



СТАТИСТИКА НЕГІЗДЕРІ

Өлшемнің үш деңгейі бойынша
жиілік үлестірімін құрастыру:
номиналды, реттік, интервалдық-
пропорционалды



Әрі қарай дәрістерімізді жалғастырамыз. Өткен дәрісімізде біз статистикадағы өлшем деңгейлеріне толық ақпарат берген едік.

Есімізге салып көрейікші қандай өлшем деңгейлері бар деп едік:

1. Номиналды
2. Реттік
3. Интервалдық-пропорционалды өлшем деңгейлері

Және есіңізде болса, статистиканың екі классы бар дедік, дескриптивті және инференциалды, бүгін біз осы бірайнымалы дескриптивті статистиканы түрлі бағытта қолдануын қарастыру және талдау, пайыздық көрсеткіштер мен пропорциялар.

Тақырыпты меңгергеннен кейін:

1. Дескриптивті статистика мәліметтерді ұғыныңқы ететінін түсіндіресіз.
2. Өлшемнің үш деңгейі бойынша айнымалыларға жиілік үлестірімін құрастырасыз және талдау жасайсыз.
3. Пайыздық көрсеткіштер мен пропорцияларды есептейсіз және интерпретациялайсыз.

Білетіміздей, зерттеу нәтижесі әрдайым түсінікті бола бермейді. Зерттеушілер деректерді реттеу және өңдеу үшін статистиканы пайдаланады, сол арқылы мағынасын оқырмандарға түсінікті етеді. Дескриптивті статистиканың мақсаты – зерттеу нәтижесін нақты және ұтымды жеткізу.

Бүгінгі дәрісте зерттеу нәтижесін шығаруға қажетті кең тараған әдістерге тоқталамыз. Оларға кестелер, пайыздық көрсеткіштер, коэффициенттер және графиктер жатады. Аталған бірайнымалы дескриптивті статистика математикалық тұрғыдан күрделі емес (тым қарапайым да емес), бірақ олар нәтиже шығару мен талдау, қорытынды жасау үшін өте пайдалы құрал.

Сандық зерттеу жобасындағы алғашқы қадам – айнымалыларды зерделеу және шамалардың қалай үлестірілгенін көру. Мұның ең тиімді жолы – барлық айнымалы үшін әрбір категория кейсінің саны туралы хабар беретін кестелерді немесе жиілік үлестірімін құрастыру. Бұл кестелер өлшеудің кез-келген деңгейінде айнымалылармен бірге пайдаланылады.

Бірінші Номиналды айнымалыларды қарастырайық, сосын жиілік үлестірімімен қатар, нәтижелердің нақтылығын арттыратын статистиканы талдаймыз.

Номиналды айнымалыларға арналған жиілік үлестірімі

Негізінен, номиналды айнымалылар үшін жиілік үлестірімін құрастыру өте оңай. Ол үшін әр категорияның қайталануын немесе айнымалының орындалу көрсеткішін санап, ал жиіліктерді кесте үлгісінде көрсетеміз. Мысалы, 2.1-кестеде 113 респонденттің гендерлік үлестірімі берілген.

Кестенің атауы және нақты белгіленген категориялары бар, жиілік бағанының төменгі жағындағы кейстердің жалпы саны (N) туралы хабардар етеді. Бұл элементтер барлық жиілік үлестірімінде болуы керек.

2.2-кестеде ҚР-дағы білім алушылар санын көрсететін «стандартты» 4 категория пайдаланылған. 2.3-кестеде «Басқа» білім беру деңгейлерін қосу арқылы толығырақ ақпарат ұсынады. Егер тағы да «Басқа» білім беру деңгейлерін қоссақ немесе білім беру деңгейінің әрқайсысындағы жеке білім беру орталықтары мен университеттерге бөлсек, категориялар саны одан әрі арта түседі.

Сонда қай межеде мағлұмат «жеткілікті» болады? Қандай жағдайда кестеде берілген мәліметтер жүйесіз, тым күрделі және түсініксіз? Бұл сұрақтардың жауабы зерттеу мақсатына сәйкес берілуге тиіс. Егер Қазақстанда кәсіби білім алушылардан гөрі жоғары білім алушылар саны басым екенін көрсеткіңіз келсе, 2.2-кестені таңдаған абзал. Ал егер білім беру деңгейлерінің әралуандығына баса назар аударғыңыз келсе, 2.3-кесте орынды. Бұл жерде қатып қалған ереже жоқ. Сондықтан барлық деңгейдегі айнымалылармен жұмыс жасау барысында детализация (категориялар саны көбірек) мен нақтылықтың (категориялар саны азырақ) біреуін таңдауға болады.

2.1-кесте. Жынысы (ойдан шығарылған мысал)

Жынысы	Жиілік
Ер	53



Әйел	60
	113

2.2-кесте. ҚР-дағы білім алушылар саны, 2018 ж.

Білім беру деңгейі	Саны
Мектепке дейінгі білім беру	862 305
Орта білім беру	3 050 800
Кәсіби білім беру	64 814
Жоғары білім беру	542 458
Басқа	537 917
Барлығы	5 058 294

Дереккөз – ҚР статистика бойынша Комитеті http://stat.gov.kz/faces/wcnav_externalId/homeNumbersEducation?_afzLoop=5688560415843310#%40%3F_afzLoop%3D5688560415843310%26_adf.ctrl-state%3Ddbysid1pf_145

2.3-кесте. ҚР-дағы жоғары білім алушылар саны, 2018 ж.

