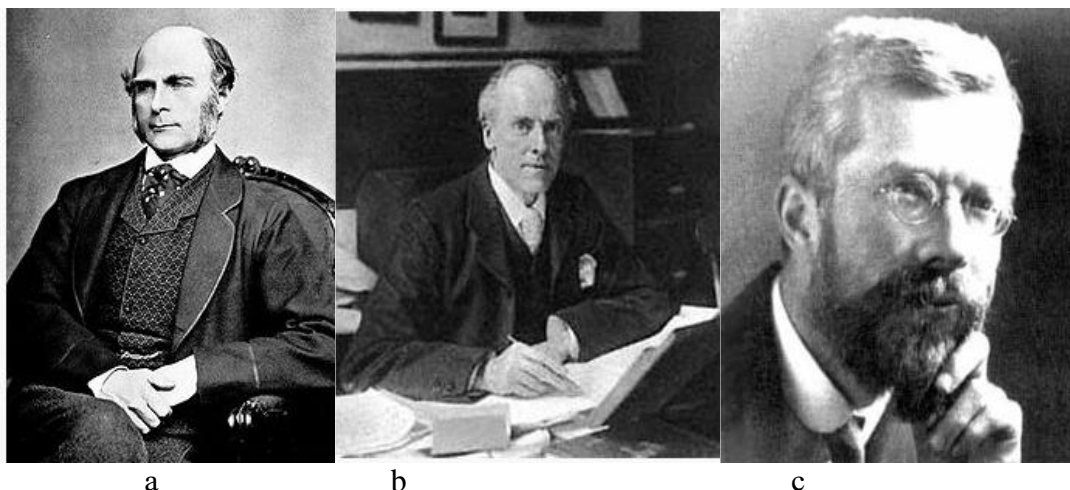
Берілгендерді жинау, оларды жалпылау, бақылау нәтижелерінің негізінде талдау және қорытындысын шығаруда биостатистика түрлі әдістерді қолданады.

Статистикалық талдау ақпарат табуға және алынған ақпараттың сапасын бағалауға көмектеседі.

Биостатистиканың міндеттері:

- сандық түрде ұсынылған биологиялық факт (өлшеу) – бұл биологиялық нысан қасиетінің жеке сан, нұсқа немесе айнымалылардың мәні түрінде берілуі;
- көп фактілердің жалпылама сипаттамасы (статистикалық бағалау) – бұл бір текті нысандардың немесе белгілердің қасиетін толық сипаттайтын параметрлердің және көрсеткіштердің есебі.
- заңдылықтарды іздестіру (статистикалық болжамдар тексеру) – бұл салыстырып отырған жиындардың, нысандардың, олардың сипаттамаларының сыртқы және ішкі себептерге тәуелділігінің кездейсоқ еместігін дәлелдеу.
- медициналық талдаулар үшін, классикалық статистикалық әдістерді қолдану;
- медициналық талдаулар үшін, заманауы статистикалық әдістерді қолдану;
- медициналық талдаулар үшін жаңа әдістер дайындау.

Биометрия негізі Фрэнсис Гальтоннан (1822-1911) басталады. (1.1 а- сурет).



a

b

c

1.1 сурет - а – Ф. Гальтон, б – К.Пирсон, с – Р.Фишер

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 4 беті	

Алғашқыда Гальтон дәрігер болғысы келген.

Бірақ Кембридж университетінде оқып жүріп, ол жаратылыстанумен, метеорологиямен, антропологиямен, тұқым қуалаушылық және даму теориясымен айналыса бастады.

1889 жылы жарық көрген тұқым қуалаушылық туралы кітабында алғаш рет *biometry* сөзі енгізілді.

Осы уақытта корреляциялық талдау негізі қаланды. Гальтон жаңа ғылымның негізін қалады және оған ат қойды. Бірақ оны ғылыми салаға математик Карл Пирсон (1857-1936) айналдырды (1.1, b - сурет). 1884 жылы Пирсон Лондон университетінде қолданбалы математика кафедрасын басқарды, ал 1889 жылы Гальтонмен және оның еңбектерімен танысты.

Пирсонның өмірінде ағылшын зоологы, биометрик, «Биометрика» журналының алғаш ұйымдастырушысы В.Уэлдон үлкен рөл атқарды. 1893 жылы Пирсон Уэлдонның зоологиялық талдауларына көмектесе отырып, орташа квадраттық ауытқу және вариациялық коэффициент ұғымын енгізді.

1898 жылы Пирсон Гальтонның тұқым қуалаушылық теориясын математикалық тұрғыдан сипаттау үшін көп регрессиялық талдаудың негізін қалады.

1903 жылы Пирсон белгілердің кездесушілігі (сопряженности признаков) теориясын, ал 1905 жылы сызықты емес корреляция және регрессия негізін жариялады.

Биометрияның келесі даму кезеңі статистик Рональд Фишермен (1890-1962) байланысты (1.1 с- сурет).

Фишер Кембридж университетінде оқып жүргенде Г.Мендель және К. Пирсон еңбектерімен танысты.

1913-1915 жылы Фишер өндіріс орынында статистик болып жұмыс жасады, ал 1915-1919 жылы орта мектепте физика, математика пәндерінен сабақ берді.

1919-1933 жылдарда Фишер ауылшаруашылық тәжірибе станциясында статистик болып жұмыс істеді.

Содан соң 1943 жылға дейін Лондон университетінің профессоры, ал 1943 жылдан 1957 жылға дейін Кембриждегі генетика кафедрасының меңгерушісі болып қызмет атқарды.

Ол таңдамалылардың үлестірім теориясын, дисперсиялық және дискриминанттық талдау әдісі, тәжірибені жоспарлау теориясын, ең жоғарғы шындыққа ұқсас әдісі және заманауи қолданбалы статистика мен математикалық генетиканың негізін құрайтын көптеген теориялардың негізін салушы болып табылады.

Орташа шамалар туралы ұғымының негізін қалаушы белгия ғалымы А. Кетле статистикалық әдістерді биология, медицина және әлеуметтанудағы мәселелерді шешуде қолданды.

Кез келген статистикалық зерттеудің бірінші кезеңі талдау жасалатын нысанның немесе үрдістің айнымалыларының нақты шамалары жөнінде мәліметтер жинау.

Статистикалық берілгендерді жинау дегеніміз зерттелетін жиынтықтың элементтері туралы мәлімет алу үдерісі және олардың қасиеттері. Бұл берілгендер статистикалық өңдеу және талдау жасауға негіз болады.

Екінші кезеңі берілгендерді типіне талдау жасау.

Берілгендердің типтері: сандық, сапалық және мерзімі (1.2. сурет).

Берілгендердің негізгі типтері сандық және сапалық және мерзімі деп бөлінеді.

Сандық берілгендер – дискретті (үзілісті) және үзіліссіз болып бөлінеді.

Дискретті берілгендер – тек бүтін сан яғни бөлшек сан болуы мүмкін емес түрде алынған сандық берілгендер. Мысалы: балалар саны және т.б.

Үзіліссіз берілгендер – теориялық жағынан алғанда бөлшек бөлігі болуы мүмкін, үзіліссіз шкалада өлшегенде алынатын берілгендер. Мысалы: дене массасы, бойы, артериялық қысымы және т.б.

Үзіліссіз берілгендер аралық және салыстырмалы болып бөлінеді.

Аралық берілгендер – физикалық мәні бар, абсолюттік шамалары өлшенетін үзіліссіз берілгендердің түрі.

Салыстырмалы берілгендер – белгінің алғашқы күйіне салыстырғанда мәнінің өзгеру (өсу немесе кему) үлесі үзіліссіз берілгендердің түрі. Бұл берілгендер өлшемсіз шама немесе пайызбен өлшенеді.

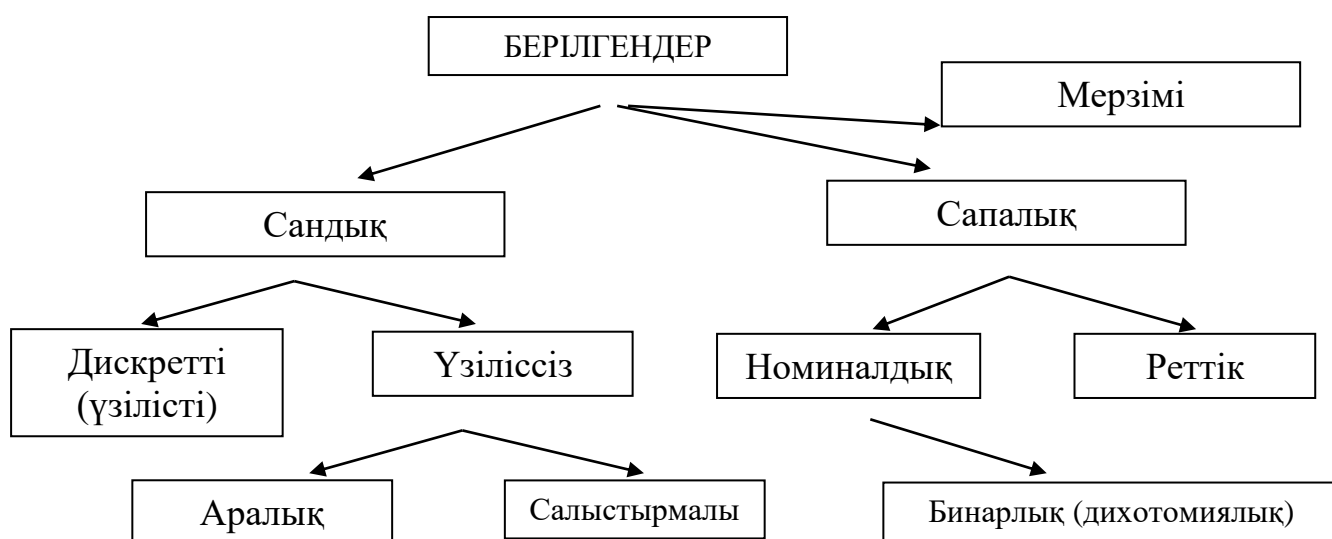
Сапалық берілгендер – номиналды және реттеуші болып бөлінеді.

Номиналды берілгендер – өлшенбейтін категорияларды шартты түрде бейнелейтін код түріндегі (диагноз коды) берілгендер.

Реттік берілгендер – бір белгінің анық байқалу дәрежесін шартты түрде бейнелейтін (қатерлі ісік ауруларының кезеңі, жүрек талмасының дәрежесі және т. б.) сапалық берілгендердің түрі.

Бұлардың дискретті берілгендерден айырмашылығы аурудың белгісін өлшеу үшін пропорционалды шкалалардың болмауы.

Бинарлық (дихотомиялық) берілгендер – сапалық мәліметтердің ерекше түрі. Мұндай типті белгінің мүмкін болатын екі мәні бар (жынысы, қандайда бір ауру түрінің болуы немесе болмауы).



1.2 сурет. Берілгендер типтерінің классификациясы

Берілгендердің ерекше түрі *мерзімі* болып табылады. Кейде олармен қандайда бір арифметикалық амалдар орындау (екі оқиғаның арасындағы осы оқиғалардың күндері бойынша абсолютті уақыт кезеңін есептеу) қажет болады.

Кейде жоғарыда келтірілген түрлердің дербес жағдайы болып табылатын берілгендердің кейбір ерекше дербес түрлерін бөліп қарастырады: шенін, ұпайлар, визуалды ұқсас шкалалар, цензуралы берілгендер және т.б.

Тереңдетілген статистикалық зерттеу жасаудан бұрын берілгендерге алдын-ала талдау жасау маңызды. Бұл кезеңде теріліп алынған берілгендерді сығу және жүйелеу үшін сызбалы әдістер қолданылады. Бұлар теріліп алынған берілгендерді ерекшелігін бағалау

және ауытқуларды анықтау яғни алдағы уақытта талдау жасау үшін қажетті статистикалық әдістерді таңдауға мүмкіндік береді.

Дискреті берілгендер кесте түрінде, бағанды диаграмма, пиктограмма, шеңберлік диаграмма, нүкте суреттері бойынша беріледі.

Үздіксіз берілгендер топтастырылған таңдама, гистограмма, «жапырақты бұтақ» немесе «мұртты жәшік» диаграммасы, Лоренц қисығы және т.б. түрде берілуі мүмкін.

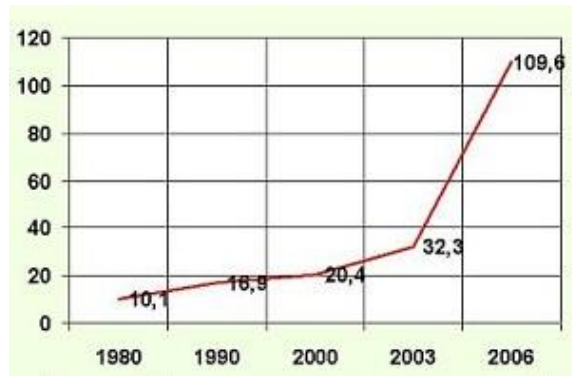
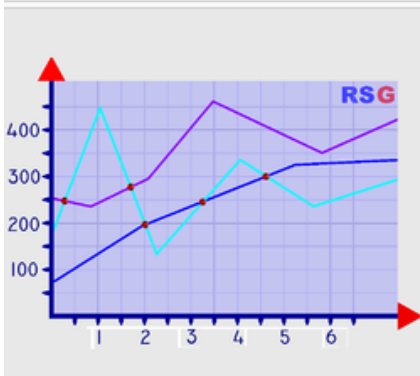
Аралас берілгендер шашырау диаграммасы түрінде берілуі мүмкін.

Берілгендердің сызба түріндегі әдістері.

Статистикалық берілгендері әртүрлі геометриялық фигура түрінде бейнеленген сызбаны *диаграмма* деп атайды.

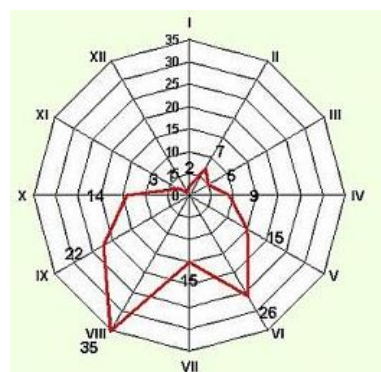
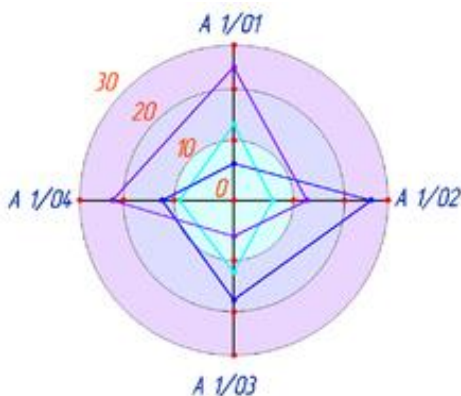
Жиі қолданылатын диаграммалардың түрлері:

- Құбыластардың жітілігінің (интенсивтілігі), қатынасының, көрнекілігінің, орташа немесе абсолютті шамаларының өзгеруін көрсеткіш ретінде бейнелейтін диаграмманы *сызықты* деп атайды (1.3 сурет).



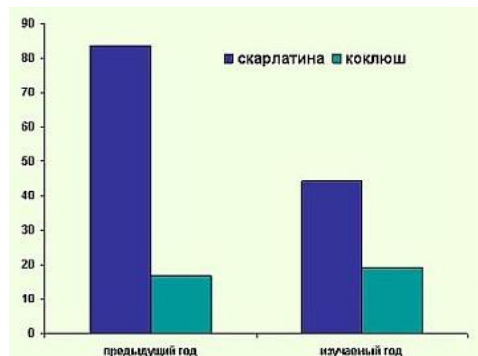
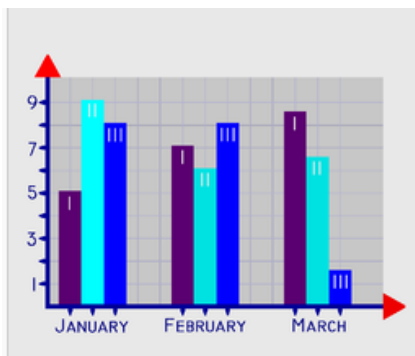
1.3 сурет. Сызықтық диаграммалар (сызба)

- Құбылыстардың өзгеруін уақыттың тұйық тізбегі (сөтке, апта, ай, жыл) ретінде бейнелеу үшін қолданылатын сызықты диаграмманың түрін *радиалды* деп атайды (1.4-сурет).



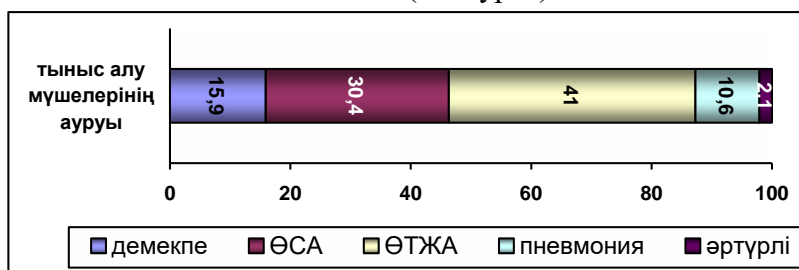
1.4 сурет. Радиальды диаграммалар

- Таңдап алынған масштабқа сәйкес құбыластардың динамикасын немесе статикасын бейнелейтін диаграмманы *бағаналы* деп атайды (1.5-сурет).



1.5 сурет. Бағаналы диаграммалар

- Құбыластардың құрылымын айқын экстенсивті көрсеткіштермен және меншікті салмағына сәйкес оның құрамды бөліктерінің түспен бөлуін бейнелейтін тік төртбұрыш түрдегі диаграмманы *бағанаішілік* деп атайды (1.6 сурет).



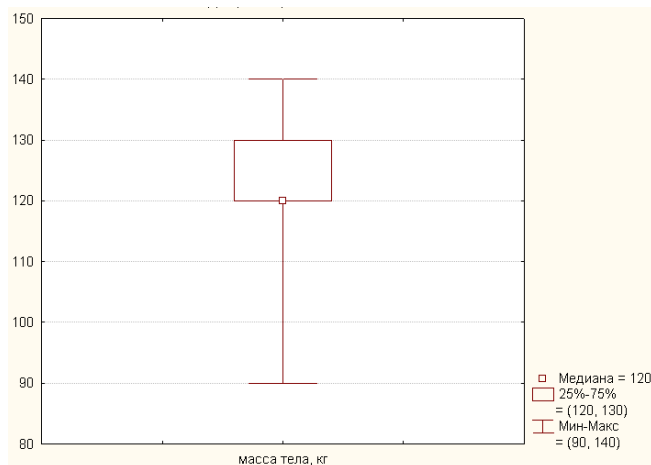
1.6 сурет. Бағанаішілік диаграмма

- Берілгендерді шамаларының өсуіне қарай кескіндеу үшін және диаграмма мен кестеден тұратын ыңғайлы сызбаны «*өркен және жапырақтар*» сызбасы деп атайды. (1.7 сурет).

3	1,0	04
665	1,1	39
53	1,2	99
9751	1,3	1135677999
955410	1,4	0148
987665	1,5	00338899
9531100	1,6	0001355
731	1,7	00114569
99843110	1,8	6
654400	1,9	01
6	2,0	
7	2,1	19
10	2,2	

1.7 сурет. «Өркен және жапырақтар» сызбасы.

- Екі параллель жақтары берілгендердің жоғарғы және төменгі кватилине сәйкес келетін, ал тік төртбұрыштың соңынан басталатын сызық ең аз және ең көп мәнін көрсететін тік төртбұрыш түрдегі сызбаны «*мұрттары бар жәшік*» сызбасы деп атайды (1.8 сурет).



1.8 сурет. «Мұрттары бар жәшік» сызбасы.

Өлшеу – нысандарды белгілі бір көрсеткіштері немесе сипатына (белгілері, атрибуттары) қарай салыстыру рәсімі.

Шкала – міндетті, қажетті өлшеу рәсімінің элементі.

Медициналық-биологиялық зерттеулерде қолданылатын өлшеу шкалаларының негізгі типтері:

- *номиналды* немесе *атаулар шкаласы* нысанның қасиеттерін классификациялауда, оларға сандық, әріптік және басқа символдық сипаттамалар беру үшін (жынысы, ұлты, көзінің, шашының түсі, диагнозы т.б.) қолданылады.
- *реттік* немесе *шендік шкаласы* – белгілердің шамасына қарай (Мясников шкаласы бойынша гипертониялық аурулардың кезеңі, Стражеско-Василенко-Лангу шкаласы бойынша жүрек талмасының дәрежесі, Фогельсон бойынша жүректің жеткіліксіздігінің айқындық дәрежесі шкаласы және т.б.) реттейді.
- *аралық шкаласы* – белгілердің жеке өлшемдерінің ауытқуын (уақыт, температура, тест ұпайлары) көрсетеді.
- *қарым-қатынас шкаласы* – өлшенген белгілердің қатынасын (бойы, салмағы, реакция мерзімі, тесттегі орындалған тапсырмалар саны) көрсетеді.

Өлшеу барысында оның *сенімділігі* және *ақиқаттылығы* бойынша сұрақ туындайды.

Өлшеу *сенімділігі* өлшеудің:

- дұрыстығына (шкала дұрыс таңдалды ма, көрсеткіш дұрыс жазылды ма, жүйелік қателіктер ескерілді ме және т.с.с);
- орнықтылығына (қайта өлшеу барысында мәліметтер сәйкес келе ме);
- қисындылығына (өлшеу керек нысанның дәл өзі ме, әлде оған ұқсас түрі алынды ма) тәуелді болады.

Өлшеу *ақиқаттылығы* нақты жағдайдағы шамаға салыстырғанда өлшеу дәлдігін сипаттайды.

