**Кіріспе**

«Статистика» терминінің көптеген мағынасы бар. Бұл әрі ғылым, әрі қызмет түрі, әрі сндық көрсеткіштер жиыны және т.б.

Статистика (statistics) – бұл деректерді (бақылау нәтижелерін) жинау, топтау, жүйелеу, талдау және түсіндіру жөніндегі ғылым. Бұл бас жиынтықты және олардың өзгергіштігін оқып білу.

Биостатистика (biostatistics) – жанды әлемге қолданылатын статистикалық ғылым. Ол биологиялық нысандарды табиғи және эксперименттік зерттеулерді жоспарлау және олардың нәтижелерін талдау ережелері жөніндегі білімдер жүйесі.

Статистикалық қорытындылар теориясы математика ғылымының маңызды бөлімдерінің бірі ықтималдықтар теориясымен тығыз байланысты. Себебі, статистикалық әдістерді қолдану нәтижесінде біз соңғы ақиқат нәтижені емес, айтылған пікірдің ықтималдық бағасын аламыз. Сондықтан оқу құралының кіріспе бөліммінде ықтималдылықтар теориясының негізгі ұғымдарымен таныс боламыз.

**Ықтималдықтар теориясының негізгі ұғымдары**

**Ықтималдық –** қандай – да бір шарттар жиынтығының жүзеге асырылуы нәтижесінде оқиғаның пайда болу мүмкіндігінің дәрежесін сипаттайтын сан.

А оқиғасының ықтималдығы р(А) арқылы белгіленеді және бірліктер үлесімен немесе пайызбен өрнектеледі.

**Ықтималдық өлшемі** – 0-ден 1-ге дейін немесе 0% - дан 100%-ға дейінгі оның сандық мәндерінің ауқымы.

**Кездейсок оқиға** – белгілі бір шарттар жиынтығының жүзеге асырылуы нәтижесінде пайда болуы немесе пайда болмауы мүмкін оқиға.

Оның ықтималдығы 0<Р(А)<1 немесе 0<Р(А)<100% арасында жатады.

Ақиқат оқиға белгілі бір шарттар жиынтығының жүзеге асырылуы нәтижесінде міндетті түрде пайда болатын оқиға. Оның ықтималдығы 1-ге немесе 100% - ға тең.

**Мүмкін емес оқиға** – белгілі бір шарттар жиынтығының жүзеге асырылуы нәтижесінде ешқашан пайдаболмайтын оқға. Оның ықтималдығы 0-ге тең болады.

Медициналық зерттеулерде 0,95 немесе 95%-дан кем емес ықтималдық *жеткілікті* деп саналады. Алайда, аса маңызды медициналық әлеуметтік салдары бар аурулар мен ахуалдарды немесе өлім-жітімділіктің және мүгедектіктің жоғары көрсеткіштерін зерттегенде, сол сияқты фармакологиялық зерттеулерде оқиғаның пайда болу ықтималдығы 0,99 (99%)-дан кем болауы тиіс.

**Оқиғаның пайда болу жиілігі** (статистикалық ықтималдық) – бұл белгілі шарттар жиынтығы жүзеге асқандағы жағдайлар санының (m) барлық жағдайлар санына (n) қатынасы:

Оқиғаның пайда болмау ықтималдығы: q=1-p

**Кездейсоқ шама** – белгілі бір шарттар жиынтығы жүзеге асқанда әртүрлі мәнді қабылдай алатын шама.

**Үлкен сандар заңы:** бақылалар саны жеткілікті үлкен болғанда кездейсоқ ауытқулар өзара жойылады және негізгі заңдылықты көрсететін кейбір параметрлердің тұрақтылығы көрінеді. Кездейсоқ оқиғаның бақыланып отырған жиілігі жеке сынауда оқиғаның пайда болу ықтималдығынан өте аз шамаға айырмасы болады.

Статистикалық талдаудың негіздерін оқып-үйренуге кірісер алдында оның *екі негізгі кезеңін* бөліп алу қажет:

* Зерттеу барысында алынған жиынын сипаттау;
* Деректерді талдау және әртүрлі статистикалық жорамалдарды тексеру.

Медицинада және денсаулық сақтауда биологиялық статистикалық әдістерді қолданудың негізгі бағыттары:

1. Деректерді мейлінше тиімді жинау және алынған нәтижелерді жалпылау.
2. Нәтижелердің екі және екіден көп топтарын салыстыру және айырмашылықтың шынайылығын бағалау.
3. Факторлар мен құбылыстардың арасындағы өзара байланысты зерттеу.

**1-Бөлім. Биостатистикадағы деректер құрылымы**

* 1. ***Деректер және олардың түрлері***

Зерттеу жұмыстарының негізгі мақсаты деректерді жинау және олардың көмегімен зерттеу жұмысының қандай да бір саласына қатысты шынайы ақарат алу.

Деректер жиынтықтағы әрбір нысанға тән белгілерді немесе қасиеттерді анықтайтын айырмалылар терминдерімен анықталады.

Жиынтықтағы әр нысан үшін тіркелген ақпараттар (мысалы, жынысы, бойының ұзындығы, дене салмағы, қан қысымы және т.б.) **айнымалы** деп аталады. "Айнымалы» термині өзгеріп отыруға қабілетті мөлшерлік көрсеткішті білдіреді.

Қарастыратын айнымалылардың санына қарай деректер жиынтығы *бір өлшемді, екі өлшемді және көп өлшемді* болуы мүмкін.

**Бірөлшемді** (бір айнымалы) деректер жиынтығы әрбір нысан үшін тіркелген тек бір ғана белгі жөніндегі ақпараттан тұрады. Мысалы, тек қана бойдың ұзындығы бойынша тіркелген ақпараттар. Бұл жағдайда статистикалық әдістер осы белгінің негізгі қасиеттерін жалпылау үшін қолданылады. Мысалы, орта мән есептеу, нысандар арасындағы айырмашылықты анықтау.

**Екіөлшемді** (екі айнымалы) деректер жиынтығы нысандардың әрқайсысы үшін тіркелген екі белгі жөніндегі ақпараттардан тұрады. Мысалы, бойдың ұзындығы және дене салмағы жөніндегі ақпараттар. Бұл жағдайда статистикалық әдістер әр белгінің қасиеттерін жалпылаумен қатар, екі белгі арасындағы байланысты зерттеу үшін қолданылады. Мысалы, сызықты байланысты анықтау, корреляция коэффициентін есептеу, регрессиялық талдау жүргізу.

**Көпөлшемді** (көп айнымалылар) деректер жиынтығы әр нысан үшін тіркелген үш немесе одан көп белгілер жөніндегі ақпараттардан тұрады. Мысалы, бойдың ұзындығы, дене салмағы және систолалық қан қысымы сияқты үш айнымалы бойынша тіркелген ақпараттар. Бұл жағдайда әр белгінің қасиеттерін жалпылау және жұп айнымалылар арасындағы байланысты анықтаумен қатар статистикалық әдісті барлық айнымалылар арасындағы өзара байланысты зерттеу үшін қолдануға болады. Мысалы, көпөлшемді талдау.

***Нақты бір жағдай үшін қандай статистикалық әдісті қолдану қажеттігін шешпей тұрып, біз әрбір айнымалының деректердің қай түріне жататындығын алдын – ала анықтап алуымыз қажет.***

Деректер **сапалық** немесе **сандық** болып екі түрге бөлінеді.

**Сапалық деректер** – жиынтықтағы әрбір нысанға тән белгілі бір сапаны сипаттайды.

Сапалық белгілерді **номинальды (категориальды)** және **реттік (ординальды)** деп екіге бөлуге болады.

* **Номинальды немесе категориальды деректер.**

Деректер атаулары белгінің мәнін немесе классификациялық нөмірін бейнелейтін категорияларға бөлінген. Әрбір категорияға түскен бақылаулар саны есептеледі. Категориялар қиылыспайды, яғни әрбір мән тек өана бір категорияға жатады.

Деректердің бұл түрін өлшеуге немесе мәндері бойынша реттеуге болмайды, олардың тек атаулары бар, оларға арифметикалық амалдар қолдануға болма йды. Категориялардың арасындағы қашықтықты анықтауға болмайды. Мысалы: қан тобы (А, В, АВ және 0), жыныс, көздің түсі, диагноз, отбасылық жағдай (үйленген, үйленбеген, жесір және т.б.) Бұл жағдайда «үйленген» «үйленбегеннен» жақсы (немесе жаман) деп айтға негіз жоқ.

* **Реттік (ординальды, рангілік) деректер.**

Деректер қиылыспайтын категорияларға жатады. Кейбір жағдайларда категориялар (градациялар, деңгейлер) реттеледі. Қатар тұрған категориялар арасындағы айырмашылық бірдей еме және олардың арасындағы арақашықтықты өлшеуге болмайды. Бұл түрдегі деректермен арифметикалық амалдар орындауға болмайды. Мысалы: аурудың кезеңдері (аурудыңасқынған сатылары, аурудың орта, бастапқы кезеңдері немесе аурудың болмауы); ауру дәрежесі (қатты ауыру, орташа, әлсіз, ауырудың болмауы).

* Номинальды деректердің тек қана екі қиылыспайтын категориядан тұратын түрін **бинарлы немесе дихотомиялық деректер** деп атайды.

Мысалы: «иә/жоқ», «пациент ауру/пациент сау», «1/0».

**Сандық деректер** – мәндері санмен өрнектелген айнымалылар.

Сандық деректер *екі түрге* бөлінеді: **дискертті** және **үздіксіз.**

* **Дискретті сандық деректер.**

Деректер тек қана белгілі бір анықталған сандық мәндер (көбіне оиғалар саны саналады) қабылдайды, әдетте олар бүтін сандар болады. Олардың арасында аралық мән болмайды. Бұл деректер үшін шама да, реттілік те маңызды.

Мысалы: балалар саны, пациенттер саны, дәрігерді жылына қатынау саны немесе соңғы бес жылдағы адамның ауру саны, жүкітлік саны, тамыр соғуының саны. Деректер – әрқашан санау нәтижесі болып табылады.

* **Үздіксіз сандық деректер**

Деректер берілген диапазондағы (кесінді) кез келген сандық мәні қабылдай алады. Мәндер саны шексіз, мәндер өлшеу дәлдігіне байланысты бөлшек құрамдас болуы мүмкін. Үтірден кейінгі бөлшек құрамындағы таңбалар саны өлшеу әдісінің дәлдігіне байланысты. Мысалы, дене салмағын килограммға дейінгі дәлдікпен немесе микрограммға дейінгі дәлдікпен анықауға болады. Үздіксіз сандық деректер өлшеу нәтижесінде аынады.

Мысалдар: дене температурасы, дене салмағы, бойдың ұзындығ, қан қысымы.

**Ескерту.** Кей жағдайларда (көбінесе) үздіксіз айнымалылар дискретті айнымалылар ретінде берілуі мүмкін. Мәселен, жас мөлшерін әдетте дискретті шама ретінде сипаттады, ал негізінде ол үздіксіз деректерге жатады. Әдетте біз «соңғы туған күндегі жас мөлшеріне» сүйенеміз, толығымен алсақ, «бүгінгі күнгі жас мөлшері» бойынша анықтауымыз қажет, сондықтан, егер әйел адам 50-демін десе, оған 50 жас және бірнеше ай, бірнеше күн немесе 51 жас дерлік болуы мүмкін. Мұндай ерекшеліктерді статистикалыық талдауға арналған компьютерлік бағдарламалардағы алгоритмдермен жұмыс істегенде ескеру қажет.

