СӘТБАЕВ УНИВЕРСИТЕТІ

SATBAYEV UNIVERSITY

Статистика және деректерді талдау

Ержан Асел Ануаркызы,Ph.D., «Ақпараттық жүйелер» кафедрасының қауымдастырылған профессоры

Assel.yerzhan@satbayev.university

Іріктеме және статистикалық бағалау

Деректен шешімге дейін — статистикалық зерттеудің теориялық негіздері, іріктеу әдістері және бағалау логикасы



Статистикалық зерттеудің басталуы: мақсат пен нысан

Кез келген статистикалық зерттеу нақты мақсатты анықтаудан басталады. Мақсат айқын болмаса, жиналған деректер өз мәнін жоғалтады және шешім қабылдауға негіз болмайды. Зерттеудің бастапқы кезеңінде екі іргелі сұрақ қойылады.

Зерттеу нысанын анықтау

Зерттеу нысаны — мәлімет алынатын жиынтық. Бұл жиынтық нақты шектелуі тиіс: мысалы, «Алматы қаласының 2024 жылғы 18–60 жас аралығындағы жұмыс істейтін тұрғындары». Нысанды дұрыс анықтау іріктеме шеңберін қалыптастырудың алғышарты болып табылады. Жиынтық неғұрлым нақты анықталса, соғұрлым дерек сапасы жоғары болады.

Зерттеу бірлігін белгілеу

Зерттеу бірлігі — деректің тікелей алынатын элементі. Бірлік жеке тұлға, үй шаруашылығы, кәсіпорын немесе аймақ болуы мүмкін. Бірлікті дұрыс таңдау зерттеудің логикасын айқындайды. Мысалы, халықтың кірісін зерттегенде бірлік ретінде үй шаруашылығын немесе жеке адамды таңдауға болады — бұл таңдау нәтижеге айтарлықтай әсер етеді.

Нысан мен бірлікті айқындаудан кейін зерттеушілер бақылау бағдарламасын, деректер жинау тәсілін және талдау логикасын жоспарлауға кіріседі. Осы кезеңдегі қателіктер кейінгі барлық нәтижеге өз ізін қалдырады.

Жоспар және бағдарлама: өлшемнің «ережесін» бекіту

Статистикалық зерттеудің жоспары мен бағдарламасы — зерттеушінің «конституциясы». Бұл құжат барлық кейінгі қадамдарды реттейді және деректер жинаудың стандартын белгілейді.

Бақылау бірлігі мен жинау материалдары

Бақылау бірлігі анықталғаннан кейін деректерді жинаудың нақты құралдары таңдалады. Анкеталар — сауалнама зерттеулерінің негізгі құралы. Олар стандартталған болуы тиіс: сұрақтардың реті, тұжырымы және жауап нұсқалары алдын ала бекітіледі. Ресми статистикада бастапқы есеп құжаттары (форматтар) немесе әкімшілік деректер базалары қолданылады. Жинау материалдарының сапасы деректердің сапасын тікелей анықтайды.

Талдау бағдарламасын алдын ала құру

Деректерді жинаудан бұрын талдау бағдарламасын жасау — озық тәжірибе. Бұл тәсіл «жинап алайық, сонан соң не істерімізді ойлайық» деген жол жоқтығын білдіреді. Алдын ала жасалған талдау жоспары мыналарды қамтуы тиіс: қандай өлшемдер есептелетіні, қандай кесте форматтары пайдаланылатыны, қандай статистикалық тестілер жүргізілетіні. Бұл тәсіл зерттеушіні нәтижеге байланысты «кейіннен деректерді бейімдеу» (p-hacking) тәуекелінен қорғайды.

Жоспар мен бағдарлама зерттеу барысында өзгерісе ұшырауы мүмкін, бірақ кез келген өзгеріс негізделіп, тіркелуі тиіс.

Уақыт пен орын: бақылауды дұрыс «қою»

Зерттеу орнын нақтылау

Зерттеу аймағын немесе мекемесін нақты анықтау деректердің географиялық өкілдігін қамтамасыз етеді. Аймақты таңдауда негізгі критерийлер: зерттеу мақсатына сәйкестігі, логистикалық қолжетімділік, тарихи деректер базасының болуы. Мысалы, ауылдық кедейлікті зерттеуде аудандар мен ауылдардың іріктемесі нақты әдіснамамен жүргізілуі тиіс.

Мекемелік зерттеулерде (мысалы, денсаулық сақтау саласы) ауруханалар, клиникалар немесе дәрігерлік амбулаториялар тізімінен іріктеме жүргізіледі. Кеңістіктік шеңберді дұрыс анықтамау нәтижелерді жалпылаудың қателікке ұшырауына алып келеді.