Ақиқаттылыққа тексерудің басты бағыты, ақпараттарды әртүрлі ақпарат таратушы көздерден алуына негізделген.

4. Иллюстрациялық материал: көрме, слайдтар

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.
2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казакстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені		№ 35-11(Б)-2024 60 беттің 9 беті

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013
2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)
2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чуудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы: Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары:

1. «Биостатистика» деген не?
2. Биометриканьң дамуындағы Ф.Гальтон, К.Пирсон, Р.Фишер ғалымдарының ролі қандай?
3. Статистикалық берілгендер жинау деген не?
4. Статистикалық берілгендердің қандай типтерін білесіздер?

Медициналық-биологиялық зерттеулерде өлшну шкаласының қандай типтері қолданылады?

№2 дәріс

1. Тақырыбы: Сипаттамалы статистика.

2. Мақсаты: Студенттерді бас және таңдаулы жиынтық түсінігімен, сол сияқты жиынтықтар параметрлерін бағалау рәсімімен таныстыру

3. Дәріс тезистері: *Статистикалық жиынтық* – уақытпен және кеңістікпен шектелген қандайда бір нысан белгісінің бірдей жиынтығы (елдегі белгілі бір жылда туылған балалар саны; бір қаладағы тұрғындар саны; берілген елдегі қатерлі ісікпен ауырғандар саны және т.б.).

Медициналық-биологиялық, клиникалық, фармацевтикалық және басқа да зерттеулерде, зерттеуші тәжірибесінде ешқашан толық топ нысанын алып, оған жаппай зерттеу жүргізу мүмкін емес, сондықтан *таңдамалы әдісті* қолданады.

Таңдамалы әдіс – статистикалық жиынтықтағы белгілі бір шекті нысанды таңдап және оны зерттеудің статистикалық тексеру әдісі.

Таңдамалы әдіс медицинада, денсаулық сақтауда және фармацияда кеңінен қолданыс табуда. Мысалы: белгілі бір патологиялық өзгерісі бар барлық ауруларды зерттеу мүмкін емес, сондықтан олардың кейбіреулерін ғана зерттейді; барлық дәрілік препараттардың стандартқа сай болуын тексеру мүмкін емес, сондықтан оларға таңдамалы бақылау жүргізеді және т.б.

Бас статистикалық жиынтық – бұл шексіз бірліктердің үлкен сандарынан тұратын элементтердің жиынтығы. Мысалы: белгілі бір патологиялық өзгерісі бар барлық аурулар, сол аймақтың барлық тұрғындары және т.б.

Таңдамалы жиынтық (таңдау) – бұл бақылаудың таңдап алынған белгілі бір бөлігінен тұратын, бұл бөлек бас жиынтықты сипаттай алатын элементтердің жиынтығы.

Жиынтық көлемі – бұл бақылау элементтерінің жалпы саны. Бас жиынтықтың көлемі «N», ал таңдамалы жиынтықтың көлемі – «n» мен белгіленеді. Егер $n \leq 30$ болса, онда таңдаманы кіші деп есептеледі.

Бақылаудың элементі - бұл зерттелетін құбылыстың жеке дербес жағдайы.

Таңдамалы зерттеу әдісі бас жиынтық шексіз немесе зерттеу нысандарды жою

жағдайына байланысты болғанда (мысалы, дәрілік препараттарды тексеру) жалғыз мүмкіндігі бар әдіс болып табылады. Сонымен қатар, ол қор шығындарын едеуір үнемдеуге мүмкіндік береді. Бұл әдістің кемшілігі тек нысанның бір бөлігі зерттелетіндіктен қателіктің пайда болуында.

Таңдаманың негізгі қасиеті – *үлгілігі (репрезентативтілігі)*, яғни таңдама сипатамасының бас жиынтықтың сипаттамасына толықтай сәйкес келуі. Ол таңдау арқылы зерттеу нәтижелері қаншалықты таңдау жүргізілген бас жиынтықты сипаттауға қолдануға болатындығын анықтайды.

Үлгілі таңдаманың түрі:

- *сапалық* – бұл бас және таңдамалы жиынтықтағы бақылау элементтері белгілерінің сәйкес келуі.
- *сандық* – бұл жеткілікті бақылау саны.

Жүйелік материалдардың статистикалық бақылауының бірінші қадамы *статистикалық үлестірімнің таңдамасын* анықтау болып табылады.

Статистикалық үлестірімнің таңдамасы (немесе *вариациялық қатар*) екі бағаннан тұратын кестеден тұрады (1 кесте).

Бірінші бағанда *нұсқа* деп аталатын вариациялық белгінің мәні жазылады және ол « x_i » белгіленеді, ал екінші бағанда *жиілік* деп аталатын және « v_i » белгіленген, нұсқаның қанша рет кездескенін көрсететін сандар жазылады.

1 кесте

Нұсқалары (x_i)	Жиіліктері (v_i)
...	...
...	...
...	...
Барлығы:	$n = \sum v_i$

Вариацияланатын белгілер дискретті және үздіксіз сипатта болуы мүмкін.

Бүтін сандар болатын белгі нұсқаларын *дискретті* деп атайды. Мысалы, отбасындағы балалар саны, науқастар саны, қораптағы ампула саны және т.б.

Егер вариациялық қатардағы нұсқалар дискретті шама түрінде берілсе, онда бұндай вариациялық қатарды *дискретті* деп атайды.

1 мысал. Жеке зерттелген тетрациклин гидрохлориді белсенділігінің алынған мәндері x_i (в ED/мг): 925, 940, 760, 905, 995, 965, 940, 925, 940, 940, 905.

Мәндердің белсенділігін және жиілікті өсу ретімен орналастыра отырып, дискретті вариациялық қатар аламыз (2 кесте):

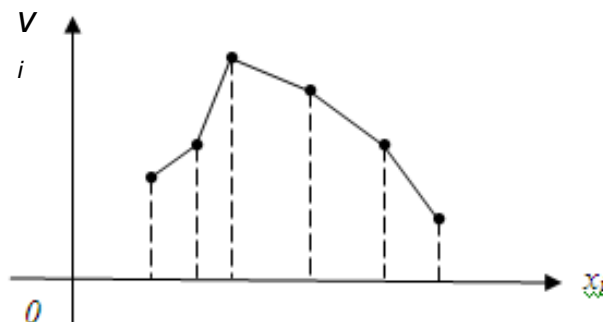
2 кесте

Нұсқалары (x_i)	Жиіліктері (v_i)
760	1
905	2
925	2
940	3
965	1
995	1
Барлығы:	$n = 10$

Дискретті вариациялық қатарды *полигон* деп аталатын көпбұрышты сызбалық түрінде көрсетуге болады (1 сурет).

Белгілі-бір аралықтағы кез-келген мәнді қабылдай алатын белгілер нұсқасы *үздіксіз* деп аталады. Мысалы, артериалді қан қысым, бойы, салмағы, жалақысы және т.б.

Үздіксіз белгілер үшін *аралық* вариациялық қатар құрылады. Берілген аралықағы белгінің ең аз мәні - төменгі аралық шегі, ал ең көп мәні - жоғарғы аралық шегі деп аталады.

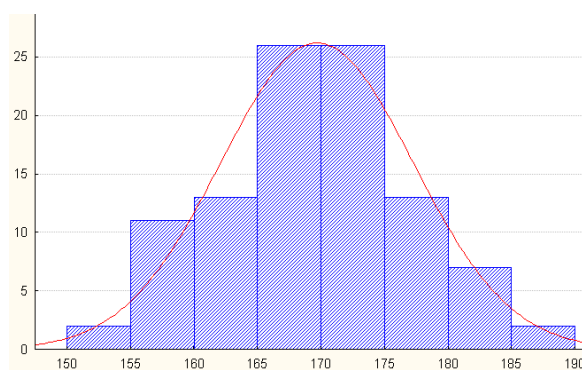


1 сурет. Полигон

2 мысал. Ер адамның (x_i) бойынның өсуінің **аралық** вариациялық қатары.

Нұсқалары (x_i), см	Жиіліктері (v_i)
150-155	1
155-160	11
160-165	14
165-170	26
170-175	26
175-180	13
180-185	8
185-190	1

Аралық вариациялық қатарды гистограмма деп аталатын сатылы тұрпат түрдегі сызбамен көрсетуге болады. (2 сурет).



2 сурет. Гистограмма.

Аралық вариациялық қатар бақылау саны үлкен болған кезде дискретті белгілер үшін де құрылады.

Вариациялық қатар *орталық үрдіс (тенденция) көрсеткіштерімен және әр түрлілік (разнообразия) көрсеткіштерімен* сипатталады.

Орталық үрдіс көрсеткіштеріне орташа және құрылымды шамалар жатады.

Орташа шамалар:

- Қарапайым орташа арифметикалық шама $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$,

ұндағы n - қатардың жалпы саны;

- Өлшенген орташа арифметикалық шама $\bar{x} = \frac{\sum_{i=1}^n x_i v_i}{\sum_{i=1}^n v_i}$,

мұндағы v_i – жиіліктер;

- Қарапайым орташа геометриялық шама $\bar{x} = \sqrt[n]{x_1 \cdot x_2 \dots x_n}$

- Қарапайым орташа квадраттық шама $\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$

- Өлшенген орташа квадраттық шама $\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2 v_i}{\sum_{i=1}^n v_i}}$

3 мысал. Келесі берілгендер бойынша 18 жастағы ер балалардың орташа салмағын анықтау керек:

Салмағы, x_i (кг)	64	63	62	61	60	59
---------------------	----	----	----	----	----	----

Шешуі: $\sum x_i = 64 + 63 + 62 + 61 + 60 + 59 = 369$, $n = 6$, $\bar{x} = \frac{369}{6} = 61,5$ - қарапайым арифметикалық орташа.

4 мысал. Келесі берілгендер бойынша 18 жастағы ер балалардың орташа салмағын анықтау керек:

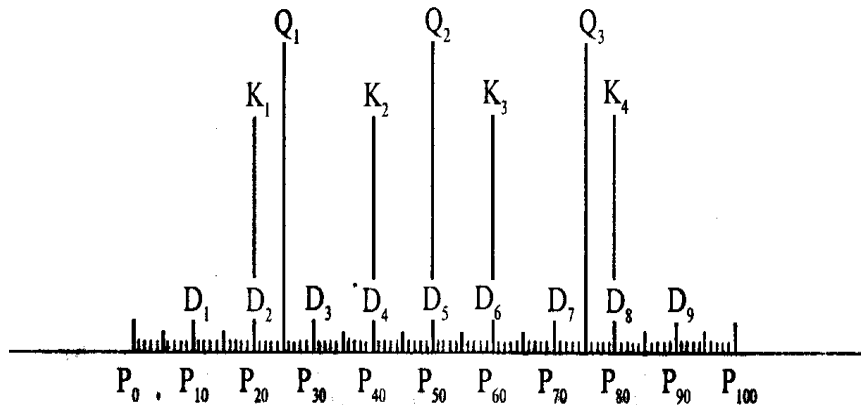
Салмағы, x_i (кг)	64	63	62	61	60	59
Бақыланғандар саны, v_i	2	3	9	6	4	1

Шешуі: $\sum x_i v_i = 64 \cdot 2 + 63 \cdot 3 + 62 \cdot 9 + 61 \cdot 6 + 60 \cdot 4 + 59 \cdot 1 = 1540$,

$\sum v_i = 2 + 3 + 9 + 6 + 4 + 1 = 25$, $\bar{x} = \frac{1540}{25} = 61,6$ - өлшенген арифметикалық орташа.

Құрылымды шамалар:

- *мода (Mo)* – ең үлкен жиіліктегі нұсқа;
- *медиана (Me)* – қатар ортасында орналасқан нұсқа;
- *квантилдер* – вариациялық қатарды теңдей жеке бөліктерге бөлетін шама:
 - *квартилдер* – вариациялық қатарды теңдей төрт бөлікке бөлетін шама;
 - *квинтилдер* - вариациялық қатарды теңдей бес бөлікке бөлетін шама;
 - *децилдер* - вариациялық қатарды теңдей он бөлікке бөлетін шама;
 - *процентилдер* - вариациялық қатарды теңдей жүз бөлікке бөлетін шама (3 сурет).



3 сурет. Вариациялық қатардың құрылымдық сипаттамасы.

Төменгі квартиль (Q_1) немесе 25-ші процентиль (P_{25}) – бұл таңдаманың 25%-пайызынан төмен жататын кездейсоқ шаманың мәні.

Төменгі квартилдің нөмірін анықтайтын формула: $N_{Q_1} = \frac{n+1}{4}$.

Жоғарғы квартиль (Q_3) немесе 75-ші процентиль (P_{75}) – бұл таңдаманың 25%-пайызынан жоғары жататын кездейсоқ шаманың мәні.

Жоғарғы квартилдің нөмірін анықтайтын формула: $N_{Q_3} = \frac{3(n+1)}{4}$.

Егер квартилдің нөмірі бөлшек болса, онда оны жақын жатқан бүтін санға.

5 мысал. Мода және медиананы келесі берілгендер бойынша анықтау керек:

Салмағы, x_i (кг)	64	63	62	61	60	59
Бақыланғандар саны, v_i	2	3	9	6	4	1

Шешуі: $M_O=62$, $M_E=61,5$.

6 мысал. Мода және медиананы келесі берілгендер бойынша анықтау керек:

1 мин. тамырдың соғу жиілігі, x_i	66	68	69	70	72	73	76
Бақыланғандар саны, v_i	1	2	3	5	3	1	1

Шешуі: $M_O=70$, $M_E=70$.

7 мысал. Жеке зерттелген тетрациклин гидрохлориді белсенділігінің алынған мәндері x_i (ED/m_2): 925, 940, 760, 905, 995, 965, 940, 925, 940, 940, 905. Квартилдерді анықтау керек.

Шешуі: Таңдаманы реттеу керек:

760, 905, **905**, 925, 925, **940**, 940, 940, **940**, 965, 995

$$n = 11, \quad N_{Q_1} = \frac{n+1}{4} = \frac{11+1}{4} = 3, \quad Q_1 = 905, \quad N_{Q_3} = \frac{3(n+1)}{4} = \frac{3(11+1)}{4} = 9, \quad Q_3 = 940,$$

$$N_{Q_2} = \frac{n+1}{2} = \frac{11+1}{2} = 6, \quad Q_2 = M_E = 940,$$

Әр түрлілік көрсеткішіне жататындар:

- *вариациялық қатардың құлашы* - вариациялық қатардың ең үлкен және ең кіші мәндерінің арасындағы айырым: $R = x_{max} - x_{min}$, мұндағы x_{max} , x_{min} – ең үлкен және ең кіші мәндерге байланысты нұсқа;
- *дисперсия* – кездейсоқ шаманың оның орташа мәнінен шашырау өлшемі:
- егер таңдама вариациялық қатар түрінде берілсе, онда дисперсия

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \text{ формуласымен анықталады;}$$

- егер таңдама кесте түрінде берілсе, онда дисперсия

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot v_i}{\sum_{i=1}^n v_i} \text{ формуласымен анықталады;}$$

- *орташа квадраттық ауытқу* – кездейсоқ шаманың нұсқадағы өлшем берлігімен берілген, оның орташа мәнінен шашырау шарасы:

$$S = \sqrt{S^2};$$

- *вариация коэффициенті* - кездейсоқ шаманың пайызбен берілген шашырау шарасы:

$$V = \frac{S}{\bar{x}} \cdot 100\%$$

Егер $V \leq 33\%$ болса, онда таңдама біртекті деп есептеледі.

8 мысал. Вариациялық қатар үшін әр түрлілік көрсеткіштерін анықтау керек.

САҚ, мм сын. бағ.	160	165	170	175	180
Бакыланатын адам саны	4	6	20	12	5

Шешуі.

1. Таңдамалы орташа: $\bar{x} = \frac{\sum_{i=1}^n x_i v_i}{\sum_{i=1}^n v_i} = \frac{8030}{47} \approx 171.$

2. Таңдамалы дисперсия: $S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot v_i}{\sum_{i=1}^n v_i} = \frac{1317}{47} \approx 28.$

3. Таңдамалы орташа квадраттық ауытқу: $S = \sqrt{S^2} = \sqrt{28} \approx 5,3.$

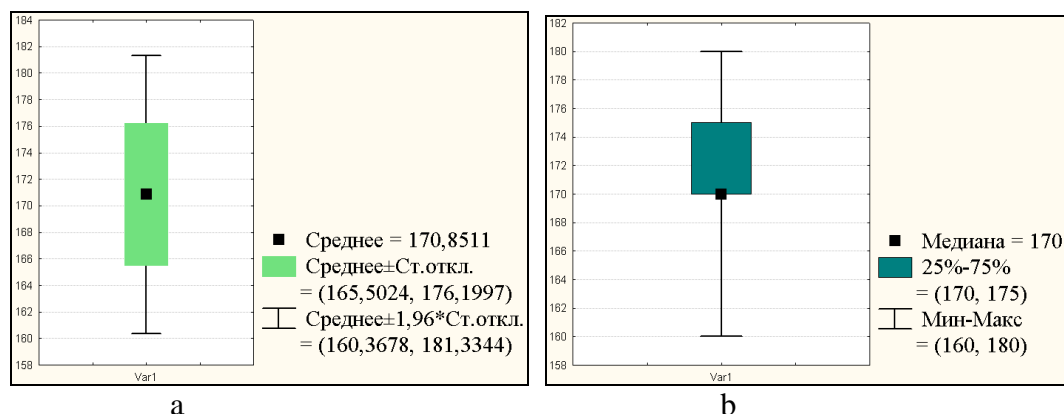
4. Вариация коэффициенті: $V = \frac{S}{\bar{x}} \cdot 100\% = \frac{5,3}{171} \cdot 100\% \approx 3,1\%.$

5. Вариациялық қатардың құлашы: $R = x_{\max} - x_{\min} = 180 - 160 = 20.$

x_i	v_i	$x_i \cdot v_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot v_i$
160	4	640	-11	121	484
165	6	990	-6	36	216
170	20	3400	-1	1	20
175	12	2100	4	16	192
180	5	900	9	81	405
Қосындысы	47	8030		255	1317

Статистикалық сипаттамаларды «мұртты жәшік» сызбасы арқылы көрсету ыңғайлы. 8-ші мысалдың шарттары үшін орындалған сызбалар (сурет 4 а, b).

Бұндай сызбаларды талдаған кезде міндетті түрде «легендағы» яғни сызбаның төменгі жағында берілген шартты белгілерге көңіл аудару керек.



4. сурет.

Бірінші сызбада (сурет 4 а) орташа, ең аз және ең көп мәндер, сонымен қатар стандартты ауытқу келтірілген. Екінші сызбада (сурет 4 b) медиана мәні, 25-ші және 75-ші процентиллер келтірілген.

Бас жинақтың сипаттамаларын (« X » - бас орташа, « D » - бас дисперсия, « σ » - орташа квадраттық ауытқу) *параметрлер* деп атайды.

Әдетте параметрлер белгісіз және оларды берілген таңдамалар негізінде жуықтап бағалауға болады. Бұл жуық мәндер *бас жинақ параметрлерінің бағасы* деп аталады.

« X » бас орташаның бағасы, « \bar{x} » таңдаманың орташасы болып табылады.

Зерттелетін белгінің өзінің « \bar{x} » орташа мәнінің айналасында шашырауын сипаттау үшін « S^2 » - таңдамалы дисперсия деп аталатын сипаттама енгізіледі.

« D » - бас дисперсияның бағасы болып « s^2 » - түзетілген таңдамалы дисперсия табылады.

Түзетілген таңдамалы дисперсия $s^2 = \frac{n}{n-1} S^2$ формуласымен анықталады.

« σ » - бас жинақтың орташа квадраттық ауытқуының бағасы болып, « s » - орташа квадраттық ауытқудың түзетілген таңдамасы табылады.

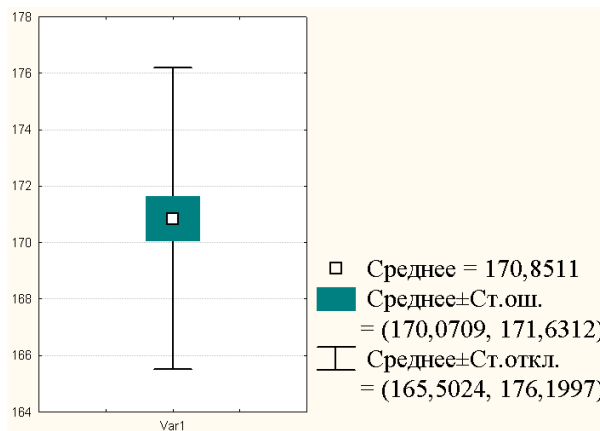
Орташа квадраттық ауытқудың түзетілген таңдамасы $s = \sqrt{s^2} = \sqrt{\frac{n}{n-1} S^2}$

формуласымен анықталады.

« $S_{\bar{x}}$ » - орташа қателік немесе орташа квадраттық қателік немесе стандартты

орташа қателік деп аталын шама $S_{\bar{x}} = \frac{s}{\sqrt{n}}$ формуласымен анықталады. Бұл шама « n »

таңдама көлемімен есептелінген бас жинақтың таңдама орташасының стандартты ауытқуын сипаттайды.



8-ші мысалдың шарттары үшін орындалған сызба:

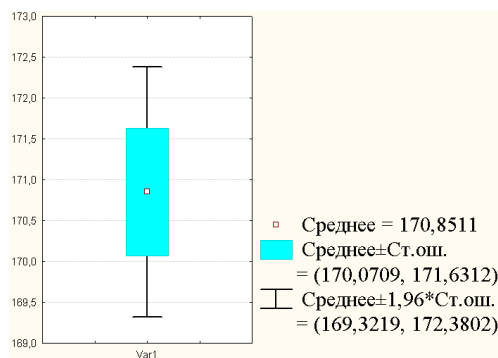
Жеке параметрлердің кейбір бағалауы *нүктелік бағалауды* береді.