Деректер құрылымы

**Деректер**

**Сандық (мөлшерлік)**

**Сапалық (категориальдық)**

*Дискретті*

Бүтін санды мәндер

**Мысалы:** бір жылда ауру болған күндер саны

*Үздіксіз*

Мәндердің белгілі бір аралығындағы кез келген мәндер

**Мысалы:** дене салмағы, бойдың ұзындығы

*Реттік*

Бірін-бірі өзара жоққа шығаратын (альтернативті), реттелген атегориялар

**Мысалы:** аурудың кезеңдері (жеңіл/орта/ауыр)

*Номинальды*

Бірін-бірі өзара жоққа шығаратын (альтернативті), реттелмеген категориялар

**Мысалы:** жыныс (ер/әйел) қан тобы А/В/АВ/0

**1-сурет.** *Айнымалылардың әртүрлі түрлері*

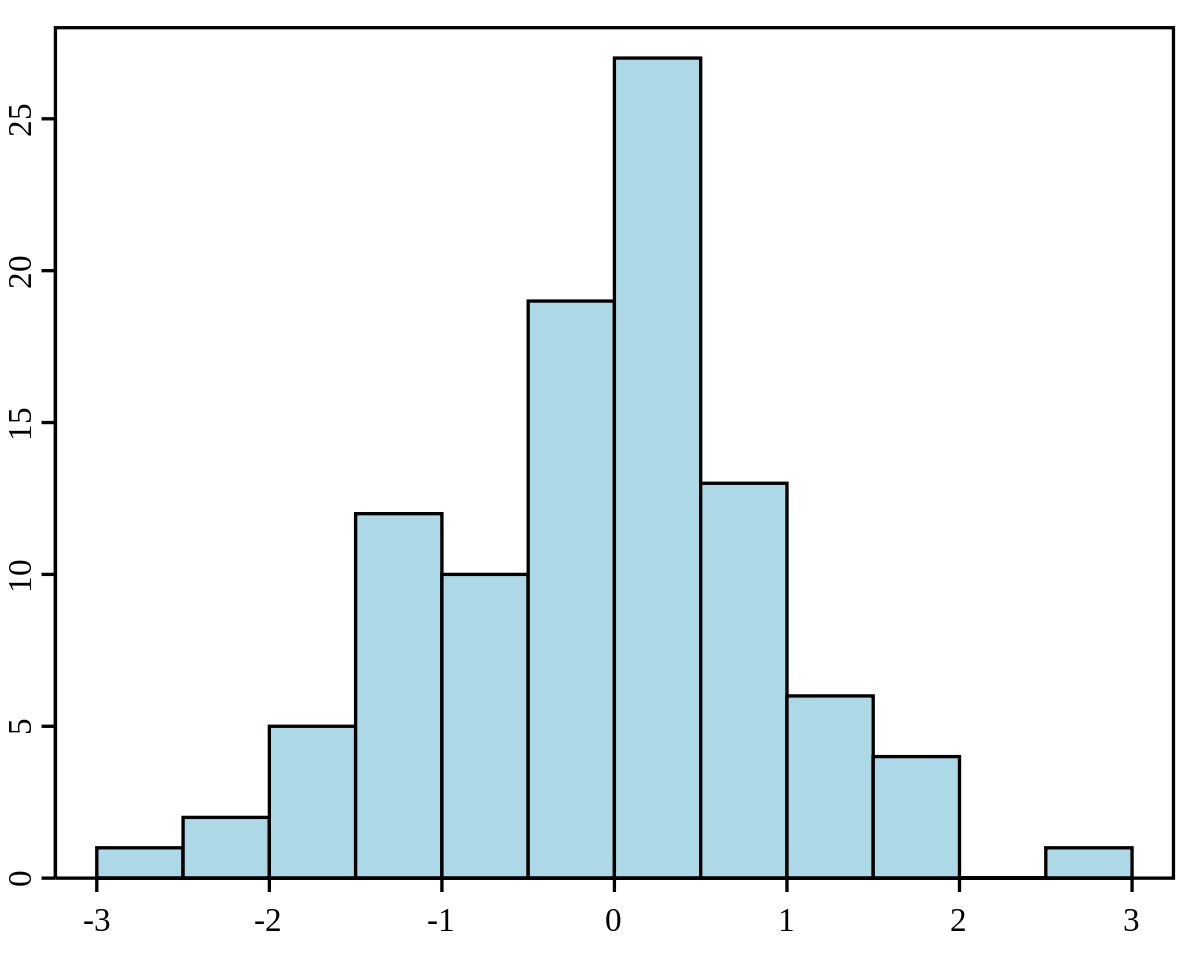
**Нүктелік график**

Нысандардың жеке мәндері бейнеленетін график. Оны бақылау саны 20-дан аз болатын үлкен емес деректер жиыны үшін қолданған орынды. Бір немесе бірнеше категориялар,дискретті айнымалының градациялары х осінде белгіленеді.

Бақыланған мәндердің диапазоны у осінде белгіленеді. Әр бақылаудың сандық мәні нүктемен бейнеленеді. Мәндері бірдей бақылаулар бір көлденең сызықтың бойында көрші орналасады. График деректердің орналасуы мен олардың өзгергіштігін көрсетеді.

**Гистограмма**

Гистограмма- бұл деректердің жиілікті тарралуының графиктік түрде кескнделуі. Ол бір-біріне іргелес орналасқан тіктөртбұрыштар түрінде белгіленеді. Әр тіктөртбұрыштың табанының ені топтау жиілігіне,яғни топтау аралығындағы бақылау санына пропорционал. Бір шыңнан тұратын гистограмма біртөбелі, екі шыңнан-екітөбелі үш немесе одан көп шыңнан-көп төбелі деп аталады.



**«Жапырақты сабақша» диаграммасы**

Бұл диаграмма мен кестенің қоспасы және ол деректердің таралуын суреттейді. Диаграмма бойынша деректер жиынының симметриялылығы жөнінде жоруға болады. Деректердің біртектілік дәрежесін сипаттайды. Әдетте тік сабақша тұрғызылады, ол деректердің ретімен келтірілген бірнеше бірінші цифрларынан тұрады.

Бұл сабақшадан сыртқа шығып тұрған жапырақтар- барлық деректердің өсу ретімен көлденең орналастырылған соңғы цифрлары.

1-мысал: Егер 2 жасар ұл балалардың бойларының ұзындығы жөнінде деректер берілсе: 87,89,90,91,91,92,92,93,93,94,94,94,94 онда мына түрде бейнелеуге болады:

8 79

9 011222334444

**«Мұртты жәшік» диаграммасы**

Бұл екі параллель қабырғалары деректердің жоғарғы және төменгі квартильдеріне сәйкес келетін тік немесе көлденең орналасқан тіктөртбұрыш.Жәшіктің ішіне көлденең жүргізілген сызық медианаға сәйкес келеді.Жәшік квартильаралық диапазонды қамтиды. Шеткі нүктелер ең төменгі және ең жоғарғы мәндерді білдіреді. Барлық шығарылып тасталған мұрттардан әрі орналасады және бірыңғай аномальды мәндер үлкен қашықтықта орналасады.

**Жиіліктер полигоны**

Дискретті қатарларды графиктік түрде кескіндеу үшін жиілікер полигоны қолданылады.Полигон-белгілердің мәндері мен олардың жиынтықтағы кездесу жиіліктері арасындағы сәйкестікті бейнеледі.Полигонды тұрғызу үшін абсцисса осіне зерттеліп отырған белгінің мәндері,ордината осіне жиіліктер салынады және сәйкес нүктелер кесінділер арқылы қосылады,нәтижесінде кесінділерден тұратын сынық сызық пайда болады.Жиіктер полигондарын бір-біріне үстемелеп салуға болады,ол екі немесе одан да көп деректер жиынын салыстаруға мүмкіндік береді.

**Екі өлшемді шашырау диаграммасы**

Бұл айнымалылар осьтері бір-біріне перпендикуляр орналасқан екі өлшемді график.Бір ғана нысанның жұп көрсеткіштері графикте координаталары х және у болатын нүктелермен бейнеледі.Диаграмма екі координатасы да үздіксіз өлшеу шкаласында берілген екі өлшемді сандық шамалар үшін қолдануға болады.Диаграмма трендтерді,шамалар арасындағы тәуелділікті,шамалардың біреуінің немесе екеуінің де көп жинақталған орнын,бір шаманың екінші шамаға қатысты шашырау өзгерттігін байқауға мүмкіндік береді.

**2-Бөлім.Таңдама әдіс**

***2.1.Бас жиынтық және таңдама***

Қандайда бір сапалық немесе сандық белгілермен сипатталатын нысандар жиыны статистикалық жиынтық деп аталады.

Статистикада ең маңызды ұғымдардың бірі **бас жиынтық** болып саналады.

**Бас жиынтық немесе ппопуляция-**бұл зерттелуге жататын барлық деректерді,құбылыстарды,нысандары немесе адамдарды қамтитын жиын.Бас жиынтық шектелген немесе шектелмеген,нақты немесе жормладаған болуы мүмкін,Мысалы,нәтижелері бойынша қант диабетімен ауыратын адамдарда қандағы фибриногеннің деңгейінің жоғарылауы дами ма деген сұраққа жауап беретын зерттеу жүргізу қажет болсын.Бұл тәрізді мәселе өте кең қойылған,сондықтан бұл жағдайда бас жиынтық ретінде жер бетіндегі қант диабетімен ауыратынбарлық аурулардағы фибриногеннің деңгейін анықтау мүмкін емес,бірақ бұл көрсеткіштікті эндокринологиялық бөлімде қандай-да бір уақыт аралығында емделген ауруларда анықтауға олады.Яғни, жаппай бақылау жүргізуге мумкін емес.

Сондықтан,таңдама жиынтықтарды зерттеу арқылы бас жиынтық жөнінде қортытынд жасауға мүкіндік беретін статистикалық әдістер қолданылады.Мұндай әдісті **таңдама әдіс** деп атайды.

Таңдаманы зеррттеу нәтижелерін барлық популяцияға жалпылау үрдісі **статистикалық қорытынды** деп аталады.

**Таңдама жиынтық** немесе жай таңдама деппбас жиынтықтан кездейсоқ таңдалынып алынған нысандар бөлігін айтады.

Бас жиынтықтың немесе таңдаманың нысандарр саны-бас **жиынтықтың** немесе **таңдаманың көлемі** деп аталады және сәйкес N,**n** арқылы белгіленеді.

Мысалы:Егер 10 000 студенттерден флюорографиялық тексеру үшін 100 студент кездейсоқ таңдалынып алынған болса,онда бас жиынтықтын көлемі 10 000, алтаңдаманың көлемі-100 болады.

Таңдама көлемі неғұрлым үлкен болса, соғұрлым таңдама жиынтықтың қасиеттері бас жиынтықтың сәйкес қасиеттерін шынайы көрсетеді.Мысалы,белгілі құрылымды аппараттың көмегімен шығатын дәрінің құрамындағы қандай да бір заттың коонцентрациясын білу керек болса,онда қаншалықты көп кездейсоқ алынған дәріні зерттесек,алынатын ақпарат соншалықты дәлірек болады.

Статистикалық әдістердің көмегімен таңдама қасиеттері бойынша бас жиынтықтың параметрлері мен қасиеттері туралы анық тұжырым жасау үшін,таңдама ***репрезентативті*** болу керек,яғни мүмкіндігінше бас жиынтықтың қасиеттерін дұрыс бейнелеу керек.

Таңдама репрезентативті болу үшін төмендегі шарттар орындалуы керек:

* Бас жиынтықтан нысандарды таңдау кездейсоқ жүргізіледі, яғни нысандардың әрқайсысының таңдамаға түсу ықтималдықтары бірдей;
* Таңдамадағы бақылаулар нәтижелері бір-бірінен тәуелсіз;
* Барлық нақты шарттарды есепке ала отырып таңдама көлемін дұрыс анықтау;
* Таңдамадағы пропорциялар бас жиынтықтың пропорцияларына сәйкес.

Мысалы,сәйкесінше 500,200 және 300 студент оқитын А, Б және С медициналық университеттерінде қалыпты физиология әні бойынша үлгерімді зертту үшін көлемі 100 болатын таңдаманы,оған А университетінің кездейсоқ таңдалған 50 студенті,Б университетінің 20 студенті және С университетінің 30 студенті кіретіндей етіп құру керек.

Статистикалық жиынтықтың бақылаудың жеке бірліктеріне қарағанда өзіне ғана тән ерекше қасиеттері бар екендігін білу қажет,олардың қатарына:зерттеліп отырған құбылыстың таралу сиппаты;зерттеліп отырған құбылыстың жалпылама сипаттамасы болатын оның орта деңгейі;жиынтықты құрайтын бірліктердің әртүрлілігі;зерттеліп отырған белгілер арасындағы байлныс және бас жиынтыққа қатысты таңдаманың репрезентативтілігі жатады.

Арнайы статистикалық әдістер арқылы осы қасиеттерді айқындайды және жалппылама сипаттамаларды алады.

***2.2.Таңдаманың статистикалық таралуы***

Таңдама деректерін статистикалық талдау барысында таралудың дискретті статистикалық қатары және аралық статистикалық қатарлары қолданылады.

***Вариация*** деп зерттеліп отырған топтағы белгінің әртүрлі мәнін айтады.Мысалы: дене салмағы,бойдың ұзындығы,лейкоциттер саны,тромбоциттер саны,фибриногеннің деңгейі және т.б әрқашан бір топтағы адамдардың әрқайсысында әртүрлі болады,яғни вариация биологиялық түрленгіштіктің көрінуі болып саналады. Түрленгіштік-барлық тірі ағзаға тән қасиет.

***Варианта*** деп зерттеліп отырған белгінің жеке өлшенген мәнін айтады. Вариантаны хі арқылы белгілейді.Таңдамадағы варианттың кездесу санын онң жиілігі деп атайды, және ni арқылы белгілейді.

***Вариациялық қатар*** деп варианталары өсу ретімен жазылған таңдаманы айтады.

Бүтін сандар түрінде өрнектелген вариациялық қатар дискретті,ал аралықтар түрінде берілген вариациялық қатар аралықты деп аталады.

***Таңдаманың статистикалық таралуы*** де варианталар мен оларға сәйкес жиіліктердің тізімін айтады.

***Аралық статистикалық қатар***

Аралық статистикалық қатарды,әдетте таңдамада бір-бірінен өзгеше варианталар саны кө болған жағдайда немесе Х белгісі үздіксіз болған жағдайда қолданады.

Көлемі n таңдамадағы деректерді тотастырғанда барлық варианталарды қамтитын кесіндісін k аралықтарға бөледі.Осыдан соң әрбір аралықтың жиілігін,яғни і-ші аралыққа түсетін ni бақылаулар санын анықтау керек.