Зерттеу мерзімін таңдау

Уақыт факторы статистикалық зерттеуде шешуші рөл атқарады. Мерзімді таңдауда мыналар есепке алынады: маусымдық ауытқулар (мысалы, егін жинау маусымындағы жұмыспен қамту статистикасы), экономикалық циклдер, мереке күндері мен ерекше оқиғалар. Деректерді жинаудың «нүктелік» (бір күн/аптадағы жағдайды тіркеу) немесе «кезеңдік» (ұзақ уақыт аралығындағы өзгерісті бақылау) болуы зерттеу сипатына байланысты.

Панельдік зерттеулерде бір топ бірліктер ұзақ уақыт бойы қайта-қайта бақыланады. Бұл тәсіл себептік байланыстарды анықтауға мүмкіндік береді.

Статистикалық бақылау: дерек жинаудың кезеңі

Статистикалық бақылау — нысан бірліктерінің сипаттамаларын есеп құжаттарында тіркеудің жүйелі процесі. Бұл кезең теориялық жоспарды нақты деректерге айналдырады.

1

Бірліктерді тіркеу принципі

Статистикалық бақылаудың негізгі міндеті — нысанның әрбір бірлігін белгіленген нысан бойынша есепке алу. Тіркеу кезінде деректердің толықтығы мен дәлдігі қатаң қадағаланады. Бір бірлікті екі рет немесе мүлде тіркемеу — бақылаудың сапасын анықтайтын маңызды шарт.

2

Толықтық пен уақыт факторы

Деректердің толықтығы — жиынтықтың барлық (немесе іріктеп алынған) бірліктері туралы мәлімет жинақталды дегенді білдіреді. Уақыт факторы деректердің «өткірлігін» (актуалдылығын) анықтайды. Тым ерте немесе тым кеш жиналған деректер зерттелетін құбылысты дұрыс бейнелемеуі мүмкін. Мысалы, жылдық өнеркәсіп есебін жылдың басында емес, аяғында жинау дұрыс нәтиже береді.

3

Деректер сапасын бақылау

Бақылау кезінде жинақталған деректерді логикалық тексеруден өткізу маңызды. Ішкі қайшылықтарды, мүмкін емес мәндерді (мысалы, жасы 150 деп тіркелген адам) және бос жауаптарды анықтап, дер кезінде түзету — деректер сапасын қамтамасыз етудің негізгі элементі.

Іріктеме не үшін керек?

Статистикалық зерттеуде барлық бірліктерді бақылаудың — «жаппай бақылаудың» — мүмкін болмайтын немесе тиімсіз жағдайлары жиі кездеседі. Дәл осы жерде іріктеме (выборка) тәсілі шешім ретінде алға шығады.

Жаппай бақылаудың шектеулері

Миллиондаған халқы бар елде барлық тұрғынды сауалнамаға тарту мүмкін емес. Сонымен қатар кейбір зерттеулер объектіні жоятын бақылауды талап етеді — мысалы, жарылатын затты тексеру. Мұндай жағдайларда тек іріктеме тәсілі ғана баламасы жоқ шешім болып табылады.

Ресурстарды үнемдеу

Іріктеме тәсілі уақытты, еңбекті, материалдық және қаржылық ресурстарды айтарлықтай үнемдейді. Мысалы, 100 000 үй шаруашылығының ішінен 3 000-ын ғана зерттеп, бүкіл жиынтық туралы сенімді тұжырым жасауға болады. Бұл статистиканың ең күшті тетіктерінің бірі.

Деректер сапасының артуы

Парадоксалды болып көрінгенмен, іріктеме кейде жаппай бақылауға қарағанда жоғары сапалы деректер береді. Өйткені шағын іріктемеде зерттеушілер әрбір бірлікке жеке назар аударып, дерек жинаудың сапасын қатаң бақылай алады. Жаппай бақылауда бір мезгілде тым көп мәлімет жинауға тырысқанда сапа шығындары артады.

Жаппай емес бақылау мағынасы: «шағын масштабтағы әлем»

Іріктеменің мәні

Іріктеме — бас жиынтықтың «шағын көшірмесі». Дұрыс жүргізілген іріктеме кезінде алынған бөлік бүкіл жиынтықтың негізгі сипаттамаларын сақтайды. Бірліктер кездейсоқ тәртіппен алынады — бұл принцип іріктеменің ғылыми негізі болып табылады. Кездейсоқтық принципі сақталмаса, іріктеме ешқандай ғылыми мағынаға ие болмайды.

Жалпылама қорытынды жасаудың логикасы

Іріктеме тәсілінің мақсаты — бас жиынтықтың параметрлерін (орташа мән, үлес, дисперсия) ең аз санмен бағалау. Математикалық статистика бізге іріктеме деректері негізінде жасалған бағалаудың дәлдігін алдын ала есептеуге мүмкіндік береді. Мысалы, 1 000 адамнан тұратын іріктеме арқылы 1 миллион халықтың ойын ±3% дәлдікпен бағалауға болады — бұл статистиканың ғажайып күші.

Бұл мүмкіндік ықтималдық теориясына, үлкен сандар заңына және орталық шекті теоремаға негізделеді.

Бас жиынтық

Барлық бірліктер мен сипаттар жиынтығы.