Бас жинақтың параметрлерінің *аралық бағалауы* деп « γ » ықтималдылығымен берілген параметрді екі жағынан жабатын нақты мәнінің аралығын айтады.

Аралық бағалауды *сенімділік аралығы*, ал соған байланысты « γ » ықтималдықты *сенімділік ықтималдылығы* немесе *сенімділік* (медицина және биологияда $\gamma=0,95$) деп атайды.

Бас орташа үшін сенімділік аралығы $\bar{x} - t \cdot S_x \leq X \leq \bar{x} + t \cdot S_x$ арақатынасынан алынуы мүмкін. Мұндағы « n » бақылаудағы \bar{x} - таңдамалы орташа, $t(\gamma; n-1)$ – « γ » және « n »-ге тәуелді кестелік шама, S_x - орташаның стандартты қателігі

Кез келген медициналық-биологиялық зертеулердің ең бір маңызды кезеңдерінің бірі - бұл қарастырылып отырған таңдама көлемін анықтау. *Таңдама көлемін анықтау* – медициналық-биологиялық зертеулердің маңызды кезеңі.



8-ші мысалдың шарттары үшін орындалған сызба.

Егер таңдама көлемі жеткіліксіз болса, онда таңдама сипаттамаларының қателіктерінің артуына және дұрыс емес қорытынды жасауға алып келеді.

Таңдама көлемі зерттелетін шаманың орташа квадраттық ауытқуына (σ), қолданылатын белгінің қуатына және маңыздылық деңгейіне (p). тәуелді болады.

Таңдама көлемін есептейтін формула: $n = \left(\frac{t \cdot \sigma}{\varepsilon} \right)^2$, мұндағы n – таңдама көлемі, σ – орташа квадраттық ауытқу, $t = 1,96$ – $p = 0,05$ үшін стандартты бірқалыпты үлестірімнің

ОҢТҮСТІК-ҚАЗАҚСТАН MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені		№ 35-11(Б)-2024 60 беттің 17 беті

кестелік мәні, $\varepsilon = 0,01$ – МЕМСТ-қа сәйкес алынған баға дәлдігі.

9 мысал. Таңдамалы тексеру үш апта ішінде жаңа препаратты қолдану вирустық жүктемені төмендететінін көрсетті.

Тексерілетін топтағы адамдардың вирустық жүктемесінің орташа квадраттық ауытқуы 0,3-ке (\log_{10} МЕ/мл) тең. Вирустық жүктемені 0,01 ($p=0,05$ болғанда) дәлдікке дейін өлшеу үшін зерттеуге қанша адам қосу керек?

Шешуі.

$$n = \left(\frac{t \cdot \sigma}{\varepsilon} \right)^2 = \left(\frac{1,96 \cdot 0,3}{0,01} \right)^2 \approx 3457.$$

Егер оқиға болатын оқиғаның ықтималдығы « p » белгісіз тәуелсіз зерттеу жүргізілсе, онда таңдама көлемін анықтайтын формула: $n = \frac{t^2}{\varepsilon^2} \tilde{p}(1 - \tilde{p})$, мұндағы $\tilde{p} = \frac{m}{n}$ - « p » ықтималдықтың бағасы, m – оқиғаның пайда болу саны, « n » - зерттеудің жалпы саны.

Таңдама көлемі « n » үлкен болған сайын зерттеу нәтижесі дәлірек болады.

10 мысал. Таңдамалы тексеру дәрілік заттың 100 ампуласының 87-сі стандартқа сай келетінін көрсетті. Біз стандартқа сай емес бұйымды пайыздық бағалауымыздың дұрыстығына 95%, сенімді болғымыз келеді. Бұл пайыз қандай шектерде жатыр? Ақаудың пайызын 0,01 дәлдікке дейін бағалағанда, таңдаманың көлемі қандай болу керек (егер $p=0,05$ болса)?

Шешуі.

$$\tilde{p} = \frac{m}{n} = \frac{87}{100} = 0,87, \quad n = \frac{t^2}{\varepsilon^2} \tilde{p}(1 - \tilde{p}) = \frac{1,96^2}{0,01^2} \cdot 0,87 \cdot (1 - 0,87) \approx 4345.$$

4. Иллюстрациялық материал: көрме, слайдтар

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.
2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013
2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)
2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы: Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары

1. «Биостатистика» деген не?
2. Биометриканың дамуындағы Ф.Гальтон, К.Пирсон, Р.Фишер ғалымдарының ролі қандай?
3. Статистикалық берілгендер жинау деген не?
4. Статистикалық берілгендердің қандай типтерін білесіздер?
5. Медициналық-биологиялық зерттеулерде өлшну шкаласының қандай типтері

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 18 беті

қолданылады?

№3 дәріс

1. Тақырыбы: Статистикалық болжамды тексерудің негізгі теориясы. Кездейсоқ шамалардың бірқалыпты үлестірімінің болжамын тексеру.

2. Мақсаты: Студенттерді статистикалық болжамды тексерудің негізгі теориясымен және Стьюденттің t-белгісімен таныстыру.

3. Дәріс тезистері:

1. Қолданбалы есептерде бақыланған таңдамалар бойынша зерттеу жүргізушіні қызықтыратын таңдама алынған бас жинақтың сипатына қатысты кейбір пікірді **болжам** деп айтуға болады. Яғни, *статистикалық болжамды тексеру* туралы мәселе айтылады.

Кездейсоқ шамалардың белгісіз үлестірім заңының түрі туралы кейбір болжамдарды (параметрлік емес) немесе белгілі үлестірімнің параметрлері туралы алдынала шарт ретінде ұсынған түсіндіруді (параметрлік) *болжам* деп атайды.

Статистикалық болжамды тексеру теориясы сезіну арқылы (интуиция) емес, дәлелді медицинаның негізгі құралы болып табылады.

Медициналық және биологиялық зерттеулердегі есептерді шешу үшін статистикалық болжамдар түзу қажет:

- зерттелетін топта қандай да бір заңға (қалыпты заңға сәйкес келетін үлестірімді талдау) сәйкес келетін үлестірімді талдау;
- белгінің үлестірім параметрлері бойынша (орташа мәндері, дисперсиясы бойынша) топтарды салыстыру.

1 мысал. Тұмаумен ауырған екі біркелкі топқа: біріншісі - «А», ал екіншісі «В»-ға дәрілік құрал егілген, олардың орташа сауығу уақыты әртүрлі. Бұл жағдай тұмауға қарсы құралдың біреуінің екіншісіне қарағанда әсерінің артық болуынан ба, немесе айырмашылық кездейсоқ анықталды ма?

Болжам: топтардағы сауығудың орташа уақыты бірдей.

2 мысал. Жаңа гипогликемикалық препараттың әсерін бағалау үшін сусамыр ауруымен ауыратын науқастың пепаратты қабылдағанға дейінгі және кейінгі қан құрамындағы глюкозаның мөлшеріне өлшеу жүргізді. Препаратты қабылдағаннан кейін қанындағы глюкозаның деңгейі төмендейді деп есептеуге бола ма?

Болжам: науқастың қанындағы глюкозаның препаратты қабылдағанға дейінгі және кейінгі орташа денгейі бірдей.

Мысал 3. Шоктың бастапқы сатысында орташа артериялық қысым (мм сын. бағ.) зерттелді. 50 адам бақыланды. Бас жиынтық бір қалыпты үлестірім заңына бағынады деп есептеуге бола ма?

Болжам: бас жиынтық бір қалыпты үлестірілген.

Болжам – бұл кездейсоқ шамалардың белгісіз үлестірім заңының түрі туралы жорамалдар (параметрлік емес) немесе белгілі үлестірімнің параметрлері туралы алдынала шарт ретінде ұсынылған түсіндірулер (параметрлік).

Әр қашан екі статистикалық болжам ұсынылады:

- H_0 нөлдік болжам – екі топ арасында айырмашылық болмауының, немесе параметрлердің нақты мәні туралы, немесе үлестірімнің қалыпты заңға сәйкестігі туралы болжам.
- H_1 баламалы болжам – екі топ арасында айырмашылық болуының, немесе параметрлердің нақты мәнінің айырмашылығы туралы, немесе үлестірімнің қалыпты заңға сәйкес еместігі туралы болжам.

Негізі нөлдік болжам зерттеу жүргізуге себепші болған, зерттелетін (медициналық, биологиялық) болжамға қарама-қарсы түрде қалыптасады.

Нөлдік болжамды тексеру үшін статистикалық әдістер (тесттер, белгілер) қолданылады.

Статистика – бұл таңдалған бақылаулардың негізінде нөлдік болжам қабылданатын немесе қабылданбайтын функция.

Статистикалық белгі деп қарастырылып отырған болжамның тәжірибеде алынған мәндерге сәйкестігің немесе сәйкес еместігін анықтайтын ережені айтады.

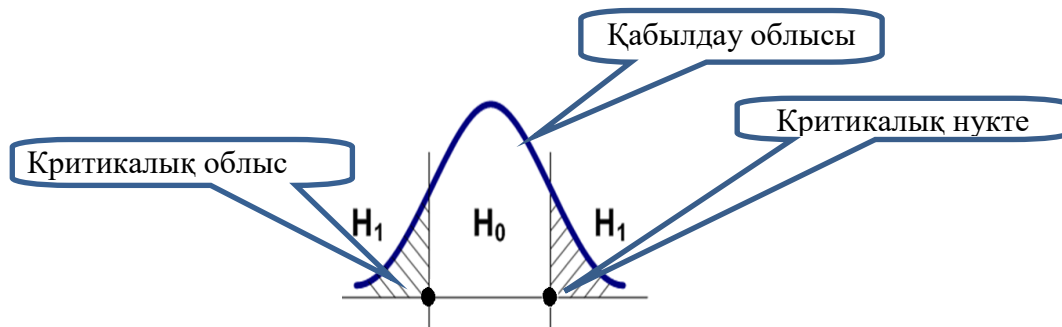
Статистикалық белгі – бұл кең түрде қолданылатын статистикалық құрал.

Бақыланатын мәнінің белгісі – үлестірім заңына бағынатын таңдамалы жиынтық бойынша есептелетін мән.

Негізгі болжам қабылданатын статистикалық белгінің көптеген мүмкін болатын мәндерін *қабылдау облысы* деп атайды.

Негізгі болжам қабылданабайтын статистикалық белгінің көптеген мүмкін болатын мәндерін *критикалық облыс* деп атайды.

Болжамды қабылдау және критикалық облысты бөліп тұратын нүктені *критикалық нүкте* деп атайды.



Статистикалық болжамдарды тексеру кезінде қателіктің келесі түрлері орын алады:

- *1-ші ретті қателік* – бұл дұрыс нөлдік болжамды қабылдамаудың ықтималдылығы.
- *2-ші ретті қателік* – бұл дұрыс емес нөлдік болжамды қабылдаудың ықтималдылығы.

Маңыздылық деңгейі – бұл зерттеуші үшін дұрыс нөлдік болжамды қате деп қабылдамаудың ең тиімді ықтималдығы, яғни зерттеушіге ұсынылатын бірінші ретті қателіктің шамасы.

Фармация, медицина және биологиялық зерттеулерде маңыздылық деңгейінің шамасын $p=0,05$ -ке тең деп алады. Стандарттар үшін маңыздылық деңгейінің шамасын $0,01$ -ке тең деп алады.

Маңыздылық деңгейі немесе бірінші ретті қатенің ықтималдығын « p » деп, ал екінші ретті қатенің ықтималдығын « γ » деп белгілейді.

Сенімділік ықтималдылығы (γ) – бұл бірінші ретті қатені болдырмайтын ықтималдылық және дұрыс болжамды H_0 ($\gamma=1-p$) деп қабылдау керек.

Кез келген статистикалық белгінің негізгі қасиеті оның *қуаттылығы*. Белгінің *қуаттылығы* деп оның жалған болжамды шығарып тастау қабілетін айтады.

Қуаттылық $1-\gamma$ ықтималдылығымен бағаланады, мұндағы γ – екінші ретті ықтималдылық.

Статистикалық болжамды тексерудің тізбесі:

- 1) Екі болжам ұсынылады: негізгі (нөлдік) H_0 және H_1 баламалы (бәсекелес).
- 2) « p » маңыздылық деңгейі беріледі. Статистикалық қорытынды ешқашан жүз

пайыздық сенімділікпен жасалмайды. Барлық уақытта дұрыс емес шешім қабылдау тәуекелділігі болуы мүмкін.

Статистикалық болжамды тексеруде мұндай тәуекелділіктің өлшемі «р» маңыздылық деңгейі болып табылады.

3) Берілгендер, яғни таңдама бойынша бақыланатын (эмперикалық, есептелінетін) белгінің мәні есептеледі.

4) Арнайы статистикалық кестеден белгінің кестелік (критикалық) мәндері анықталады.

5) Бақыланған және кестелік мәндерді салыстыру арқылы сол және басқа болжамдардың дұрыстығы туралы қорытынды жасалады.

Биостатистикада кездейсоқ шамалардың үлестірім түрі туралы болжамдар жиі қолданылады.

Көптеген биологиялық және медициналық көрсеткіштер (физикалық даму, қан плазмасын түзетін құрылымдар және т.б.), сонымен қатар оларды өлшеу кезінде жіберілген қателіктер қалыпты үлестірімге бағынады.

Сондықтан қалыпты үлестірілген кездейсоқ шамалардың параметрлері туралы болжамды тексере білеуді үйрену қажет.

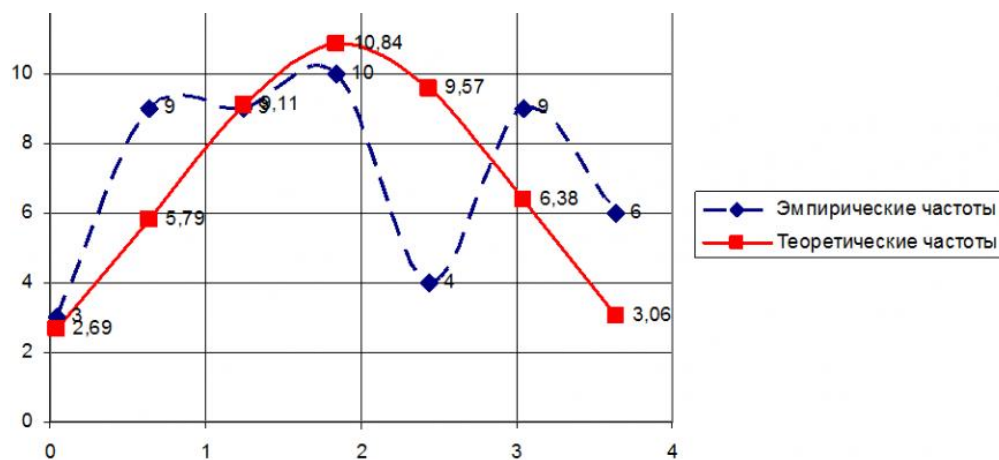
Осы және басқада үлестірулердің сипаттамасы туралы барлық жорамалдар – болжам болады. Сондықтан олар *келісім белгісінің* көмегімен статистикалық тексерілуі керек. Бұл белгілер теориялық және эмперикалық жиіліктер арасындағы алшақтықтың мағынасы жоқ, яғни кездейсоқ, ал мағынасы бар кезде, яғни кездейсоқ емес деген байланысты анықтауға мүмкіндік береді.

Осылайша келісім белгілері эмперикалық қатардағы үлестірім сипаты туралы болжамды түзетудің дұрыстығын теріске шығаруға немесе растауға мүмкіндік береді.

Келісім белгілерінің ішіндегі кең тарағандары χ^2 -Пирсон және Колмогоров–Смирнов белгілері болып табылады.

Бұл белгілер екі жағдайда қолданылады:

- белгілердің тәжірибелік үлестірімі мен теориялық үлестірімін салыстыру үшін (бір қалыпты, көрсеткіштік, біркелкі және т.б.) (6.1. сурет);
- бір белгінің екі тәжірибелік үлестірімін салыстыру үшін.



Эмперикалық және теориялық жиіліктерді салыстыру.

5. χ^2 –Пирсонның келісім белгісін қолданудың тізбесі.

1) H_0 : «X» кездейсоқ шама $F(x)$ үлестірім функциясы болады.

H_1 : «X» кездейсоқ шама $F(x)$ үлестірім функциясы болмайды.

1) $p=0,05$ - маңыздылық деңгейі.

$$2) \chi^2_{есен} = \sum_{i=1}^k \frac{(v_i - v_i^*)^2}{v_i^*} \quad (6.1)$$

мұндағы, k – эмпирикалық бөлінуге бөлінген топтар саны, v_i – i тобындағы бақылау жиілігінің белгісі, v_i^* - теориялық жиілік.

3) $\chi^2_{кесте}(p, f)$,

мұндағы $f = k - 1 - r$ - еркіндік дәрежелерінің саны (кестелік мәні), k - таңдама тобының саны, r – болжалатын үлестірім параметрлерінің саны (бір қалыпты үлестерім үшін $r=2$).

4) Егер $\chi^2_{есен} \leq \chi^2_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $\chi^2_{есен} > \chi^2_{кесте}$ болса, онда « H_0 » қабылданбайды

Пирсонның келісім белгісі жиынның көлемі үлкен ($n > 30$) болса қолданылады, бұл жағдайда әр топтың жиілігі бестен кем болмау керек.

6. Колмогоров–Смирнов келісім белгісін қолданудың тізбесі.

1) H_0 : « X » кездейсоқ шама $F(x)$ үлестірім функциясы болады.

H_1 : « X » кездейсоқ шама $F(x)$ үлестірім функциясы болмайды.

2) $p=0,05$ - маңыздылық деңгейі.

$$3) \lambda_{есен} = d_{\max} \sqrt{n}, \quad (6.2)$$

мұндағы $d_{\max} = \max |F_n(x) - F(x)|$ - бақыланатын « $F_n(x)$ » және теориялық « $F(x)$ » үлестірім функцияларының абсолютті шамасының ең үлкен мәнінің айырымы, n – статистикалық қатардағы бақыланатын сан.

4) $\lambda_{кесте} = 1,36$ ($p=0,05$ болғандағы кестелік мәні).

5) Егер $\lambda_{есен} \leq \lambda_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $\lambda_{есен} > \lambda_{кесте}$ болса, онда « H_0 » қабылданбайды.

Колмогоров-Смирнов белгісі бақылау саны үлкен ($n > 30$) болғанда қолданылады.

4. Иллюстрациялық материал: көрме, слайдтар

5. Әдебиет:

• Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.

2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

• Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013

2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

• Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 22 беті

2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары

1. Статистикалық болжам деп нені айтады? Қандай статистикалық болжамдарды білісіздер?
2. Сіздер статистикалық болжамның қандай түрлерін білесіздер?
3. Бірінші және екінші текті қателіктер деп нені айтады?
4. Не үшін Стьюденттің t-белгісі медициналық-биологиялық берілгендерді талдауда үлкен қолданысқа ие?
5. Стьюденттің t-белгісін қолданғанда қандай шарттар орындалуы керек?
6. Екі таңдамалы және жұптасқан Стьюдент белгісінің ерекшеліктері неде?

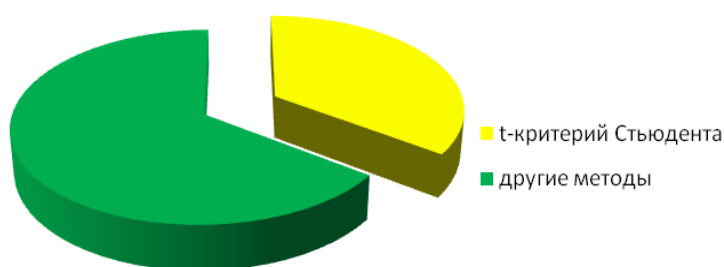
№4 дәріс

1. **Тақырыбы:** Екі топтың белгілерінің орташа мәндерін салыстыру.
2. **Мақсаты:** Студенттерді статистикалық болжамды тексерудің негізгі теориясымен және Стьюденттің t-белгісімен таныстыру.
3. **Дәріс тезистері:**

Медициналық-биологиялық зерттеу үдерісінде (мысалы, бақылау және зерттелуші топтарда немесе зерттеуге дейін және кейін) көбіне тексеру нәтижелерін салыстыруда мәселерер туындайды.

Бұл мәселелерді шешу үшін көптеген статистикалық белгілердің түрлері бар. Олардың әрқайсысының бір –бірінен өзіндік ерекшелігі бар (мысалы, бкрілгендердің түрлері, таңдама көлемі, салыстырылған таңдаманың санымен, салыстырылған таңдаманың сапасымен (тәуелді және тәуелсіз) және т.б.).

Мұндай белгілердің ең әйгілісі шамамен 30-40% ғылыми медициналық жұмыстарда қолданылатын Стьюденттің t-белгісі болып табылады. (7.1 сурет).



1-сурет. Ғылыми медициналық жұмыстарда қолданылатын статистикалық әдістердің арақатынасы.

Стьюденттің t-белгісі – ($H_0 : \bar{x}_1 = \bar{x}_2$) екі таңдаманың орташа шамаларының теңдігі туралы болжамды қабылдауға немесе қабылдамауға мүмкіндік беретін таңдаманың біркелкілігін тексеру әдісі.

Бұл белгіні (1876-1936 ж.) «Гиннес» компанияда сыра сапаларын бағалау үшін ағылшын химигі Уильям Госсеттпен өңдеген болатын.

Госсеттің мақаласы 1908 жылы «Биометрика» журналында «Student» (Студент) деген бүркеншек атпен шықты.