Білім беру деңгейі	Саны
Мектепке дейінгі білім беру	862 305
Орта білім беру (күндізгі)	3 039 100
Орта білім беру (кешкі)	75 000
Кәсіби білім беру	64 814
Колледж	489 337
Жоғары білім беру	542 458
Жоғары білімнен кейінгі (поствузовское)	48 580
Барлығы	5 121 594

Дереккөз – ҚР статистика бойынша Комитеті http://stat.gov.kz/faces/wcnav_externalId/homeNumbersEducation?_afzLoop=5688560415843310#%40%3F_afzLoop%3D5688560415843310%26_adf.ctrl-state%3Ddbysid1pf_145

Нақтылықты арттыру: пайыздық көрсеткіштер мен пропорциялар

2.2 және 2.3-кестелердің кемшіліктерін байқаған боларсыз: қамтылған жиіліктер тым үлкен, сол себепті оларды түсіну қиын. Мысалы, осы кестелерді жиіліктер тұрғысынан түсіндіріп көрейік: «Қазақстандағы 5 121 594 білім алушылардың 3 039 100-ы орта білім алушылар». Бұл мәлімдемеде ешқандай қателік жоқ, ақиқаты сол. Бірақ осы дерек пайыздық көрсеткішке аударылса анағұрлым анығырақ болар еді: «Америкалықтардың шамамен 59,3%-ы орта білім алушылар» десек.

Пайыздық көрсеткіш – өте пайдалы статистика, өйткені олар өңделмеген жиіліктерді 100-дік негізде стандарттау арқылы санақ жүйесін қалыптастырады. Пайыздық көрсеткіштің математикалық формуласы экранда берілген:

2.1-формула

$$\text{Пайыздық көрсеткіш: } \% = \left(\frac{f}{N}\right) \times 100 \% = \left(\frac{f}{N}\right) \times 100$$

Мұндағы f = жиілік немесе арнайы категория бойынша кейстер саны

N = барлық категория бойынша кейстер саны

Пайыздық есептеулерді көрсету үшін 2.1-кестеде келтірілген деректерді қараңыз. Үлгідегі ер адамдардың пайыздық көрсеткішін қалай табуға болады? Мұндағы ер адамдар саны 53 ($f=53$) және жалпы кейстер саны 113 ($N=113$) екенін есте сақтаңыз. Ендеше, экрандағы өрнекке көңіл бөлейік:

$$\% = \left(\frac{f}{N}\right) \times 100 = \left(\frac{53}{113}\right) \times 100 = (0.4690) \times 100 = 46.9\%$$



Дәл осындай процедураны қолдану арқылы әйелдердің пайыздық үлесін таба аламыз:

$$\% = \left(\frac{f}{N}\right) \times 100 = \left(\frac{60}{113}\right) \times 100 = (0.5310) \times 100 = 53.10\%$$

Өңделмеген жиіліктерге қарағанда пайыздық көрсеткіштерді оқу және ұғыну жеңілдірек. Сонымен бірге пайыздық көрсеткіштер әдетте өлшемдердің барлық деңгейіндегі айнымалылар үшін жиілік үлестірімімен қоса беріледі. Мысалы, 2.4 пен 2.3-кестеде бірдей ақпарат ұсынылған, алайда пайыздық көрсеткіш оны оқуды айтарлықтай жеңілдетіп тұр.

Көлемі әртүрлі топтарды салыстыру үшін пайыздық көрсеткішті пайдаланудың артықшылығы айқын байқалады. Мысалы, 2.5-кестеде келтірілген ақпаратты қараңыз. Қай колледжде салыстырмалы түрде әлеуметтік ғылымдар мамандығының саны көп? Оқуға қабылдау ауқымы әртүрлі болғандықтан, өңделмеген жиіліктерді салыстыру қиын. Үлестірімді 100-дік деңгейде стандарттау арқылы пайыздық көрсеткіштер екі кампустың көлеміне негізделген айырмашылықты жояды. Дәл сондай деректер 2.6-кестеде пайызбен берілген.

2.4-кесте. 2018 жылы Қазақстандағы білім алып жатқан адамдар саны (%)

Білім беру деңгейі	Саны	Пайызбен (%)
Мектепке дейінгі білім беру	862 305	16,8
Орта білім беру (күндізгі)	3 039 100	59,3
Орта білім беру (кешкі)	75 000	1,5
Кәсіби білім беру	64 814	1,3
Колледж	489 337	9,6
Жоғары білім беру	542 458	10,6
Жоғары білімнен кейінгі (поствузовское)	48 580	0,9
Барлығы	5 121 594	100,0

Источник – Комитет по статистике Республики Казахстан http://stat.gov.kz/faces/wcnav_externalId/homeNumbersEducation?_afzLoop=5688560415843310#%40%3F_afzLoop%3D5688560415843310%26_adf.ctrl-state%3Ddbysid1pf_145

2.5-кесте. Екі колледждегі мамандықтар бойынша білім алушылар саны (ойдан шығарылған мысал)

Негізгі мамандық	А колледжі	В колледжі
Бизнес	103	3120
Жаратылыстану ғылымдары	82	2799
Әлеуметтік ғылымдар	137	1884
Гуманитарлық ғылымдар	93	2176
	N= 415	N= 9979

2.6-кесте. Екі колледждегі мамандықтар бойынша білім алушылар саны (ойдан шығарылған мысал)

Негізгі мамандық	А колледжі	В колледжі
Бизнес	24,82%	31,27%
Жаратылыстану ғылымдары	19,76%	28,05%
Әлеуметтік ғылымдар	33,01%	18,88%
Гуманитарлық ғылымдар	22,41%	21,81%
	N = 100,00%	N = 100,01%
	(415)	(9979)

2.6-кесте айырмашылықтар мен ұқсастықтарды анықтауды жеңілдетеді. (әлеуметтік ғылымдар



мамандығы бойынша жалпы А колледжінде білім алушылар саны В колледжіне қарағанда аз болса да), А колледжінде әлеуметтік ғылымдар мамандығы бойынша білім алушылардың пайыздық көрсеткіші В колледжіне қарағанда артығырақ, гуманитарлық мамандықтардың үлесі де осындай.