Қорытынды

Іріктеме негізінде жалпы тұжырым жасау.



Кездейсоқ іріктеу

Бұрын анықталған әдіспен таңдауды жүргізу.

Іріктеме көлемі мен сенімділік идеясы

Іріктеме зерттеулерінің орталық сұрағы: «Қанша бірлік таңдау керек?» Бұл сұрақтың жауабы бірнеше факторға байланысты.

1

Іріктеме көлемі мен сенімділік

Іріктеме көлемі ұлғайған сайын нәтижелер сенімдірек болады — бұл аксиома. Алайда бұл байланыс сызықты емес: іріктемені 100-ден 200-ге дейін ұлғайту нәтижені айтарлықтай жақсартса, 10 000-нан 10 100-ге дейін ұлғайту іс жүзінде өзгеріс бермейді. Осыдан келіп «оңтайлы іріктеме» ұғымы туындайды.

2

Рұқсат етілген қате деңгейі

Сенімділік интервалы зерттеуші алдын ала белгілейтін «рұқсат етілген қателік шегі». Мысалы, ±5% қате деңгейін қабылдасаңыз, іріктеме аз болады; ±1% дәлдік қажет болса, іріктеме едәуір ұлғаяды. Социологиялық сауалнамаларда әдетте 95% сенімділік деңгейі және ±3% қате шегі қолданылады.

3

Жиынтықтың вариациясы

Бас жиынтықтың ішкі алуандығы (вариациясы) да іріктеме көлемін анықтайды. Егер зерттелетін сипаттама жиынтық бойынша бірдей болса (мысалы, барлығы бір жаста), аз іріктеме жеткілікті. Алуандылық жоғары болған сайын іріктеме ірірек болуы тиіс.

Іріктеме көлемін анықтайтын стандартты формулалар бар: $n = z^2 \cdot p \cdot (1-p) / e^2$, мұнда z — сенімділік деңгейіне сәйкес мән, p — күтілетін үлес, e — рұқсат етілген қате.

«Өкілдік қателер» түсінігі: неге қате шығады?

Іріктеме зерттеулерінің нәтижесі бас жиынтықтың нақты мәнінен ауытқуы мүмкін. Бұл ауытқу — өкілдік қате (репрезентативтілік қатесі) деп аталады. Өкілдік қатені толығымен жою мүмкін емес, бірақ оны басқаруға болады.

Іріктеу алшақтығының табиғаты

Іріктеме деректері мен бас жиынтық арасындағы алшақтық екі түрлі себептен туындайды. Біріншіден, іріктеменің өзіне тән «кездейсоқ ауытқу» болады — бұл табиғи және математикалық жолмен бақыланады. Екіншіден, іріктеу процесінде жіберілген жүйелік қателіктер нәтижені бір бағытқа «қарайту» тудырады. Бұл екі түрлі қате принципиалды жолмен ажыратылады және түзетілу тәсілі де әр түрлі болады.

Кездейсоқ және жүйелік қателердің айырмашылығы

Кездейсоқ қате — статистикалық іріктеменің ажырамас серігі. Оны азайтуға болады, бірақ мүлде жоюға болмайды. Жүйелік қате керісінше — принципі жоқ таңдау немесе методологиялық кемшіліктен туындайды. Жүйелік қате нәтижені жүйелі түрде «дұрыс» мәннен алшақтатады. Мысалы, телефон арқылы жүргізілген сауалнама телефоны жоқ үй шаруашылықтарын автоматты түрде шетте қалдырады — бұл жүйелік ауытқу.

Жүйелік қате: кездейсоқтық принципі бұзылса

Жүйелік қате (систематикалық bias) — статистикалық зерттеудің ең қауіпті жауы. Ол нәтижелерді бір бағытқа ауытқытып, зерттеушінің оны байқамауына ыңғайлы болуы мүмкін.

Іріктеу кезіндегі жүйелік қате

Популяция бірліктерін таңдауда кездейсоқтық принципі бұзылуы ең жиі кездесетін себеп. Мысалы, зерттеуші «қолайлы» респонденттерді таңдаса, «қиын» бірліктерді өткізіп жіберсе, іріктеме жиынтықты дұрыс бейнелемейді. Тарихи мысал: 1936 жылы Literary Digest журналы АҚШ президенттік сайлауын болжаған кезде автомобиль иелері мен телефон абоненттерін ғана сауалнамаға тартқан — бай аудиторияны ғана қамтып, нәтижесі өрескел қате болып шықты.

Жоспарлау кезіндегі аралас факторлар

Зерттеу жоспарлау кезеңінде бақыланбаған үшінші айнымалылар (confounders) жүйелік қате туғызуы мүмкін. Мысалы, білім деңгейі мен кіріс арасындағы байланысты зерттеген кезде «отбасы әлеуметтік-экономикалық жағдайын» бақыламасаңыз, алған нәтижеңіз шындықты бұрмалайды.