Стьюденттің t-белгісі қолданылуы:

- Екі *тәуелсіз* таңдама мәндерінің орташа шамаларының теңдігі (*екі таңдамалы t-белгісі*) туралы болжамды тексеруде. Бұл жағдайда саны әртүрлі бақыланатын және тәжірибелік таңдамалылар талданады. Мысалы, қант диабетімен ауыратын аурулар тобы және денсаулығы жақсы адамдар тобы;
- Екі *тәуелді* таңдама мәндерінің орташа шамаларының теңдігі (*жұпталған t-белгісі*) туралы болжамды тексеруде. Бұл жағдайда тәжірибеге дейінгі және кейінгі сол таңдама талданады. Мысалы, антиаритмдік дәріні қабылдағанға дейінгі және қабылдағаннан кейінгі аурулардың орташа қан тамырының соғуы.

Стьюдент белгісін келесі екі шарт орындалғанда қолдануға болады:

- қарастырылатын таңдама қалыпты үлестірілімді
- қарастырылатын таңдаманың дисперсиясы тең.

Зерттеушілердің анықтағаны бойынша, медициналық-биологиялық зерттеулерде, екі шарт бірдей тек қана оқиғаның 4-5% ғана орындалады.

Сонымен қатар диссертацияларды, ғылыми жұмыстарды, медициналық журналдарда жарияланған мақалаларды талдау кезінде Стьюденттің t-белгісі 50% пайыз жұмыстардағы статистикалық есептеулерде қолданылған.

Стьюденттің екі таңдамалы t-белгісін қолдану реті:

$$1) H_0: \bar{x}_1 = \bar{x}_2.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$ -маңыздылық деңгейі.

$$3) t_{расч} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2} \cdot (n_1 + n_2 - 2)},$$

мұндағы n_1, n_2 – қарастырылатын таңдамалардың көлемі, s_1^2, s_2^2 – қарастырылатын таңдамалардың дисперсиясы, \bar{x}_1, \bar{x}_2 – салыстырылатын таңдамалылардың орташа.

4) $t_{табл}(p; f)$, мұндағы $f = n_1 + n_2 - 2$ – еркіндік дәреже саны (кестелік мәндер).

5) Егер $t_{расч} > t_{кр}$ болса, онда « H_0 » қабылданады.

Егер $t_{расч} > t_{кр}$ болса, « H_0 » қабылданбайды.

Стьюденттің белгісі таңдам көлемі ($n_{1,2} \leq 30$) аз болған кезде қолданылады.

1 мысал. Егер босану кезінде аналық жатыр мойны ұзақ ашылмаса, онда босану ұзақтығы артады және кесер тілігін жасау қажет. Ғалымдар E_2 простагландин гелінің әсерін аналық жатыр мойнының ашылу жылдамдағын анықтамақшы болды. Зерттеуде босанатын 2 топ қатыстырылды. Бірінші топтағы босанушылардың аналық жатырының мойнына E_2 простагландинген гелін, ал екінші топтағы босанушыларға плацебо гелін енгізді. Екі топтағы 21 босанушының жастары, бойы және босану уақыттары шамамен бірдей. Простагландин гелін пайдаланған топтағылардың босану уақыты орташа 8,5 сағатқа (орташа ауытқуы 4,7 сағат), ал бақылау тобындағылар 13,9 сағатқа (орташа ауытқу 4,1 сағат) созылды. E_2 простагландин гелі босану ұзақтығын қысқартады деп тұжырым жасауға бола ма?

Шешуі.

$$1) H_0: \bar{x}_1 = \bar{x}_2.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$.

$$3) t_{расч} = \frac{13,9 - 8,5}{\sqrt{(21-1) \cdot 4,1^2 + (21-1) \cdot 4,7^2}} \cdot \sqrt{\frac{21 \cdot 21}{21+21}} (21+21-2) \approx 4.$$

$$4) t_{кесте}(0,05;40) = 2,02.$$

5) $t_{есен} > t_{кесте}$, болғандықтан, « H_0 » қабылданбайды, демек E_2 простагландин гелибосану ұзақтығын қысқартады.

Стьюденттің жұпталған t-белгісін қолдану реті:

$$1) H_0: \bar{x}_1 = \bar{x}_2.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$ - маңыздылық деңгейі.

$$3) t_{есен} = |\bar{d}| \cdot \sqrt{\frac{n(n-1)}{\sum_{i=1}^n d_i^2 - n\bar{d}^2}},$$

мұндағы $d = x_i - y_i$ - сәйкес жұп айнымалылар арасындағы мәндерінің айырымы, \bar{d} - осы айырымының орташа мәні, n – таңдама көлемі.

4) $t_{кесте}(p; f)$, мұндағы $f = n-1$ - еркіндік дәреже саны (кестелік мәндер).

5) Егер $t_{есен} > t_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $t_{есен} > t_{кесте}$ болса, « H_0 » қабылданбайды.

2 мысал. Жаңа гипогликемикалық препараттың әсерін бағалау үшін сусамыр ауруымен ауыратын науқастың пепаратты қабылдағанға дейінгі және кейінгі қан құрамындағы гликозаның мөлшеріне өлшеу жүргізді:

Науқастың номері	Қандағы глюкоза мөлшері, моль/л	
	Дәріні қолданғанға дейінгі	Дәріні қолданғаннан кейінгі
1	9,6	5,7
2	8,1	4,2
3	8,8	6,4
4	7,9	5,5
5	9,2	5,3
6	8,0	5,2
7	8,4	5,1
8	10,1	5,9
9	7,8	7,5
10	8,1	5,0
Орташа мәні	8,6	5,6

Препаратты қабылдағаннан кейін қанындағы глюкозаның деңгейі төмендейді деп есептеуге бола ма?

Шешуі.

$$1) H_0: \bar{x}_1 = \bar{x}_2.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$ - маңыздылық деңгейі.

3)

Науқас №	Қандағы глюкоза мөлшері, моль/л		Мәндердің айырмашылығы $d = x_i - y_i$	d^2
	Дәріні қолданғанға дейінгі, x_i	Дәріні қолданғаннан кейінгі, y_i		
1	9,6	5,7	3,9	15,21
2	8,1	4,2	3,9	15,21
3	8,8	6,4	2,4	5,76
4	7,9	5,5	2,4	5,76
5	9,2	5,3	3,9	15,21
6	8,0	5,2	2,8	7,84
7	8,4	5,1	3,3	10,89
8	10,1	5,9	4,2	17,64
9	7,8	7,5	0,3	0,09
10	8,1	5,0	3,1	9,61
Қосындысы			30,2	103,22

$$\bar{d} = \frac{30,2}{10} = 3,02 \quad t_{есен} = 3,02 \cdot \sqrt{\frac{10(10-1)}{103,22 - 10 \cdot 3,02^2}} \approx 8,3.$$

$$4) t_{кесте}(0,05;9) = 2,26.$$

5) $t_{есен} > t_{кесте}$ болғандықтан, « H_0 » қабылданбайды яғни, дәріні қолданғаннан кейін қандағы глюкоза мөлшері кемиді, демек жаңа құрал тиімді.

Кейде таңдамалы орташалардың айырымының сенімділік бағасы келесі формуламен есептеледі:

$$t_{есен} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{m_1^2 + m_2^2}},$$

мұндағы \bar{x}_1, \bar{x}_2 – салыстырылатын орташа шамалар, m_1 және m_2 – салыстырылатын орташа шамалардың қателері.

Егер $t_{есен} \geq 2$ болса, онда « H_0 » қабылданбайды.

3 мысал. Медик-студенттердің емтиханға дейін және емтиханнан кейін артериялық қысымына зерттеу жүргізді. Емтихан тапсырғанға дейін ең жоғары қысым $98,8 \pm 4,0$, ал емтихан тапсырғаннан кейін $84 \pm 5,0$ мм сын.бағ. құрады. Емтихан тапсырғанға дейін студенттердің артериялық қан қысымының шынында да жоғарылауы орын алды деп айтуға болама?

Шешімі.

$$1) H_0: \bar{x}_1 = \bar{x}_2.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$ – маңыздылық деңгейі.

$$3) t_{расч} = \frac{98,8 - 84}{\sqrt{4^2 + 5^2}} \approx 2,3.$$

4) $t_{есен} > 2$ болғандықтан, « H_0 » қабылданбайды, яғни емтиханнан кейін қан қысымы төмендеп, қалыпты жағдайға жақындайды.

4. Иллюстрациялық материал: көрме, слайдтар

5. Әдебиет:

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SKMA -1979-	SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казakhstanская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 26 беті	

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.
2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.
3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015
 - Қосымша
 - 1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013
 - 2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.
 - Электрондық ресурстар:
 - 1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)
 - 2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
 1. 3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары

7. Статистикалық болжам деп нені айтады? Қандай статистикалық болжамдарды білісіздер?
8. Сіздер статистикалық болжамның қандай түрлерін білесіздер?
9. Бірінші және екінші текті қателіктер деп нені айтады?
10. Не үшін Стьюденттің t -белгісі медициналық-биологиялық берілгендерді талдауда үлкен қолданысқа ие?
11. Стьюденттің t -белгісін қолданғанда қандай шарттар орындалуы керек?
12. Екі таңдамалы және жұптасқан Стьюдент белгісінің ерекшеліктері неде?

№5 дәріс

1. Тақырыбы: Параметрлік емес балама.

2. Мақсаты: Студенттерді орташаның біркелкілік болжамын тексеру үшін кейбір параметрлік емес белгілермен таныстыру.

3. Дәріс тезистері:

Статистикалық белгілер параметрлік және параметрлік емес болып бөлінеді.

Параметрлік белгі – салыстырылатын таңдамалылардағы бірқалыпты үлестірілгендігін болжайды. және есептеу үдерісі кезінде үлестірім параметрлерін (орташа, дисперсия, орташа квадраттық ауытқу) қолданады. Мысалы: Стьюденттің t -белгісі, Фишердің F -белгісі және т.б.

Параметрлік емес белгі – салыстырылатын таңдамалылардағы бірқалыпты үлестірілгендігін болжамайды және есептеу үдерісі кезінде белгі мәндерінің шенін қолданады. Мысалы: Манна-Уитни, белгісі, Уилкоксон белгісі, таңбалар белгісі және т.б.

Шен –мәндік белгінің реттік нөмері.

Әрбір параметрлік белгіге, кем дегенде бір параметрлік емес белгі балама болады.

Екі таңдамалы Стьюденттің t -белгісінің баламасы, Манна-Уитнидің U –белгісі. Жұптасқан Стьюденттің t -белгісінің баламасы Уилкоксонның W - белгісі.

Манн-Уитнидің U -белгісі - екі тәуелсіз таңдаманы салыстыру үшін белгінің сандық өлшеу деңгейі бойынша алынған параметрлік емес статистикалық белгі.

Таңдамалылар арасындағы айырмашылықты анықтайтын бұл әдісті 1945 жылы америкалық химик және статистик Ф. Уилкоксоном (сурет 3.1, а) ұсынды.

1947 жылы әдісті математик Х.Б. Манном және Д.Р. Уитни кеңейтіп қайта өңдеді.

OÑTÜSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»	№ 35-11(Б)-2024 60 беттің 27 беті
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені		

Манн-Уитнидің U -белгісі аз таңдаманы салыстыру үшін қолданады. Әр таңдамада үштен кем емес белгінің мәні болуы керек. Бір таңдамада екі мән, ал екіншісінде бестен кем емес болған жағдай рұқсат етіледі ($n_1, n_2 \geq 3$ немесе $n_1=2, n_2 \geq 5$)

Салыстырылатын топтарда белгінің сәйкес мәндерінің болмауы (барлық сандар-эртүрлі) немесе сәйкес келетін мәндердің аз болуы Манна-Уитнидің U -белгісін қолданудың шарты болып табылады.

Уилкоксонның W -белгісі - параметрлік емес статистикалық белгісі, санжағынан өлшенген сыналатын қандайда бір белгінің деңгейінде екі тәуелді таңдаманы салыстыру үшін қолданылады.

Уилкоксон белгісі, егер « n » таңдаманың көлемі $5 \leq n \leq 50$ теңсіздігі қанағаттандырғанда қолданылады.

Манн-Уитни белгісін қолдану сызбасы:

1) $H_0: \bar{x}_1 = \bar{x}_2.$

$H_1: \bar{x}_1 \neq \bar{x}_2.$

2) $p=0,05$ маңыздылық деңгейі

3) Екі салыстырылатын таңдамадан бірыңғай реттілген қатар құрылады. Бірінші және екінші таңдамадан тұратын бірыңғай реттелген қатар екіге бөлінеді, бұл жағдайда әрбір бірліктің шенінің мәні белгіленеді.

Бірінші таңдаманың элементінің үлесіне тиетін, және жеке - екінші таңдаманың элементінің үлесіне тиетін шеннің қосындысы жеке есептеледі.

$$U_{расч} = n_1 \cdot n_2 + \frac{n_x \cdot (n_x + 1)}{2} - T_x$$

мұндағы T_x – екі шендік қосындының үлкені, n_x – T_x –қа сәйкес таңдама көлемі, n_1, n_2 – қарастырылатын таңдама көлемі.

4) $U_{табл}(p; n_1; n_2).$

5) Егер $U_{есеп} > U_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $U_{есеп} \leq U_{кесте}$ болса, онда « H_0 » қабылданбайды.

1 мысал. Семіздікке шалдыққан аурулардың артық салмақтан арылуға мүмкіндік беретін препараттың тиімділігі зерттелуде. Осыған орай еріктілер тобына белгілі бір диета тағайындалды.

Бір айдан кейін диетаның сақталу режімі және препаратты үздіксіз қабылдауды тексеру мақсатында жоғалтқан салмақ (кг) тіркелді. Тәжірибе жүргізу үшін 8 адамнан тұратын топ таңдап алынды. Оның ішінде үшеуі зерттелетін препаратты (зерттелуші топ), ал бесеуі плацебо (бақланатын топ) қабылдаған. Сегізден үшді зерттелуші тапқа таңдау кездейсоқ түрде жүргізілді. Тәжірибеге қатысушылардың барлығы препаратты қабылдадық деп есептеді.

Зерттелуші топ	Бақыланатын топ
Жоғалтқан салмағы, кг	Жоғалтқан салмағы, кг
6,2	4,0
3,0	-0,5
3,9	3,3
	1,5
	3,0

Шешуі.

1) $H_0: \bar{x}_1 = \bar{x}_2.$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p=0,05$ – маңыздылық деңгейі

3) Біріңғай қатар құрамыз.

Жоғалтқан салмағы, кг	6,2	3,0	3,9	4,0	-0,5	3,3	1,5	3,0
Шені	8	3,5	6	7	1	5	2	3,5

Біріңғай шендік қатарды бірліктен тұратын бірінші және екінші таңдамалар деп екіге бөлеміз.

Зерттелуші топ		Бақыланушы топ	
Жоғалтқан салмағы, кг	Шен	Жоғалтқан салмағы, кг	Шен
6,2	8	4,0	7
3,0	3,5	-0,5	1
3,9	6	3,3	5
		1,5	2
		3,0	3,5
	$T_1=17,5$		$T_2=18,5$

T_1 және T_2 – шендер қосындысы; $T_1 < T_2$, демек $T_2 = T_x$, $n_x = n_2 = 5$.

$$U_{есен} = 3 \cdot 5 + \frac{5 \cdot (5+1)}{2} - 18,5 = 11,5.$$

4) $U_{кесте} (0,05; 3; 5) = 1$.

5) $U_{есен} > U_{кесте}$ болады, « H_0 » қабылданады, демек дәрі тиімді емес.

Уилкоксон белгісін қолдану сызбасы:

1) $H_0: \bar{x}_1 = \bar{x}_2.$

$$H_1: \bar{x}_1 \neq \bar{x}_2.$$

2) $p \approx 0,05$ – маңыздылық деңгейі.

3) Екінші және бірінші өлшемдер арасындағы жеке мәндер есептелінеді.

Айырмашылықтардың абсолютті шамалары шеніне қарай (аз шамаға аз шен беріледі) реттеледі.

Әрбір шенге айырмашылыққа сәйкес «+» немесе «-» таңбасы қойылағандықтан таңбалы шен алынады.

4) $W_{есен}$ белгісінің есептелетін мәні, шендер мәнінің қосындысымен анықталады. $W_{кесте} (p; n)$, мұнда n – таңдама көлемі.

5) Егер $|W_{есен}| \leq W_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $|W_{есен}| > W_{кесте}$, болса, онда « H_0 » қабылданбайды.

2 мысал. Ультрадыбыстық қондырғыда жұмыс істейтін 12 жұмысшының ашқарынмен жұмысқа дейінгі және жұмыстан кейінгі үш сағаттан кейін қанындағы қант құрамында өзгерістер бола ма?

№	Жұмысқа дейінгі қант құрамы	Жұмыстан кейінгі қант құрамы
1	112	54
2	82	67
3	101	96

4	72	59
5	79	79
6	82	76
7	64	66
8	70	66
9	88	48
10	81	50
11	66	61
12	88	61

Шешуі.

1) $H_0: \bar{x}_1 = \bar{x}_2.$

$H_1: \bar{x}_1 \neq \bar{x}_2.$

2) $p \approx 0,05$ – маңыздылық деңгейі.

3)

№	X	y	x-y	x-y	$r_{ x-y }$	Белгілер шені
1	112	54	58	58	12	12
2	82	67	15	15	8	8
3	101	96	5	5	4,5	4,5
4	72	59	13	13	7	7
5	79	79	0	0	1	1
6	82	76	6	6	6	6
7	64	66	-2	2	2	-2
8	70	66	4	4	3	3
9	88	48	40	40	11	11
10	81	50	31	31	10	10
11	66	61	5	5	4,5	4,5
12	88	61	27	27	9	9
Белгілер шенінің қосындысы						$W_{есен} = 74$

$W_{есен} = 74.$

4) $W_{кесте} (0,052; 12) = 50.$

5) $|W_{есен}| > W_{кесте}$ болса, онда « H_0 » қабылданбайды, демек жұмысшылардың жұмысқа дейінгі және жұмыстан кейінгі қандағы қант құрамында өзгерістер бар.

4. Иллюстрациялық материал:

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.

2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013

2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые

OÑTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 30 беті	

дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)

2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары:

1. Параметрлік және параметрлік емес статистикалық белгілердің айырмашылығы неде?
2. Неліктен Манна-Уитни белгісін, екі таңдамалы Стьюденттің t-белгісінің баламасы деп атайды?
3. Неліктен Уилкоксон белгісін, жұптасқан Стьюденттің t-белгісінің баламасы деп атайды?

№ 6 дәріс

1. Тақырыбы: Бір факторлы дисперсиялық талдау.

2. Мақсаты: Студенттерді дисперсиялық талдау негіздерімен таныстыру.

3. Дәріс тезистері:

1. Дисперсиялық талдау деп ағылшын математигі және генетигі Р.Фишер ХХ ғасырдың 20-жылдарында биология мен ауыл шаруашылығының бірқатар эксперименттік тапсырмалары үшін әзірленген статистикалық әдістер топтамасын атайды.

Есептің қойылуы. X_1, X_2, \dots, X_k бас жинақтары берілсін, мұнда:

- барлық «к» бас жинақ қалыпты үлестірілген;
- барлық бас жинақтардың дисперсиялары бірдей.

Осындай шарттар орындалғанда және «р» маңыздылық деңгейінде орташа шамалардың теңдігінің нөлдік болжамын (гипотезасын) тексеру қажет, яғни $H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k$.

Бас жиынтықтың әрбіріне бір немесе бірнеше факторлар әсер етіп, олардың орташа мәндерін өзгертуі мүмкін.

Соңғы нәтижеге әсерін тигізетін көрсеткішті *фактор* деп атайды. Фактор бір немесе бірнеше болуы мүмкін. Фактордың нақты іске асуын *фактор деңгейі* деп атайды.

Өлшенетін белгінің мәнін фактордың әсеріне берілетін *жауап* (үн қату) деп атайды.

Мысалы, гипертония ауруына шалдыққан кейбір науқастар санына қарай кездейсоқ түрде «к» топтарына бөлінген, олардың әрқайсысына белгілі бір дәрі-дәрмек қабылдау тағайындалған. Нәтижесінде артериалдық қысым көрсеткішінің орташа мәнінің өзгеруі бақылауға алынады.

Бұл мысалда:

- « n_i » науқастардан құрылған « i » тобындағы көрсеткіштердің мәндері – бұл « n_i » көлемдегі « i »-ші таңдама;

- дәрі-дәрмек – бұл бақылаудағы көрсеткіш шамасына ықпалын тигізетін *фактор*;
- артериялық қысым өзгерісінің көрсеткіші – бұл фактор ықпалына деген *үн қату*.

Топтар бойынша қабылданатын дәрі-дәрмектер не түрімен, не мөлшерімен, не қандай да бір басқа жағдайларымен ажыратылады деген болжам бар. Онда ықпал етуші фактор, *фактор деңгейлері* деп аталатын бірқатар құрамдарға бөлінеді.

Егер фактор шаманың үн қатуына әсер етсе, онда шамалардың орташаларының теңдігі туралы $H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k$ нөлдік болжам қабылданбайды.

Зерттеліп отырған факторлар санына байланысты дисперсиялық талдау *бірфакторлы* және *көпфакторлы* болып бөленеді.

Мысалдағы артериалдық қысымның өзгеруін зерттеу жолдары:

- фактор – жыл маусымы (деңгейлері: қыс, көктем, жаз, күз);
- фактор – тәжірибе жүргізілетін орын (оның деңгейлері: ауруханада немесе үйде емделу);
- фактор – режим (оның деңгейлері: төсекте, әдеттегідей немесе таза ауада үнемі жаяу жүру) және т.б.

Бір факторлы дисперсиялық талдау үшін таңдалған деректерді әдетте кесте түрінде көрсетеді.