Қалған екі мамандықтар бойынша айырмашылықтарды қалай сипаттар едіңіз?

Әлеуметтік ғылымдарда пайыздық көрсеткішпен қатар пропорция да пайдаланылады. Пропорциялар 0,00-ден 1,00-ге дейін өзгереді: олар нәтижені 100 пайыздық көрсеткіш арқылы емес, 1,00 негізінде стандарттайды. Пропорция пайыздық көрсеткішпен бірдей, тек пайыздық көрсеткіш сынды 100-ге көбейтілмейді.

2.2-формула

$$\text{Пропорция} = \left(\frac{f}{N}\right) = \left(\frac{f}{N}\right)$$

Пайыздық көрсеткішті 100-ге бөлу арқылы пропорцияға айналдыруға болады, керісінше пропорцияны 100-ге көбейту арқылы пайыздық көрсеткішке конвертациялауға болады. Екі статистика да бірдей ақпараттың эквивалентті көрінісі және оларды өзара алмастыру оңай. Мысалы, 2.1-кестедегі ерлердің салыстырмалы санын пропорция арқылы түрлендіріп көрелік:

$$\text{Пропорция} = \left(\frac{f}{N}\right) = \left(\frac{53}{113}\right) = 0,47 = \left(\frac{f}{N}\right) = \left(\frac{53}{113}\right) = 0,47$$

Осы статистикалық деректердің қайсысын таңдаған жөн?

Көпшілік (статистиктер де) пайыздық көрсеткішті тез түсінеді, ол өте қолайлы. Ал пропорциялар сирек қолданылады, негізінен ықтималдықтармен жұмыс істеген кезде есептеледі. Пайыздық көрсеткіштің артықшылығы берілу жеңілдігіне байланысты: пайыздық көрсеткіш пен пропорциялар нәтиже шығарудың тең дәрежедегі дұрыс жолы.

“2.1-Статистиканы қолдан” қосымшасында көрсетілген нұсқаулармен танысыңыз. (Пайыздық көрсеткіштер мен пропорцияларды есептеу және интерпретациялауды шыңдау үшін оқулықта берілген 2.1 және 2.2-жаттығуларды орындаңыз).

Рет-ретімен орында	Пайыздық көрсеткіш пен пропорцияны табу
Реті	Орындалуы
1.	f (категория бойынша кейстер саны) және N (барлық категория бойынша кейстер саны) мәндерін табыңыз. f арнайы бір категориядағы кейстер саны (мысалы, университеттегі ер балалар) және N барлық категория бойынша кейстер саны (мысалы, университеттегі барлық білім алушылар, ер балалар мен қыз балалар) екенін, категория мен тұтас топ бірдей болған жағдайдан басқа кезде (мысалы, барлық студенттер ер балалар болғанда), f N-ге қарағанда кіші болатынын есте сақтаңыз. Пропорция үлесі 1,00-ден, ал пайыздық көрсеткіш 100%-дан аспауы керек.
2.	Пропорцияны табу үшін f-ті N-ге бөліңіз.
3.	Пайыздық көрсеткішті табу үшін 2-қадамда шыққан мәнді 100-ге көбейтіңіз.

Статистиканы қолдан.

Пайыздық көрсеткіштер мен пропорцияларды пайдалану

Бұл статистика салыстырмалы түрде қарапайым болғанымен, іс жүзінде бірнеше нұсқаулықты ескерген абзал.

1. Кейстер саны (N) аз болған жағдайда (айталық, 20-дан кем болса) пайыздық көрсеткіштер мен пропорцияларды емес, нақты жиіліктерді пайдаланған дұрыс. Кейстер саны аз болғанда, пайыздық көрсеткіштер тұрақсыз және болмашы ауытқулар нәтижені түбегейлі өзгертіп жіберуі мүмкін. Мысалы,



бастапқыда 10 ер адам және 10 әйел адамды алдыңыз (яғни әр жыныстың 50%-ы), содан кейін тағы бір әйел адамды қоссаңыз, пайыздық үлес 52,4% әйел және 47,6% ер адамға өзгереді. Әрине, бақылау саны көбейген сайын, әрбір қосымша кейстің әсері төмен болады. Егер 500 ер адам мен 500 әйел адамды алған болсақ, оған тағы бір әйел адамды қосқанда, әйелдердің пайыздық көрсеткіші оннан біріне азаймақ (мысалы, 50%-дан 50,05%-ға дейін).

2. Әрдайым бақылау санын пропорциямен және пайыздық көрсеткішпен бірге баяндап отырыңыз. Бұл оқырманға іріктеме көлемінің сәйкестігін бағалауға мүмкіндік береді, керісінше, зерттеушінің статистика арқылы ақпаратты бұрмалауына жол бермейді. «Үш адамның екеуі басқа курстарға қарағанда статистика оқуды қалайды» деген мәлімдеме әсерлі болуы мүмкін, бірақ сіз тек үш адам сұралғанын білсеңіз, тұжырым өзіндік мәнін жоғалтады. Кейстер саны көрсетілмеген есептерге аса сақтықпен қараған жөн.

3. Пайыздық көрсеткіштер мен пропорциялар бөлуді талап етсе де, номиналды айнымалы үшін есептелуі мүмкін. Бұл өлшеу жөніндегі нұсқаулықты бұзу емес (1.5-кестені қараңыз). Пайыздық көрсеткіштер мен пропорциялар айнымалының шамаларын бөлуді талап етпейді (мысалы, тест барысында орташа шаманы есептеген жағдайда). Дұрысы – нақты категория бойынша (f) кейстер саны іріктемедегі кейстердің жалпы санына (N) бөлінеді. Айталық «іріктеменің 53,10%-ы әйел адамдар» дегеніміз «113 респонденттің 60-ы әйелдер» дегенге барабар, бұл әлдеқайда ыңғайлы әрі түсінікті.