Жүйелік қатені анықтау және азайту

Жүйелік қатені анықтау үшін іріктеме сипаттамаларын бас жиынтықтың белгілі параметрлерімен салыстыру қолданылады. Постстратификация (кейінгі қабаттастыру) тәсілі арқылы іріктеменің құрылымын бас жиынтыққа жақындатуға болады.

Кездейсоқ қате: өлшеу-тіркеу процесінің «тербелісі»

Кездейсоқ қате қайдан туындайды?

Кездейсоқ қате зерттеудің бірнеше кезеңінде жинақталуы мүмкін:

- Өлшеу кезінде: сауалнама сұрақтарының екіұштылығы респонденттің жауабын болжанбаған бағытта шайқауы мүмкін
- Тіркеу кезінде: интервьюер деректерді нысанға жазған кезде жіберілген қателер
- Деректерді компьютерге көшіру кезінде: «1» мен «11»-ді шатастыру сияқты типографиялық қателер
- Кодтау кезінде: ашық жауаптарды санаттарға жатқызудағы субъективтілік

Кездейсоқ қатені бақылау

Кездейсоқ қате математикалық тұрғыдан болжанады және іріктеме көлемін ұлғайту арқылы азайтылады. Орталық шекті теорема бойынша іріктеме мөлшері артқан сайын іріктеме орташасы бас жиынтық ортасына ұмтылады.

Практикалық тәсілдер: қос тіркеу (double entry) — деректерді екі рет тәуелсіз енгізіп, нәтижелерді салыстыру; логикалық тексеру программалары арқылы мүмкін емес мәндерді автоматты анықтау; интервьюерлерді оқыту және бақылау.

Кездейсоқ қатені «нөлге» дейін азайту мүмкін емес, бірақ оның деңгейін есептеп, нәтижелерге қосымша ретінде ұсынуға болады — бұл ғылыми адалдықтың белгісі.

Дерек сапасына әсер ететін негізгі тәуекелдер

Статистикалық зерттеуде деректер сапасына қауіп төндіретін бірнеше маңызды тәуекелдер бар. Оларды алдын ала білу және басқару — сапалы зерттеудің шарты.



Іріктемелі жиынтықты қалыптастыру кемшіліктері

Іріктеме шеңберінің (sampling frame) толық еместігі — жиі кездесетін мәселе. Мысалы, халықты тіркеу тізімі ескірген болса, жаңа тұрғындар іріктемеге ену мүмкіндігінен айырылады. Немесе тізім бір бірлікті бірнеше рет тіркесе, сол бірліктің таңдалу ықтималдығы артады. Бұл екі жағдайда да іріктеме нәтижесі бас жиынтықты дұрыс бейнелемейді.



Бірліктерді алмастыру тәуекелі

Талап етілетін бірліктердің орнына «қол жетімді» бірліктерді алу — іріктеменің репрезентативтілігін бұзатын кең тараған тәжірибе. Мысалы, таңдалған адам сауалнамаға қатысудан бас тартса, зерттеуші «оңай» табылатын адаммен алмастырады — бірақ бұл адамдар жиі ұқсас профильде болады, нәтижесінде белгілі топтар жеткіліксіз өкілдікке ие болады.



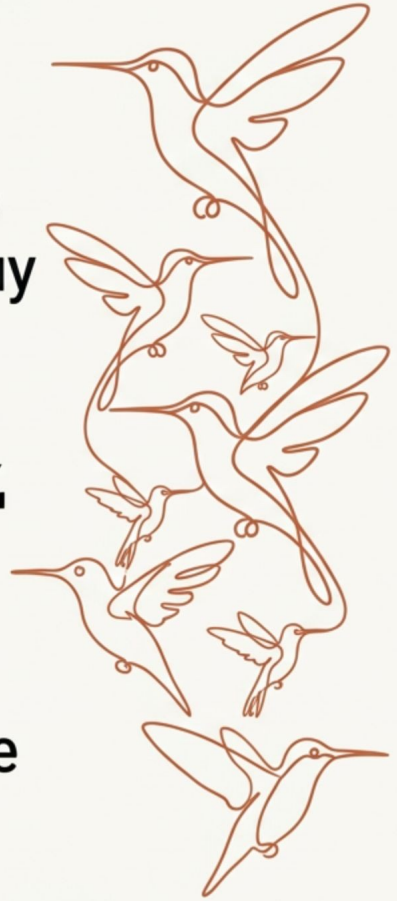
Жауапсыздық (non-response) тәуекелі

Таңдалған бірліктердің жауап бермеуі — деректер сапасының маңызды мәселесі. Жауапсыздық кездейсоқ болса, іріктеме мөлшері азайғанымен нәтиже сенімді болады. Бірақ жауапсыздық жүйелік болса (мысалы, байлар немесе кедейлер жаппай бас тартса), нәтиже мүлде бұрмаланады. Жауапсыздық деңгейін бақылап, оны азайтудың арнайы стратегиялары қолданылады.