***2.3. Таңдаманың сандық сиаттамалары***

Вариациялық қатарлар мен таңдама деректерінің графиктік кескіндеулері таңдама жиынтықтағы елгінің қалай түрленетіні жөнінде көрнекі түрде ақпарат бере алады.Бірақ олар таңдаманы толық сипаттау үшін жеткіліксіз,бұл мақсатта жалпылайтын сандық сипаттамаларды қолдану керек.

Таңдаманың сандық сипаттамалары бақылау деректері жөнінде сандық мәлімет береді және оларды өзара салыстыруға мүмкіндік береді.Практикалық маңызы ең жоғары сандық сипаттамаларға таңдама таралулардың орын және шашырау сипаттамалары жатады.Орын сипаттамаларына таңдама орта,медиана және мода жатады.

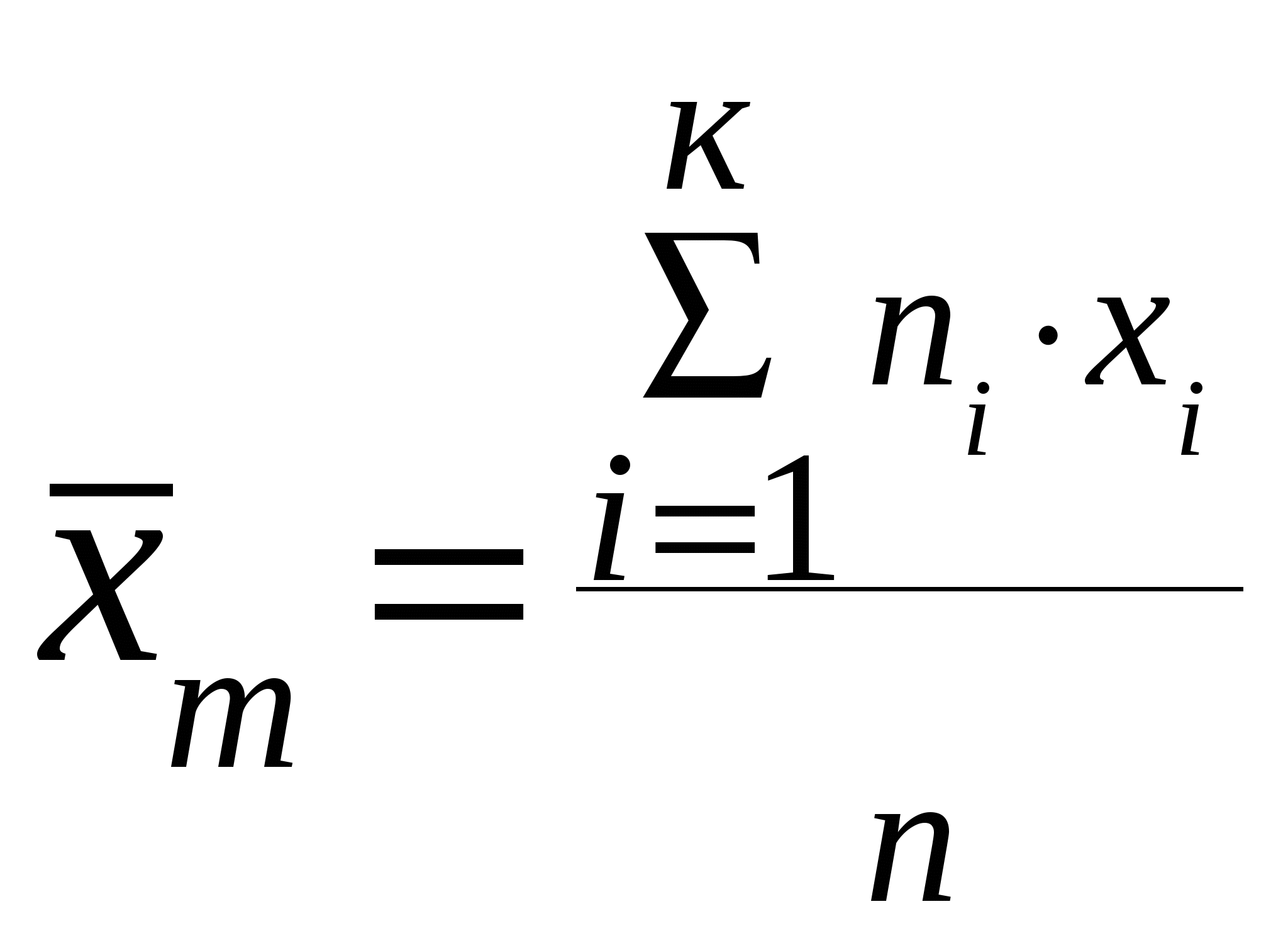
**Орын сипаттамалары**

***Таңдама орта***

Таңдама орта-таңдаманың негізгі сипаттамаларының бірі.Ол-зерттеліп отырған белгі өлшемінің жалпылама сипаттамасы.Ол бір санмен сапалы біртекті жиынтықты сандық сипаттауға мүмкіндік береді.

Х сандық белгісіне қатысты бас жиынтықты зерттеу үшін n көлемді таңдама алынады.

Х таңдама орта деп таңдама жиынтық белгісінің орта ариифметикалық мәнін айтады. Егер n көлемді таңдаманың барлық x1…x2..xnмәндері әртүрлі болса онда

 түрінде болады.

***Таңдама ортаның өлшем бірлігі белгілердің мәндерінің өлшем бірлігімен бірдей.***

***Медиана***

Медиана вариациялық қатарды тең екгі бөлетін мән,одан төмен және жоғары мәндердің бірдей саны орналасады,және медианаға дейін оданн кіші мәндердің 50%-ы одан кейін үлкен мәндердің 50%-ы жатады.

Егер бақылаулар саны n тақ болса,онда медиананы есептеу жеңіл.Бұл вариациялық қатардағы -ші нөмірлі варианта болады.Мысалы n=11 болса,онда медиана вариациялық қатардағы 6-шы нөмерлі варианта болады.Егер n жұп болса,онда медиана вариациялық қатардағы -ші және ( +1)-ші нөмірлі варианталардың арифметикалық ортасы болады.Мысалы,егер n=20 болса,онда медиана =10-шы және нөмірлі варианталардың арифметикалықортасы болады.

Егер деректердің таралуы симметриялы болса,онда медиана орта мәнге ұқсас,егер деректердің таралуы оң жаққа қарай қисайған асимметриялы болса,онда орта мәннен кіші,егер деректердің таралуы сол жаққа қарай қисайған асимметриялы болса,онда медиана орта мәннен үлкен.

**Мода (moda,mo)**

Мода (moda)-деректер жиынында ең жиі кездесетін мән,егер деректер үздіксіз болса,онда оларды топтастырып модальды топты анықтайды.Кейбір деректер жиынында мода болмайды,себебі әрбір мән бір рет кездеседі.Кейде модалар бірнешеу болуы мүмкін,ол біртекті емес жиынтықтарда кездеседі.Мода жалпылайтын сипаттама ретінде өте сирек қолданылады.

**Шашырау сипаттамалары**

Орта мәндер өзгермелі белгілер жөнінде толық ақпарат бермейді.Мәселен,екі таңдаманың таралуларының орта мәндері бірдей болулары мүмкін,бірақ олардың біреуінде белгілердің мәндері орта мәннің айналасына жақын орналасуы , ал екіншісінде алыс орналасулары мүмкін.Сондықтан орта мәндермен қатар таңдаманың шашырау сипаттамаларын да есептейді.

*Шашырау көрсеткіштеріне таңдаманың дисперсиясы,стандарттық ауытқу,таңдаманың құлашы,процентильдер (квартильдер),вариация коэффициенті,квартильаралық құлаш жатады.*

**Дисперсия(variance)**

Бұл статистиканың атауы латынның disperses-шашыраңқы,шашылған деген сөзінен шыққан.Деректердің шашырауын өлшеу тәсілдерінің бірі-әрбір бақылаудың орта мәннен ауытқуын анықтау болып табылады.Ауытқұ неғұрлым үлкен болса,бақылаудың өзгергіштігі,түрленгіштігі соғұрлым үлкен болатыны айқын.Алайда біз бұл ауытқулардың орта мәнін шашырау өлшемі ретінде ала алмаймыз,себебі оң ауытқулар теріс ауытқулардың орнын толтырады(олардың қосындылары нөлге тең).Бұл сұрақты шешу үшін,біз әрбір ауытқуды квадратқа шығарамыз және квадратталған ауытқулардың орта мәнін табамыз,бұл шама *дисперсия* деп аталады.

*Таңдама дисперсия* деп белгінің бақыланатын мәндерінің орта мәнінен ауытқу квадраттарының орта арифметикалық мәнін айтады.

Мұндағы –таңдама дисперсиясының белгіленуі (кейбір әдебиеттерде таңбалануы кездеседі).

Дисперсия ең маңызды шашырау сипаттамаларының бірі болып саналады.Ол бақылау орталық мәннің айналасында орналасуының шашырау дәрежесін сипаттайды.

Егер бізге зерттеліп отырған бас жиынтықтағы барлық мәндері белгілі болса,онда дисперсияны бас жиынтық үшін:

Формуласымен есептеген болар едік,мұндағы бас дисперсия,N-бас жиынтықтың көлемі.

Таңдама дисперсиясының формуласында бөлімінде таңдама көлемінен 1 саны алынып тасталынады.Себебі,таңдама ішінде мәндердің шашылуы бас жиынтықтағыдай соншама көп болмайды,сондықтан n-ға емес,(n-1)-ге бөлу дисперсиясының бағасының төмендеуінің орнын толтырады деп саналады.

Дисперсияны таңдама мәндерінің квадраттарының орта мәні мен орта мәннің квадратының айырымы ретінде де есептеуге болады:

D= -

**Дисперсияның өлшем бірлігі-алғашқы бақылаулардың өлшем бірліктерінің квадраты,мысалы:егер вариация килограммен өлшенсе,онда дисперсияның өлшем бірлігі «кг-квадрат» болады.**

**Стандарттық ауытқу (standart deviation,SD)**

Стандарттық ауытқу (орта квадраттық ауытқу) деректер мәндерінің түрленгіштігін бейнелейді және деректер жиынының өзгергіштігін анықтау үшін қолданады.Стандарттық ауытқу-дисперсиядан алынған оң түбір.

Бұл көрсеткіш неғұрлым үлкен болса,өлшенген айнымалының шашырауы соғұрлым үлкен.

**Стандарттық ауытқулардың өлшем бірлігі таңдама ортаның өлшем бірлігімен бірдей.**

**Стандарттық ауытқудың қолданылуы**

Педиатрияда стандарттық ауытқуды балалардың физикалыө дамуын бағалау үшін қолданады:ол нақты баланың деректерін сәйкес стандарттық көрсеткіштермен салыстыру арқылы жүзеге асады.Стандарт ретінде сау балалардың физикалық дамуларының көрсеткіштерінің орта мәні алынады.Көрсеткіштерді стандартпен салыстыруды арнайы кесте арқылы жүргізеді,кестеде стандарттар оларға сәйкес келетін сигмальді шкалалармен бірге келтіріледі.

Егер баланың физикалық дамуының көрсеткіші *стандарт ±∂* ішінде болса,онда баланың физикалық дамуы (осы көрсеткіш бойынша) қалыпқа сәйкес деп саналады.Егер көрсеткіш *стандарт ±2∂* ішінде жатса,онда қалыптан айтарлықтай емес (субнорма) ауытқу бар.Егер көрсеткіш бұл шекаралардан шығып кетсе,онда баланың физикалық дамуында қалыптан күрт айырмашылық бар (паталогия болуы мүмкін).

***Құлаш (өзгеру аралығы)(амплитуда,range)***

Құлаш- деректер жиынындағы айнымалының ең жоғарғы және ең төменгі мәндерінің арасындағы айырма:

R=

***Процентильдер (перцентильдер,percentile)***

*Процентильдер* - бұл вариациялық қатарды тең 100 бөлікке бөлетін мәндер.Бақылаудын 1%-ы орналасқанға дейінгі х шамасы (бақылаудың 99%-ы х мәнінен жоғары жатады) *бірінші процентиль* деп аталады.Бақылаудың 2%-ы орналасқанға дейінгі х шамасы екінші процентиль деп аталады және т.с.с.

Вариациялық қатарды тең 10 бөлікке бөлетін мәндер,яғни 10,20,30,...90-шы процентильдер *децильдер* деп аталады.

Вариациялық қатарды тең 4 бөлікке бөлетін мәндер,яғни 25-ші,50-ші және 75-ші процентильдер *квартильдер* деп аталады.50-ші процентиль-*медиана*.Квартильдерді Q таңбасымен белгілейді және олар екеу:*төменгі* немесе *бірінші квартиль ()* және *жоғары* немесе *үшінші квартиль ()* деп аталады.

Бірінші квартиль вариациялық қатарды екі бөлікке бөлетін сан:бақылаулардың 25%-ы бірінші квартильге дейін орналасады,75%-ы одан кейін орналасады.

Үшінші квартиль вариациялық қатарды екі бөлікке бөлетін сан:бақылаулардың 75-ы үшінші квартильден кіші,25%-ы одан үлкен.

-Вариациялық қатардың -ші нөмірлі вариантасы;

-вариациялық қатардың -ші нөмірлі вариантасы.