Сынақ нөмері	А факторының деңгейі			
	A ₁	A ₂	...	A _k
1	x ₁₁	x ₁₂	...	x _{1k}
2	x ₂₁	x ₂₂	...	x _{2k}
...		
R	x _{r1}	x _{r2}	...	x _{rk}
Топтық орташа	$\bar{x}_{\text{топ1}}$	$\bar{x}_{\text{топ2}}$...	$\bar{x}_{\text{топk}}$

2. Дисперсиялық талдаудың негізгі мақсаты таңдамалы дисперсияны екі құрамдас бөлікке бөлу:

- біріншісі – бұл орташа мәндердің өзгеруіне әсерін тигізетін факторге сәйкес келетін *факторлық дисперсия*;
- екіншісі – бұл кездейсоқ себептер әсерінен пайда болатын, орташалардың өзгергіштігіне әсерін тигізбейтін *қалдық дисперсия*.

Зерттелетін факторды сандық жағынан бағалау үшін осы құрамдас бөліктерді Фишер белгісімен салыстыру қолданылады.

Факторлық дисперсия ($S^2_{\text{факт}}$) – бұл фактордың әсерінен таңдаманың орташа шамаларының өзгеруіне сәйкес келетін дисперсия:

$$S^2_{\text{факт}} = \frac{SS_{\text{факт}}}{k-1} = \frac{r \sum_{j=1}^k (\bar{x}_{\text{топ}j} - \bar{x})^2}{k-1}$$

мұндағы $SS_{\text{факт}}$ - орташа квадраттық ауытқулардың факторлық қосындысы, k - фактор деңгейлерінің саны; r - әр топтағы мәндер саны; \bar{x} - жалпы орташа; $\bar{x}_{\text{топ}}$ - топтық орташа.

Қалдық дисперсия ($S^2_{\text{қалд}}$) – бұл орташа шамалардың өзгеруіне ықпал етпейтін, кездейсоқ себептер әсерінен ғана болатын дисперсия:

$$S^2_{\text{қалд}} = \frac{SS_{\text{қалд}}}{k(r-1)} = \frac{\sum_{i=1}^r (x_{i1} - \bar{x}_{\text{топ1}})^2 + \sum_{i=1}^r (x_{i2} - \bar{x}_{\text{топ2}})^2 + \dots + \sum_{i=1}^r (x_{ik} - \bar{x}_{\text{топk}})^2}{k(r-1)}$$

мұндағы $SS_{\text{қалд}}$ – ауытқулар квадраттарының қалдық қосындысы.

Жалпы дисперсия ($S^2_{\text{жалпы}}$) - бұл факторлық және қалдық дисперсияның қосындысы:

$$S^2_{\text{жалпы}} = \frac{SS_{\text{жалпы}}}{n-1} = \frac{\sum_{j=1}^k \sum_{i=1}^r (x_{ij} - \bar{x})^2}{n-1}$$

мұндағы $SS_{\text{жалпы}} = SS_{\text{факт}} + SS_{\text{қалд}}$.

3. Бір факторлық дисперсиялық талдауды жүргізу ретті:

1) $H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k .$

$H_1: \bar{x}_1 \neq \bar{x}_2 \neq \dots \neq \bar{x}_k .$

2) $p=0,05$ - маңыздылық деңгейі.

3) $F_{расч} = \frac{S^2_{факт}}{S^2_{осн}} .$

4) $F_{табл}(p, f_1, f_2)$, мұндағы $f_1=k-1, f_2=k(r-1)$ – еркіндік дәрежесі (кестелік мәндер), k – фактор деңгейінің саны, r - әр топтағы мәндер саны.

5) Егер $F_{есеп} \leq F_{кес}$, онда « H_0 » қабылданады.

Егер $F_{есеп} > F_{кес}$, онда « H_0 » қабылданбайды.

Мысал 1. Белгілі бір жас санатындағы ересек тұрғындар арасында екі жыл бойындағы тыныс алу жолдары ауруларының саны тіркелген. Зерттеу мақсаты – темекі тартудың тыныс алу жолдарының ауруға шалдығыуына әсерін статистикалық дәлелдеу.

Кездейсоқ түрде әрқайсысы 4 адамнан тұратын 3 топ іріктеліп алынды, олардың ішінде: 1 топ – темекі тартпайтындар, 2 топ – темекі тарту өтілі 5 жылға дейін, 3 топ – темекі тарту өтілі 5 жылдан астам.

Осылайша зерттеліп отырған «А» факторы – темекі тарту, фактор деңгейлері A_1, A_2, A_3 - темекі тарту өтілі. Темекі тарту факторына жауап - тыныс алу жолдарының ауру саны.

x_{ij} – 12 ауру санының мәні алынды, мұндағы j – фактор деңгейінің нөмірі ($j=1,2,3$), i – сәйкес таңдап алынған (топтағы) элементтің нөмірі, $i=1,2,3,4$:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 \\ 0 & 2 & 4 \\ 1 & 2 & 5 \\ 2 & 1 & 3 \end{bmatrix}$$

$\{x_{ij}\}$ - қалыпты үлестірілген жинақтан алынған таңдама деп қарастырылады.

Барлық берілгендерді кестеге енгізу керек:

Сынақ нөмірі	«А» факторының деңгейі		
	A_1	A_2	A_3
1	1	3	3
2	0	2	4
3	1	2	5
4	2	1	3
Топтық орташа	4/4=1	8/4=2	15/4=3.75

Шешуі:

1) $H_0: \bar{x}_1 = \bar{x}_2 = \bar{x}_3 .$

$H_1: \bar{x}_1 \neq \bar{x}_2 \neq \bar{x}_3 .$

2) $p=0,05$ - маңыздылық деңгейі.

3) Есептеледі:

3.1) Жалпы орташа:

$$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3} = \frac{1 + 2 + 3,75}{3} = 2,25 .$$

3.2) Квадраттық ауытқудың факторлық қосындысы:

$$SS_{\text{факт}} = r \sum_{j=1}^3 (\bar{x}_{epj} - \bar{x})^2 = 4 \cdot [(1 - 2,25)^2 + (2 - 2,25)^2 + (3,75 - 2,25)^2] = 15,5.$$

3.3) Квадраттық ауытқудың қалдық дисперсиясы:

$$SS_{\text{ост}} = \sum_{i=1}^4 (x_{i1} - \bar{x}_{ep1})^2 + \sum_{i=1}^4 (x_{i2} - \bar{x}_{ep2})^2 + \sum_{i=1}^4 (x_{ik} - \bar{x}_{ep3})^2 =$$

$$= [(1-1)^2 + (0-1)^2 + (1-1)^2 + (2-1)^2] + [(3-2)^2 + (2-2)^2 + (2-2)^2 + (1-2)^2] +$$

$$+ [(3-3,75)^2 + (4-3,75)^2 + (5-3,75)^2 + (3-3,75)^2] = 6,75.$$

3.4) Фактор дисперсия:

$$S_{\text{факт}}^2 = \frac{SS_{\text{факт}}}{k-1} = \frac{15,5}{3-1} = 7,75.$$

3.5) Қалдық дисперсия:

$$S_{\text{қалд}}^2 = \frac{SS_{\text{қалд}}}{k(r-1)} = \frac{6,75}{3(4-1)} = 0,75.$$

$$3.6) F_{\text{бақл}} = \frac{S_{\text{факт}}^2}{S_{\text{қалд}}^2} = \frac{7,75}{0,75} = 10,3.$$

$$4) F_{\text{кест}}(0,05; 2; 9) = 4,26.$$

5) $F_{\text{бақл}} > F_{\text{кест}}$, онда « H_0 » қабылданбайды, яғни темекі тарту факторы тыныс алу жолдары ауруларына елеулі әсерін тигізеді.

Крускал–Уоллис H-белгісі үш және одан көп тәуелсіз топтарды салыстыру үшін дисперсиялық талдаудың параметрлік емес баламасы.

Аталған белгі берілгендердің нақты берілген мәндері емес, олардың шендерін қолданып есептейді.

Егер топтағы үлестірім бір қалыпты болмаса, H- белгісі есептеледі.

Үш таңдаманы салыстырғанда, олардың әрқайсысында 3 бақылаудан кем болмауы, немесе 4 бақылаудан, ал басқасында -2-ден; бұл жерде қай таңдамада қанша сынақ жүргізілгені емес, олардың 4:2:2 қатынасы маңызды.

H-белгісінің критикалық мәндері таңдаманың саны $k \leq 5$, ал сынақтың саны әрбір топта $n_i \leq 8$ болған жағдайда ғана қарастырылады. Үлкен көлемді таңдама санында және әрбір таңдамадағы сынақта кестеден χ^2 -белгісінің критикалық мәнін қолдануға болады, себебі Крускал-Уоллис белгісі « χ^2 » үлестіріміне асимптоталы жақындайды.

Крускал –Уоллистің H-белгісін қолданудың тізбесі:

$$1) H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k.$$

$$H_1: \bar{x}_1 \neq \bar{x}_2 \neq \dots \neq \bar{x}_k.$$

2) $p=0,05$ - маңыздылық деңгейі.

$$3) H_{\text{есеп}} = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1),$$

мұнда $n = \sum_{i=1}^k n_i$ - барлық топ бойынша жалпы бақылау саны, R_i - шеннің қосындысы, i -ші

таңдаманың шендерінің қосындысы.

4) Таңдама саны $k \leq 5$ болған жағдайда $H_{кест} (p; n_1; n_2; \dots; n_5)$,

Мұнда n_1, n_2, \dots, n_5 – қарастырылатын таңдаманың көлемі.

Таңдама саны $k > 5$ болған жағдайда $H_{кест} = \chi^2_{кест} (p; f)$,

мұнда $f = k - 1$ – еркіндік дәрежесінің саны (кестелік мән).

5) Егер $H_{есеп} < H_{кест}$, онда « H_0 » қабылданады.

Егер $H_{есеп} \geq H_{кест}$, онда « H_0 » қабылданбайды.

2 мысал. Ішетін суды ластаған химиялық заттардың мөлшерлік жүктемесін бағалау үшін, әр түрлі жастағы тұрғындардың пайдаланатын су құбырындағы ішуге жарамды судың мөлшерін зерттеді. Нәтижесінде төмендегі мәліметтер алынды:

Балалар ($n_1=8$)	Жасөспірімдер ($n_2=7$)	Үлкендер ($n_3=9$)
Су, л/күн	Су, л/күн	Су, л/күн
1,22	1,47	1,56
1,24	1,52	1,58
1,31	1,55	1,81
1,31	1,70	1,89
1,45	1,93	2,00
1,52	2,00	2,00
1,84	3,00	2,55
2,52		2,58
		4,00
$\bar{x}_1 = 1,55$	$\bar{x}_1 = 1,88$	$\bar{x}_1 = 2,22$

Балалар, жасөспірімдер және үлкен тұрғындардың ауыз суды тұтынуының орташа мәндерінің теңдігі туралы болжамды тексеру.

Шешуі:

Зерттелетін топтар тәуелсіз, ал мәліметтер бір қалыпты емес үлестірім болғандықтан, нөлдік болжамды тексеру үшін Крускал-Уоллис белгісін пайдаланамыз.

1) $H_0: \bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k$.

$H_1: \bar{x}_1 \neq \bar{x}_2 \neq \dots \neq \bar{x}_k$.

2) $p=0,05$ - маңыздылық деңгейі.

3)

Балалар ($n_1=8$)		Жасөспірімдер ($n_2=7$)		Үлкен ($n_3=9$)	
Су, л/күн	Шен	Су, л/күн	Шен	Су, л/күн	Шен
1,22	1	1,47	6	1,56	10
1,24	2	1,52	7,5	1,58	11
1,31	3,5	1,55	9	1,81	13
1,31	3,5	1,70	12	1,89	15
1,45	5	1,93	16	2,00	18
1,52	7,5	2,00	18	2,00	18
1,84	14	3,00	23	2,55	21
2,52	20			2,58	22
				4,00	24
Шендердің	56,5		91,5		152

ҚОСЫНДЫСЫ

$$H_{есеп} = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1) =$$

$$= \frac{12}{24 \cdot (24+1)} \cdot \left(\frac{56,5^2}{8} + \frac{91,5^2}{7} + \frac{152^2}{9} \right) - 3 \cdot (24+1) = 0,02 \cdot (399 + 1196 + 2567,1) - 75 = 8,24.$$

$$4) H_{кесте} = \chi_{кесте}^2(0,05; 2) = 5,99.$$

5) $H_{кесте} > H_{кесте}$, онда « H_0 » қабылданбайды, яғни әр түрлі жастағы топтарда күнделікті әр түрлі мөлшерде ішетін су пайдаланылады.

Екі факторлы дисперсиялық талдау – белгіге екі ұйымдастырылған фактор әсер етуін зерттейтін статистикалық әдістер жүйесі.

Екі факторлық дисперсиялық талдау тек әр фактордың жеке әсерін ғана емес, сонымен қатар олардың өзара әрекеттестігінің де әсерін бағалауға мүмкіндік береді.

Екі факторлы дисперсиялық талдау жеткілікті үлкен, арнайы бағдарламалық қамтамасыздандыруды (Statistica 10, SPSS және т.б.) пайдалануды ұсынады.

3 мысал. Химиялық лабораторияда « Y » химиялық синтез өнімінің шығуына температураның (« A » фактор) және катализатордың (« B » фактор) әсерін тексереді. Алынған нәтижелер кестеде келтірілген. « A » және « B » факторларының және олардың комбинациясының көрсетілген белгіге әсері туралы болжамды тексеру талап етіледі.

	B_1	B_2	B_3
A_1	16; 19; 17; 16	18; 16; 17; 14	16; 16; 18; 13
A_2	22; 22; 19; 23	18; 19; 23; 24	18; 16; 19; 20
A_3	20; 16; 18; 19	18; 17; 19; 19	20; 20; 16; 16
A_4	23; 20; 22; 23	19; 18; 19; 22	20; 19; 20; 22

Шешуі:

1) Болжамды құрастыру.

« A » факторы үшін болжам:

H_0 : « A_i » температурасының барлық режимінде « Y » химиялық синтез өнімінің нәтижелерінің орташалары арасында айырмашылық жоқ.

H_1 : « A_i » температурасының барлық режимінде « Y » нәтижелерінің орташалары арасында айырмашылық бар.

« B » факторы үшін болжам:

H_0 : « B_j » катализатор түрлері үшін « Y » орташа нәтижелері арасында айырмашылық жоқ.

H_1 : « B_j » катализатор түрлері үшін « Y » орташа нәтижелері арасында айырмашылық бар.

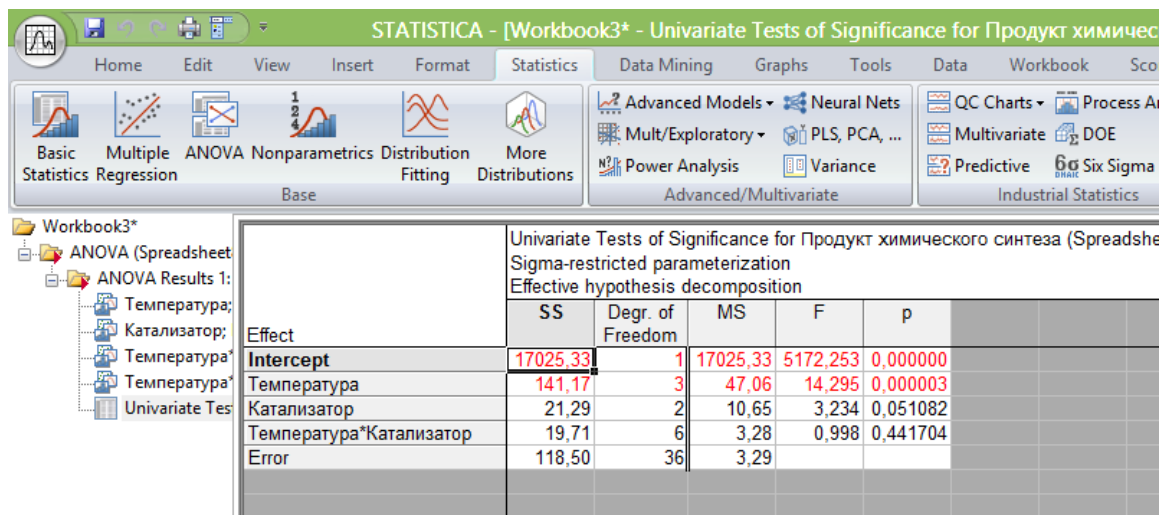
« A » және « B » факторларының өзара әсерлері үшін болжам:

H_0 : A (температура) және « B » (катализатор) факторлары « Y » (химиялық синтез өнімі) нәтижесіне тиімді өзара әсері болмайды.

H_1 : « A » және « B » факторлары « Y » (химиялық синтез өнімі) нәтижесіне тиімді өзара әсері болады.

2) «Statistica 10» бағдарламасының көмегімен екі факторлы дисперсиялық талдауды

жүргізу



The screenshot shows the Statistica 10 software interface. The main window displays the results of a univariate ANOVA test. The table below is a reproduction of the data shown in the software output.

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	17025,33	1	17025,33	5172,253	0,000000
Температура	141,17	3	47,06	14,295	0,000003
Катализатор	21,29	2	10,65	3,234	0,051082
Температура*Катализатор	19,71	6	3,28	0,998	0,441704
Error	118,50	36	3,29		

Statistica 10 бағдарламасында орындалған екі факторлы дисперсиялық талдаудың нәтижесінің қорытынды кестесі.

3) Қорытынды:

Температура режимі химиялық синтез нәтижесінің орташа қорытындысына әсер етеді.

Катализатор түрі химиялық синтез нәтижесінің орташа қорытындысына әсер етпейді.

Температура мен катализатор бірігіп химиялық синтез нәтижесінің орташа қорытындысына әсер етпейді.

4. Иллюстрациялық материал: көрме, слайдтар.

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.

2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013

2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)

2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>

3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

4. <http://www.statsoft.ru/>

6. Бақылау сұрақтары:

1. Қандай нөлдік болжам дисперсиялық талдаудың көмегімен тексеріледі?
2. Дисперсиялық талдауды қолданғанда қандай шарттар орындалуы керек?
3. Дисперсиялық талдаудың негізгі идеясы қандай?
4. Крускал–Уоллистің H-белгісі қай кезде қолданылады?

OÑTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»	№ 35-11(Б)-2024 60 беттің 37 беті
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені		

№ 7 дәріс

1. Тақырыбы: Корреляциялық талдау.

2. Мақсаты: Студенттерді корреляциялық талдаудың негізімен таныстыру.

3. Дәріс тезистері:

1. Эпидемиологияның ең маңызды міндеттерінің бірі қауіп қатер факторы

Медицинада қауіп қатер факторы – бұл аурудың пайда болуына мүмкіндік туғызатын фактор (темекі тарту- жүрек талмасы немесе қатерлі ісікке алып келетін қауіп қатер факторы, су жүйесіндегі болатын апат- іш

Ауру асқынуындағы қауіп қатер факторын сандық жағынан бағалау үшін корреляциялық талдау әдіс қолданылады.

Корреляциялық талдау – бұл екі немесе одан көп кездейсоқ шамалардың арасындағы байланыстың тығыздығын (тарлығын) және бағытын анықтайтын сандық әдіс.

Ғылыми айналымға «корреляция» терминін алғаш рет француз палеонтологы Ж. Кювье (XVIII ғ.) енгізді, ал статистикада оны Ф. Гальтон (XIX ғ.) алғаш қолданды.

Айнымалылар арасындағы байланысты сандық жағынан сипаттау үшін корреляция коэффициенті деген ұғым енгізілген.

Корреляция коэффициенті – байланыстың күшін және оның бағытын сипаттайтын, $[-1, 1]$ аралығынды мәнді қабылдайтын көрсеткіш.

Байланыстың күшін және оның бағытын сипаттау үшін корреляция теориясында аңғылшын статистигі Чеддоктың шкаласы қолданылады.

Байланыстың сандық мөлшері	Күштің сапалық сипаттамасы
0,1 - 0,3	Әлсіз
0,3 - 0,5	Қалыпты
0,5 - 0,7	Айқын
0,7 - 0,9	Жоғары
0,9 - 1	Күшті

Бағыты бойынша тура және кері корреляциялық байланыс деп бөлінеді.

Тура корреляциялық байланыс – бір айнымалының артуы басқа айнымалының артуына (су құбырындағы судағы стандартты емес сынамалар артқан кезде, іш сүзегімен аурудың артуы) байланысты болатын байланыс.

Кері корреляциялық байланыс – бір айнымалының кемуі басқа айнымалының артуына («В» сарауруына қарсы екпе жұмысының ауқымын көбейткен сайын, онымен ауыратындардың кемуі) байланысты болатын байланыс.

Тура байланыс кезінде корреляция коэффициенті «0»-ден «+1»-ге дейінгі мәндерді қабылдайды.

Кері байланыс кезінде корреляция коэффициенті «-1»-ден «0»-ге дейінгі мәндерді қабылдайды.

Егер корреляция коэффициенті «0»-ге тең болса, онда құбылыстар арасында байланыс болмайды.

Егер корреляция коэффициенті «+1» немесе «-1» болса, онда құбылыстар арасында байланыс функциялық болады.

2. Екі айнымалы арасындағы тәуелділікті талдаған кезде шашырау диаграмасын қолданады.

