

3-апта
Дәріс №5-6
Зерттеудің сандық және сапалық түрлері

1. Сандық зерттеулер туралы
2. Сапалық зерттеулер туралы

1. Сандық зерттеулердің негізгі сипаттамалары. Негізінен сандық зерттеулер жаратылыстану ғылымынан әлеуметтік ғылымдарға кеелген. Солай бола тұрса да, әлеуметтік ғылымдардағы сандық зерттеулердің бірқатар ерекшеліктері бар.

- сандарды қолдану. – сандық зерттеулердің ең маңызды белгісі сандардың айналасына құрылуы. Бұл бір жағынан зерттеушіге көп мүмкіндіктер береді, екінші жағынан, кейбір шектеулер қатарын да тудырады. Сандар математика ғылымында дәлелденгендей күшке ие. Сондай-ақ сандар әлсіз, себебі олар зерттеу контексінде «контекстік өткенісіз» (): олар мәнсіз, егер біз белгілі бір сан қолданылатын қандай да бір категорияны және белгілі бір айнымалы шеңберіндегі әртүрлі құндылықтарды анықтамайынша. Сонда бізге біз қолданатын айнымалылардың мазмұны және шекараларының нақты анықтамалары және айнымалы шеңберінде жұмсалатын түрлі құндылықтардың сипаттамалары керек. Осының бәрін қамтамасыз ете алған жағдайда ғана сандар жұмыс жасайды.

- категориялау. - әрбір зерттеу тәсілдемесіне ортақ нәрсе, ол - қоршаған дүниеден жинақталатын шексіз ақпарат оны басқара алатынға дейін қысқартылуы керек. Зерттеушілер бұл ақпаратты құрылымдау және формалау үшін «категориялар» немесе «кодтарды» пайдаланады бірақ осы жерден сапалық пен сандықтың табиғатындағы ұқсастық аяқталады. Категория мен категориялау процесінің табиғаты бұл екі зерттеу тәсілдемесінде екітүрлі. Сандық зерттеулерде зерттеу нақты кодтау кестесімен басталуы мүмкін, мысалы гендер айнымалысында ер -1 әйел -2 деп белгіленеді. (сапалық зерттеулерде кодтар вербалды, сандық емес.)

- айнымалылар – сандық зерттеушілер адамдарға не адамдар тобына көп қызықпайды, олар зерттеулерін айнымалылар төңірегіне құрады. Айнымалыларды зерттейді. Сандық зерттеулер айнымалылар арасындағы қарым-қатынасты айқындауды мақсат етеді.

- статистика және статистика тілі. – екеуінің арасындағы тығыз байланыстылықтен статистикалық терминологияның көбі сандық зерттеулердің сөздік құрамының бөлігі болып кеткен.

Сандық зерттеулердің артықшылықтары мен кемшіліктері. Сандық зерттеулердің жақтаушылары *артықшылықтары* деп мыналарды айтады: жүйелі, қатаң/нақты (rigorous), қатаң бақыланады, нақты өлшемдерді қамтиды, басқа контекстерге қолдануға болатын сенімді және қайталанатын деректермен қамтамасыз етеді. Практикалық тұрғыдан дайындық кезеңі ұзақ уақытты қажет еткенімен, зерттеудің процесі салыстырмалы түрде тез, қаражат үшін құнды, себебі деректерді талдау статистикалық компьютерлік бағдарламаны қолдану арқылы жасала алады. Кемшіліктері олар барлық бақыланған қатысушылар тобы бойынша жауаптардың орташа есебін алады, жеке өмірдің алуан түрлілігін әділетті ету мүмкін емес.

Сандық талдаудың артықшылықтары

- Ғылыми. Сандық деректер математика мен ықтималдық принциптеріне негізделген статистикалық әдістердің әртүрлі формаларына сәйкес келеді. Мұндай статистика талдауларға ғылыми беделділік аурасын береді. Талдаулар зерттеушінің құндылықтарына емес, объективті заңдылықтарға негізделген сияқты.
- Сенім. Маңыздылықтың статистикалық сынақтары зерттеушілерге олар жасаған интерпретациялар мен олардың қорытындыларына сенімділік тұрғысынан қосымша сенімділік береді.