Сөзіміз сенімді болу үшін мынадай мысал кетіріп көрелік

Ірі әлеуметтік қызмет көрсету Орталығындағы бір бөлім жетекшісі атқарушы директорға кіріп, мынадай кадр мәселесін көтерді.

Атқарушы директор: Тығыз шаруаларым боп тұр, айып етпеңіз, бірнеше минутым ғана бар. Кадр мәселесін мейлінше қысқа баяндаңыз.

Бөлім жетекшісі: Жұмыс жүктемесін орындау үшін адам саны жеткіліксіз. Менің бөлімімде Орталықта жұмыс істейтін 177 қызметкердің 50-і ғана бар. Бірақ өткен жылы агенттік атқарған 16 722 істің 6231-імен менің бөлімім айналысты.

Атқарушы директор (есінен тұрып): Қызық екен. Бұл мәселе бойынша міндетті түрде сізбен кейінірек хабарласамы дейді.

Қалай ойлайсыз Бөлім жетекшісі өз мәселесін қалай ұсынуы керек еді? Қайткенде оның мәлімдемесі тиімдірек және естуге ыңғайлы болар еді?

Белгілі болғандай Бөлім жетекшісі екі түрлі сандар жиынтығын салыстырғысы келді (атап айтқанда, біріншісі - өз бөлімінің қызметкерлерін жалпы агенттік қызметкерлерімен салыстыру, екіншісі - өз бөліміндегі жұмыс жүктемесін жалпы агенттіктегі жұмыс жүктемесімен салыстыру арқылы), бірақ нәтижелерді пропорция мен пайыздық көрсеткіш арқылы жеткізу өтімдірек болар еді. Егер Бөлім жетекшісі «Жалпы штаттың 28,25%-ы ғана менің бөлімімде тіркелген, ал біз жалпы Орталық жүктемесінің 37,26%-ын орындаймыз» деп айтқанда ше?

Иә Бұл хабарлама анығырақ болар ма еді?

Пайыздық көрсеткіштердің есептеу реті экранда ұсынылады:

Бірінші пайыздық көрсеткіш былай есептеледі:

$$\% = (f/N) \times 100 = (50/177) \times 100 = (0,2825) \times 100 = 28,25\%$$

ал екінші пайыздық көрсеткіш былай есептеледі:

$$\% = (f/N) \times 100 = (6231/16,722) \times 100 = (0,3726) \times 100 = 37,26\%$$

Енді Реттік айнымалылар үшін жиілік үлестірімін көрелік

Реттік айнымалылар үшін жиілік үлестірімі номиналды айнымалылармен бірдей құрастырылады. 2.7-кестеде университет жатақханасының сапасы бойынша жүргізілген сауалнамада студентердің жауаптарының жиілік үлестірімі бойынша ақпарат берілген. Нақтылықты арттыру мақсатында кестеге пайыздық баған қосылды.

2.7-кесте. Университет кампусындағы туу көрсеткішін бақылауды қолдау (ойдан шығарылған мысал)



Университет жатақханаларындағы студенттер өмірі мен олардың қауіпсіздігі Университет Жастары Кеңесінің бақылауында болуы керек деген тұжырымға қатысты пікіріңіз: (1) толық келісесіз; (2) келісесіз; (3) келіспейсіз немесе; (4) мүлде келіспейсіз.		
Жауап	Жиілік	Пайыздық көрсеткіш
Толық келісемін	350	25,55%
Келісемін	462	33,72%
Келіспеймін	348	25,40%
Мүлде келіспеймін	210	15,33%
	1370	100,00%

2.8-кесте. Біріктірілген категориялар бойынша университет кампусындағы туу көрсеткішін бақылауды қолдау (ойдан шығарылған мысал)

Университет жатақханаларындағы студенттер өмірі мен олардың қауіпсіздігі Университет Жастары Кеңесінің бақылауында болуы керек деген тұжырымға қатысты пікіріңіз: (1) толық келісесіз; (2) келісесіз; (3) келіспейсіз немесе; (4) мүлде келіспейсіз.		
Жауап	Жиілік	Пайыздық көрсеткіш
Толық келісемін немесе келісемін	812	59,27%
Келіспеймін немесе мүлде келіспеймін	558	40,73%
	1370	100,00%

Кестеден пікірлердің тең дәрежеде үлестірілгені байқалады. Ең көп кездесетін жауап – «келісемін» (33,72%) және студенттердің басым бөлігі (59,27%) студенттер өмірі мен олардың қауіпсіздігі Университет Жастары Кеңесінің бақылауында болуы керек деген тұжырыммен келіседі немесе толығымен келіседі. Егер зерттеуші осы үлгіні көрсеткісі келсе немесе кестені ықшамдауға тырысса, категорияларды 2.8-кестедегідей құрауға болады. Дегенмен, нәтижелерді қарапайым, ықшам берудің де өз олқылықтары бар: келісу мен келіспеудің түрлі деңгейлері арасындағы айырмашылық жоғалады. 2.2 және 2.3-кестелердегі білім беру деңгейлерін талқылағанда зерттеуші толығырақ (көп категориялар беру) және айқынырақ (аз категориялар беру) арасындағы балансты сақтауы керек. (Номиналды және реттік айнымалылар үшін жиілік үлестірімдерін құру және интерпретациялауды шыңдау үшін 2.5 және 2.12-жаттығуларды орындаңыз).