Квартильдерді табу үшін вариациялық қатарды медианамен тең екі бөлікке бөлу керек,сосын олардың әрқайсысының медианасын табу керек.Мысалы,егер таңдама көлемі n=6 болса,онда таңдаманың бірінші квартилі үшін екінші варианта,ал үшінші квартиль үшін бесінші варианта алынады.

1 3 4 6 7 8

Таңдамасында =3,=7.

Егер n=9 болса,онда бірінші квартиль үшін 2-ші және 3-ші варианталардың арифметикалық ортасы,ал үшінші квартиль үшін 7-ші және 8-ші варианталардың арифметикалық ортасы алынады:

2 4 5 6 8 9 10 11 12

Таңдамасында =4.5 және =10,5 ке тең.

*Квартильаралық құлаш* бірінші және үшінші квартильдердің,яғни 25-ші және 75-ші процентильдердің арасындағы айырма.Оған вариациялық қатардың ортаңғы 50% бақылаулары кіреді және бақылаулардың 25%-ы ортаңғы нүктеден төмен,25%-ы жоғары жатады.

*Интердецильді* құлаштың ішінде ортаңғы 80% бақылаулар,яғни 10-шы мен 90-шы процентильдердің арасында орналасқан бақылаулар жатады.

Көбіне 95% бақылаулар жататын құлашты қолданады,яғни төменгі жақтан 2,5% және жоғары жақтан 2,5% бақылаулар шығарылып тасталынады.Бұл аралықты аурудың диагностикасын жүзеге асырғанда қолдануға болады.Бұл жағдайда ол *референтті аралық* немесе *қалыпты құлаш* деп аталады.

**Процентильдердің қолданылуы**

Медиананың 25-ші және 75-ші процентильдерге қатысты орналасуына қарай таралудың қаншалықты асимметриялы екендігін пайымдауға болады.

Лабораториялық көрсеткіштердің қалыптылығының шекаралары ретінде көбінесе 5-ші және 95-ші процентильдер қолданылады.

Физикалық даму көрсеткіштерін бағалағанда белгілі бір пациенттің көрсеткіші 3-ші процентильден аз болса,ол *күрт төмен* деп бағаланады,мәндері бойынша 3-ші және 10-шы процентильдердің арасында жатса *төмен*, 25-ші мен 75-ші процентильдердің арасында (квартильаралық құлаш) жатса - *орта*,сол сияқы 75-ші мен 90-шының арасында - *ортадан жоғары*, 90-шы мен 97-шінің арасында- *жоғары*, 97-ші процентильден жоғары орналасса- *күрт жоғары* деп бағаланады.

**Қорытынды:**Егер таңдама қалыпты таралған бас жиынтықтан алынғаны белгілі болса,онда таңдама орта мен таңдама стандарттық ауытқуды қолдану керек.Егер жиынтық қалыпты таралу заңына бағынбайтындығы жөнінде негіз болса,онда медиананы және 25-ші және 75-ші процентильдерді қолдану керек.

**Мысал:**Көлемдері және медианалары бірдей екі таңдама алайық және олардың қалыпты таралмаған бас жиынтықтардан алынғаны белгілі.Орта мәнге қатысты мәндердің шашырау дәрежесін анықтау керек.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|  | 1 | 3 | 5 | 10 | 5 | 3 | 1 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|  | 8 | 2 | 1 | 4 | 1 | 8 | 4 |

*Шешуі*:X және Y таңдамалары үшін сәйкес мен -ті табамыз.

X:=9: 11 Y:

X үшін квартильаралық Y үшін квартильаралық

құлаш құлаш

**Қорытынды**

Деректердің орталық мәннің айналасында шашылуы У таңдамасында үлкен.

**Вариация коэффициенті**

Бұл өлшем бірлігіне тәуелсіз шашырау өлшемі.

Вариация коэффициенті (CV)-стандарттық ауытқудың таңдама ортаға пайыздық қатынасы:

CV=\*100%

Вариация коэффициенті- бұл вариациялық қатардың өзгергіштігінің салыстырмалы өлшемі.Вариация коэффициенті әрбір нақты вариациялық қатардың әртүрләләгән,яғни таңдаманың біртектілігін бағалау үшін қолданылады.

* CV<10% болғанда қатардың әртүрлілігі әлсіз,
* 10%<CV<20%-орта,
* CV>20%- күшті болып саналады.

**2.4.Сенім аралығы (confidence interval)**

Медициналық-биологиялық зерттеулердің көбісі популяциядан алынған таңдама үшін ақиқат нәрсе сол популяцияның барлығы үшін де қандай да бір дәрежеде ақиқат болады деген алғышартқа негізделген.Сонымен,таңдаманың сипаттамаларын есептеу сәйкес популяцияның сондай сипаттамаларын *бағалау* үшін қолданылады.Бұл бағалардың *дәлдігі* өлшеу техникасына байланысты болатын өзгергіштік дәрежесіне (өлшеу қателігі),таңдама көлеміне және оның репрезентативтілігіне (таңдама қателігі),сол сияқты барлық биологиялық сипаттамаларға тән өзгергіштікке (кездейсоқ қате) байланысты.Бағаға байланысты өзгергіштік дәрежесі *сенім аралығы* арқылы беріледі.

Сенім аралығы-бұл ішінде популяцияның ақиқат параметрі жататын,әдетте біз 95% сенімді болатын мәндер ауқымы.Параметрдің нақты мәнінің осы ауқымның шекараларынан шығып кету ықтималдығы

1-0.95=0.05(5%)-дан аспайды.

Сенім ықтималдығын толықтыратын шаманы әдетте деп белгілейді және оны *мәнділік деңгейі* деп атайды.

**Бас орта мән үшін сенім аралығы**

(-; + )

Формаһуласымен анықталады,мұндағы таңдама ортаның стандарттық қатесі,-стьюдент коэффициенті.

**Стандарттық қате (Standard Error Means,SEM)**

Орта мәннің стандарттық қатесі таңдама ортаның дәлдік өлшемін бейнелейді және = формуласымен есептеледі,мұндағы =стандарттық ауытқу,n=бақылау саны.

Үлкен стандарттық қате бағаның дәл еместігін көрсетеді,ал кіші стандарттық қате бағаның дәлдігін көрсетеді.Егер таңдама көлемі ұлғайса және деректердің шашылуы аз болса,онда стандарттық қате азаяды,яғни біз дәлірек баға аламыз.

Бас жиынтықтың барлығын бақылау мүмкіндігі әрқашан бола бермейтіндіктен,таралу параметрлерін бас жиынтықтан кездейсоқ таңдалынып алынған таңдама бойынша бағалаймыз.Ортаның стандарттық қатесі таңдама орта арқылы бас жиынтықтың ортасын бағалағанда дәлдік өлшемі қызметін атқарады.

***Стандарттық ауытқу немесе ортаның стандарттық қатесі:***

Стандарттық ауытқу деректердің мәндеріндегі құбылмалылықты бейнелейді және ол деректер жиынындағы өзгергіштікті анықтауда қажет болғанда көрсетілуі тиіс.

Стандарттық қате таңдама ортаның дәлдігін бейнелейді және ол деректер жиынындағы орта мәнді анықтау қажет болғанда көрсетілуі тиіс.

***Еркіндік дәрежелері (degree freedom)***

Еркіндік дәрежелерінің саны (df)-бұл таңдама құрамындағы еркін түрленетін бірліктер саны немесе бақылаулар саны минус бағаланатын статистикалық параметрлер саны.

Мысалы,үй қояндарының салмақтары бойынша қандай да бір вариациялық қатар бар болсын:.Бұл қатардың жалпылама сипаттамасы –орта мән.Егер бізде орта мәннен басқа сандық деректер жоқ болса,онда қатардың әрбір жеке мәнін қалай анықтауға болады?Бұл жиынтықтағы бір мәнді анықтау қалған басқа мәндерге байланысты екені белгілі.Мысалы,2 қоянның салмағы бірге алғанда 6кг,ал олардың біреуінің салмағы 2.5 кг.Екінші қоянның салмағын бірінші қоян арқылы білуге болады,яғни оның салмағы бірінші қоянның салмағымен дәл анықталған.Сонымен,тек бір ғана еркіндік дәрежесі бар (2-1=1).Егер үш қоянның салмағы бірге алғанда 5кг болса,онда бір қоянның салмағы қалған екі қоянның салмақтары арқылы дәл анықталған.Яғни бұл жағдайда 2 еркіндік дәрежелері (3-1=2) бар және т.с.с.

**Деректердің таралу түрлері**

Ең қарапайым статистикалық әдісті таңдау көбіне белгілердің таралу түріне байланысты.Деректердің таралуы көп жағдайда унимодальды болып келеді.Кейде таралу бимодальды (екі өркешті) немесе біркелкі (әр шаманың деректер жиынында болу ықтималдығы бірдей) болуы мүмкін.Унимодальды таралудағы негізгі мақсат-бұл деректердің басым бөлігінің ең үлкен және ең кіші мәндерге қарасты қай жерде орналасқанын көру.

Медициналық зерттеулерде бақыланатын таралулардың негізгі түрлері:

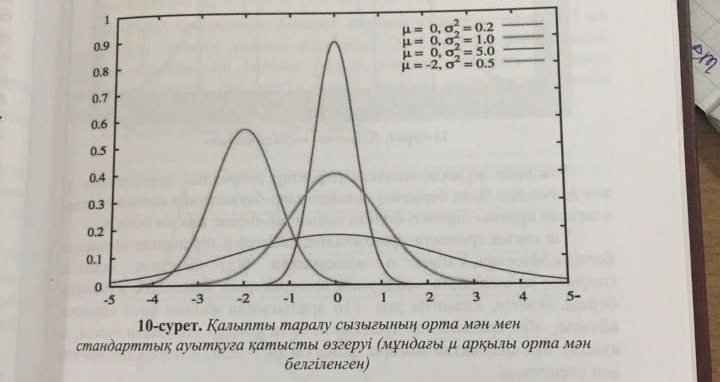
* *Симметриялық*-орталық нүктенің айналасына шоғырлану ,таралудың

бір жағы екінші жағының симметриялық бейнесі.

Оңға қарай қисайған-оң жаққа қарай созылған ұзын құйрық ( бір немесе бірнешеүлкенмәндер)  
Солға қарай қисайған - сол жаққа созылған ұзын құйрық (бір немесе бірнеше кішімәндер) Таралулардың басқа түрлеріне қарағанда жиірек кездесетін және статистикадағы ең маңызды таралулардың бірі-қалыпты таралу.  
Қалыпты таралу түрінде варианталар вариациялық қатардың ортасына қарасты симметриялы түрде орналасады және олардың саны орталықтан алыстаған сайын біртіндеп кеміп отырады,сонымен бірге бақылаудың ең көп саны қатардың ортасына жинақталады.Кездейсоқ шамалардың мұндай таралу түрі медициналық-биологиялық көрсеткіштерде жиірек кездесетін болғандықтан,онықалыптыдепатаған.  
Гаусс қисығы Қалыпты таралудың графигі орта мәнге қатысты симметриялы,қоңырау тәрізді болады және қалыпты таралу қисығы немесе Гаусс қисығы деп аталады.

**Қалыпты таралудың негізгі сипаттамалары:**

* Толығымен екі параметірмен анықталады: орта және дисперция;
* Қоңырау тәріздес(унимодальды)
* Орта мәңге қатысты симметриялы;
* Егер орта мән ұлғайса онға қарай,орта мән азайса солға қарай жылжиды(дисперция тұрақты жағдайда);
* Қисықтың түрі бас жиынтықтың стандарттық ауытқуымен анықталады;
* Стантарттық ауытқуы азтаруға жіңішке,жоғары созылған қисықтар,ал стандарттық ауытқуы үлкен таралуға жазыңқы қисықтар сәйкес келеді.(10 сурет)



**Қосымша қасиеттері:**

* Қалыпты тарудың орта мәні мен медианасы тең;

Тәжірибе жүзінде алынған деректер теориялық деректермен әр кез дәлме-дәл бола бермейді, алайда олар бақылаулар саны көп және оларды құрамы біртекті болған сайын бір-біріне жақын болады.

Үш сигма ережесін практикалық есептерді шешкенде қолдануға болады. Мәселен, х және s өлшемдерін білу белгінің қалыпты(нормаға сәйкес) орта мәндерінің шекараларын анықтауға мүмкіндік береді. Әдетте, қалыпты деп ±1s аралығында жатқан орта мәндерді айтады, ±2s аралығында ауытқыған орта мәндер нормадан төмен, ал ауытқу бұл аралықтан жоғары болса (яғни, > ±2s) – нормадан жоғары деп есептеледі.

1 – Мысал. (С.Гланц, 45 бет) Орақ тәрізді жасушалы анемияның ауырлығының клиникалық бағалары келтірілген: 0;0;0;1;1;1;1;1;1;1;1;1;1;1; 2; 2; 2; 2; 3; 3; 3; 3; 4; 4; 5; 5; 5; 5; 6; 7; 9; 10; 11. Таңдама ортаны, стандарттық ауытқуды, медиананың, 25-ші және 75-ш процентильдерді табыңдар. Таңдама қалыпты таралған бас жиынтықтан алынған деп санауға бола ма? Жауаптарыңды негіздеңіздер.