- Өлшеу. Сандық деректерді талдау сипаттау мен талдау үшін берік негіз береді. Түсіндірулер мен тұжырымдар әсерлерге емес, өлшенген шамаларға негізделеді және бұл, ең болмағанда, негізінен, түпнұсқалық үшін басқалар тексере алатын шамалар.
- Талдау. Алдын ала тиісті дайындық және жоспарлау болған жағдайда, сандық деректердің үлкен көлемін салыстырмалы түрде жылдам талдауға болады. Процедура «жұмыс істеп» болғаннан кейін, зерттеушілер өз нәтижелерін салыстырмалы түрде тез сұрай алады.
- Презентация. Кестелер мен диаграммалар сандық деректерді ұйымдастырудың және нәтижелерді басқаларға жеткізудің қысқа және тиімді әдісін қамтамасыз етеді. Кеңінен қол жетімді компьютерлік бағдарламалық қамтамасыз ету кестелер мен диаграммаларды жобалауға көмектеседі және статистикалық талдаудан ауыр жұмыстың көп бөлігін алады.

Сандық талдаудың кемшіліктері

- Деректер сапасы. Сандық деректер оларды жинау үшін қолданылатын әдістер мен қойылған сұрақтар сияқты жақсы.

- Техник. Зерттеушілер зерттеудің негізінде жатқан кеңірек мәселелер есебінен талдау әдістерімен әуестеніп кету қаупі бар. Атап айтқанда, зерттеушілердің қолындағы компьютерлердің күшімен назар зерттеудің нақты мақсатынан талдаудың техникалық аспектілеріне назар аударуы мүмкін.

- Деректердің шамадан тыс жүктелуі. Деректердің үлкен көлемі сандық талдаудың күші болуы мүмкін, бірақ абайсызда зерттеушіні шамадан тыс жүктеуі мүмкін. Тым көп жағдайлар, тым көп айнымалылар, тым көп факторларды ескеру – талдауды тым күрделілікке бағыттауға болады. Зерттеуші батпақта болуы мүмкін. Сандық талдау сырттай көрінетіндей ғылыми объективті емес. Зерттеушілерде сапалы деректер сияқты деректерді талдауды жүргізу кезінде қалау және таңдау элементі бар. Және олардың шешімдері қорытындыларға әсер етуі мүмкін. Сандық деректерді талдау техникалық және ғылыми болып көрінуі мүмкін болғанымен, іс жүзінде зерттеуші әлі де нәтижелерге нәзік жолдармен әсер ету мүмкіндігіне ие.

Сандық тәсілдемемен орындалған зерттеулерден алынған деректерді сандық деректер деп атауға болады. Сандық деректер сандар түрінде болады. Олар, ең алдымен, сауалнамалар (surveys) мен эксперименттер сияқты зерттеу стратегияларына (Денскомб стратегия деген) және анкеталар (questionnaires) мен бақылау сияқты зерттеу әдістеріне (мына екеуін әдістер деген) қатысты. Дегенмен, бұл сандық деректердің жалғыз көздері емес. Мысалы, мәтіндермен (мысалы, сұхбат транскриптітері) мазмұнды талдауды пайдалану (the use of content analysis with texts (such as interview transcripts) сандық деректерді де бере алады. Осылайша, сандық деректерді анықтау кезінде қолданылатын зерттеу әдісінің түрі маңызды емес. Негізгі мәселе - бұл әдіс тудыратын деректердің сипаты. Осыны есте ұстаған жөн және 13.1-кестеде көрсетілгендей сандық деректерді әртүрлі зерттеу әдістерін қолдану арқылы алуға болатынын білу керек.