Үшінші қарастыратын үлестірімдер - Интервалдық-пропорционалды айнымалылар үшін жиілік үлестірімдері

Жалпы интервалдық-пропорционалды айнымалылар үшін жиілік үлестірімдерін құру номиналды және реттік айнымалылардан гөрі күрделірек. Интервалдық-пропорционалды айнымалылардың шамалары сан алуан болады, демек зерттеуші түсінікті, ықшам кестелерді шығару үшін категорияларды біріктіруі немесе топтастыруға тиіс. Ендеше зерттеуші тағы да мәліметтің толықтығы мен айқындығы арасында таңдау жасауға мәжбүр болады.

Мәселен, сіз бір қаладағы «жас мөлшері» айнымалысының қалай үлестірілгенін сипаттағыңыз келеді дерлік. Көбінде жас диапазоны өте кең болып келеді – жаңа туған нәрестеден бастап 90 жастағы немесе одан асқан адамдар. Егер әр жыл сайын орын алатын жасты (немесе шама) көрсетсеңіз, сізде 80, 90 немесе одан да көп категориялар шығуы мүмкін; мұндай кестені оқу қиынға соғады. Шамалар (жылдар) түсінікті болуы үшін олар үлкен категорияларға топтастырылуы керек.

Бұл категориялар қаншалықты ауқымды болуы керек?

Кестеге қанша категория кірістірілуге тиіс?

Көбірек (саны жағынан көп шағын категориялар) немесе нақтырақ ақпаратпен (саны жағынан аз үлкен категориялардың) қамтамасыз етуіміз керек пе?



Ал ендеше Интервалдық-пропорционалды айнымалылар үшін жиілік үлестірімдерін құрастырып көрейік

Түсіндіруді жеңілдету үшін 20 студенті бар шағын колледж тобының жиілік үлестірімін мысалға аламыз. Колледж студенттерінің жас диапазоны шамалас болғандықтан, бір жылдық категорияларды пайдалана аламыз (интервалды деректермен жұмыс барысында бұл категориялар көбіне кластық интервалдар деп аталады).

2.9-кесте. Колледж дәрісханасындағы студенттер саны (ойдан шығарылған мысал)

1 жылдық жасқа негізделген интервал	
Жасы	Жиілік
18	5
19	6
20	3
21	2
22	1
23	1
24	1
25	0
26	1
N=20	

Жиілік үлестірімі жас мөлшерін ретке тізу, орын алған әрбір шаманың (жас мөлшерінің жылы) қайталану санын есептеу және әрбір категория бойынша кейс сандарының қосындысын шығару арқылы құрастырылады.

2.9-кесте 18 бен 19 кластық интервалдардағы кейстер туралы ақпарат береді және олардың шоғырлануын анықтайды.

Бұл кесте жалпылама түсінікті болғанымен, одан да ықшамдалған (егжей-тегжейлі емес) нәтиже қажет деп көрелік. Оған қол жеткізу үшін топтық шамаларды ауқымдырақ кластық интервалдарға топтастыру керек. Интервалдарды ірілендіру (айтарлық, 2 жылды топтастыру) арқылы интервалдардың санын азайтасыз және ықшам нәтиже көрсетесіз.

2.10-кестедегі шамаларды топтастыру жасырақ респонденттердің басымдығын айқын көрсетеді. Пайыздық көрсеткіш бағанын қоссақ, бұл тренд одан әрі нығая түседі.

Бұл жердегі Белгіленген шектерге тоқтылып өтелік

2.10-кестеде берілген кластық интервалдар арасында айқын алшақтық бар екеніне назар аударыңыз. Яғни белгіленген кластық шектер бір бірліктік дистанция арқылы ажыратылған. Бір қарағанда, бұл алшақтықтар, есіңізде болса, өткен дәрістерде келтірілген (1.2-кестені қараңыз) толық қамту принципі бұзатын сияқты. Бірақ жас мөлшері бүтін сандармен өлшенгендіктен, бұл «алшақтық» қиындық тудырмайды. Өлшемнің дәлдік деңгейін ескерсек (айталық, жылдың оннан бір бөлшегінің орнына бір (тұтас/бүтін) жыл қарастырылып отыр) ешбір кейс шамасы берілген интервалдар арасынан тыс қалмайды.

2.10-кесте. Колледж дәрісханасындағы студенттердің жасы (ойдан шығарылған мысал)

2 жастан тұратын интервал ауқымы		
Жасы	Жиілік	Пайыздық көрсеткіш
18–19	11	55,0%
20–21	5	25,0%
22–23	2	10,0%



24–25	1	5,0%
26–27	1	5,0%
N=20		100,0%

2.10-кестедегі кластық интервалдар толық қамту және өзара жоққа шығарушы қасиетке ие, сәйкесінше 20 респонденттің қай-қайсысы да жас категориясына сәйкес іріктеледі.

Дегенмен мұнда жас мөлшері аса нақтылықпен өлшенсе, қиындықтар болатынын ескерген жөн. Егер жас мөлшері жылдың ондық бөлшегі арқылы өлшенетін болса, 19,5 жастағы адам 2.10-кестедегі қай кластық интервалға жатады? Мұндай екіұштылықтың алдын алу үшін әрдайым деректердің дәлдік деңгейімен бірдей кластық интервалдардың нақты шегін белгілейміз. Олай болса жас мөлшері жылдың ондық бөлшегі арқылы өлшенсе, 2.10-кестедегі кластық интервалдардың шегі де жылдың ондық бөлшегімен беріледі.

Мысалы:

17,0 – 18,9 жас

19,0 – 20,9 жас

21,0 – 22,9 жас

23,0 – 24,9 жас

25,0 – 26,9 жас

Категориялардың өзара дербестігін қамтамасыз ету үшін кластық интервалдар бір-бірімен қиылыспауы тиіс. Егер кластық интервалдардың шекараларын деректердің дәлдік деңгейімен бірдей (бүтін сандармен, ондықтармен, жүздіктермен берілуі мүмкін) белгісеңіз және интервалдар арасындағы «алшақтықты» сақтасаңыз, онда сіз құрастыратын жиілік үлестірімінде әрбір кейс нақты бір категорияға ғана тиесілі болады.