Іріктеу түрлері: ықтималды және ықтималды емес

Іріктеу тәсілдерінің екі негізгі тобы бар, олардың арасындағы таңдау зерттеу мақсатына, ресурстарға және нәтиженің жалпыланатын деңгейіне байланысты жүргізіледі.

Ықтималды іріктеу (кездейсоқ). Әр бірліктің таңдалу ықтималдығы белгілі, қате математикалық есептеледі, нәтиже жалпыланады, күрделірек және қымбатырақ.



Ықтималды емес іріктеу. Таңдалу ықтималдығы белгісіз, қате есептелмейді, жалпылау шектеулі, жылдам және арзанырақ.



Таңдалған іріктеу механизмі нәтижедегі қателікті бағалау логикасын түбегейлі анықтайды. Ықтималды іріктеу ғана нәтижелерді бас жиынтыққа математикалық жолмен жалпылауға мүмкіндік береді. Ықтималды емес іріктеу пилоттық зерттеулерде, сапалық зерттеулерде немесе ресурстар шектеулі болғанда қолданылады.

Ықтималды (кездейсоқ) іріктеу: «әркімде шанс бар»

Ықтималды іріктеудің негізгі принципі: бас жиынтықтың әрбір бірлігінің іріктемеге ену ықтималдығы белгілі және нөлден үлкен болады. Бұл принцип іріктемеге ғылыми негіз береді.

1

Қарапайым кездейсоқ іріктеу

Жеребе немесе кездейсоқ сандар кестесі арқылы жүргізіледі. Мысал: 500 студенттен тұратын тізімді нөмірлеп, кездейсоқ сандар генераторы арқылы 50 нөмір таңдау. Әрбір студенттің таңдалу ықтималдығы тең — 50/500 = 10%.

2

Жүйелі іріктеу

Тізімнен әрбір k-ші бірлікті таңдау тәсілі. Мысалы, 1000 бірліктен 100-ін алу үшін k=10, яғни алғашқы бірлікті 1–10 аралығынан кездейсоқ таңдап, одан кейін 10, 20, 30... деп жалғастыру.

3

Қабаттастырылған іріктеу

Бас жиынтықты алдын ала топтарға (қабаттарға) бөліп, әрбір қабаттан бөлек іріктеу жүргізу. Мысалы, халықты жас топтарына бөліп, әр топтан пропорционал санда таңдау — жас топтарының барлығы іріктемеде дұрыс өкілдікке ие болады.

4

Кластерлік іріктеу

Алдымен «кластерлер» (топтар, мысалы ауылдар немесе мектептер) таңдалып, одан кейін таңдалған кластерлердің барлық бірліктері немесе іші іріктеледі. Логистикалық тиімділігі жоғары, бірақ дизайн эффектісі (design effect) есепке алынуы тиіс.

Ықтималды емес іріктеу: жылдам, бірақ қауіп жоғары

Ықтималды емес іріктеу түрлері

Еркін іріктеме

Зерттеуші «қолы жеткен» бірліктерді таңдайды. Жылдам және арзан, бірақ репрезентативтілігі өте төмен. Пилоттық зерттеулерге ғана жарайды.

Мақсатты (әдейі) іріктеме

Зерттеуші өз білімі негізінде «типтік» бірліктерді таңдайды. Сапалық зерттеулерде кеңінен қолданылады, бірақ субъективтілік тәуекелі жоғары.

Квоталық іріктеме

Алдын ала белгіленген квоталар (мысалы, 50% әйел, 50% ер) толтырылады. Жаппай сауалнамаларда қолданылады, бірақ квотаны толтыру тәсілі нәтижені бұрмалауы мүмкін.

Репрезентативтілік мәселесі

Ықтималды емес іріктеудің басты шектеулігі — нәтижелерді бас жиынтыққа математикалық жолмен жалпылаудың мүмкін еместігі. Іріктеме қателігін есептеу үшін ықтималдық теориясы қолданылатын болса, ықтималды емес іріктемеде бұл теория жұмыс істемейді.

Бұл дегеніміз, алынған нәтиженің бас жиынтыққа қаншалықты жақын екендігін математикалық жолмен бағалауға болмайды. Нәтижеге сенімділік тек зерттеушінің тәжірибесі мен пайымына негізделеді.

Алайда ықтималды емес іріктеме барлық жағдайда жарамсыз емес. Гипотезаларды алдын ала тексеруге, сапалық зерттеулерге және ресурстар мен уақыт шектеулі жағдайларда бастапқы мәліметтер жинауға жарайды.