| Numbers | Research method |
|--|-----------------|
| Answers to closed-ended questions | Questionnaires |
| Content analysis of transcripts Interviews | Interviews |
| Measurements from experiments | } Observation |
| Observation schedule used with events | |
| Official statistics (health, education, trade, etc.) | } |
| Business data (performance, employment, etc.) | |

Content analysis of, for example, company reports

Статистика дереккөздерін көрсету

<http://en.unesco.kz/> Resources-Statistics

<https://stat.gov.kz/> Ресми сатистика-Салалар

Сандық деректердің түрлері

Сандар заттардың әртүрлі түрлерін көрсете алады және сандық деректерді пайдаланғысы келетін әлеуметтік зерттеушілер сандық деректердің қандай түрін қолданатынына толық сенімді болуы керек (Мартин Денскомб. The Good Research Guide For small-scale social research projects Fourth Edition 2010). Бұл деректермен не істеуге болатыны, статистика тұрғысынан және талдаудан шығатын қорытындылар үшін өте маңызды. Кейбір деректер түрлерінде жұмыс істейтін, бірақ басқаларында жұмыс істемейтін белгілі бір статистикалық техникалар бар. Осыған байланысты сандық деректерді алты тақырып бойынша қарастыруға болады. Алғашқы төртеуі Стивенс (1946) белгілеген өлшемдердің төрт түрі арасындағы айырмашылықты қарастырады. Олар номиналды nominal, реттік ordinal, интервалдық interval және қатынас ratio деп аталады. Бұған қоса, зерттеушілер «дискретті» және «үздіксіз» деректер арасындағы айырмашылықты да білуі керек.

Номиналды деректер

Номиналды деректер заттарды санау және оларды категорияға орналастыру арқылы алынады. Бұл басқа түрлермен салыстырғанда статистикалық манипуляцияға аз мүмкіндік беретін сандық деректердің ең төменгі деңгейі. Әдетте, еркек/әйел немесе Ақ/Оңтүстік Азия/Африкалық Кариб сияқты белгілі бір категориядағы мүшелердің жалпы саны бар. Бұл категориялар жай ғана атауларға негізделген; атауларда негізгі тәртіп жоқ.

Реттік деректер

Номиналды деректер сияқты, реттік деректер белгілі бір санаттарға тағайындалған заттардың санына негізделген, бірақ бұл жағдайда санаттар қандай да бір анық, реттелген, деңгейленген қатынаста тұрады. Санаттар - «ретімен». Бұл әрбір санаттағы деректерді басқа санаттардағы деректерден жоғары немесе төмен, артық немесе аз және т.б. сияқты басқа санаттардағы деректермен салыстыруға болатындығын білдіреді. Реттік деректердің ең айқын мысалы респонденттерге бес балдық шкала бойынша жауап беру сұралатын сауалнаманы пайдаланудан келеді, мысалы:

Statement

Academic detailing is a useful form of education that aligns providers' prescribing behavior with evidence-based practice.

| | | | | |
|--------------------------|-----------------|----------------|--------------|-----------------------|
| Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 1 | 2 | 3 | 4 | 5 |

Бұл шкала Likert шкаласы (Likert scale) ретінде белгілі. Айта кету керек, тек реттік тәртібі - бұл қорытынды жасауға болады. - Стоит подчеркнуть, что порядок рангов — это все, что можно вывести. Реттік деректермен біз реттіктің себебін немесе олардың қаншалықты ерекшеленетінін білмейміз.

Аралық деректер

Аралық деректер реттік деректер сияқты, бірақ санаттар шкала бойынша реттелген. Бұл категориялар арасындағы «қашықтықтың» белгілі фактор екенін және оның пропорционалды екенін білдіреді. Бұл зерттеуші деректермен «артық» немесе «кем» терминімен жұмыс істеп қана қоймай, қанша көп немесе қанша аз екенін айта алады дегенді білдіреді.