Кластық интервалдар туралы айтқанда Ортаңғы нүктелер мәселесімен кезігеміз

Кей жағдайда нақты бір графиктерді құрастыру барысында, кластық интервалдардың ортаңғы нүктелерімен жұмыс істеу қажет болады. Ортаңғы нүктелер – кластық интервалдың жоғарғы және төменгі шектерінің дәл ортасы. Оны табу үшін жоғарғы және төменгі шектердің қосындысын 2-ге бөлу керек. Көрнекілік ретінде 2.1-кестеде кластық интервалдардың екі түрлі жиынтығының ортаңғы нүктелері көрсетілген. Толық нұсқаулық алу үшін «Рет-ретімен орында» қосымшасын қараңыз. (Ортаңғы нүктелерді табу үшін 2.8b және 2.9b-жаттығуларын қараңыз).

2.11-кесте. Ортаңғы нүктелерді анықтау

3 бірлікті кластық интервал	
Кластық интервал	Ортаңғы нүкте
0–2	1,0
3–5	4,0
6–8	7,0
9–11	10,0

6 бірлікті кластық интервал	
Кластық интервал	Ортаңғы нүкте
100–105	102,5
106–111	108,5
112–117	114,5
118–123	120,5

Рет-ретімен орында	Ортаңғы нүктелерді анықтау
--------------------	----------------------------

**Реті Орындалуы**

1. Жиілік үлестіріміндегі ең төменгі интервалдың жоғарғы және төменгі шектерін табыңыз. 2.11-кестедегі интервалдардың үстіңгі жинағы үшін ең төменгі интервал (0–2) 0, 1 және 2 деген шамалардан тұрады. Осы интервалдағы жоғарғы шек – 2, ал төменгі интервал – 0.
2. Жоғарғы және төменгі шектердің қосындысын 2-ге бөліңіз. 0–2 интервалы үшін: $(0+2)/2 = 1$. Бұл интервалдың ортаңғы нүктесі – 1.
3. Басқа интервалдардың ортаңғы нүктелері әрбір интервал үшін 1 және 2-қадамды қайталау арқылы анықталады. Балама ретінде кез-келген интервалдың ортаңғы нүктесін табу үшін келесі төменгі интервалдың ортаңғы нүктесіне интервал қашықтығының мәнін қосыңыз. Мысалы, 2.11-кестенің үстіңгі жолындағы төменгі интервал – 0–2, ал ортаңғы нүктесі – 1. Интервал қашықтығы – 3 бірлік (яғни олардың әрқайсысы 3 шамадан тұрады), ендеше келесі жоғарғы интервал (3–5) үшін ортаңғы нүкте $1 + 3$, яғни 4 болады. 6–8 интервалы үшін ортаңғы нүкте $4 + 3$, яғни 7 және тағы сол сияқты.

Кумулятивті жиілік және кумулятивті пайыздық көрсеткіш

Интервалды-пропорционалды деректердің бастапқы жиілік үлестіріміне жиі қолданылатын екі қосымша – кумулятивті жиілік және кумулятивті пайыздық көрсеткіш бағандары. Бұл бағандар үлестірімде берілген шамаларға немесе кластық интервалдарға қанша кейстен келетінін бір қарағаннан білуге мүмкіндік береді. Осы бағандарды қосу үшін «Рет-ретімен орында» қосымшасын қараңыз. 2.10-кестеге қосылған кумулятивті жиілік және кумулятивті пайыздық бағандар 2.12-кестеде көрсетілген.

Кумулятивті бағандар зерттеушінің кейстердің шамалар диапазонында қаншалықты таралғандығы туралы тұжырым жасауы қажет болған жағдайларда пайдалы. Мысалы, 2.12-кесте сыныптағы студенттердің басым бөлігі 21 жастан кіші екенін анық көрсетеді. Егер зерттеуші осы фактіні атап көрсеткісі келсе, кумулятивті бағандарды қолданған дұрыс.

Шын мәнінде, зерттеулерде 20-дан астам кейс пен кестеде көрсетілген мысалдан әлдеқайда көп категория болады, сондықтан кумулятивті жиіліктерге қарағанда кумулятивті пайыздық көрсеткіштер тиімдірек.

2.12-кесте. Колледж дәрісханасындағы студенттер жасы

Жасы	Жиілік	Кумулятивті жиілік	Пайыздық көрсеткіш	Кумулятивті пайыз
18–19	11	11	55,0%	55,0%
20–21	5	16	25,0%	80,0%
22–23	2	18	10,0%	90,0%
24–25	1	19	5,0%	95,0%
26–27	1	20	5,0%	100,0%
N=	20		100,0%	

Рет-ретімен орында**Жиілік үлестіріміне кумулятивті жиілік және пайыздық көрсеткіш бағандарын қосу****Реті Орындалуы****Кумулятивті жиілік бағаны**

1. Ең төменгі кластық интервалдан бастаңыз (ең төмен шамалардан тұратын интервал). Кумулятивті жиілік бағанының басындағы шама осы интервал аралығындағы кейстер санымен бірдей.
2. Келесі кластық интервалға өтіңіз. Бұл интервал үшін кумулятивті жиілік осы интервалдағы кейстер саны мен оның алдындағы (кіші) интервалдағы кейстер санының қосындысына тең.
3. Кумулятивті жиілік N-ге тең, ең үлкен шамалары бар интервалға жеткенше, бір интервалдан келесі интервалға дейінгі аралықтағы кейстер санын қосуды (немесе жинақтауды) жалғастырыңыз.