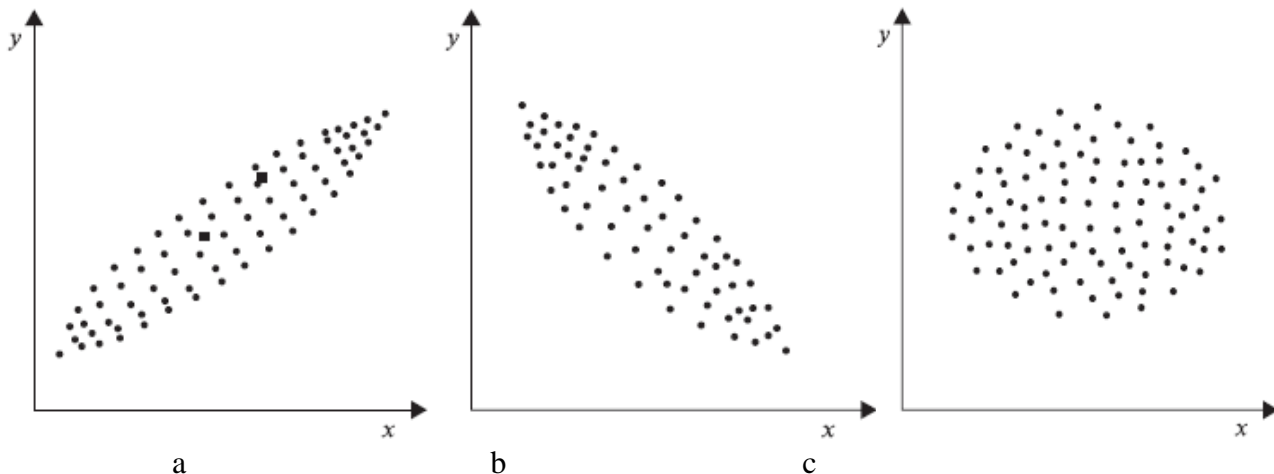
Шашырау диаграмасы - екі айнымалы арасындағы корреляциялық тәуелділікті

көрсететін көрнекі әдіс.

Шашырау диаграмасы – бұл тәжірибелік бақланған нүктелерді белгілі бір масштабпен түсірілген сызба түріндегі дәл диаграмма.

Сызбадағы нүктелердің координаталары қарастырылатын шамалардың және оған әсер ететін факторлардың мәндеріне сәйкес келеді.

Нүктелердің орналасуы екі айнымалы арасындағы байланыстың бар екенін және сипатын көрсетеді.



Шашырау диаграмасы: a – тура байланыс; b – кері байланыс; c – байланыс жоқ.

3. Сызықты (жұптасқан *Пирсонның*) корреляция коэффициенті – байланыстың күшін және бағытын сипаттайтын көрсеткіш:

$$r_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}},$$

мұндағы r_{xy} – корреляция коэффициенті; «x» және «y» – корреляцияланатын қатар; \bar{x} , \bar{y} – орташа шамалар.

Жұптасқан корреляция коэффициенті *параметрлік* коэффициент болып табылады.

Егер келесі шарттар орындалса, онда Пирсонның жұптасқан корреляция коэффициенті қолдануға болады:

- салыстырылатын айнымалылар аралық (интервалдық) немесе қатынас шкаласынан алынған болуы керек;
- айнымалылардың таралуы бір қалыпты үлестірімге жақын болуы керек;
- қарастырылатын айнымалылардың саны бірдеу болуы керек.

4. Корреляция коэффициентінің сенімділігі оны есептелетін орташа қателікпен салыстыру арқылы анықталады.

$$m_r = \pm \frac{1 - r_{xy}^2}{\sqrt{n}},$$

Корреляция коэффициентінің орташа қателігі: мұндағы r_{xy} – корреляция коэффициенті; n – бақылау саны.

Егер корреляция коэффициентіні орташа қателіктен 3 есе артық болса, онда сенімді деп есептелінеді. Олай болмаса бақылау санын көбейту керек.

Корреляция коэффициентінің сенімділігін арнайы кестеден анықталады.

Мысал. Келесі мәндер үшін Пирсонның сызықты корреляция коэффициентін есептеу керек:

1000 тұрғынға шаққанда ОРЗ-мен ауыру, x	352	228	340	300	196	258	237
1000 тұрғынға шаққанда пневмониямен ауыру, y	64	60	52	48	46	41	32

Шешуі:

1) Есептеу кестесін құру:

№	x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x}) \cdot (y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
1	352	64	79	15	1185	6241	225
2	228	60	-45	11	-495	2025	121
3	340	52	67	3	201	4489	9
4	300	48	27	-1	-27	729	1
5	196	46	-77	-3	231	5929	9
6	258	41	-15	-8	120	225	64
7	237	32	-36	-17	612	1296	289
Сумма	1911	343	0	0	1827	20934	718
Среднее	273	49					

2) Корреляция коэффициентін есептеу: $r_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} = \frac{1827}{\sqrt{20934 \cdot 718}} = 0,47$.

3) Алынған нәтижені талдау: қарастырылған белгілердің арасындағы байланыс тура, бір қалыпты қалыпты.

4) Корреляция коэффициентінің орташа қателігін есептеу:

$m_r = \pm \frac{1 - r_{xy}^2}{\sqrt{n}} = \pm \frac{1 - 0,47^2}{\sqrt{7}} = 0,3$, корреляция коэффициенті өзінің орташа қателігінен 3 есе артық болмағандықтан сенімді емес.

4. Клиникалық және фармацевтикалық құбылыстарды талдауда жиі келесі параметрлік емес коэффициенттер қолданылады:

- Спирменнің шендік корреляциясы;
- Кендалдың « τ » (тау);
- Юла ассоциациясы;
- Пирсонның контингенциясы;
- Чупровтың түйісуі;
- « γ » (гамма) және т.б..

1904 жылы корреляциялық талдау жүргізу үшін ағылшын психологы, Лондон және Честерфилд университеттерінің профессоры Ч.Э. Спирмен әзірлеген және ұсынған *шендік корреляция коэффициентін* қарастырайық.

Шендік корреляция коэффициенті – бұл әр түрлі белгілердің берілген нұсқасындағы шендердің арасындағы байланысты өлшейтін коэффициент. *Спирменнің шендік корреляция*

коэффициентін сандық белгілер сияқты сапалық белгілердің, егер олардың мәндерін өсу немесе кему дәрежесіне қарай реттеуге болса, олардың арасындағы байланыстардың тығыздығын анықтау үшін қолданылады.

Спирменнің шендік корреляция коэффициенті:

$$\rho = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (r_{x_i} - r_{y_i})^2,$$

мұндағы n – жиындық көлемі, $r_{x_i} - r_{y_i}$ - разность между рангами i -ші нысан шендерінің арасындағы айырмашылық.

Байланыс тығыздығының сапалық сипаттамасының шендік корреляция коэффициентін, басқа корреляция коэффициенттері сияқты Чеддок шкаласымен бағалауға болады.

Спирменнің шендік корреляция коэффициенті егер « n » таңдама көлемі $5 \leq n \leq 40$ теңсіздігін қанағаттандырған кезде қолданылады.

Мысал. Бір ауылдық жерде созылмалы індет Флекснер дизентериясының бар екендігі тіркелді. Алданала жүргізілген талдаулар және зертханалық зерттеулер су ішетін жүйедегі құбырларда көрсеткіші бойынша стандартты емес үлгілері бар екендік бактериологиялық көрсеткіштері бойынша (қауіп қатер факторы) көрсетті. Бұл екі белгі арасындағы байланыстың бар екендігі туралы болжамды тексеру керек.

Айлар	Дизентериямен ауырған науқастар саны (x)	Судағы стандартты емес үлгілердің үлесі (y)
Январь	10	0
Февраль	9	0,5
Март	2	1,1
Апрель	7	2,0
Май	6	1,8
Июнь	11	2,9
Июль	26	6,7
Август	32	4,5
Сентябрь	46	8,7
Октябрь	38	7,1
Ноябрь	8	3,2
Декабрь	5	0

Шешуі:

1) Есептеу кестесін құру:

№	x	y	r_x	r_y	$r_x - r_y$	$(r_x - r_y)^2$
1	2	0	7	1,5	5,5	30,25
2	9	0,5	6	3	3	9
3	2	1,1	1	4	-3	9
4	7	2,0	4	6	-2	4
5	6	1,8	3	5	-2	4
6	11	2,9	8	7	1	1

7	26	6,7	9	10	-1	1
8	32	4,5	10	9	1	1
9	46	8,7	11	12	-1	1
10	38	7,1	12	11	1	1
11	8	3,2	5	8	-3	9
12	5	0	2	1,5	0,5	0,25
Сумма						70,5

2) Корреляция коэффициентін есептеу:

$$\rho = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (r_{x_i} - r_{y_i})^2 = 1 - \frac{6}{12^3 - 12} \cdot 70,5 \approx 0,75.$$

3) Алынған нәтижені талдау: қарастырылған белгілердің арасындағы байланыс тура, жоғары.

4) Корреляция коэффициентінің орташа қателігін есептеу:

$$m_r = \pm \frac{1 - r_{xy}^2}{\sqrt{n}} = \pm \frac{1 - 0,75^2}{\sqrt{12}} \approx 0,12, \text{ корреляция коэффициенті өзінің орташа қателігінен } 3 \text{ есе артық болғандықтан сенімді.}$$

4. Иллюстрациялық материал: көрме, слайдтар.

5. Әдебиет:

• Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.

2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

• Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013

2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

• Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)

2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>

3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Акнұр баспасы, 2019. <http://elib.kaznu.kz>

4. <http://matstats.ru/>

6. Бақылау сұрақтары:

1. Эпидемиологиялық талдауда корреляциялық талдау не үшін қолданылады?
2. Корреляция коэффициенті қандай шектік аралықта өзгереді?
3. Зачем нужны Шашырау диаграмасы не үшін керек?
4. Пирсонның жұпталған корреляция коэффициенті қандай формуламен есептелінеді?
5. Корреляция коэффициентінің сенімділігі қалай анықталады?
6. Спирменнің шендік корреляция коэффициенті қандай жағдайда қолданылады?
7. Спирменнің жұпталған корреляция коэффициенті қандай формуламен есептелінеді?

O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ		SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 42 беті	

№ 8 Дәріс

1. Тақырыбы: Регрессиялық талдау.

2. Мақсаты: Студенттерді регрессиялық талдаудың негіздерімен таныстыру.

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

3. Дәріс тезистері:

«Регрессия» термині алғаш рет биометрияның негізін салушы Ф. Гальтон (XIX ғ.) енгізген, оның ойын ізбасары К. Пирсон дамытқан.

Регрессиялық талдау- бір немесе бірнеше белгілердің «факторлық белгілердің) және салдардың (нәтижелі белгілердің) арасындағы байланысты өлшеуге мүмкіндік беретін статистикалық өңдеу әдісі.

Белгі – бұл негізгі ерекшелік, зерттелген құбылыстың немесе үдерістің ерекшелігі.

Нәтижелік белгі – зерттелелетін көрсеткіш.

Факторлық белгі - нәтижелік белгінің мәніне әсер ететін көрсеткіш.

Регрессиялық талдаудың мақсаты регрессия теңдеуі түрінде берілген орташа мәннің (y) нәтижелік белгісінің, (x_1, x_2, \dots, x_n), факторлық белгіге функционалдық байланысын бағалау болып табылады.

$$y = f(x_1, x_2, \dots, x_n)$$

Регрессия екі түрге бөлінеді: жұпталған және көпше.

Жұпталған (қарапайым) регрессия:

$$y = f(x) - \text{түрдегі теңдеу}$$

Жұпталған регрессия бойынша нәтижелі белгі бір аргументтен алынған функция сияқты, яғни бір факторлы белгіден алынады.

Регрессиялық талдау келесі кезеңдерден тұрады:

- функция түрін анықтаудан;
- регрессия коэффициенттерін анықтаудан;
- нәтижелі белгінің теориялық мәндерін есептеуден;
- регрессия коэффициентінің статистикалық маңыздылығын тексеруден;
- регрессия теңдеуінің статистикалық маңыздылығын тексеруден.

Көпше регрессия:

$$y = f(x_1, x_2, \dots, x_n) - \text{түрдегі теңдеу.}$$

Нәтижелі белгі бірнеше аргументтен алынған функция сияқты, яғни бірнеше факторлы белгіден алынады.

2. Функцияның түрін дұрыс анықтау үшін теориялық берілгендер негізінде байланыстың бағытын табу керек.

Байланыстың бағыты бойынша регрессияның бөлінуі:

- *түзу регрессия*, тәуелсіз «x» шамасының артуы немесе кемуіне сәйкес тәуелді «y» шамасы артатын немесе кемитін шартқа байланысты пайда болады.
- *кері регрессия*, тәуелсіз «x» шамасының артуы немесе кемуіне сәйкес тәуелді «y» шама сәйкесінше кемитін немесе артатын шартқа байланысты пайда болады.

Байланысты сипаттау үшін келесі жұпталған регрессия теңдеулерінің түрлерін қолданады:

- $y=a+bx$ – **сызықтық**;
- $y=e^{ax+b}$ – **экспоненциалды**;
- $y=a+b/x$ – **гиперболалық** ;
- $y=a+b_1x+b_2x^2$ – **параболалық** ;
- $y=ab^x$ – **көрсеткіштік** және т.б.

мұнда a, b_1, b_2 – теңдеудің коэффициенттері (параметрлері); y – нәтижелі белгі; x – факторлық белгі.

3. Регрессия теңдеуін құру оның коэффициенттерін (параметрлерін) бағалауға әкеп соқтырады, ол үшін ең кіші квадраттар әдісін қолданады (ЕКӨ).

ЕКӨ - нақты нәтижелік белгі «у» мәнінің « y_x » теориялық мәнінен ауытқу квадратының қосындысы ең аз болатын параметрлерді бағалауға мүмкіндік береді, яғни:

$$\sum (y - y_x)^2 \rightarrow \min$$

Ең кіші квадраттар әдісі бойынша $y=a+bx$ регрессия теңдеуінің параметрлерін анықтау формуласы:

$$a = \bar{y} - b\bar{x}, \quad b = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}$$

мұндағы a – еркін коэффициент, b – регрессия коэффициенті, бірлік өлшемде факторлық белгі « x » өзгергенде, нәтижелі белгі «у» қаншаға өзгертіндігін көрсетеді.

4. Регрессия коэффициентінің маңыздылығын бағалау үшін Стьюдент t -белгісі қолданылады.

Регрессия коэффициентінің маңыздылығын тексерудің сызбасы:

1) $H_0: a=0, b=0$ - регрессия коэффициенттері елеусіз, нөлден өзгеше. $H_1: a \neq 0, b \neq 0$ - регрессия коэффициенттері елеулі, нөлден, өзгеше.

2) $p=0,05$.

$$3) \quad t_{b \text{ расч}} = \frac{b}{m_b}, \quad t_{a \text{ расч}} = \frac{a}{m_a},$$

мұндағы m_b, m_a - кездейсоқ қателіктер.

$$m_b = \sqrt{\frac{\sum (y - y_x)^2}{n-2} \cdot \frac{1}{\sum (x - \bar{x})^2}}; \quad m_a = \sqrt{\frac{\sum (y - y_x)^2}{n-2} \cdot \frac{\sum x^2}{n \sum (x - \bar{x})^2}}.$$

4) $t_{табл}(p; f)$, мұндағы $f=n-k-1$ – еркіндік дәрежесінің саны (кестелік мән), n – бақылау саны, k – теңдеудегі « x » айнымалы болғандағы параметр саны.

5) Егер $t_{есен} > t_{кестм}$ болса, онда H_0 қабылданбайды, яғни коэффициент маңызды.

Егер $t_{есен} < t_{кестм}$ болса, онда H_0 қабылданады, яғни коэффициент маңызды емес.

5. Тұрғызылған регрессия теңдеуінің дұрыстығын тексеру үшін Фишер белгісі қолданылады:

Регрессия теңдеуінің маңыздылығын тексерудің сызбасы:

1) H_0 : регрессия теңдеуі маңызды емес.

H_1 : регрессия теңдеуі маңызды.

2) $p=0,05$.

$$3) F_{расч} = \frac{\sum (y_x - \bar{y})^2}{n - k - 1} = (n - 2) \frac{r_{xy}^2}{1 - r_{xy}^2},$$

мұнда n - бақылау саны; k - теңдеудегі « x » айнымалы болғандағы параметр саны; y – нәтижелі белгінің нақты мәні; y_x – нәтижелі белгінің теориялық мәні; r_{xy} - жұпталған корреляция коэффициенті.

$$4) F_{кест}(p; f_1; f_2),$$

мұнда $f_1=k$, $f_2=n-k-1$ - еркіндік дәрежесінің саны (кестелік мән).

5) Егер $F_{есеп} > F_{кест}$, онда регрессия теңдеуі дұрыс таңдалған және тәжірибеде қолдануға болады.

Егер $F_{есеп} < F_{кест}$, онда регрессия теңдеуі дұрыс таңдалмаған.

6. Регрессия талдаудың сапалық өлшемінің негізгі көрсеткіші *детерминация коэффициенті* (R^2) болып табылады.

Детерминация коэффициенті « y » айнымалының қандай бөлігі талдауда ескерілгендігін және талдауға енгізілген фактордың туғызатын әсерін көрсетеді.

Детерминация коэффициенті (R^2) [0,1] аралығында мәндерді қабылдайды. Егер $R^2 \geq 0,8$ болса, регрессия теңдеуі сапалы болып табылады.

Детерминация коэффициенті корреляция коэффициентінің квадратына тең, яғни $R^2 = r_{xy}^2$

7. Мысал 1. Төмендегі берілгендер бойынша регрессия теңдеуін тұрғызу және талдау:

1000 тұрғынға келетін тұмаумен ауыру, x	352	228	340	300	196	258	237
1000 тұрғынға келетін пневмониямен ауыру, y	64	60	52	48	46	41	32

Шешуі:

1) Корреляция коэффициентін есептеу: $r_{xy} = 0,47$. Белгілер арасындағы байланыс түзу және бір қалыпты.

2) Жұпталған сызықты регрессия теңдеуін құру.

2.1) Есептелген кесте құру.

№	X	y	xy	x^2	y_x	$(y_x - \bar{y})^2$	$(y - y_x)^2$	$(x - \bar{x})^2$
1	352	64	22528	123904	55,89	47,54	65,70	6241
2	228	60	13680	51984	45,07	15,42	222,83	2025
3	340	52	17680	115600	54,85	34,19	8,11	4489
4	300	48	14400	90000	51,36	5,55	11,27	729
5	196	46	9016	38416	42,28	45,16	13,84	5929
6	258	41	10578	66564	47,69	1,71	44,77	225
7	237	32	7584	56169	45,86	9,87	192,05	1296
Қосынды	1911	343	95466	542637	343	159,45	558,55	20934
Орташа	273	49	13638	77519,6	49	22,78	79,79	2990,6

2.2) Регрессия коэффициенттерін есептеу:

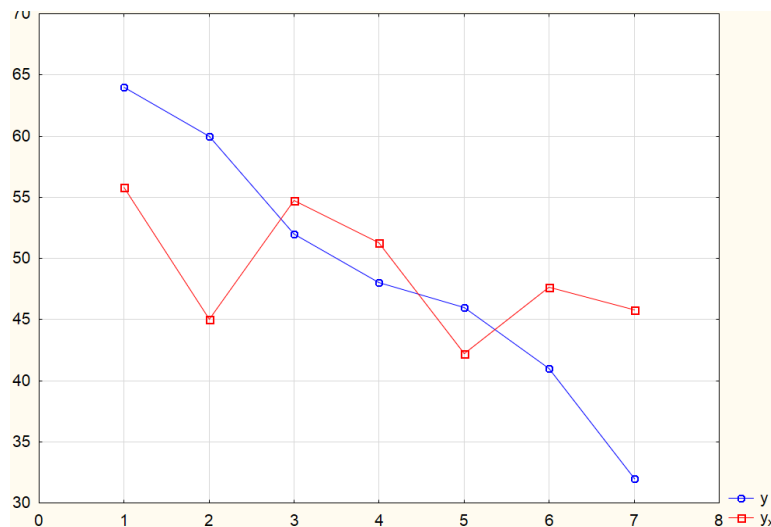
$$b = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{x^2 - \bar{x}^2} = \frac{13638 - 49 \cdot 273}{77519,6 - 273^2} = 0,087,$$

$$a = \bar{y} - b\bar{x} = 49 - 0,087 \cdot 273 = 25,17.$$

Жұпталған сызықты регрессия теңдеуі: $y = 25,17 + 0,087x$

3) «x» нақты мәндерді регрессия теңдеуіне қою арқылы «y_x» теориялық мәндерін табу.

4) Нәтижелі белгінің нақты «y» және теориялық «y_x» мәндері бойынша сызба тұрғызу:



Нәтижелі белгінің нақты «y» және теориялық «y_x» мәндері бойынша графигі.

5) Регрессия коэффициенттерінің статистикалық маңыздылығын тексеру:

5.1) Кездейсоқ қателіктерді есептеу:

$$m_b = \sqrt{\frac{\sum (y - y_x)^2}{n - 2}} = \sqrt{\frac{558,55}{7 - 2}} \approx 0,073;$$

$$m_a = \sqrt{\frac{\sum (y - y_x)^2}{n - 2} \cdot \frac{\sum x^2}{n \sum (x - \bar{x})^2}} = \sqrt{\frac{79,8}{7 - 2} \cdot \frac{542637}{7 \cdot 2990,6}} \approx 20,34.$$

$$5.2) t_{b \text{ есен}} = \frac{b}{m_b} = \frac{0,087}{0,073} \approx 1,19,$$

$$t_{a \text{ есен}} = \frac{a}{m_a} = \frac{25,17}{20,34} \approx 1,24.$$

$$5.3) t_{кесм}(0,05; 5) = 2,57$$

5.4) $t_{b \text{ есен}} < t_{кесм}$, яғни b коэффициенті – маңызды емес,

$t_{a \text{ есен}} < t_{кесм}$, яғни a коэффициенті – маңызды емес.