Күнтізбелік жылдар мұндай деректердің лайықты мысалын береді. 1966, 1976, 1986, 1996 және 2006 жылдар үшін жиналған деректер бір-бірінен ерте немесе кеш болуы тұрғысынан ғана ерекшеленбейді, сонымен қатар белгілі уақыт аралығы бойынша ерте немесе кеш болады. Бұл зерттеушіге әртүрлі кезеңдер арасындағы айырмашылықты салыстыру үшін қосу мен азайтуды (бірақ көбейту немесе бөлуді емес) қолдануға мүмкіндік береді: 1976 және 1986 жылдар арасындағы айырмашылықты 1996 және 2006 жылдар арасындағы айырмашылықпен тікелей салыстыруға болады және т.б.

Қатынас деректері

Бұл қатынастар аралық деректерге ұқсас, тек санаттар «шын нөл» немесе абсолютті сілтеме нүктесі бар шкалада бар. Санаттар табыс, қашықтық және салмақ сияқты нәрселер туралы болғанда, олар қарым-қатынас деректерін береді, себебі таразыларда нөлдік нүкте бар. Алдыңғы мысалдағы күнтізбелік жылдар бұл шкалада жоқ, себебі 0 жыл барлық уақыт пен тарихтың басын белгілемейді. Шынайы нөлдік шкаланың маңызды ерекшелігі – зерттеуші интервалдағыдай қосу және алуды қолданумен шектелмей, көбейту мен бөлуді пайдалана отырып, қатынас тұрғысынан әрбір санат бойынша деректерді салыстыра және салыстыра алады. Қатынас деректері математикалық манипуляцияға қаншалықты қолайлы екендігі тұрғысынан деректердің ең жоғары деңгейі болып табылады.

Дискретті деректер

Белгілі бір деректер табиғи түрде тұтас бірліктермен келетін құбылыстарға негізделген. Бір отбасына шаққандағы балалар саны, шын мәнінде, бүтін сандар болуы керек. Мүмкін сандарды біріктіріп, орташа есеппен 1,9 немесе кез келген нәрсеге жету үшін, бірақ біз отбасында 0,9 бала бар деп ойламаңыз.

Бұл мысалда әрбір отбасына шаққандағы балаларды өлшеуіміз нақты болуы мүмкін және әрбір үй шаруашылығында, жуықтауларды немесе дәл нәтижелерді қажет етпейді ең жақын бөлшек. Дискретті деректер бөліктерге бөлінеді: 1, 2, 3, 4 және т.б.

Үздіксіз деректер

Бұған қарама-қарсы деректердің белгілі бір түрлері бар, олар практикалық мақсаттарда сөзсіз «ең жақын бірлікке дейін» өлшенеді, өйткені олар ұқыпты, дискретті бөліктерге келмейтіндіктен. Мұндай заттар ең жақын шағын бірлікке дейін өлшенеді, өйткені олар айнымалы ретінде үздіксіз болады. Мұнда адамдардың бойы, жасы және салмағы айқын мысалдар болып табылады. Негізінде, бұл элементтерді өлшеуге тырысатын дәлдіктің шегі жоқ. Біз биіктікті миллиметрмен, салмақты грамммен, жасты секундпен өлшей аламыз. Алайда практикалық мақсаттар үшін мұндай дәлдік кері нәтиже береді, өйткені ол талдауға мүмкіндік бермейтін деректердің үлкен массивтерін шығарады.

Айнымалылар үздіксіз болған кезде, зерттеуші, шын мәнінде, деректерді санаттауы керек. Деректер бірлігі нақты бой, нақты салмақ, нақты жас емес; керісінше, бұл «бойдың 2 см диапазонға түсуі», «салмағы 0,5 келіге дейінгі» немесе «соңғы туған күніндегі жасы». Үздіксіз деректердің маңыздылығы жылжымалы шкала бойынша нүктелерді/бірліктерді өлшеу қиындықтарын жеңу үшін зерттеуші қалыптастыратын жасырын категорияларға байланысты. Зерттеуші мұндай категориялардың шекараларын және категорияның орта нүктесі қай жерде орналасқанын анық білуі керек. Үздіксіз деректерге негізделген статистиканың дәлдігі деректер санаттарының шекаралары мен ортасын білуге байланысты.