Реті Орындалуы

Кумулятивті пайыздық көрсеткіш бағаны

1. Әрбір категория бойынша кейстердің пайыздық көрсеткішін есептеңіз және кумулятивті жиілік үшін берілген қадамдарды қайталаңыз. Ең төменгі кластық интервал осы интервалдағы кейстердің пайыздық көрсеткішімен бірдей басталады.
2. Келесі жоғарғы интервал үшін кумулятивті пайыздық көрсеткіш осы интервалдағы кейстердің пайыздық көрсеткіші мен төменгі (кіші) интервалдағы кейстердің пайыздық көрсеткішінің қосындысына тең.
3. Кумулятивті пайыздық көрсеткіш 100%-ға тең, ең үлкен шамалары бар интервалға жеткенше, бір интервалдан келесі интервалға дейінгі аралықтағы кейстердің пайыздық көрсеткішін қосуды (немесе жинақтауды) жалғастырыңыз.

Тең емес кластық интервал

Жалпы ереже ретінде жиілік үлестірімінің кластық интервалдары нақты және ұғынуға жеңіл болуы үшін көлемі жағынан тең болуға тиіс. Мысалы, 2.12-кестеде берілген кластық интервалдардың арасы 2 жылға тең. Кластық интервалдарды тұжырымдаудың түрлі жолдары бар, әр жағдайды жеке-жеке қарастырайық.

Ашық интервалдар. 2.12-кестеге 47 жастағы бір студентті қоссақ, үлестірім жиілігі қандай болар еді? Енді 21 кейс болды, ең үлкен респондент (енді 47 жасар) пен екінші үлкен респондент (26 жасар) арасында елеулі алшақтық бар. Егер сол жасы үлкен студентті 2.12-кестеге сай қосатын болсақ, 46-47 интервалына жеткенше нөлдік кейсті тағы да тоғыз кластық интервалды (28-29, 30-31, 32-33 және тағы сол сияқты) енгізуіміз керек болар еді. Бұл артық орын алады, түсініксіз әрі шатастыруы мүмкін. әл жоғары (немесе төмен) шамаларды өңдеудің балама жолы – үлестірім жиілігіне ашық интервал қосу (2.13-кестедегідей).

2.13-кестедегі ашық интервал ақпаратты ықшамдап ұсынуға мүмкіндік береді. Ең төмен кластық интервал (мысалы, 17 жаста және одан кіші) ретінде ашық интервалды қосу арқылы өте төмен шамаларды өңдеуге болады. Әрине, мұндай тиімділіктің де кемшілігі бар – кесте ашық интервалдағы нақты шамалар туралы ақпаратты елемейді, сондықтан бұл техника кездейсоқ қолданылмауға тиіс.

Тең емес көлемдегі интервалдар. Кейбір айнымалылар бойынша шамалардың басым бөлігі тығыз топтастырылып, ал басқалары шамалардың кең диапазоны бойынша таралып отыр. АҚШ-тағы табыс үлестірімін мысалға алайық.

2.13-кесте. Колледж дәрісханасындағы студенттердің жасы (N = 21)

Жасы	Жиілік	Кумулятивті жиілік
18–19	11	11
20–21	5	16
22–23	2	18
24–25	1	19
26–27	1	20
28 және одан үлкен	1	21
N=	21	

2012 жылы отбасылардың көбінің (шамамен 54%) жылдық кірісі 25 000\$ - 100 000\$ аралығында деп көрсеткен, ал көпшіліктің (24,4%) кірісі одан төмен. Кірісі 100 000 доллардан жоғары ауқатты отбасыларға қатысты мәселе туындайды (статистикалық тұрғыдан алғанда). Бұл категориялардағы отбасылар аз (21,6%), дегенмен оларды есепке алу міндетті.

Егер 2.14-кестеде 10 000 долларға тең интервалды пайдаланған болсақ, барлық ауқатты отбасыны қамту үшін 30-40 не одан да көп интервал қажет. Ал жоғары кіріс диапазонындағы интервалдардың біразында, әсіресе кірісі 150 000 доллардан асатын отбасыларда кейстер саны аз не жоқ болар еді. Соның салдарынан кесте өте үлкен және түсінуге күрделі болады.

Мұндай жағдайда айнымалыларды тиімдірек қорытындылау үшін 2.14-кестедегідей тең емес интервалдарды пайдалана аламыз. 2.14-кестеде кейбір интервалдардың қашықтығы - 10 000\$,



Кітап: Статистика негіздері

Дәріс: Өлшемнің үш деңгейі бойынша жиілік үлестірімін құрастыру: номиналды, реттік, интервалдық-пропорционалды

басқаларының қашықтығы - 25 000\$ не 50 000\$ және екі ашық интервал (ең төменгі және ең жоғарғы интервалдар) бар. Аралас қашықтықты интервалдарды пайдалану оқырманды сәл шатастыруы мүмкін, дегенмен ықшамдылық пен тиімділіктің өзара үйлесімі маңызды. (Интервалдық-пропорционалды айнымалылар бойынша жиілік үлестірімдерін құрастыру мен интерпретациялауды шындау үшін 2.5-2.9 және 2.12-жаттығуларды қараңыз).