6) Регрессия теңдеуінің статистикалық маңыздылығын тексеру:

$$6.1) F_{есен} = \frac{\sum (y_x - \bar{y})^2}{\sum (y - y_x)^2} = \frac{159,45}{558,55} = 1,43.$$

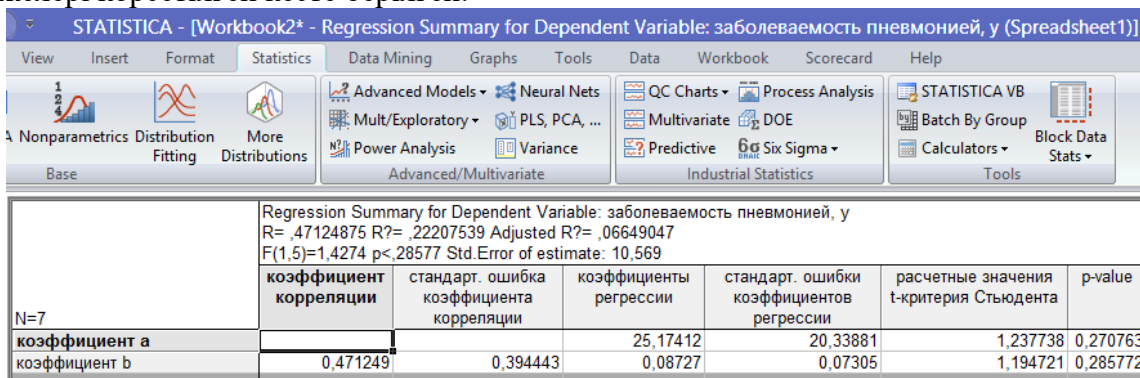
$$6.2) F_{кесм} (p; k; n-k-1) = (0,05; 1; 5) = 6,61.$$

6.3) $F_{есен} < F_{кесм}$, онда регрессия теңдеуі дұрыс таңдап алынбаған. Бұл нәтижені орташа тәуелділік ($r_{xy} = 0,47$) және бақылау санының аз болуымен түсіндіруге болады.

7) Детерминация коэффициентін есептеу: $R^2 = (0,47)^2 = 0,22$. Тұрғызылған теңдеу сапалы емес.

Себебі, регрессиялық талдау жүргізгенде есептеу көлемді болғандықтан, арнайы бағдарламаларды қолдану ұсынылады (Statistica 10, SPSS және т.б.).

Суретте «Statistica 10» бағдарламасының көмегімен регрессиялық талдаудың нәтижелері көрсетілген кесте берілген.



Regression Summary for Dependent Variable: заболеваемость пневмонией, y
 R= ,47124875 R²= ,22207539 Adjusted R²= ,06649047
 F(1,5)=1,4274 p<.28577 Std.Error of estimate: 10,569

	коэффициент корреляции	стандарт. ошибка коэффициента корреляции	коэффициенты регрессии	стандарт. ошибки коэффициентов регрессии	расчетные значения t-критерия Стьюдента	p-value
N=7						
коэффициент а			25,17412	20,33881	1,237738	0,270763
коэффициент b	0,471249	0,394443	0,08727	0,07305	1,194721	0,285772

4. Иллюстрациялық материалдар: көрме, слайдтар.

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы. -Эверо, 2014.
2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.
3. Бөлешов М.Ә. Медициналық статистика: оқулық. - Эверо, 2015

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013
2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)
2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары:

1. Регрессиялық талдау қандай кезеңдерден тұрады?
2. Регрессияның қандай түрлерін білесің?

3. Жұпталған сызықтық регрессия теңдеуінің коэффициенттері қалай анықталады?
4. Регрессия коэффициентінің маңыздылығы қалай тексеріледі?
5. Регрессия теңдеуінің маңыздылығы қалай тексеріледі?

№9 дәріс

1. Тақырыбы: Сапалы белгілерді талдау.

2. Мақсаты: Түйіндес кестелерді және екі тәуелсіз популяцияның пропорциясын қолдануды ережесімен таныстыру

3. Дәріс тезистері:

Белгі – байқау нысанының өлшенген қасиеті.

Белгілер:

- 1) сандық – бойдың ұзындығы; салмақ; науқастың жасы; жүрек соғу жиілігі және т.б.
- 2) сапалық – диагноз; шаштың түсі; науқастың жынысы; нәтиже (тірі, өлі) және т.б.

Сапалық белгілерді талдағанда **түйіндес кестелерді** пайдаланады.

Екі «*A*» және «*B*» сапалы белгілері болсын. «*A*» белгісінің сатылары (градация) - *r*, ал «*B*» белгісінің сатылары - *s*.

Көлемі *n* таңдама берілген.

Таңдаманың әрбір нысаны «*A*» белгісінің деңгейлерінің біреуіне және бір мезгілде қандайда бір «*B*» белгісінің деңгейіне ие болуы мүмкін. Осы таңдама арқылы *A_i* және *B_j* оқиғаларының жиіліктерін жеке және кез келген комбинацияда анықтауға болады.

A_iB_j. Оқиғасының жиілігі *v_{ij}* белгіленген. *A_i* белгісінің пайда болу саны (*A_i* белгісінің жиілігі) $v_{i.} = \sum_{j=1}^s v_{ij} = v_{i1} + v_{i2} + \dots + v_{is}$ тең болады.

B_j белгісінің пайда болу саны $v_{.j} = \sum_{i=1}^r v_{ij} = v_{1j} + v_{2j} + \dots + v_{rj}$ тең болады.

Жалпы бақылау саны, яғни таңдама көлемі: $v_{..} = \sum_{i=1}^r v_{i.} = \sum_{j=1}^s v_{.j} = \sum_{i=1}^r \sum_{j=1}^s v_{ij}$

Жиіліктердің (*v_{i.}*, *v_{.j}*, *v_{..}*) индексдерін нүктелерімен алмастыру осы индекс бойынша қосындының нәтижесін білдереді. Алынған жиіліктер түйіндес кесте түрінде берілген:

	<i>B₁</i>	<i>B₂</i>	...	<i>B_j</i>	...	<i>B_s</i>	
<i>A₁</i>	<i>v₁₁</i>	<i>v₁₂</i>	...	<i>v_{1j}</i>	...	<i>v_{1s}</i>	<i>v_{1.}</i>
<i>A₂</i>	<i>v₂₁</i>	<i>v₂₂</i>	...	<i>v_{2j}</i>	...	<i>v_{2s}</i>	<i>v_{2.}</i>
...
<i>A_i</i>	<i>v_{i1}</i>	<i>v_{i2}</i>	...	<i>v_{ij}</i>	...	<i>v_{is}</i>	<i>v_{i.}</i>
...
<i>A_r</i>	<i>v_{r1}</i>	<i>v_{r2}</i>	...	<i>v_{rj}</i>	...	<i>v_{rs}</i>	<i>v_{r.}</i>
	<i>v_{.1}</i>	<i>v_{.2}</i>	...	<i>v_{.j}</i>	...	<i>v_{.s}</i>	<i>v_{..}=n</i>

Түйіндес кестелерді талдау кезінде нөлдік болжам келесі түрде беріледі: $H_0: A$ және B белгілер арасында байланыс жоқ.

Түйіндес кестенің әрбір ұяшығы үшін теориялық жиіліктер есептеледі (яғни $A_i B_j$ комбинацияларының әрқайсына).

$$v_{ij}^* = v_{.i} \cdot \frac{v_{.j}}{v_{..}}$$

	B1	B2	...	Bj	...	Bs	
A1	v_{11}^*	v_{12}^*	...	v_{1j}^*	...	v_{1s}^*	$v_{1.}^* = v_{1.}$
A2	v_{21}^*	v_{22}^*	...	v_{2j}^*	...	v_{2s}^*	$v_{2.}^* = v_{2.}$
...
Ai	v_{i1}^*	v_{i2}^*	...	v_{ij}^*	...	v_{is}^*	$v_{i.}^* = v_{i.}$
...
Ar	v_{r1}^*	v_{r2}^*	...	v_{rj}^*	...	v_{rs}^*	$v_{r.}^* = v_{r.}$
	$v_{.1}^* = v_{.1}$	$v_{.2}^* = v_{.2}$...	$v_{.j}^* = v_{.j}$...	$v_{.s}^* = v_{.s}$	$v_{..}$

H_0 болжамы орындалған кезде бақыланатын v_{ij} жиіліктері v_{ij}^* теориялық жиіліктерінен айырмашылығы көп болмау керек.

Теориялық және бақыланатын жиіліктерді салыстыру үшін *Пирсонның χ^2 -белгісі* қолданылады.

Пирсонның χ^2 - белгісі

Бұл белгі, егер түйіндес кестесінің ұяшықтарында жиіліктер 5-тен жоғары немесе тең болғанда сапалық сипаттамаларды талдау үшін қолданылады.

Егер байқау саны аз болса және кесте торындағы ұяшықтардың жиілігі 5-тен аз болса, онда χ^2 өлшемі белгісі қолданылмайды. Оның орнына Фишердің нақты белгісі қолданылады.

Пирсонның χ^2 – келісім белгісін қолданудың тізбесі

1) H_0 : белгілердің арасында байланыс жоқ.

H_1 : белгілердің арасында байланыс бар.

2) $p=0,05$ – маңыздылық деңгейі

3) $\chi^2_{есеп} = \sum_{i=1}^r \sum_{j=1}^s \frac{(v_{ij} - v_{ij}^*)^2}{v_{ij}^*}$, мұндағы v_{ij} – бақыланатын жиіліктер, v_{ij}^* – теориялық жиіліктер

4) $\chi^2_{кесте(p, f)}$ мұндағы $f = (r-1)(s-1)$ - еркіндік дәреже саны (2 қосымшаны қара).

5) Егер $\chi^2_{есеп} \leq \chi^2_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $\chi^2_{есеп} > \chi^2_{кесте}$ болса, онда « H_0 » қабылданбайды.

Мысалы: Жедел ірінді өкпенің зақымдалуының төрт формасы үшін бақыланған және қайтыс болғандар саны берілген.

Пирсонның χ^2 –белгісі арқылы топтар арасындағы қайтыс болғандар саны бойынша айырмашылықтың маңыздылығын бағалау керек.

Топ нөмері	Арудың формасы	Оқиға саны		Ауырғандар саны
		қайтыс болғандар	сауыққандар	
1	Ірінді абсцесс	5	136	141



2	Гангреналық абсцесс	11	37	48
3	Бөлікті гангрена	7	8	15
4	Жаппай гангрена	6	5	11

Шешуі.

1) H_0 : белгілердің арасында байланыс жоқ.

H_1 : белгілердің арасында байланыс бар.

2) $p=0,05$ – маңыздылық деңгейі

$$3) v_{ij}^* = v_{.i} \cdot \frac{v_{.j}}{v_{..}}$$

3.1) Теориялық жиіліктерді есептеу

	B1	B2	Барлығы
A1	5	136	141
A2	11	37	48
A3	7	8	15
A4	6	5	11
Барлығы	29	186	215

	B1	B2	Барлығы
A1	$29 \cdot 141 / 215 = 19$	$186 \cdot 141 / 215 = 122$	141
A2	$29 \cdot 48 / 215 = 6,5$	$186 \cdot 48 / 215 = 41,5$	48
A3	$29 \cdot 15 / 215 = 2$	$186 \cdot 15 / 215 = 13$	15
A4	$29 \cdot 11 / 215 = 1,5$	$186 \cdot 11 / 215 = 9,5$	11
Барлығы	29	186	215

3.2) Шамасын есептеу
$$\frac{(v_{ij} - v_{ij}^*)^2}{v_{ij}^*}$$

	B1	B2
A1	$(5-19)^2/19=10,3$	$(136-122)^2/122=1,6$
A2	$(11-6,5)^2/6,5=3,1$	$(37-41,5)^2/41,5=0,5$
A3	$(7-2)^2/2=12,5$	$(8-13)^2/13=1,9$
A4	$(6-1,5)^2/1,5=13,5$	$(5-9,5)^2/9,5=2,1$

3.3) Есеп
$$\chi^2_{есеп} = \sum_{i=1}^r \sum_{j=1}^s \frac{(v_{ij} - v_{ij}^*)^2}{v_{ij}^*} = 10,3 + 3,1 + 12,5 + 13,5 + 1,6 + 0,5 + 2 + 2,1 = 45,5$$

4) $\chi^2_{кесте}(p, f)$, мұндағы $f=(r-1)(s-1)=(4-1)(2-1)=3$ – еркіндік дәрежесі саны $\chi^2_{кесте}(0,05;3) = 7,8$

5) Егер $\chi^2_{есеп} > \chi^2_{кесте}$ болса, онда « H_0 » қабылданбайды, яғни топтар арасындағы қайтыс болғандар саны бойынша айырмашылық маңызды.

Белгіні есептеу рәсімі өте күрделі, сондықтан талдауды арнайы (мысалы, «STATISTICA») бағдарламасы арқылы орындау тиімді.

Биология және медицинада 2×2 өлшем түйіндес кестелері (төрт өрісті кесте) өте жиі қолданылады.

Тексеруден өткен тұлғаларды сипаттайтын екі сапалық белгі бар (олардың әрқайсысында екі баға бар).

	Бірінші белгі (бірінші сатысы)	Бірінші белгі (екінші сатысы)	Барлығы
Екінші белгі (бірінші сатысы)	a -жиілігі	b -жиілігі	$a+b$
Екінші белгі (екінші сатысы)	c -жиілігі	d -жиілігі	$c+d$
Барлығы	$n_1=a+c$	$n_2=b+d$	$n=a+b+c+d$

Пирсонның χ^2 – белгісін 2×2 кестесі үшін қолданудың тізбесі

1) H_0 : бір белгінің кездесу жиілігі екінші белгінің кездесу жиілігіне байланысты емес.

H_1 : бір белгінің кездесу жиілігі екінші белгінің кездесу жиілігіне байланысты.

2) $p=0,05$ – маңыздылық деңгейі.

3) $\chi^2_{есеп} = \frac{(ad - bc)^2 n}{(a+b)(c+d)(a+c)(b+d)}$, мұндағы a, b, c, d - түйіндес кестесіндегі жиіліктер.

4) $\chi^2_{кесте}(p, f)$, мұндағы f - еркіндік дәрежесі. Ол 2×2 түйіндес кестелер үшін $f=1$.

5) Егер $\chi^2_{есеп} > \chi^2_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $\chi^2_{есеп} < \chi^2_{кесте}$ болса, онда « H_0 » қабылданбайды.

Мысалы: Контрацепті таблеткалар қабылдайтын және балалары сарыаурумен ауыратын, емшек сүтін емізетін аналар арасындағы байланыс зерттелді. Зерттеуге арналған деректер кестеде келтірілген.

Аналардың таблеткаларды қабылдауы	Сарыауру бар	Сарыауру жоқ	Барлығы
Таблетканы қабылдады	33	24	57
Таблетканы қабылдамады	14	45	59
Барлығы	47	69	116

Шешуі.

1. H_0 : балалардың сарыаурумен ауыруы анасының контрацепті препаратты қабылдауына тәуелді емес.

H_1 : балалардың сарыаурумен ауыруы анасының контрацепті препаратты қабылдауына тәуелді.

2. $p=0,05$ – маңыздылық деңгейі

3. $\chi^2_{есеп} = \frac{(ad - bc)^2 n}{(a+b)(c+d)(a+c)(b+d)} = \frac{(33 \cdot 45 - 24 \cdot 14)^2 \cdot 116}{57 \cdot 59 \cdot 47 \cdot 69} = 14,04$

4. $\chi^2_{кесте}(0,05;1) = 3,8$

5. $\chi^2_{есеп} > \chi^2_{кесте}$ болғандықтан, балалардың сарыаурумен ауыруы анасының контрацепті препаратты қабылдауына тәуелді болады.

Йетс түзетуі

2×2 кестесіндегі χ^2 үшін келтірілген формула жоғары мәндер береді. Іс жүзінде, бұл нөлдік болжамның тым жиі қабылданбауына алып келеді. Бұл әсерді болдырмау үшін формулаға Йетс

түзету енгізеді: $\chi^2_{есеп} = \frac{n \left(ad - bc - \frac{n}{2} \right)^2}{(a+b)(c+d)(a+c)(b+d)}$

Жоғарыда қарастырылған мысал үшін Йетс түзетуімен белгінің есептелген мәні:

$$\chi^2_{есен} = \frac{n \left(ad - bc - \frac{n}{2} \right)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{116 \left(33 \cdot 45 - 24 \cdot 14 - \frac{116}{2} \right)^2}{57 \cdot 59 \cdot 47 \cdot 69} = 12,66$$

Пирсонның χ^2 -белгісі **тәуелсіз** таңдамаларды салыстыру үшін қолданылады. Егер таңдамалар **тәуелді** болса, онда Макнемардің χ^2 -белгісі қолданылады.

Егер бірдей нысандарда екі байқау жасалса: «**дейін**» және «**кейін**», онда таңдамалар **тәуелді** деп аталады.

Макнемардің χ^2 -белгісі тек 2×2 өлшемді түйіндес кестелер үшін қолданылады.

«дейінгі» белгі	«кейінгі» белгі	
	«кейінгі» (-) екінші сатысы	«кейінгі» (+) бірінші сатысы
«дейінгі» (+) бірінші сатысы	a (+) тен (-) ке өзгеру саны	b (+) тің сақталу саны
«дейінгі» (-) екінші сатысы	c (-) тің сақталу саны	d (-) тен (+) ке өзгеру саны

Макнемардің χ^2 – белгісін қолданудың тізбесі

1) H_0 : фактор әсер еткеннен кейін белгі сатысының кездесу жиілігі өзгерген жоқ.

H_1 : фактор әсер еткеннен кейін белгі сатысының кездесу жиілігі өзгерді.

2) $p=0,05$ – маңыздылық деңгейі

$$3) \chi^2_{есен} = \frac{(|a-d|-1)^2}{(a+d)}$$

4) $\chi^2_{кесте}(p, f)$, мұндағы f - еркіндік дәрежесінің саны. Ол 2×2 ұқсас кестелер үшін $f=1$.

5) Егер $\chi^2_{есен} > \chi^2_{кесте}$ болса, онда « H_0 » қабылданады.

Егер $\chi^2_{есен} < \chi^2_{кесте}$ болса, онда « H_0 » қабылданбайды.

Мысалы: Асқынған аралас респираторлы вирустық инфекцияны кешенді емдеуде метаболикалық түрдегі пробиотиктің тиімділігі және оның ішек микробиоценозына әсері зертелді. Зерттеуге 32 науқас қатысты. Зерттеуге арналған деректер кестеде келтірілген.

Пробиотикалық емдеуге дейін	Пробиотикалық емдеуден кейін	
	Дисбактериоз жоқ	Дисбактериоз бар
Дисбактериоз бар	9	5
Дисбактериоз жоқ	18	0

Шешуі.

1) H_0 : пробиотикті қолданғаннан кейін дисбактериоздың жиілігі өзгерген жоқ.

H_1 : пробиотикті қолданғаннан кейін дисбактериоздың жиілігі өзгерді.

2) $p=0,05$ – маңыздылық деңгейі

$$3) \chi^2_{есен} = \frac{(|a-d|-1)^2}{(a+d)} = \frac{(|9-0|-1)^2}{(9+0)} = 7,11$$

$$4) \chi^2_{кесте}(0,05;1) = 3,8$$

5) $\chi^2_{есен} > \chi^2_{кесте}$ болғандық пробиотикті қолданғаннан кейін дисбактериозбен ауыратын науқастардың саны өзгермегені туралы болжам қабылданбайды.

4. Иллюстрациялық материал: Көрме, слайдтар.

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.

2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.

3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдістемелік құрал.–Алматы: Эверо, 2013

2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)

2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>

3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Ақнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары

1. Сапалы белгілерді талдаудың ерекшелігі неде?

2. 2x2 түйіндес кесте дегеніміз не?

3. Пирсонның χ^2 – белгісін қолдану кезінде қандай шарттарды сақтау керек?

4. Не үшін Йетс түзетулерін қолдану керек?

5. Макнемардің χ^2 – белгісі қандай жағдайларда қолданылады?

№ 10 дәріс

1. **Тақырыбы:** Динамикалық қатарларды талдау.

2. **Мақсаты:** Қандайда бір белгімен біртекті емес жиынтықтарды стандарттаудың тікелей әдісі бойынша салыстыру. Құбылыстардың заңдылықтары және үрдістері туралы қорытындыларға арналған динамикалық қатар деңгейлерін талдау.

3. **Дәріс тезистері:**

Динамикалық (уақытты) қатар – бірнеше реттелген кезеңнің немесе периодтардың кез-келген көрсеткіштерінің мәндер жиыны.



O'NTUSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	 SKMA -1979-	SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Медициналық биофизика және ақпараттық технологиялар кафедрасы «Биостатистика» пәні бойынша дәріс кешені	№ 35-11(Б)-2024 60 беттің 53 беті	

Кез келген уақытты қатар екі элементтен тұрады:

- келтірілген статистикалық берілгендер жататын **уақыт моментінен** немесе **кезеңдерінен** (t_i) тұрады;
- зерттелетін нысанды белгілі бір моментте немесе көрсетілген уақыт кезеңінде сипаттайтын **статистикалық көрсеткіштерінен** немесе **қатардың деңгейінен** (y_i) тұрады.

1. Динамикалық қатарлардың (уақыт бойынша) түрлері

- *Моменттік динамикалық қатарлар* зерттелетін құбылыс өлшемдерінің белгілі мерзімдегі (моментте) өзгеруін көрсетеді.

Күні	1.01.2017	1.04.2017	1.07.2017	1.10.2017	1.01.2018
Емхана қызметкерлерің саны (адам).	192	190	195	198	200

- *Аралық динамикалық қатарлар* зерттелетін құбылыс өлшемдерінің жеке кезеңдегі (уақыт аралығындағы) өзгеруін көрсетеді.