Сандық әдістер туралы айту міндетті түрде айнымалылар туралы айтуды білдіреді сондықтан сандық зерттеулерде айнымалылар дегеніміз және сандық зерттеулердегі басқа да маңызды концептілер қандай екенін анықтау маңызды, әсіресе бұл терминдерді қате түсіну кез келген сандық зерттеу кезінде елеулі проблемаларға әкелуі мүмкін. Oxford English Dictionary (OED) айнымалы (зат есім) дегенді былайша анықтайды: «Something which is liable to vary or change; a changeable factor, feature, or element». //«Құбылтуға немесе өзгертуге болатын нәрсе; өзгертін фактор, қасиет-белгі немесе элемент».

Басқаша сөзбен айтқана айнымалы белгілі бір жағдайдың белгісі болып табылады. Мысалы, лингвистикадағы (және басқа әлеуметтік психологиялық ғылымдардағы) жиі қолданылатын айнымалы – гендер. Гендер – айнымалы, оның екі ықтимал мәні бар: ер немесе әйел.

Айнымалылар категориялық (categorical) және үздіксіз (continuous) болады. Тәуелді айнымалы категориялық болғанда біз қолданатын статистикалық тест *chi-square test* (sometimes abbreviated as χ^2) деп аталады. Тәуелді айнымалы үздіксіз болғанда қолданатын статистикалық тест *t-test* деп аталады.

Бұл үшін ең икемді және, әрине, ең жиі қолданылатын статистикалық сынақ хи-квадрат тесті болуы мүмкін. Хи-квадрат сынағы номиналды деректермен, сондай-ақ реттік, интервалдық және қатынас деректерімен жұмыс істейді - бұл оның танымалдылығын түсіндіру үшін көп нәрсе жасайды. Бағандар мен жолдар әлеуметтік зерттеушілер жинаған деректерде ерекше орын алатын жыныс, этникалық және кәсіп сияқты факторларға негізделген номиналды деректерді пайдалана алады (13.6 кестені қараңыз).

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

χ^2 = the test statistic \sum = the sum of

O = Observed frequencies E = Expected frequencies

Хи-квадрат тесті //Chi-square tests. Біз хи-квадрат тестін сипаттаудан бастаймыз, өйткені оларды табиғи тіл деректеріне қолдануға болады. Бұл иллюстрация үшін біз Шарманың (2005) үнділік ағылшын тілінде сөйлеушілер арасында белгілі және белгісіз

артиклдердің қолданылуын зерттеуден алынған деректерді қолданамыз. Біз қарастыратын деректерде Шарма сөйлеушілердің артиклдерді келесідей сөйлемдерде қолдануын зерттейді (Платт және басқалар, 1984)

1) I want to spend some time in *a* village, definitely if I get *a* chance.

Бізді қызықтыратын артикл - *a* белгісіз артиклі 'a village' and 'a chance' дегендердегі. Хинди тілінде Шарма қарастыратын L1 спикерлері 'a village' and 'a chance' сияқты зат есімді фразаларды ешқандай артиклсіз жұмсайды. Үнді ағылшын тілінде сөйлейтіндердің арасында жиі кездесетін нәрсе - бұл L1 аудару үлгісі, онда хинди тілінің белгісіз артиклдер жүйесі ағылшын тіліне аударылады. (1) -дегі сөйлем (2) -дегідей болады:

2) I want to spend some time in \emptyset village, definitely if I get \emptyset chance.