2.14-кесте. АҚШ-тағы отбасылардың табыс үлестірімі, 2012 жыл

Кірісі	Отбасылардың пайыздық көрсеткіші	Кумулятивті пайыздық көрсеткіш
10 000 доллардан төмен	7,7%	7,7%
10 000 доллардан 14 000 долларға дейін	5,6%	13,3%
15 000 доллардан 24 000 долларға дейін	11,1%	24,4%
25 000 доллардан 34 000 долларға дейін	10,4%	34,8%
35 000 доллардан 49 000 долларға дейін	13,8%	48,6%
50 000 доллардан 74 000 долларға дейін	18,0%	66,6%
75 000 доллардан 99 000 долларға дейін	11,9%	78,5%
100 000 доллардан 149 000 долларға дейін	12,4%	90,9%
150 000 доллардан 199 000 долларға дейін	4,6%	95,5%
200 000 доллардан жоғары	4,6%	100,1%
	100,1%	
	N = 115 969 540	

Дереккөз:

U.S. Bureau of the Census. 2014. American Community Survey, 2012. Retrieved from <http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS12YRDP03&prodType=tabl>

Қорытынды: Интервалдық-пропорционалды айнымалылардың жиілік үлестірімдері

Интервалдық-пропорционалды айнымалылар үшін жиілік үлестірімдерін құрастырудың қатып қалған ережелері жоқ. «Егер алшақтығы» бар кластық интервалды белгілесеніз және әрбір шама үшін интервалыңыз болса, онда толықтық (категориялар саны көбірек) пен нақтылық (категориялар саны азырақ) арасындағы ең тиімді балансты сақтау туралы шешім қабылдай аласыз. «Рет-ретімен орында» қосымшасынан жиілік үлестірімін құрастыруға пайдалы жалпы нұсқауларды алуға болады.

Сонымен қатар кестелерді құрастыру үшін компьютер және SPSS сияқты бағдарламалар болуы керек. «SPSS бағдарламасын пайдалану» секциясы жиілік үлестірімін құрастыруға көмектесетін бағдарламаны қолдану нұсқаулықтарын ұсынады.

Р е т - р е т і м е н о р ы н д а	Интервалдық-пропорционалды айнымалылар үшін жиілік үлестірімін құрастыру
Реті	Орындалуы
1	Қанша кластық интервалды (k) пайдаланғыңыз келетінін анықтап алыңыз. Шамамен 10 (k = 10) интервалы өте ыңғайлы. Көптеген зерттеу талабына сай интервалдар саны 10-нан кем болады, сонымен қатар 15 немесе 20 интервалды жиілік үлестірімін де кездестіруге болады. 20-дан артық интервалдар өте сирек пайдаланылады, өйткені жиілік үлестірімінің нәтижесін түсіну қиынға соғады.
2	Үлкен шамадан кіші шаманы азайту арқылы шамалардың диапазонын (R) табыңыз.



3	(Екінші қадамдағы) R-ді (бірінші қадамдағы) k-ға бөлу арқылы кластық интервалдардың (i) көлемін (size) анықтаңыз: $i = R/k$ i-дің мәнін ыңғайлы бүтін санға дөңгелектеп алыңыз. Бұл – интервалдың көлемі немесе қашықтығы.
4	Кіші шекарасы ең төмен шамаға тең не төмен болатындай ең кіші интервалды белгілеңіз. Ең үлкен интервал – ең үлкен шамаға тең немесе одан жоғары. Жалпы интервалдар көлемі жағынан тең болуы керек, дегенмен ыңғайлы болған жағдайда тең емес немесе ашық интервалдардың да қолданылуы мүмкін.
5	Мәліметтерді өлшеуде пайдаланған дәлдік деңгейінде кластық интервалдардың шегін белгілеңіз. Интервалдарды бір-біріне қабағтастырмаңыз. Сәйкесінше әрбір кейс тек нақты бір категорияға ғана тиесілі болатындай кластық интервалдарды іріктеп алыңыз.
6	Әрбір кластық интервалдағы кейстерді санаңыз, әрі қарай аралық нәтижелерді «жиілік» деген бағанда көрсетіп отырыңыз. Осы бағанның төбесінде кейстердің жалпы санын (N) көрсетіңіз. Сонымен қатар кесте пайыздық көрсеткіш, кумулятивті жиіліктер және кумулятивті пайыздық көрсеткішке арналған бағандардан тұруы мүмкін.
7	Жиілік үлестірімін мұқият тексеріп шығыңыз. Ақпараттың басым бөлігі анық емес пе? Анық болмаса, кестені кластық интервалдардың санын көбейту арқылы қайта құраңыз (немесе шағын көлемді интервалдарды пайдаланыңыз). Кестеде ақпарат тым көп пе? Онда кестедегі интервалдар санын азайтыңыз (немесе ірі көлемді интервалдарды пайдаланыңыз). Кейстері жоқ интервалдар саны көп пе? Көп болса, ашық немесе тең емес интервалдарды қарастырыңыз. Жиілік үлестірімі сіздің өз еркіңізбен қабылдаған шешімдеріңіздің нәтижесі екенін есте сақтаңыз. Зерттеу мақсатын ескере отырып, кесте тиімді болмаса, толық және ықшам болғанына көз жеткізгенше, кестені қайта жасай беріңіз.
8	Кестеге нақты, қысқаша атау беріңіз. Сондай-ақ есебіңіз бірнеше кестеден тұрса, кестелерге нөмір қою керек. Барлық категория мен баған анық жазылуға тиіс.

Сөзіміздің ТҮЙІНІ

1. Біз жалғыз айнымалының үлестірімін сипаттаудың әртүрлі тәсілдерін, жалпылай алғанда, зерттеудің нәтижелері туралы есеп беруді қарастырдық. Негізгі екпін – нәтижелерді анық және қысқа беруге бағытталды.
2. Жиілік үлестірімдері – кейбір айнымалылардың тұтас үлестірімін сипаттайтын кестелер. Статистикалық талдау әрбір айнымалы үшін осындай кестелерді құрастыру мен шолу жасаудан басталады. Пайыздық көрсеткіш бағандары жиілік үлестірімдерінің оқылуын жеңілдетеді.
3. Пайыздық көрсеткіштер мен пропорциялар – салыстырмалы жиілік тұрғысынан нәтиже шығарудың түрлі тәсілдері. Тұтас үлестірімге қарағанда пайыздық көрсеткіштер мен пропорциялар айнымалы категорияларының салыстырмалы орнын білдіреді.