Жыл	2010	2011	2012	2013	2014
Қызылшаға қарсы вакцинацияланған балалар саны (мың адам)	88,5	93,2	98,0	102,8	108,8

2. Динамикалық қатарлардың (уақыт ішінде қамту толықтығы бойынша) түрлері

- **Толық** динамикалық қатарлардың аралығы бірдей болады.
- **Толық емес** динамикалық қатарлардың аралығы бірдей болмайды.

3. Динамикалық қатарлардың (қатар деңгейінің түрі бойынша берілуі) түрлері

- *Абсолютті мәндердің қатарлары* – қатар деңгейі сәйкес (кг, л, км, сағ, тг және т.б.) өлшем бірліктерімен беріледі.
- *Салыстырмалы мәндердің қатарлары* – қатар деңгейі сәйкес пайыз, бөліктер, промилелер және т.б. түрінде беріледі.
- *Орташа мәндердің қатарлары* – қатар деңгейі орташа көрсеткіштер болып табылатын сандармен беріледі.

Тренд – бұл дамудың уақыт ішіндегі негізгі үрдісін көрсететін уақыттан алынған функция.

Трендті орнату үшін динамикалық қатар түзетіледі.

Түзету келесі тәсілдермен іске асады:

Кезеңді ірілеу, топтық орташаны есептеу, жылжымалы орташаны есептеу, ең кіші квадраттар әдісі.

Зерттелетін құбылыстың динамикасын дәл сапалы бағалау үшін ең кіші квадраттар әдісі қолданылады: $\sum (y - y_t)^2 \rightarrow \min$, мұндағы, y - нақты (эмпирикалық) қатар деңгейі, y_t - қатар деңгейін түзететілген, яғни « t » уақыт кезеңінде аналитикалық теңдеуіне сәйкес есептелінген мәні.

Әдетте сызықты тренд құрады – бұл уақыт қатарын өзгерту үрдісін көрсететін түзу сызықты теңдеу $y_t = a + bt$ түрінде беріледі, мұндағы « a » және « b » – коэффициенттер, төмендегі формулалар бойынша есептелінеді:

$$a = \bar{y} - b\bar{t}, \quad b = \frac{\overline{yt} - \bar{y}\bar{t}}{\bar{t}^2 - \bar{t}^2}$$

$y_t = a + bt$ теңдеуге « t »-нің мәнін қойып, қатар деңгейінің түзетілген мәнін табады.

Құбылыстың уақыт ішіндегі дамуының жылдамдығын және жітілігін талдау деңгейлерді өзара салыстыру нәтижесінде алынатын статистикалық көрсеткіштер арқылы іске асады.

Сызықты тренд теңдеуі

Егер зерттелетін динамикалық қатардың есептеу уақытын, уақыт көрсеткіштерінің қосындысын $\sum t = 0$ нөлге тең болатындай етіп алса, онда теңдеулердің (a және b) параметрлерін анықтайтын жеңілдетілген формулалар алуға болады:

- егер қатарда тақ сандар болса $t_{max} = k - \frac{n+1}{2}$
- егер қатарда жұп сандар болса $t_{жұп} = 2k - (n+1)$

мұндағы k – жылдың реттік нөмірі, n – кезеңдегі жылдардың саны.

$\sum t = 0$ болғанда сызықты тренд теңдеуінің коэффициенттерін табу үшін $a = \frac{\sum y_i}{n}$,

$b = \frac{\sum y_i \cdot t_i}{\sum t_i^2}$ формулалары қолданылады.

Трендтік үлгі болып табылатын анықталатын функцияның есептелген параметрлеріне сәйкес, қажетті динамикалық қатар үшін түзу сызықты теңдеудің жазылуы: $y_t^* = a + bt$.

Бұл теңдеуге «t»-ның мәндерін қою арқылы **теүзетілген деңгейлерді (y_t^*) табуға болады.**

Егер есептеулер дұрыс орындалса, онда эмпирикалық қатардың мәндерінің қосындысы, түзетілген қатардың есептелген деңгейлерінің қосындысымен сәйкес келуі керек, яғни $\sum y_i = \sum y_t^*$.

Динамикалық қатардың **болжанатын мәндерінің** деңгейін анықтау үшін экстраполяция әдісі қолданылады.

Экстраполяция зерттелген қатардан тыс деңгейлерді табу, яғни өткен кезеңде байқалып отырған үрдістің болашағын кеңейту.

Іс жүзінде болжанатын құбылыстарды экстраполяциялаудың нәтижесі әдетте аралық бағалар түрінде алынады - бұл болжамның сенімді аралығы.

Аралықтардың шекараларын анықтау үшін $y_t^* \pm t_p S$, формуласы қолданылады.

Мұндағы y_t^* - «t» уақыт кезіндегі қатар деңгейінің болжанған мәнінің дәл бағасы, S – трендтен алынған қалдық орташа квадраттық ауытқу.

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n - m}}$$
, мұндағы n – динамикалық қатар деңгейлерінің саны, m – тренд

үлгісінің параметрлерінің саны (сызықты үшін $m=2$), t_p - маңыздылық деңгейі $p=0,05$ болғандағы Студенттің үлестіріміне сәйкес және еркіндік дәрежесі $f=n-m$ болғандағы сенімділік коэффициенті.

1 мысал.

Берілген динамикалық қатар негізінде:

- 1) ауруға шалдығу үрдісін көрсететін сызықты тренд теңдеуін құрастыру;
- 2) динамикалық қатардың түзетілген деңгейі бойынша сызбаға теориялық қисық тұрғызу және жалпы ауруға шалдығу үрдісінің сипаты туралы қорытынды жасау;
- 3) сенімділік ықтималдығы 95% болатын 2018 жылы жел шешекпен ауыратындардың

болжамын анықтау керек.

4)

Жыл	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
10 мың тұрғынға шаққанда жел шешекпен ауыру	3,5	4,9	3,6	5,7	6,5	5,5	8,1	7,2	5,0	7,3

Шешуі.

Есептеу кестесін құру

Жыл	Кезең (k)	Көрсеткіші (y _i)	Кезең (t _i)	y _i t _i	t ₂
2008	1	3,5	-9	-31,5	81
2009	2	4,9	-7	-34,3	49
2010	3	3,6	-5	-18	25
2011	4	5,7	-3	-17,1	9
2012	5	6,5	-1	-6,5	1
2013	6	5,5	1	5,5	1
2014	7	8,1	3	24,3	9
2015	8	7,2	5	36	25
2016	9	5,0	7	35	49
2017	10	7,3	9	65,7	81
Қосындысы		57,3	0	59,1	330

1) Сызықты тренд теңдеуінің коэффициенттерін анықтау:

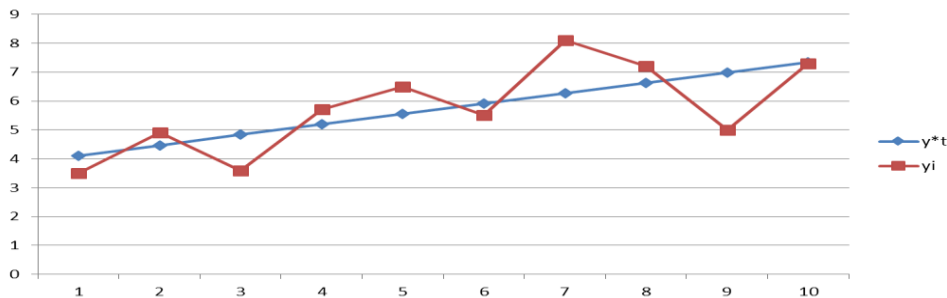
$$a = \frac{\sum y_i}{n} = \frac{57,3}{10} = 5,73 \quad b = \frac{\sum y_i \cdot t_i}{\sum t_i^2} = \frac{59,1}{330} = 0,18$$

Сызықтық тренд теңдеуі: $y_i^* = a + bt \Rightarrow y_i^* = 5,73 + 0,18 \cdot t$

2) Бұл теңдеуге «t» мәндерін қою арқылы (yt*) түзетілген деңгейлер анықталады.

Жыл	Кезең (k)	Көрсеткіші (y _i)	Кезең (t _i)	y _i t _i	t ₂	yt*
2008	1	3,5	-9	-31,5	81	4,11
2009	2	4,9	-7	-34,3	49	4,47
2010	3	3,6	-5	-18	25	4,83
2011	4	5,7	-3	-17,1	9	5,19
2012	5	6,5	-1	-6,5	1	5,55
2013	6	5,5	1	5,5	1	5,91
2014	7	8,1	3	24,3	9	6,27
2015	8	7,2	5	36	25	6,63
2016	9	5,0	7	35	49	6,99
2017	10	7,3	9	65,7	81	7,35
Қосындысы		57,3	0	59,1	330	57,3

Сызбада динамикалық қатардың деңгейі бойынша түзетілген теориялық қисық сызықты құру.



Қорытынды: 10 жыл ішіндегі алынған мәліметтер негізінде осы аймақтағы жел шешек ауруына шалдығу үрдісінің артқандығы туралы қорытынды жасауға болады.

3) 95% сенімділік ықтималдығымен 2018 жылға болжам құрастыру.

$$3.1) \text{ Мәнін анықтау: } S = \sqrt{\frac{\sum_{i=1}^n (y_i - y_t^*)^2}{n - m}}, \quad S = \sqrt{\frac{11,04}{10 - 2}} = 1,17$$

Жыл	Кезең (k)	Көрсеткіші (y _i)	Кезең (t _i)	y _i t _i	t ₂	yt*	y _i - y _t *	(y _i - y _t *) ²
2008	1	3,5	-9	-31,5	81	4,11	-0,61	0,37
2009	2	4,9	-7	-34,3	49	4,47	0,43	0,18
2010	3	3,6	-5	-18	25	4,83	-1,23	1,51
2011	4	5,7	-3	-17,1	9	5,19	0,51	0,26
2012	5	6,5	-1	-6,5	1	5,55	0,95	0,90
2013	6	5,5	1	5,5	1	5,91	-0,41	0,17
2014	7	8,1	3	24,3	9	6,27	1,83	3,35
2015	8	7,2	5	36	25	6,63	0,57	0,32
2016	9	5,0	7	35	49	6,99	-1,99	3,96
2017	10	7,3	9	65,7	81	7,35	-0,05	0,00
Қос-сы		57,3	0	59,1	330	57,3		11,04

t=11 уақытында динамикалық қатар деңгейінің болжанатын мәнінің нүктелік бағасын есептеу: $y_{11}^* = 5,73 + 0,18 * 11 = 7,71$

Кестеден сенімділік коэффициенті табу: $t(0,05; 8) = 2,31$

Болжанатын аралықтың шекараларын формула бойынша анықтау:

$$y_t^* \pm t_p S \Rightarrow 7,71 \pm 2,31 \cdot 1,17 \Rightarrow 5,01 \leq y_{\text{болжам}} \leq 10,41$$

Қорытынды: 95% ықтималдықпен, 2018 жылы осы аймақтығы жел шешекпен науқастану 10000 мың адамға шаққанда 5,01-ден кем емес және 10,41 адамнан артық болмайды деп айтуға болады.

Динамикалық қатардың көрсеткіштері

Құбылыстың уақыт ішіндегі дамуының жылдамдығын және жітілігін талдау, олардың деңгейлерін өзара салыстыру нәтижесінде алынатын статистикалық көрсеткіштер арқылы іске асады.

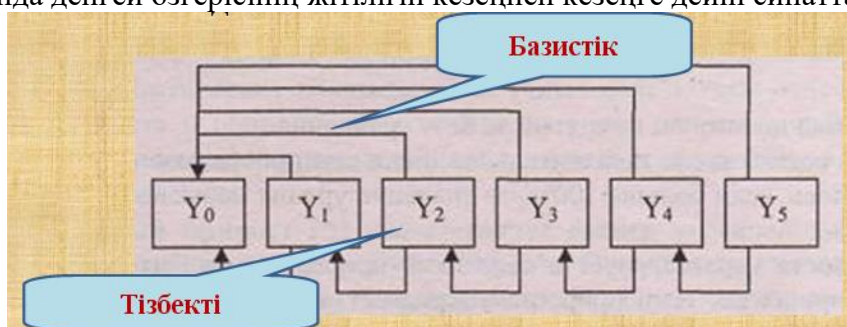
Мұндай көрсеткіштерге жататындар:

- абсолюттік өсу;
- өсудің шапшаңдығы;
- бір пайыз өсудің абсолюттік мәні және т.б.

Осы бойынша салыстыратын деңгейді **есептік**, ал салыстыру жүргізілетін деңгейді – **базалық (базистік)** деп атау қабылданған.

• **Тұрақты базалы (базисті)** динамикалық көрсеткіштер базалық деңгей жататын кезеңнен i -інші кезеңге дейінгі деңгейдегі барлық өзгерістердің соңғы нәтижелерін сипаттайды.

• **Айнымалы базалы (тізбекті)** динамикалық көрсеткіштер шектеулі зерттелетін уақыт аралығында деңгей өзгерісінің жітілігін кезеңнен кезеңге дейін сипаттайды.



Абсолютті өсу (Δ_i) – динамикалық қатардың екі деңгейінің айырмашылығы ретінде анықталынатын көрсеткіш. Ол қатардың берілген деңгейін салыстыру үшін база ретінде қабылданған деңгейден қаншалықты артық немесе кем екенін көрсетеді: $\Delta_i^{\sigma} = y_i - y_0$, мұндағы Δ_i^{σ} - абсолюттік базистік өсу; y_i – салыстырмалы кезеңнің деңгейі, y_0 – базистік кезеңнің деңгейі.

Айнымалы базамен салыстырғанда абсолюттік өсу $\Delta_i^T = y_i - y_{i-1}$ тең болады, мұндағы y_{i-1} – тікелей алдыңғы кезеңнің деңгейі.

Жыл	Көрсеткіші (y_i)	Δ_i^{σ} (2008 жылдың базасы)	Δ_i^T
2008	3,5	-	-
2009	4,9	4,9-3,5=1,4	4,9-3,5=1,4
2010	3,6	3,6-3,5=0,1	3,6-4,9=-1,3
2011	5,7	5,7-3,5=2,2	5,7-3,6=2,1
2012	6,5	6,5-3,5=3	6,5-5,7=0,8
2013	5,5	5,5-3,5=2	5,5-6,5=-1
2014	8,1	8,1-3,5=4,6	8,1-5,5=2,6
2015	7,2	7,2-3,5=3,7	7,2-8,1=-0,9
2016	5,0	5,0-3,5=1,5	5,0-7,2=-2,2
2017	7,3	7,3-3,5=3,8	7,3-5,0=2,3

Ұлғаю коэффициенті (k_i) екі салыстырмалы деңгейдің қатынасы ретінде анықталынады және берілген деңгейдің қаншалықты базистік кезең деңгейінен қанша рет артық екендігін көрсетеді: базистік $k_i^{\sigma} = \frac{y_i}{y_0}$, тізбекті $k_i^T = \frac{y_i}{y_{i-1}}$.

Жыл	Көрсеткіші (y_i)	k_i^{σ} (2008 жылдың базасы)	k_i^T
-----	----------------------	-------------------------------------	---------

2008	3,5	-	-
2009	4,9	4,9/3,5=1,40	4,9/3,5=1,40
2010	3,6	3,6/3,5=1,03	3,6/4,9=0,73
2011	5,7	5,7/3,5=1,62	5,7/3,6=1,58
2012	6,5	6,5/3,5=1,86	6,5/5,7=1,14
2013	5,5	5,5/3,5=1,57	5,5/6,5=0,85
2014	8,1	8,1/3,5=2,31	8,1/5,5=1,47
2015	7,2	7,2/3,5=2,06	7,2/8,1=0,89
2016	5,0	5/3,5=1,43	5/7,2=0,69
2017	7,3	7,3/3,5=2,09	7,3/5=1,46

Егер өсудің коэффициенттері пайызбен өрнектелген болса, онда оларды **ұлғаю шапшаңдығы** деп атайды, яғни олар көрсеткіштің пайызбен берілген жылдамдық өзгерісінің уақыт бірлігімен сипатталады: $T_{\text{үлғаю}} = k \cdot 100\%$

Жыл	Көрсет. (y_i)	k_i^{σ}	$T_{\text{лаю}}^{\sigma}$ %	k_i^T	$T_{\text{үлғаю}}^T$ %
2008	3,5	-	-	-	-
2009	4,9	4,9/3,5=1,40	140	4,9/3,5=1,40	140
2010	3,6	3,6/3,5=1,03	103	3,6/4,9=0,73	73
2011	5,7	5,7/3,5=1,62	162	5,7/3,6=1,58	158
2012	6,5	6,5/3,5=1,86	186	6,5/5,7=1,14	114
2013	5,5	5,5/3,5=1,57	157	5,5/6,5=0,85	85
2014	8,1	8,1/3,5=2,31	231	8,1/5,5=1,47	147
2015	7,2	7,2/3,5=2,06	206	7,2/8,1=0,89	89
2016	5,0	5/3,5=1,43	143	5/7,2=0,69	69
2017	7,3	7,3/3,5=2,09	209	7,3/5=1,46	146

Өсу шапшаңдығы берілген кезең деңгейінің базистік деңгейден қанша пайызға көп (немесе аз) екендігін көрсетеді. Бұл көрсеткіш екі түрде есептеледі.

- абсолютті өсудің базистік деңгейге қатынасы ретінде:

$$\text{базистік } T_{rcy}^{\sigma} = \frac{y_i - y_0}{y_0} \cdot 100\% ; \quad \text{тізбекті } T_{rcy}^T = \frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100\%$$

- ұлғаюдың шапшаңдығы (%) және 100% арасындағы айырмашылық ретінде:

$$T_{rcy} = T_{\text{үлғаю}} - 100\%$$

Өсудің бір пайыздық абсолюттік мәні (A_i) алынған өсудің шапшаңдығының мәнін дұрыс бағалау үшін пайдаланылады.

Тек тізбекті әдіспен есептеуге болады.

$$A_i = \frac{y_i - y_{i-1}}{T_{rcy}} = \frac{\Delta_i^T}{T_{rcy}}$$

Зерттелетін құбылыстың динамикасын жалпы сипаттауда кезең қатары үшін түрлі орташа көрсеткішті анықтайды.

- Орташа абсолюттік өсу** – уақыт аралығындағы көрсеткіштің орташа шамасының өзгеруі.

$$\bar{\Delta} = \frac{\sum_{i=1}^n \Delta_i^T}{n-1} = \frac{y_n - y_1}{n-1}$$

- **Ұлғаюдың орташа шапшандығы** – бұл динамикалық қатар деңгейлері жітілігінің өзгеруі сипаттамасы. Ол бірлік уақыт ішінде динамикалық қатар деңгейінің орташа қанша рет өзгертіндігін көрсетеді: $\bar{T} = \bar{k} \cdot 100\%$, мұндағы $\bar{k} = n-1 \sqrt{\frac{y_n}{y_1}}$ - ұлғаюдың орташа коэффициенті; n - қатар деңгейінің саны.
- **Өсудің орташа шапшандығы** келесі формуламен есептелінеді: $\bar{T}_{\text{ocy}} = \bar{T}_{\text{yлеаю}} - 100\%$

4. Иллюстрациялық материалдар: көрме, слайдтар.

5. Әдебиет:

- Негізгі:

1. Койчубеков Б.К. Биостатистикаға кіріспе курсы: оқу құралы.-Эверо, 2014.
2. Раманқұлова А.А. Биостатистика. –Ақ-Нұр, 2013.
3. Бөлешов М.Ә. Медициналық статистика: оқулық.- Эверо, 2015

- Қосымша

1. Мысалдар мен тапсырмалардағы биостатистика: оқу-әдіс. құрал.–Алматы: Эверо, 2013
2. Койчубеков Б.К., Букеева А.С. Мысалдар мен тапсырмалардағы биостатистика. /Оқу-әдістемелік құрал/ - Алматы, «Эверо», 2013.

- Электрондық ресурстар:

1. Биостатистика [Электронный ресурс]: оқулық/ Қ.Ж. Құдабаев [ж/б.].- Электрон. текстовые дан. (85,7Мб). - Шымкент: ОҚМФА, 2015. - 185 б. эл. опт. диск (CD-ROM)
2. Медициналық-биологиялық деректерді статистикалық талдауда excel және spss statistics бағдарламаларын қолдану. Чуудиновских В.Р. [ж/б.]. <https://aknurpress.kz/login>
3. Раманқұлова А.А. Биологиялық статистика. [Мәтін]: оқу құралы/ А. Раманқұлова.- 2-бас.- Алматы : Акнұр баспасы, 2019. <http://elib.kaznu.kz>

6. Бақылау сұрақтары:

1. Динамикалық қатар дегеніміз не? Ол қандай элементтерден тұрады?
2. Қандай динамикалық қатарларды білесіздер?
3. Тренд дегеніміз не?
4. Динамикалық қатар қандай жолымен түзетіледі?
5. Сызықты трендтің коэффициентері қалай анықталады?
6. Базистік және тізбекті көрсеткіштерін арасында қандай айырмашылық бар?
7. Қандай салыстырмалы динамикалық көрсеткіштерді білесіздер?
8. Қандай орташа динамикалық көрсеткіштерді білесіздер?

ОҢТҮСТІК-ҚАЗАҚСТАН

**MEDISINA
AKADEMIASY**

«Оңтүстік Қазақстан медицина академиясы» АҚ



SOUTH KAZAKHSTAN

**MEDICAL
ACADEMY**

АО «Южно-Казахстанская медицинская академия»

Медициналық биофизика және ақпараттық технологиялар кафедрасы

«Биостатистика» пәні бойынша дәріс кешені

№ 35-11(Б)-2024

60 беттің 60 беті