Хинди артиклдер жүйесін ағылшын тіліне көшіру - бұл айнымалылар процесі, және бұл айнымалылар Шарманың талдауының негізгі көзі (фокус). өз зерттеуінде Шарма хинди тілінен алынған артиклдерді Hindi-derived articles ағылшын тілінде қолданудың әртүрлілігі сөйлеушілердің білім деңгейіне және ағылшын тілін функционалды қолдануға байланысты деп болжам жасайды (гипотезасы) (бұл жерде ағылшын тілінде білімнің төмен деңгейі мен ағылшын тілінде сөйлеу тәжірибесі хинди тілінен алынған артиклдерді қолданудың артуымен байланысты болады). Эксперименттік гипотезасын тексеру үшін (яғни ағылшын тілін оқу және функционалды қолдану хинди тілінен алынған артиклдерді қолдануға әсер етеді) Шарма 12 үнділік ағылшын тілінде сөйлейтіндерден алынған деректерді зерттейді. Ол сөйлеушілерді ағылшын тілін функционалды және білу шеберліктеріне қарай 3 топқа бөледі, онда 1-топ ең төменгі шеберлік және 3-топ ең жоғарғы шеберлігі бар сөйлеушілерден тұрады. Бұл сценарийде тәуелді айнымалы **хинди тілінен шыққан артиклдерді қолдану** болып табылады, және тәуелсіз айнымалы ол **сөйлеушінің ағылшын тілінде білім және функционалды деңгейі** (яғни 1,2,3 топтар) болып табылады. Тәуелді және тәуелсіз айнымалылардың екеуі де категориялық болып табылатындықтан, хи-квадрат тестісі қолайлы.

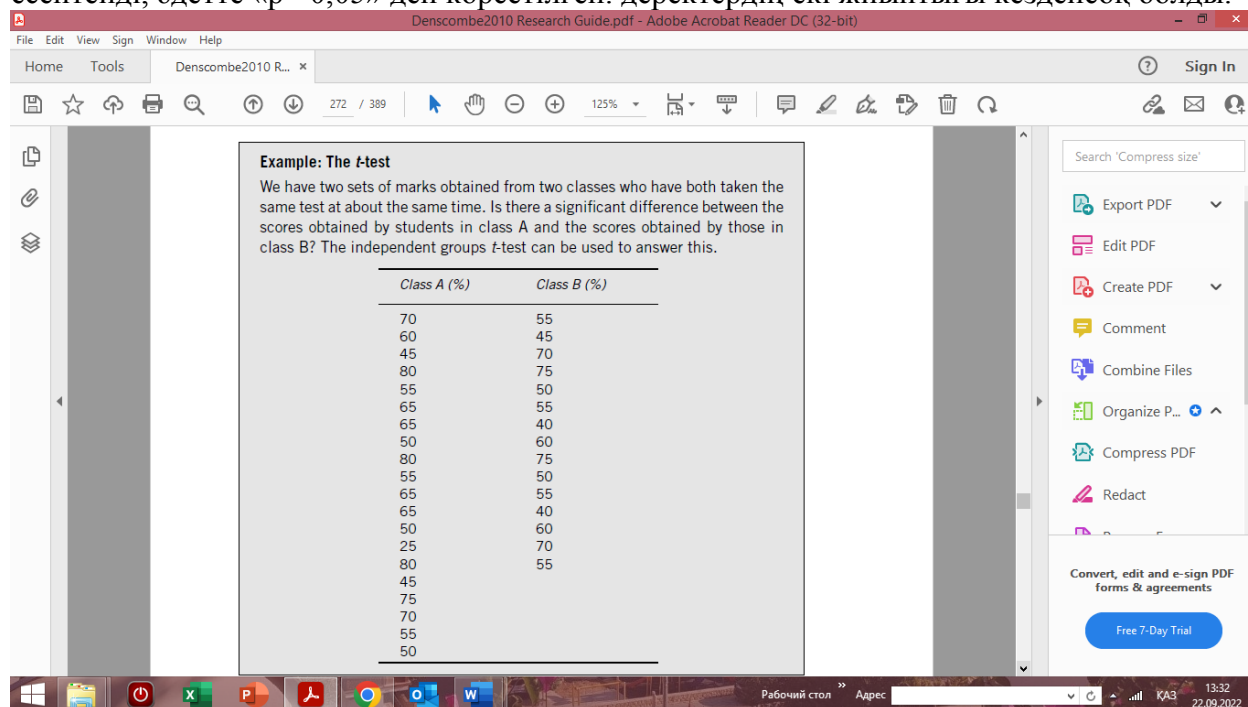
the *term* stands for *convergence*
of findings and *corroboration* of research results.