Статистикалық бағалау: Болжамнан шындыққа

Іріктеме деректерін пайдаланып бас жиынтықтың параметрлерін бағалаудың математикалық негіздері

Нүктелік бағалау

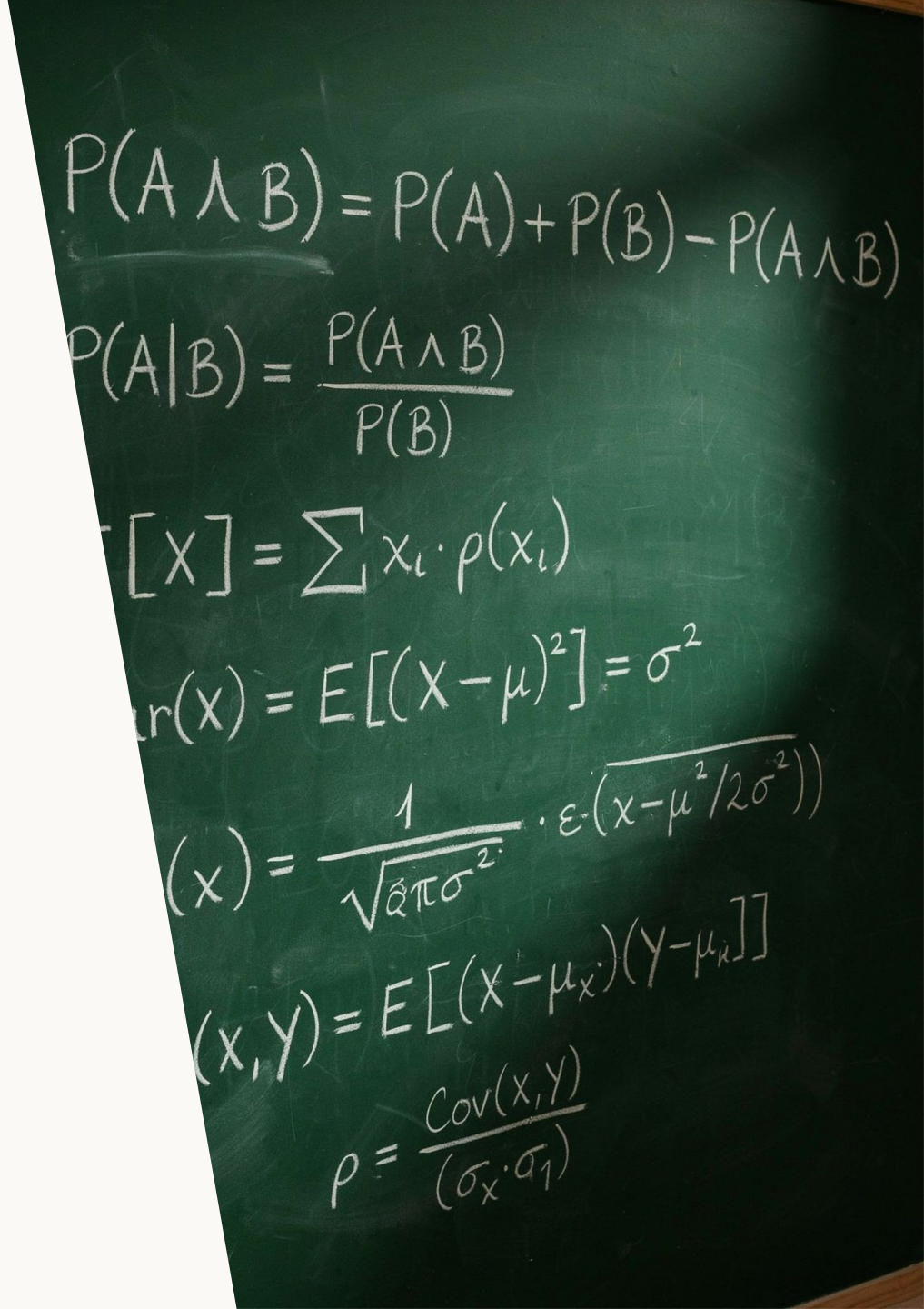
Бас жиынтық параметрін бір нақты санмен бағалау. Мысалы, іріктеме орташасы бас жиынтық орташасының нүктелік бағасы болып табылады.

Интервалдық бағалау

Белгілі сенімділік деңгейімен параметрдің қандай аралыққа түсетінін анықтау. Сенімділік интервалы нәтиженің тиімділігін бейнелейді.

Гипотезаларды тексеру

Теориялық болжамдарды деректер негізінде растау немесе жоққа шығару. Статистикалық маңыздылық (p-мән) шешімнің математикалық негізі.



Статистикалық ықтималдық: Тәжірибе не дейді?

Ықтималдық — болашақ оқиғаның мүмкіндігін сандық өлшейтін математикалық ұғым. Статистикалық бағалаудың бүкіл аппараты ықтималдық теориясына негізделеді.

Жиілікті ықтималдық

Жиілікті ықтималдық тәжірибелік тәсілге негізделеді: егер n рет тәжірибе жасалса және оқиға M рет болса, онда оқиғаның ықтималдығы P(A) = M/n болады. Тәжірибе саны артқан сайын бұл қатынас нақты (теориялық) ықтималдыққа жақындайды.

Мысалы: тиын лақтырудан 1000 рет жүргізілсе және «жазу» 498 рет түссе, P(жазу) ≈ 0.498 ≈ 0.5. Бұл «үлкен сандар заңының» практикалық бейнесі.