**Шешуі.**

1. Деректердің статистикалық таралуын жазамыз (дискретті статистикалық қатар):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 10 | 11 |
|  | 3 | 11 | 4 | 4 | 2 | 4 | 1 | 1 | 1 | 1 | 1 |

=33

1. Таңдама орта ==3,09
2. Стандартты ауытқу ===2,43
3. Медиана =
4. 25-ші процентиль (немесе бірінші квартиль) , 75- ші процентиль (немесе үшінші квартиль) =5, себебі вариациялық қатардың медиана екі бөлікке бөледі, бірінші бөліктегі қатардың медианасы 1-ге тең, сол сияқты екінші бөліктегі қатардың медианасы 5-ке тең.
5. Жиілік полигонын тұрғызамыз.

**12-сурет.** *Орақ тәрізді жасушалы анемияның ауырлығының клиникалық бағалары*

**3-Бөлім. Сатистикалық жорамалдар және оларды тексеру (hypothesis testing)**

***3.1. Статистикалық жорамал ұғымы.***

***Статистикалық жорамал*** деп –бас жиынтықтың белгісіз таралу түрі немесе белгілі таралудың параметрлерінің шамасы жөніндегі ұғымдарды айтады.

Статистикалық жорамалдарды тексеру медициналық зерттеулерде қолданылатын ең маңызды статистикалық әдістердің бірі болып саналады.

Медициналық зерттеулерде ағзаның қалыпты жағдайы мен патологиялық жағдайындағы, емдеуге дейінгі және емдеуден кейінгі немесе әртүрлі емдеу тәсілі қолданған кездегі көрсеткіштерін салыстыру маңызды орын алады. Басқаша айтқанда, статистикалық жорамалдарды тексеру теориясы дәлелді медицинаның негізгі құралы болып табылады.

Статистикалық жорамал таңдама бақылаудың нәтижелері бойынша статистикалық әдістермен тексеріледі.

Жорамалдарды статистикалық тексеруге *салыстыру арқылы тексеру* және *әртүрлі үрдістерді бағалау* жатады: емдеу тиімділігі, аурудың ұзақтылығы, ауру-сырқаулықтың ауырлығы, емдеу әдістерін салыстыру, препараттардың және медициналық техниканың сипаттамаларын салыстыру және т.б.

Салыстырылып отырған топтардың қайсысының артықшылығы жөнінде әдетте, орта мәндердің, үлестердің немесе басқа таңдама көрсеткіштерінің арасындағы айырмашылық бойынша тұжырымдайды. Айырмашылықтың шынайылығы жөніндегі сұрақ таңдама сипаттамалары бойынша тексеру негізіндегі шешіледі.

*Статистикалық жорамалдарды тексерудің мағынасы бақыланып отырған (немесе тәжірибе негізінде алынған) белгілер арасындағы айырмашалық кездейсоқ жағдайдың нәтижесі болу ықтималдығы қандай екендігін анықтау.*

Тексеруге жататын негізгі жорамалды **нөлдік жорамал** деп атайды және **Н0** арқылы белгілейді.

**Нөлдік жорамал Н0** деп –таңдама деректерін салыстырғанда олардың арасындағы айырмашылықтың жоқтығы немесе ұқсастығы жөніндегі ұйғарымдарды айтады. Оның мағынасы салыстырылып отырған топтардың бас параметрлері арасындағы айырма нөлге тең және таңдама сипаттамалары арасындағы кез-келген айырмашылық мүлде кездейсоқ сипатқа ие дегенді білдіреді. Мысалы, егер бір таңдама қалыпты таралған бас жиынтықтан алынған және параметрлері 1 және 1 , ал екінші таңдама параметрлері 2 және 2 болатын бас жиынтықтан алынған болса, онда нөлдік жорамал 1 = 2 , яғни 1 - 2 =0 түрінде ұйғарылған болар еді.

Біз әрқашан әсерді жоққа шығаратын нөлдік жорамалды (Н0)тексереміз. Мысалы, егер бізге популяциядағы ер кісілер мен әйелдердің шылым шегу көрсеткіштерін салыстыру қажет болса, онда нөлдік жорамал төмендегідей түрде бола еді:

Н0 – популяциядағы ерлер мен әйелдердің шылым шегу көрсеткіштері бірдей.

Нөлдік жорамалға қарама-қарсы жорамал **балама жорамал** деп аталады және **Н1** арқылы белгіленеді. Ол айырмашылықтың бар болуы жөніндегі жорамал, яғни орта мәндер тең емес 1 - 2 ≠0, жоғарыдағы мысалға қатысты балама жорамал Н1 –популяциядағы ерлер мен әйелдердің шылым шегу көрсеткіштері әртүрлі деп ұйғарылар еді.

Зерттеулерде нөлдік жорамалды тұжырымдау – аса маңызды, әрі қажетті кезең, себебі балама жорамал нөлдік жорамалды жоққа шығарғаннан кейін ғана қабылданады.

***3.2. Статистикалық жорамалдарды тексеру барысында жіберілетін қателер.***

Статистикалық жорамалды жоққа шығару немесе қабылдау жөніндегі шешім таңдама деректері негізінде қабылданады. Сондықтан нөлдік жорамалға қатысты қате шешім қабыдануы мүмкін. Болуы мүмкін қателерді І текті және ІІ текті деп бөлінеді. Көрнекілік үшін болуы мүмкін қателерді кестеде келтірейік:

*Жорамалды тексергенде болуы мүмкін қателер*

5 кесте

|  |  |  |
| --- | --- | --- |
| Шындығында | Н0 қабылдау | Н0 жоққа шығару |
| Н0 дұрыс | Қате жоқ  Ықтималдық: 1-α | І текті қате  Ықтималдық: α |
| Н0 дұрыс емес | ІІ текті қате  Ықтималдық: β | Қате жоқ  Ықтималдық: 1-α |

Көріп отырғанымыздай әртүрлі төрт жағдай болады.

І текті қателік – дұрыс нөлдік жорамалды жоққа шығару, яғнишындығында жоқ әсерді бар деп қорытынды жасау.

ІІ текті қателік – дұрыс емес нөлдік жорамалды қабылдау, яғни шындығында бар әсерді жоқ деп қорытынды жасау.

Таңдама деректері бойынша ғана ақпаратқа ие бола отырып (толық емес ақпарат) бас жиынтық жөнінде дәл емес, ал қандай да бір белгілі ықтималдықпен тұжырым айтуға болады.

Нөлдік жорамалды екі түрлі қателердің болу мүмкіндігі мейлінше аз болатындай тексеру қажет.

Бірінші текті қатені жіберу ықтималдығын α арқылы белгілейміз, **α** саны мәнділік деңгейі деп аталады. **Мәңділік деңгейі** – айырмашылықтың кездейсоқ пайда болуының ықтималдығы.

ІІ текті қатені жіберу ықтималдығын **β** арқылы белгілейік.

ІІ текті қатені жібермеу ықтималдығын **критерийдің қуаттылығы** деп атайды. Критерий қуаттылығы 1-β тең. Қуаттылық неғұрлым үлкен болса, соғұрлым ІІ текті қатенің ықтималдылығы аз.

**Мысалы:** В вирустық гепатиттің бар болуына диагностикалық тексеру жүргізілді.

Н0: тексерілуші адам сау.

Н1 тексерілуші адам гепатит жұқтырған.

І текті қате – тесттің нәтижесі бойынша тексерілуші адам гепатитпен ауырады, ал шындығында ол сау (жалған позитивті нәтиже).

ІІ текті қате – шындығында гепатит жұқтырған адамды қателікпен сау деп санады (жалған негативті нәтиже).

Мәнділік деңгейі немесе қабылданған жорамалды бағалағанда жіберілуі мүмкін І текті қатенің ықтималдық 5% , 1%, 0,1% болуы мүмкін, алайда медициналық – биологиялық зерттеулерде, егер арнайы басқа мән ескертілмесе, ол әдетте 5%-ға тең деп қабылданады, яғни 0,05-ке тең мәнділік деңгейі нәтижелерді дұрыс түсіндіру үшін жеткілікті деп саналады. Егер нәтижелер 1-5% деңгейде мәнді болса, онда маңызды статистикалық мәнділік бар деп айтылады, 1% -дан аз деңгейдегі мәнділік - жоғары статистикалық мәнділік көрсетеді.

Ғылыми жұмыстардағы р – тәжірибеде алынған мәнділік деңгейі.

Статистикалық талдау үшін компьютерлік бағдарламалардың кең түрде ендірілуі қазіргі кезде мәнділік деңгейін дәл есептеуге мүмкіндік береді. Сондықтан, замануи зерттеулерде р мәнін «» немесе «» белгілерінсіз нақты көрсетуге болады. Мысалы, р=0,004 жазуы тәжірибедегі бақыланған айырмашылық көрсетілген деңгейде статистикалық мәнді дегенді білдіреді.

Биология мен медицина ықтималдығы төмен жорамал деп жүзеге асу ықтималдығы 5% -дан кем болатын жорамалды айтады.Оқиға ықтималдығы мәнділік деңгейімен тығыз байланысты. Оны зерттеуші тәжірибелік деректерді алғанға дейін анықтау тиіс.

Мәнділік деңгейі берілгеннен кейін нөлдік жорамалды кабылдайтын немесе жоққа шығаратын ережені табады. Ондай ереже **статистикалық критерий** деп аталады.

**Зерттеу нәтижелерінің статистикалық мәнділігін бағалау.**

Деректердің *статистикалық мәнділігі* деп олардың бейнелеп отырған болмысқа сәйкестік дәрежесін айтады, яғни *статистикалық мәнді* деректерге объективті шынайылықты бұрмаламайтын және дұрыс бейнелейтін деректер жатады.

***Зерттеу нәтижелерінің статистикалық мәнділігін бағалау- таңдама жиынтықта алынған нәтижелерді қандай ықтималдықпен барлық бас жиынтыққа көшіруге болатынын анықтауды білдіреді.***

Статистикалық мәнділікті бағалау жиынтықтың бөлігі бойынша қаншалықты барлық жиынтықты және оның заңдылықтарын тұтас пайымдауға болатының түсіну үшін қажет.

***3.3. Статистикалық критерийлер және олардың түрлері***

**Статистикалық критерий** – нөлдік жорамалды қабылдауға немесе жоққа шығаруға мүмкіндік беретін ереже.

Статистикалық критерийлер ең алдымен таңдама деректерінің қасиеттері негізінде және олардың таралу түрлеріне қарай таңдалынады.

**Статистикалық критерийді таңдау**

Деректерді талдаудың қолайлы статистикалық әдісін таңдауға ықпал ететін жағдайлар:

1. Талдау мақсаты.

Зертелетін өзгерістер немесе айырмашылықтар бір бағытта немесе екі бағытта болуы мүмкін, деректер арасындағы айырмашылық немесе өзара байланыс зерттелуі мүмкін.

1. Дереутер түрлері.

Зерттелетін белгінің мәні сандық немесе сапалық айнымалы болуы мүмкін.

1. Деректердің таралу түрлері: қалыпты немесе қалыпты емес.
2. Жұптасқан деректер (бір ғана субъектілер үшін қайталап өлшеу, «дейін-кейін» есебі).
3. Тәуелсіз деректер.
4. Салыстырылатын топтар саны.

Сәйкес критерийді (жалпы жағдайда оны К арқылы белгілейік) таңдағаннан кейін К қабылдайтын барлық мәндер жиынын екі қиылыспайтын ішкі жиындарға бөлеміз: олардың біреуінде критерийдің нөлдік жорамал жоққа щығарылатын мәндері жатады, ал екіншісінде – нөлдік жорамал қабылданатын мәндері жатады, яғни *сыни аймақ* және *жорамалды қабылдау аймағы.*

**Сыни аймақ** – нөлдік жорамалды жоққа шығаратын критерий мәндерінің жиынтығы.

**Жорамалды қабылдау аймағы** – нөлдік жорамалды қабылдайтын критерий мәндерінің жиынтығы.

**Жорамалдарды тексерудің негізгі принципі:** егер К критерийдің таңдама бойынша есептелген бақылау мәні сыни аймаққа жатса – нөлдік жорамалды жоққа шығарады, егер критерийдің бақылау мәні жорамалды қабылдау аймағында жатса – нөлдік жорамалды қабылдайды.

Сыни аймақты жорамалды қабылдау аймағынан бөліп тұратын нүктелерді *сыни нүктелер* деп атайды. Оларды Ксыни деп белгілейік.

Нөлдік жорамалды жоққа шығару басталатын критерийдің мәні **сыни мән** деп аталады.

Сыни нүктелердің әрбір критерий үшін оның таралу түріне сәйкес келетін арнайы кестелерден табады.

Нөлдік жорамалды қабылдаған жағдайда бақылау мәндері мен ақиқат (бас жиынтық бойынша) мәндердің арасындағы айырмашылық кездейсоқ себептерден туған және мәнді емес (принципті емес) деп есептелінеді.

Нөлдік жорамалды теріске шығарған жағдайда, бақылау мәндері мен теориялық (нөлдік жорамалға сәйкес) мәндердің арасындағы айырмашылық мәнді, яғни принципті себептерден туған деп есептеледі.