Кейс

Екі топтың немесе санаттың айтарлықтай деңгейде айырмашылығы бар-жоғын қалай көреміз? Жоба зерттеушілері екі деректер жинағын салыстырып, олардың арасында айтарлықтай айырмашылық бар-жоғын көргісі келеді. Мысалы, мұғалім оқушылардан алынған сынақ ұпайларының екі жинағын салыстырғысы келуі мүмкін: бірі оқу жылының басында алынған жиынтық, екіншісі оқу жылының соңында сол сынақ бойынша сол оқушылардан алынған. Немесе ол бір сыныптағы оқушылардан алған нәтижелерді екіншісінің нәтижелерімен салыстырғысы келуі мүмкін. Шағын ауқымды зерттеулерде осындай екі топты немесе категорияларды осы жолдармен салыстыру қажеттілігі өте жиі кездеседі

Осы мақсат үшін ең қолайлы статистикалық тест t-тест болып табылады. t-тест зерттеушіге маңыздылықтың статистикалық тестін береді. Ол зерттеушіге деректердің екі жинағы арасындағы кез келген айырмашылықтар кездейсоқтыққа байланысты болуының нақты ықтималдығын көрсететін цифрға жету үшін екі деректер жиынтығының құралдарын және олардың стандартты ауытқуларын пайдаланады. t-тестін қолданатын зерттеушілер нөлдік гипотезаны бастапқы нүкте ретінде қабылдайды. Олар статистикалық тест арқылы әйтпесе көндірмейінше, нақты айырмашылық жоқ деп есептейді, бұл олардың ықтималдығы өте күшті екенін айтады.

деректердің екі жинағының арасында табылған кез келген вариация таза кездейсоқтықтан басқа нәрсенің нәтижесі болды. Хи-үлестіру сынағы сияқты, зерттеушілер әдетте 20-дан 1-ден аз ықтималдық бар екенін анықтаған кезде, әдетте айырмашылықтарды «нақты» деп есептейді, әдетте « $p < 0,05$ » деп көрсетілген. деректердің екі жиынтығы кездейсоқ болды.



Example: The t-test

We have two sets of marks obtained from two classes who have both taken the same test at about the same time. Is there a significant difference between the scores obtained by students in class A and the scores obtained by those in class B? The independent groups t-test can be used to answer this.

| Class A (%) | Class B (%) |
|-------------|-------------|
| 70 | 55 |
| 60 | 45 |
| 45 | 70 |
| 80 | 75 |
| 55 | 50 |
| 65 | 55 |
| 65 | 40 |
| 50 | 60 |
| 80 | 75 |
| 55 | 50 |
| 65 | 55 |
| 65 | 40 |
| 50 | 60 |
| 25 | 70 |
| 80 | 55 |
| 45 | |
| 75 | |
| 70 | |
| 55 | |
| 50 | |

t-тестінің екі маңызды күшті жақтары бар, олар кішігірім зерттеулер үшін ерекше пайдалы. t-тест шағын іріктеу өлшемдерімен (30-дан аз) жақсы жұмыс істейді – бұл шағын ауқымды зерттеулердің айқын артықшылығы – және топтар дәл бірдей мөлшерде болуы міндетті емес. Мысалы, 20 элементтен тұратын деректер жиынын 15 деректер жиынымен салыстыруға болады.

Presenting the data – tables and charts

Tables

Bar charts

Histograms

Line graphs

Pie charts