Үлкен сандар заңы және оның маңызы

Үлкен сандар заңы (ҮСЗ) — статистиканың іргетасы: тәжірибе саны (n) шексіздікке ұмтылған сайын іріктеме орташасы бас жиынтық орташасына ықтималдықпен жақындайды. Бұл заң іріктеме зерттеулерінің жарамдылығын математикалық тұрғыдан негіздейді.

Теориялық ықтималдық пен нақты деректер арасындағы байланыс: теориялық модельдер (мысалы, қалыпты үлестіру) тәжірибелік деректерді жақсы болжайды — бұл математикалық статистиканың күші. Алайда «барлық модельдер дұрыс емес, бірақ кейбіреулері пайдалы» (Джордж Бокс) — нақты деректер ешқашан теориялық модельге дәл сәйкес келмейді.

Орталық шекті теорема (ОШТ): Кез келген үлестірімнен алынған іріктемелердің орташа мәндері $n \to \infty$ кезінде қалыпты үлестірімге ұмтылады. Бұл теорема статистикалық қорытындының негізі болып табылады.

Оқиғалар тобы: Толық жүйені қалай құру керек?

Ықтималдық теориясында оқиғаларды жүйелі топтастыру негізгі операциялардың бірі болып табылады. Толық жүйені құра білу — статистикалық бағалаудың маңызды элементі.

Толық топ

Толық топ — бір-бірін жоққа шығаратын және барлық мүмкін нәтижелерді қамтитын оқиғалар жиыны. Математикалық шарт: $P(A_1) + P(A_2) + \ldots + P(A_n) = 1$. Мысал: зар лақтырғанда 1, 2, 3, 4, 5, 6 нүктелер — толық топ, олардың ықтималдықтарының қосындысы бірге тең.

Қарама-қарсы оқиғалар

А оқиғасы болмаса, Ā оқиғасы болатын жұп оқиғалар. Мысал: нысанаға тигізу (A) мен тигізбеу (Ā). $P(A) + P(\bar{A}) = 1$, демек $P(\bar{A}) = 1 - P(A)$. Бұл қарапайым қатынас есептеулерді жеңілдетеді: тигізу ықтималдығы 0.7 болса, тигізбеу ықтималдығы автоматты түрде 0.3 болады.

Тәуелсіз оқиғалар

Бір оқиғаның нәтижесі екінші оқиғаның болу ықтималдығына әсер етпейтін жағдай. $P(A \cap B) = P(A) \cdot P(B)$. Мысал: екі тиынды бір мезгілде лақтыру — бірінің нәтижесі екіншісіне тәуелсіз. Тәуелсіздікті болжану — статистикалық модельдердің жиі қабылдайтын жорамалы.

Шартты ықтималдық: Контексттің маңызы

P(A|B) формуласы

Шартты ықтималдық — B оқиғасы орындалған жағдайда A оқиғасының болу ықтималдығы. Формула: P(A|B) = P(A∩B) / P(B), мұнда P(B) > 0.

Бұл формула мынаны айтады: B болды деп білгеннен кейін A-ның ықтималдығы қалай өзгереді? Бұл ақпарат қосылған кезде болжамымызды жаңарту — байесиандық статистиканың негізі.

Мысал: Қандайда бір ауруды анықтайтын тест 95% дәлдікпен жұмыс істейді делік. Алайда ауру сирек кездессе (мысалы, 1000-дан 1), оң нәтиже шыққанда шынымен ауру болу ықтималдығы 95% емес — шартты ықтималдық есебі бойынша бұл мән айтарлықтай төмен болуы мүмкін.

Өмірлік мысал: Ауа райы мен егін өнімділігі

Ауа райының егін өнімділігіне шартты әсерін бағалайық. B оқиғасы = «жазда жауын-шашын жеткілікті болды», A оқиғасы = «егін өнімділігі жоғары болды».

P(A) — жауын-шашын туралы ақпаратсыз өнімділіктің жоғары болу ықтималдығы (тарихи деректер бойынша, мысалы 0.55). P(A|B) — жауын-шашын жеткілікті болғанда өнімділіктің жоғары болу ықтималдығы (мысалы 0.82). Айырмашылық айтарлықтай — бұл ауа райының егін өнімділігіне статистикалық маңызды шартты әсерін дәлелдейді.

Шартты ықтималдық деректерді талдаудағы себептік байланыстарды анықтаудың математикалық тілі болып табылады.

Статистикалық бақылаудың эволюциясы

Қазақстан статистикасының цифрлық трансформациясы — тәуелсіздік жылдарындағы ең маңызды институционалдық өзгерістердің бірі. Бұл өзгеріс тек технологиялық емес, методологиялық та революция болды.

1991–1993: Кеңестік жүйе

Қағаз жүзіндегі есептілік, қолмен есептелетін агрегаттар, орталықтандырылған деректер жинау жүйесі. Деректер айлар, тіпті жылдар бойы жиналып, талданатын.