**Біржақты (one-sided) және екіжақты (two-sided) критерийлер**

Егер екі таңдамалы салыстырған кезде бағаланып отырған айырмашылықты бағыты алдын-ала белгілі болса (мысалы, х1=х2 нөлдік жорамалымен х1х2 (бірақ х1х2 емес!)),онда *біржақты критерийді* қолданады. р(х1х2) ықтималдығымен жұмыс істейміз.

Егер айырмашылықтың бағыты белгісіз, көрсетілмеген болса (мысалы, х1х2 немесе х1х2 баламалары есепке алынбаса), онда *екіжақты критерий* қолданылады.

Көп жағдайларда екіжақты критерийді қолдану ұсынылады.

**Параметрлік және параметрлік емес критерийлер**

Статистикалық жорамалдарды тексеру үшін қолданылатын критерийлер екі түрге бөлінеді: **параметрлік** және **параметрлік емес** критерийлер.

***Параметрлік критерийлер*** деп бас жиынтықтағы зерттелетін белгілердің қалыпты таралуына негізделген және олардың негізгі параметрлік (орта мән, дисперсия) есептеуді қажет ететін деректерді өңдеудің статистикалық талдау әдістерін айтады.

Егер бақылаулар саны аз және белгілердің таралу түрі белгісіз, немесе нәтижелер жартылай сандық және сапалық (сырқаттың ауырлығы, реакция қарқындылығы, емдеу нәтижелері) болған жағдайларда параметрлік әдістер жарамайды. Бұл жағдайларда шынайылықты бағалаудың параметрлік емес әдістерін қолдану керек.

**Параметрлік емес критерийлер** деп, бас жиынтықтағы зертелетін белгілердің қалыпты таралуына негізделмеген және олардың негізгі параметрлерін есептеуді талап етпейтін (жиіліктермен немесе рангілермен амалдар орындауға негізделген) деректерді өңдеудің статистикалық талдау әдістерін айтады.

Параметрлік емес критерийлер белгінің таралу түрі қалыпты таралудан ерекше болғанда, таңдама көлемі аз болған жағдайда, сол сияқты реттелген деректерді талдау үшін қолданылады. Егер өзгермелі белгілер санмен емес, ал шартты белгілермен өрнектелген болса бұл жағдайда тек қана параметрлік емес критерий қолданылады. Оларды кейде рангілік әдіс деп те атайды. Бұл критерийлер деректерді олардың рангілерімен алмастырады (яғни реттелген деректер жиынтығындағы олардың орнын сипаттайтын 1,2,3 және т.с.с. сандар).

Параметрлік емес критерийлер әсіресе таңдама көлемі аз (деректердің таралу заңын бағалау мүмкін емес) және деректер категориялық шкалада өрнектелген жағдайда қолдануға ыңғайлы. Алайда, егер деректер саны көп болса онда параметрлік емес критерийлерді қолдану тиімсіз (мағынасы жоқ).

Әрбір параметрлік критерий үшін ең болмағанда бір параметрлік емес аналог болады. Бұл критерийлерді мына топтардың біріне жатқызуға болады:

* Тәуелсіз топтар арасындағы айырмашылық критерийлері;
* Тәуелді топтар арасындағы айырмашылық критерийлері;
* Айнымалылар арасындағы тәуелділік критерийлері.

**Тәуелсіз топтар арасындағы айырмашылық критерийлері**

Егер зертеліп отырған белгілі бір айнымалының орта мәніне қатысты екі таңдамалы салыстыру қажет болса, онда тәуелсіз таңдамалар үшін Стьюденттің жұптаспаған t- критерийі қолданылады. Бұл критерийге параметрлік емес балама критерийлер: Манн-Уитнидің U-критерийі және Колмогоров – Смирновтың екі таңдамалы критерий болып саналады.

**Тәуелді топтар арасындағы айырмашылық критерийлері**

Егер бір ғана таңдама қатысты екі айнымалыны салыстыру қажет болса (мысалы, А гепатитімен ауыратын аурулардың жұқпалы аурулар клиникасына түскен кездегі және одан шығар алдындағы биохимиялық көрсеткіштері), онда Стьюденттің тәуелді таңдамалар үшін жұптасқан t-критерийі қолданылады. Бұл критерийге параметрлік емес балама критерийлер: таңбалардың Z-критерийі және Вилкоксонның Т критерийі болып табылады.

**Айнымалылар арасындағы тәуелділік критерийлері**

Екі айнымалылар арасындағы тәуелділікті бағалау үшін Пирсонның корреляция коэффициенті есептелінеді. Пирсонның корреляция коэффициентінің *параметрлік емес баламасы* Спирменнің рангілік корреляция коэффициенті, Кендаллдың тау және Гамма коэффициенттері болып саналады. Егер екі қарастырылып отырған айнымалылар табиғаты бойынша категориялық болса,онда тәуелділікті бағалау үшін қолайлы параметрлік емес критерийлер Хиквадрат және Фишердің дәл критерийі болады.

**Екіден көп топтардың арасындағы айырмашылық критерийлері**

Егер екіден көп топтардың орта мәндерін салыстыру қажет болса (мысалы, әртүрлі этникалық топқа жататын әйелдерде тромбоциттер саны өзгеріп тұра ма деген сұрақ),онда дисперсиялық талдау қолданылады. Дисперсиялық талдаудың *параметрлік емес баламасы* Крускал- Уоллис критерийі.

**Статистикалық жорамалдарды тексеру кезеңдері.**

1. Нөлдік жорамалды ұйғару және міндетті түрде оған сәйкес балама жорамалды ұйғару.
2. Нөлдік жорамалды тексеру үшін қолайлы статистикалық критерийді таңдау.
3. Таңдама деректері бойынша критерийдің бақылау мәнін есептеу.
4. мәнділік деңгейін таңдау.
5. Берілген мәнділік деңгейі үшін статистикалық критерий таралуының арнайы статистикалық кестесінен сыни аймақты анықтау. Сыни аймақтың өлшемі мәнділік деңгейімен анықталады.
6. Статистиканың есептелген мәнін сыни мәнмен салыстыру.
7. Қорытынды жасау: егер статистиканың есептелген мәні сыни аймақта жатса, нөлдік жорамал жоққа шығарылады және балама жорамал қабылданады; егер критерийдің статистикасы сыни аймаққа түспесе,онда берілген мәнділік деңгейінде нөлдік жорамалды жоққа шығаруға негіз жоқ.

Жорамалды тексеру барысында алынған нәтижелерді дұрыс түсіндіру өте маңызды. Критерий мәнінің мәнді болмауы нөлдік жорамалдың дұрыстығына қатаң дәлел болып табылмайды. Бұл тек қана алынған деректердің оған қайшы келмейтінін көрсетеді.Статистикалық жорамалдарды тексере отырып, біз бас жиынтықтан алынған көлемі шектеулі таңдама деректерімен ғана жұмыс істейтінімізді ұмытпауымыз қажет. Сондықтан статистикалық жорамалдарды тексергенде жасалатын барлық қорытындылар ықтималдық сипатта болады. I және II текті қателердің ықтималдық деңгейлерінің маңызды болатыны да осы себепті.

**4-Бөлім. Сандық деректерге статистикалық талдау жүргізу**

**4.1. Сандық деректер екі тәуелсіз топтар**

Тәуелсіз таңдамалар деп әрқайсысында әртүрлі нысандар бақыланатын таңдамаларды айтады.Мысалы, бірінші бақылау тобы сау адамдар және екінші тәжірибелік топ қандай да бір белгілі емдеу түрін алып жатқан ауру адамдар болуы мүмкін.

Айталық және тәуелсіз бас жиынтықтары бар болсын.Олардан алынған репрезентативті екі тәуелсіз таңдамалар деректері негізінде орта мәндерді салыстыру қажет делік. Мұндай есеп қандай да бір белгілі әсерге душар болған екі аурулар тобын салыстыру – емдеу түрінің ықпал ету дәрежесі жөнінде, ықпалдың мүмкін мәнділігі жөнінде немесе,керісінше, олардың болмауы жөнінде пікір айтуға мүмкіндік береді.

**Жұптаспаған немесе тәуелсіз екітаңдамалы**

**Стьюденттің t-критерийі**

Британ математигі К.Пирсонның оқушысы Вильям Госсет 1908жылы Биометрика журналында бас жиынтықтың көрсеткіштерін олардың таңдама аналогтарымен алмастыруға болатынын көрсететін жұмысын жариялады.Бұл жұмыс баспада жарияланғаннан кейін бақылаулар саны аз таңдамаларды салыстыруға мүмкіндік туды.Уақыт өте келе Стьюдент критерийі ең кең тараған,танымал критерий болды және осы күнге дейін ол медициналық және биологиялық зерттеулерде кең қолданылады.

Стьюденттің жұптаспаған t-критерийі екі тәуелсіз таңдамалар бойынша бас жиынтықтың орта мәндерінің бағалары арасындағы айырмашылықты статистикалық мәнділікке тексеруге мүмкіндік береді.

Стьюдент критерийін қолдануға қойылатын талаптар

1. Салыстырылатын таңдамалардың екеуі де қалыпты таралған бас жиынтықтардан алынған.
2. Тек қана екі топты салыстыруға болады.
3. Бас жиынтықтардың дисперсияларының теңдігін [D(] ескеру қажет. Дисперсиялардың теңдігін анықтау үшін Фишердің F- критерийін қолдану қажет.
4. Стьюденттің t-критерийін тәуелді топтар үшін есептеуге *айырмалар әдісіне* негізделген басқа тәсіл қолданылады.

Стьюдент критерийін есептеу үшін әртүрлі төрт формула қолданылады. Олардың қайсысын таңдау зерттеліп отырған жиынтықтардың дисперсияларының біртектілігіне және бақылаулар санына байланысты.

1. Дисперсиялар тең D()=D() және салыстырылатын таңдамалардағы бақылаулар саны әртүрлі,

t

Бұл жағдайда еркіндік дәрежесінің санын есептеу үшін

dƒ=

формуласы қолданылады.

1. Дисперсиялар тең D()=D() және салыстырылатын таңдамалардағы бақылаулар саны бірдей,

t

формуласы қолданылады.

1. Дисперсиялар тең емес D()D() және салыстырылатын таңдамалардағы бақылаулар саны әртүрлі,

t

Бұл жағдайда еркіндік дәрежесінің санын есепеу үшін күрделірек формуланы қолдану ұсынылады:

dƒ=

1. Дисперсиялар тең емес D()D() және салыстырылатын таңдамалардағы бақылаулар саны бірдей,

t

Бұл жағдайда еркіндік дәрежесінің санын есептеу үшін

dƒ= (n - 1)

формуласы қолданылады.

Келтірілген формулалардағы және - екі таңдамалардың дисперсиялары,- таңдамалардың орта мәндері, –таңдамалардағы бақылаулар саны.

*Орта мәндердің айырмаларының таңбасы есепке алынбайды.*

Тәуелсіз екі таңдамалар үшін жұптаспаған t-критерийін қолдану әдісі:

1. Нөлдік және балама жорамаларды анықтау:

екі популяцияның орта мәндері тең,яғни салыстырылып отырған жиынтықтардың көрсеткіштері арасында айырмашылық жоқ.

екі популяцияның орта мәндері тең емес,яғни салыстырылып отырған жиынтықтардың көрсеткіштері арасында мәнді айырмашылық бар.

1. Зерттеліп отырған екі тәуелсіз таңдамалардың сандық сипаттамаларын және орталарды, дисперсияларды және стандарттық ауытқуларды есептеу.
2. Дисперсиялардың біртектілігін бағалау.
3. Дисперсиялардың теңдігі/тең еместігі жөнінде алынған деректер негізінде t статистикасының мәнін және dƒ еркіндік дәрежесінің санын есептеу.
4. Берілген мәнділік деңгейіне () ; және еркіндік дәрежесіне сәйкес , арнайы кестеден t-критерийінің сыни мәнін табу.
5. жорамалға қатысты шешім қабылдау.

**Егер t-критерийін қолдануға қойылатын талаптар қанағаттандырылмаса, онда оның параметрлік емес баламасын қолдану қажет.**

**Фишердің F критерийі (екі дисперсияларды салыстыру;F-test)**

Фишердің F критерийін есептеу үшін дисперсиялардың қатынасын есептейді :

F=

Алымында **дисперсиялардың үлкен мәні тұру керек, яғни** , сондықтан Фишердің F критерийін есептегенде 1-ден кіші мән алу мүмкін емес.Алынған шаманы кестелік мәнімен салыстырады.

Кестелік мән таңдалынып алынған мәнділік деңгейіне () және еркіндік дәрежесінің санына байланысты. Еркіндік дәрежелерінің санын Формулалары арқылы есептейді.Мұндағы - алымының үлкен дисперсиясының , -бөліміне сәйкес кіші дисперсияның еркіндік дәрежелерінің саны.Егер дисперсиялардың қатынасының есептелген мәні() кестелік мәннен үлкен немесе оған тең болса, , онда берілген мәнділік деңгейінде дисперсиялар тең емес деген қорытынды жасалынады. Егер дисперсиялардың қатынасының есептелген мәні() кестелік мәннен кіші болса, , онда дисперсиялардың теңдігі жөнінде қорытынды жасалынады.

Сиал қышқылының мөлшерін анықтау нәтижелері

|  |  |
| --- | --- |
| 1-топтағы сиал қышқылының мөлшері | 2-топтағы сиал қышқылының мөлшері |
| 240 | 314 |
| 235 | 270 |
| 270 | 220 |
| 280 | 226 |
| 185 | 230 |
| 287 | 305 |
| 148 | 278 |
|  | 210 |
|  | 228 |
|  | 335 |
|  | 305 |
|  | 335 |

2. Семіздіктен зардап шегетін аурулардың артық салмағын тастауға мүмкіндік беретін препараттың тиімділігі зерттелді. Ол үшін еріктілер тобына белгілі бір емдәм тағайындалды. Емдәм ұстаған және препаратты үнемі қабылдаған бір ай уақыттан кейін жоғалтылған салмақтың шамасы белгіленді. Тәжірибе жүргізу үшін 8 адамнан тұратын топ таңдалынып алынды және олардың 3-еуі зерттеліп отырған препаратты (тәжірибелік топ), ал 5-еуі плацебо (бақылау тобы) қабылдады.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Тәжірибелік топ (жоғалтылған салмақ, кг) | 6.2 | 3.0 | 3.9 |  |  |
| Бақылау тобы (жоғалтылған салмақ, кг) | 4.0 | -0.5 | 3.3 | 1.5 | 3.0 |

Препараттың тиімсіздігі жөніндегі жорамалды тексеріңіз.  
Екі топ арасындағы айырмашылықтың статистикалық мәнділігін анықтандар.  
**3.** Қандағы бос гепарин мөлшері:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Қалыпты жағдай | 5.7 | 5.9 | 6.3 | 6.6 | 5.0 | 3.7 | 4.0 | 4.5 | 5.7 | 5.6 |
| Стоматит | 13.9 | 13.5 | 12.0 | 10.3 | 13.0 | 15.7 | 14.7 |  |  |  |

Екі топ арасындағы айырмашылықтың статистикалық мәнділігін анықтаңдар.  
**4.** Солтүстік тұрғындарының гемоглобин көрсеткіштері (г/л):

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Сау балалар (қыздар) | 121 | 124 | 125 | 127 | 131 | 121 | 118 | 122 | 124 |
| Сау балалар (ұлдар) | 125 | 126 | 131 | 136 | 125 | 122 | 119 | 136 | 141 |

Екі топ арасындағы айырмашылықтың статистикалық мәнділігін анықтаңдар.  
 **5.** Манни-Уитни критерийі көмегімен тәуелсіз таңдамалар бір бас жиынтықтан алынғандығы туралы жорамалды тексеріңіз.  
 Дәрмек затының қандағы мөлшері, ммоль/г

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 топ | 8 | 8 | 9 | 10 | 7 | 7 | 9 | 9 | 11 | 6 |
| 2 топ | 8 | 9 | 9 | 11 | 12 | 12 | 13 | 13 | 12 | 11 |

**4.2. Сандық деректер: екі тәуелді топтар**

Тәуелді (өзара байланысқан) таңдамалар деп екі әртүрлі жағдайларда бір таңдама бақыланған (өлшенген) екі айнымалылардан тұратын таңдаманы айтады. Мысалы, зерттеліп отырған топтың қандай-да бір дәрілік препаратты қабылдағанға дейін және қабылдағаннан кейін өлшенеді. Көбінесе, мұндай есептерді «әсерге дейін және әсерден кейін» есебі деп те атайды. Ал деректерді *жұптасқан* деп атайды.  
 *Деректерге талдау жасағанда екі таңдаманың арасындағы тәуелділікті ескеру аса маңызды, олай болмаған күнде жұптасқан байланыстың артықшылығы жоғалады.*  
 Біз мұны әр жұп үшін мәндердің айырмасын есептеп, екі таңдаманы бір айырмалар таңдамасы ретінде қарастыру арқылы шешеміз. Яғни, емдеудің зерттеліп отырған көрсеткішке ықпалы жөніндегі сұраққа жауап беру үшін, емдеуге дейінгі көрсеткіштердің орта мәнін емдеуден кейінгі көрсеткіштердің орта мәнімен **айырмалар әдісі** арқылы салыстырамыз.  
 Егер зерттеліп отырған жиынтықтағы жұптасқан көрсеткіштер, сәйкес олардың айырмалары қалыпты таралған және дисперсиялары бірдей болса, онда деректерге статистикалық талдау жүргізу үшін Стьюденттің жұптасқан t- критерийін қолданамыз.   
 Егер жұптасқан көрсеткіштер, сәйкес олардың айырмалары қалыпты таралу заңына бағынбаса онда жұптасқан t- критерийінің *параметрлік емес баламасы (аналогы)* Вилкоксонның Т- таңбалы рангілер критерийін қолдану қажет.

**Стьюденттің жұптасқан t- критерийі**

Қарастырып отырған деректер жұптасқан болғандықтан екі таңдама көлемі бірдей және n- ге тең. Жұптасқан көрсеткіштер арасындағы айырманы d арқылы белгілейміз.  
 Статистикалық талдауды жүргізу барысы:  
 **1.** Зерттеудің әр жұбы үшін көрсеткіштердің айырмаларын есептеу:  
 d1=x1-x2i; i=1,…, n;

**2.** Алынған d1 мәндерінің орта мәнін, стандарттық ауытқуын және орта мәннің стандарттық қатесін төменгі формулалар бойынша есептеу:

Стандарттық ауытқу: **σd =**

Орта мәннің стандарттық қатесі:

**3.** Критерий статистиканың шамасын есептеу: t =

**4.** Еркіндік дәрежесінің санын есептеу: df=n-1  
 **5.** Арнайы статистикалық кестеден α мәнділік деңгейіне сәйкес сыни нүктені табу.  
 **6.** Критерий статистикасының шамасын сыни мәнмен салыстырып Н0 нөлдік жорамалға қатысты шешім қабылдау:  
 Егер tбак>tсыни болса, онда тәуелді таңдамалардың орта мәндерінің теңдігі жөніндегі Н0 жорамалы жоққа шығарылады.  
 Егер tбак<tсыни болса, онда Н0 жорамалын жоққа шығаруға негіз жоқ.  
 **7.** Статистикалық талдау нәтижелеріне қорытынды жасау.  
 p мәнін түсіндіру және бас жиынтықтағы орта мәндердің шынайы айырмашылығы үшін сенім аралығын есептеу.  
 95% сенім аралығы ( – t0,05 · өрнегімен анықталады.  
 **Тапсырмалар  
 1.** Жаңа туылған 10 нәрестенің дене температурасы қолтық астынан (Х) және тік ішектен (Ү) өлшенеді. Төмендегідей нәтижелер алынды.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| хi | 36,8 | 37,1 | 37,3 | 37 | 37,1 | 36,9 | 36,7 | 37,2 | 37 | 36,9 |
| yi | 36,9 | 37,2 | 37,2 | 37,2 | 37,3 | 37 | 36,8 | 37,1 | 37,2 | 37,1 |

Тік ішектегі температура қолтық астындағы температураға қарағанда жоғары деп санауға бола ма?  
 **2.** Жүректің ишемиялық ауруы кезінде темекі шегу стенокардия ұстамасын тудыруы мүмкін. Бұл никотин миокардтың оттегіге мұқтаждығын арттырады, көміртек тотығы гемоглобинмен байланысып, оттегінің келуін төмендетеді. Алайда ұстаманың дамуына темекі түтінінің басқа компоненттері себепші бола ма? Оны анықтау үшін жүректің ишемиялық ауруымен ауыратын адамдардың стенокардия ұстамасының дамуына дейінгі денеге түсірілген физикалық салмақтың ұзақтығы белгіленеді. Әрбір науқас үшін тәжірибені бес никотинсіз темекіні шеккенге дейін және шеккеннен кейін жүргізді. Келесі деректер алынды:

*Стенокардия ұстамасының дамуына дейінгі денеге түсірілген физикалық   
 салмақтың ұзақтығы.*

|  |  |  |
| --- | --- | --- |
| Науқастар | Никотинсіз темекі шегу | |
| Дейін | Кейін |
| А | 289 | 155 |
| Б | 203 | 117 |
| В | 359 | 187 |
| Г | 243 | 134 |
| Д | 232 | 135 |
| Е | 210 | 119 |
| Ж | 251 | 145 |
| З | 246 | 121 |
| И | 224 | 136 |
| К | 239 | 124 |
| Л | 220 | 118 |
| М | 211 | 107 |

Бұл деректер бойынша қандай қорытынды жасауға болады?

**3.** Балаларда ынтасыз сілекей бөліну жылдамдығы (мл/мин):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Операцияға дейін | 0,18 | 0,19 | 0,14 | 0,22 | 0,15 | 0,17 |
| Уранопластикадан кейін | 0,2 | 0,22 | 0,16 | 0,26 | 0,15 | 0,18 |

Екі көрсеткіштің орта мәндері арасындағы айырмашылықтың статистикалық мәнділігін анықтаңдар.  
 **4.** Жыл сайынғы тексеру кезінде денсаулық сақтауды ұйымдастыруды және әлеуметтік қамтамасыз етуді жақсарту үшін тұрғындардың 1000 адамына шаққандағы аурудың таралу көрсеткіші зерттелді. Тексерудің бастапқы және соңғы кезеңдеріндегі сұрау нәтижесінде 10 жыл ішіндегі көрсеткіштер мәндері алынды.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Тексерудің бастапқы кезеңі | 46,1 | 43,8 | 37,9 | 41,6 | 40,7 | 46,3 | 45,9 | 46,9 | 50,3 | 53,7 |
| Тексерудің соңғы кезеңі | 50,5 | 40,3 | 42,2 | 40,1 | 36,9 | 43,4 | 52,5 | 54,7 | 42,1 | 52,2 |

Cтьюденттің жұптасқан критерийін қолданып, тексеру уақытында қолданған шаралардың тиімділігін тексеру қажет.  
 **5.** Жүректің ишемиялық ауруы кезіндегі несептегі норадреналин мөлшері (мкг/тәулік):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Емдеуге дейін | 36,9 | 38,2 | 36,1 | 33,5 | 34,8 | 37,0 | 35,1 | 40,0 | 38,5 | 38,3 |
| Емдеуден кейін | 40,5 | 43,8 | 56,7 | 49,8 | 50,4 | 45,4 | 39,9 | 38,7 | 40,3 | 51,6 |

Екі көрсеткіштің орта мәндері арасындағы айырмашылықтың статистикалық мәнділігін анықтаңдар.

**6.** Тиреотоксикоз кезіндегі трийодиронин мөлшері (мг/мл):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Емдеуге дейін | 4,35 | 5,01 | 6,03 | 4,35 | 5,40 | 5,12 | 2,02 | 2,65 | 4,56 | 5,25 |
| Емдеуден кейін | 2,25 | 4,40 | 3,15 | 2,35 | 1,35 | 3,40 | 2,31 | 1,72 | 3,75 | 4,21 |

**7.** Тіс тасы пайда болуының алдын алу мақсатында екі дәрі салыстырылды. Әрбір зерттелуші 48 сағат бойы тісін дәрінің бір түрімен шайған соң тіс тасы бағаланып отырды. Біршама уақыттан кейін тәжірибені дәрінің басқа түрімен қайталады. Нәтижесінде төмендегідей деректер алынды:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Аммоний хлориді | 32 | 60 | 25 | 45 | 65 | 60 | 68 | 83 | 120 | 110 |
| Хлоргексидин | 14 | 39 | 24 | 13 | 9 | 3 | 10 | 14 | 1 | 36 |

Аммоний хлоридімен тіс шаюдың тиімділігін бағалаңыз.

**Вилкоксонның Т-таңбалы рангілер критерийі** *(Wilcoxon test)*

Вилкоксонның Т-критерийі жұптасқан деректердің «плюс» және «минус» таңбалы айырмаларына рангілер беруге негізделген және бас жиынтықтағы жұптасқан бақылаулар арасында айырмашылықтың жоқтығы, яғни жұптасқан деректер үшін медианалардың айырмасының нөлге теңдігі жөніндегі Н₀ нөлдік жорамалды тексеру үшін қолданылады.

Алғаш рет Френк Уилкоксон ұсынған.

Басқаша атаулары: Вилкоксонның W критерийі, Уилкоксонның рангілер қосындысы критерийі, Тәуелді таңдамалар үшін Уилкоксон критерийі. (таңдама көлемі: 5≤ n ≤ 50)

Статистикалық талдауды жүргізу барысы:

1. Нөлдік және балама жорамалдар ұйғарылады және мәнділік деңгейі тағайындалады.
2. Жұп мәндердің әрқайсысы үшін олардың арасындағы айырмалар есептеледі және оларға сәйкес «плюс» және «минус» таңбалары беріледі.
3. Айырмалары нөлге тең жұптар (нөлдік өзгерістер) әрі қарай талдаудан шығарылып тасталады және n мәні сәйкес азаяды.
4. Айырмалар мәндердің өсу ретіне қарай таңбаны есепке алмай реттеліп жазылады ( абсолюттік шамалары бойынша вариациялық қатар құру).
5. Айырмалардың вариациялық қатары ранжирленеді, бірдей айырмаларға сәйкес рангілердің орта мәндері беріледі.
6. Оң таңбалы айырмалардың рангілерінің қосындысы (Т₊) және теріс таңбалы айырмалардың рангілерінің қосындысы (Т₋) табу.
7. Т критерийінің статистикасы қосындылардың кішісіне тең, яғни сирек кездесетін таңба – (типтік емес өзгерістер) сәйкес келетін рангілер қосындысына тең.
8. Т критерийінің есептелген мәнін кестелік сыни мәнмен салыстыру және Н₀ нөлдік жорамалға қатысты шешім қабылдау.

Егер > болса, онда тәуелді таңдамалар арасындағы айырмашылықтың жоқтығы жөніндегі Н₀ жорамалы қабылданады.

**Мысал**. Ультрадыбысты қондырғыда жұмыс істейтін адамның қандарындағы қант мөлшері аш қарынға жұмысқа дейін және 3 сағат жұмыс істегеннен кейін өлшенеді. Өлшеу нәтижелері кестеде келтірілген.

*Жұмысқа дейін аш қарынға және ультрадыбысты қондырғыда 3 сағат жұмыс істегеннен кейін тексерілгендердің қандарындағы қант мөлшері (мг%).*

|  |  |
| --- | --- |
| Жұмысқа дейінгі қант мөлшері | Жұмыстан кейінгі қант мөлшері |
| 112 | 54 |
| 82 | 67 |
| 101 | 96 |
| 72 | 59 |
| 79 | 79 |
| 82 | 76 |
| 64 | 66 |
| 70 | 66 |
| 88 | 48 |
| 81 | 50 |
| 66 | 61 |
| 88 | 61 |

Ультрадыбысты қондырғыда жұмыс істеу қандағы қант мөлшеріне әсер етеді деп айтуға бола ма?

**Шешуі:**

1. Н₀ нөлдік жорамал мен Н₁ балама жорамалдары ұйғарамыз.

Н₀: Екі таңдама көрсеткіштерінің арасындағы айырмашылық статистикалық мәнді емес, яғни ультрадыбысты қондырғыда жұмыс істеу қандығы қант мөлшеріне әсер етпейді.

Н₁: Екі таңдама көрсеткіштерінің арасындағы айырмашылық статистикалық мәнді, яғни ультрадыбысты қондырғыда жұмыс істеу қандағы қант мөлшеріне маңызды әсер етеді.

1. Вилкоксонның жұптасқан Т статистикасын есептеу нәтижелерін кестеге енгізейік. Өзгерген шамаларды абсолюттік шамалары бойынша ранжирлейміз. Рангілердің қосындысын есептеп, оның формула бойынша есептелген мәнмен бірдей болуын тексереміз.

9-кесте

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Дейін | Кейін | Өзгерістер  (айырмалардың  мәндері) | Айырмалардың  абсолюттік  шамалары | Рангілер  (R) | Типті емес өзгерістердің белгісі |
| 1 | 2 | 3 | 4 | 5 | 6 |
| 112 | 54 | -58 | 58 | 11 |  |
| 82 | 67 | -15 | 15 | 7 |  |
| 101 | 96 | -5 | 5 | 3,5 |  |
| 72 | 59 | -13 | 13 | 6 |  |
| 79 | 79 | 0 | 0 | - |  |
| 82 | 76 | -6 | 6 | 5 |  |
| 64 | 66 | 2 | 2 | 1 | \* |
| 70 | 66 | -4 | 4 | 2 |  |
| 88 | 48 | -40 | 40 | 10 |  |
| 81 | 50 | -31 | 31 | 9 |  |
| 66 | 61 | -5 | 5 | 3,5 |  |
| 88 | 61 | -27 | 27 | 8 |  |
|  |  |  |  | *=*66 | =1 |

Деректерді Вилкоксонның жұптасқан Т-критерийі бойынша өңдеу төмендегідей түрде жүргізіледі:

1. Бақыланған белгілердің өзгерген шамаларын есептейміз. Ол үшін «кейінгі» көрсеткіштерден сәйкес «дейінгі» көрсеткіштерді азайтамыз (3- баған). Нөлдік өзгеріс болған (өзгеріс болмаған) көрсеткіштерді әрі қарай талдаудан алып тастаймыз.
2. Өзгерістердің әрбір мәніне сәйкес абсолюттік шамасын қоямыз (4- баған).
3. Өзгерістердің абсолюттік шамаларын ранжирлейміз (5- баған).
4. Рангілердің қосындысын есептейміз. Біздің мысалда ол 66- ға тең.
5. Рангілер қосындысын формула бойынша есептейміз:

(N=11, себебі бір көрсеткіште өзгеріс болған жоқ, сәйкес таңдама көлемі азайды).

1. Екі тәсілмен алынған рангілер қосындыларын салыстыру арқылы ранжирлеудің дұрыстығын тексереміз. Біздің жағдайда екі шама бірдей 66=66, демек, ранжирлеу дұрыс жүргізілді.
2. Кестеде (6-баған) типтік емес өзгерістерді (\*) белгісімен белгіледік. Біздің жағдайда бұл – бір оң өзгеріс.
3. Типтік емес өзгерістердің рангілерін қосамыз. Сол қосынды ізделініп отырған шамасы болады. Біздің жағдайда бұл қосынды: .
4. және n=11 үшін сыни мәнді анықтаймыз. (3-қосымша, 3.4- кесте) .
5. Критерий статистикасын сыни мәнмен салыстырамыз

демек Н₀ нөлдік жорамалы жоққа шығарылады (р<0.05).

**Қорытынды.** Ультрадыбысты қондырғыда жұмыс істеу қандағы қант мөлшеріне маңызды әсер етеді, яғни қант мөлшерін азайтады.

**Z таңбалар критерийі (sign test)**

Z таңбалар критерийімен жұптасқан бақылауларды (мысалы, емдеуге дейін және кейін) бағалағанда қолданылатын Т критерийінен өзгешелігі ол өзгеріске ұшыраған шамаларды емес, ал тек қана олардың бағыттарын ескереді. Сондықтан бұл өзгерістердің сипаты балама түрде ескіріледі (ұлғайған - кеміген, нашарлаған – жақсарған және с.с., қысқару үшін әдетте «+» және «-» таңбаларымен белгіленеді, осыдан келіп критерийдің атауы шыққан). Айырмасы нөлге тең жұптасқан бақылаулар (= таңбасымен немесе 0-мен белгілеуге болады) әрі қарай салыстырылудан шығарылып тасталынады. Осыған байланысты нөлдік айырмалар санын мейлінше аз болдыруға тырысу қажет (сандық және жартылай сандық бақылаулардың өлшеу дәлдігін жоғарылату арқылы таңдама деректердің үзіліссіздігін қамтамасыз ету).

Егер оң өзгерістер саны теріс өзгерістер санына жақын болса, онда салыстырылып отырған таңдама жиынтықтар арасындағы айырмашылық статистикалық мәнді бола алмайтыны айқын. Керісінше, мәнді айырмашылықтар ықтималдығы өзгерістер бір жаққа қарай елеулі бағытталғын жағдайда, яғни таңбалардың біреуі басым болған жағдайда артады.

Таңбалар критерийін практикада қолдану төмендегідей түрде жүзеге асады:

1. Салыстырылып отырған жұптасқан бақылаулардағы өзгерістердің бағыттары анықталады және бақылаудың әр жұбы үшін «+» немесе «-» таңбаларымен, ал өзгерістер жоқ болған жағдайда 0-мен белгіленеді;
2. Айырмашылығы бар (яғни, «+» және «-» таңбаларымен белгіленген) жұп бақылаулардың жалпы саны (n) саналады;
3. Салыстырудың бірдей нәтижелерінің (яғни, + немесе – таңбалары санының) аз саны есептеледі де Z әрпімен белгіленеді;
4. Алынған Z саны берілген жұп бақылаулар саны үшін арнайы кестедегі сыни мәндермен (салыстырылады;
5. Егер Z саны (5% мәнділік деңгейіне сәйкес келетін) сыни кестелік мәнге тең немесе одан үлкен болса, онда орын алған өзгерістер кездейсоқ, статистикалық мәнді емес ( нөлдік жорамал дұрыс) деген қорытынды жасалады.

Егер Z саны немесе сыни мәндерінен кіші болса, онда айырмашылық 5%-дан кем (1%-дан кем) қате жіберу ықтималдылығымен мәнді деп саналады.

**Мысал.** 10 науқас адамның өттеріндегі билирубиннің мөлшері антибиотик енгізгенге дейін және енгізгеннен кейін өлшенеді.

Билирубин мөлшерінің айырмашылығының статистикалық маңыздылығын анықтау қажет.

10-кесте

*Өттегі билирубиннің антибиотиктер енгізгенге*

*дейінгі және кейінгі мөлшерлері*

|  |  |  |
| --- | --- | --- |
| Науқастар | Билирубиннің мөлшері  Енгізгенге дейін Енгізгеннен кейін | Әсердің бағыты |
| А | 68 110 | + |
| Ә | 83 101 | + |
| Б | 70 120 | + |
| Г | 100 180 | + |
| Д | 110 100 | - |
| Е | 100 100 | 0 |
| Ж | 180 240 | + |
| З | 60 120 | + |
| И | 200 160 | - |
| К | 210 300 | + |

**Шешуі.**

1. Билирубиннің көбеюі 7 науқаста, ал азаюы 2 науқаста байқалды.
2. Жұп бақылаулар саны 10-ға тең болады. Бір жұпта нәтижелер теңдей болады.
3. Бағалау бір бағыттағы нәтижелердің ең аз саны бойынша жүргізіледі. Берілген жағдайда бұл қандағы билирубин мөлшері азайған 2 науқаста болды.
4. Өзара байланысқан екі жиынтықтың арасындағы айырмашылықтың шынайылығы кесте бойынша жүргізіледі. Кестені пайдаланғанда жалпы жұп бақылаулар санынан дейінгі және кейінгі әсерде нәтижелері бірдей болған, яғни өзгерістер болмаған жағдайлар алып тасталынады. Сонымен, біздің мысалда n=9 (10-1), Z=2 (бір бағыттағы нәтижелердің аз саны). Кесте бойынша (3 қосымша, 3.5- кесте).

**Қорытынды.** Билирубин мөлшеріндегі өзгерістер статистикалық мәнді емес.

**Тапсырмалар**

1. Қанның ұю уақытын анықтаудың екі әдісін салыстыру үшін, әрбір сынама төмендегідей екі әдіспен бағаланады:

Бюркер бойынша – бөлме температурасында фибрин жіптерінің пайда болуы;

Ли - Уайт бойынша- 37 градус Цельсияда термостаттағы шыны түтікті аударғанда қан төгілмейді.

Екі әдіспен анықтағандағы қанның ұю уақыты

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BURKER | 10 | 9 | 8 | 8 | 7 | 7 | 5 | 5 | 6 | 6 | 7 |
| LIWITE | 10 | 8 | 9 | 10 | 6 | 10 | 6 | 6 | 7 | 7 | 9 |

Вилкоксон критерийін қолданып, екі әдістің қанның ұю уақыттарындағы айырмашылықтарының мәнділігін бағалаңдар.

1. Созылмалы уремиямен ауыратын аурулардың арнайы рационмен емделгенге дейінгі және емнен кейінгі гемоглабин көрсеткіштері арасындағы айырмашылықтардың шынайылығын екі тәуелді таңдама жиынтықтар үшін параметрлік емес критерийді қолданып тексеру қажет.

|  |  |
| --- | --- |
| Ауру | Гемоглабин %  Емдеуге дейін Емдеуден кейін |
| А  Б  В  Г  Д  Е  Ж  З | 60 59  74 68  78 70  46 49  40 40  74 80  62 60  65 67 |

1. Артық салмақты тастауға мүмкіндік беретін арнайы емдәмнің тиімділігін анықтау керек. Зерттеуге қатысушылардың әрқайсысының салмағы тәжірибеге дейін және емдәмді қолданғаннан кейін өлшенеді. Бес қатысушыдан тұратын топ үшін алынған деректер төмендегідей.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Тәжірибеге дейінгісалмақ, кг | 93,2 | 98,2 | 105,6 | 86,8 | 95,5 |
| Тәжірибеден кейінгі салмақ, кг | 88,9 | 94,5 | 106,1 | 84,3 | 92,5 |

* 1. ***Сандық деректер: екіден көп топтар***

Медициналық-биологиялық зерттеулерде бір топты екінші топпен салыстыру тәсілдерімен қатар бір мезгілде бірнеше таңдамаларды салыстыруды қажет ететін есептер жиі кездеседі. Бұл жағдайда топтардың орта мәндерін жұп-жұбымен (әр топты әрқайсысымен) салыстыру тиімсіз болуы мүмкін, себебі салыстыру санының көптігінің салдарынан туатын I текті қателіктің жоғары мәні нәтижелер қорытындысының дұрыс жасалмауына әкеліп соғар еді. Сондықтан, салыстырылып отырған барлық топтардың арасында айырмашылықтың бар-жоғын бірден анықтау үшін *бір ортақ статистикалық критерий* қолданамыз.