1994–1999: Алғашқы цифрлануы

FoxPro және MS Access жүйелеріне көшу — алғашқы технологиялық төңкеріс. 1994 жылы БҰҰ стандарттарына негізделген жаңа статистикалық модель қабылданды. Деректер өңдеу уақыты айтарлықтай қысқарды.

2000–2010: Методологиялық реформа

Халықаралық стандарттарға (СНС, ХВҚ ССРД) толық көшу. Шетелдік сарапшылармен бірлескен жобалар іске асырылды. Ұлттық шоттар жүйесі жаңартылды.

2015–2024: Цифрлық экожүйе

Электрондық есептілікке, ашық деректерге және геоақпараттық технологияларға толық көшу кезеңі. Статистиканы халықаралық аудиторияға жеткізу мүмкіндігі артты.



Халықаралық стандарттарға көшу

Қазақстан статистикасының халықаралық стандарттарға сәйкестендірілуі — деректердің ғаламдық салыстырылымдылығын қамтамасыз еткен маңызды қадам.

СНС 2008 жүйесі

Ұлттық шоттар жүйесінің 2008 жылғы нұсқасы — макроэкономикалық статистиканың ғаламдық стандарты. ЖІӨ, инфляция, сауда балансы сияқты көрсеткіштер осы стандарт бойынша есептеледі. Бұл Қазақстан экономикасын дүние жүзіндегі кез келген елмен тікелей салыстыруға мүмкіндік береді. СНС 2008-ге толық көшу Қазақстанды әлемнің жетекші статистикалық мүшелері қатарына қосты.

ХВҚ-ның ССРД стандарты

Халықаралық Валюта Қорының Деректерді Таратудың Арнайы Стандарты (ССРД/SDDS) деректердің ашықтығы мен сапасына қатаң талаптар қояды. Деректерді уақтылы жариялау, метадеректерді ашу, ревизия саясатын мөлдір ету — ССРД-ның негізгі элементтері. Инвесторлар мен халықаралық ұйымдар үшін ССРД сәйкестілігі — сенімділіктің белгісі.

Халықаралық сараптама

Түркия, Италия, Германия статистика агенттіктерімен бірлескен техникалық ынтымақтастық жобалары жүргізілді. Шетелдік сарапшылар Қазақстан статистиктерін оқытты, методологиялық нұсқаулықтар аударылды және бейімделді. Бұл ынтымақтастық халықаралық тәжірибені тез игеруге мүмкіндік берді.

Деректерді дезагрегациялау: Ешкімді шетте қалдырмау

«Орташа» деген ұғым кейде шындықты жасырады. Орташа табыс өскені жалпы халықтың жағдайының жақсарғанын білдіре ме? — Міндетті емес.

Жалпы орташа көрсеткіштің артындағы шындық

Елдің орташа табысы өскенде байлардың кірісі тез өсіп, кедейлердікі баяу өссе де жалпы орташа жоғарылайды. Бірақ бұл жағдайда теңсіздік үлкейген. Осы себептен ТМД елдерінің статистикалық агенттіктері «дезагрегациялау» — деректерді бөлшектеп талдау тәсіліне барған сайын назар аударады.

Деректерді жынысы, жасы, географиялық орны, білім деңгейі, ұлты бойынша бөлу — теңсіздіктің нақты бетпердесін ашудың бірден-бір жолы. «Ешкімді шетте қалдырмау» (Leave No One Behind) — БҰҰ-ның ТДМ жүйесінің негізгі ұстанымы.

2024 жылғы MICS зерттеуі

Мемлекеттерге қатысты бала денсаулығы мен білім беру саласындағы Көп индикаторлы кластерлік зерттеу (MICS) — ЮНИСЕФ үйлестіретін ірі іріктемелі сауалнама. Қазақстанда 2024 жылы жүргізілген MICS зерттеуі мыналарды анықтауға мүмкіндік берді:

- Ауылдық және қалалық балалардың оқуға қолжетімділігіндегі айырмашылық
- Ер балалар мен қыз балалардың денсаулық сақтау қызметін пайдалану деңгейі
- Аймақтар арасындағы балалар кедейлігінің таралуы
- Жас аналардың білімі мен баланың жетілу арасындағы байланыс

Бұл деректер теңсіздікті анықтап, мақсатты саясат шешімдерін негіздеуге қызмет етеді.

Қазақстанның ОДМ көрсеткіштері

Орнықты Даму Мақсаттары (ОДМ) — 2015 жылы БҰҰ-ның барлық мүше мемлекеттері бекіткен 17 жаһандық мақсат. Қазақстан осы мақсаттардың орындалуын ұлттық статистика жүйесі арқылы мониторингтейді.

17 мақсат және 230 индикатор

Қазақстанның ОДМ мониторинг жүйесі 17 мақсатқа сәйкес 230-дан астам индикатордан тұрады. Бұл индикаторлар жыл сайын есептеліп, халықаралық деректер базасына жіберіледі. Ұлттық индикаторлардың бір бөлігі ғаламдық деңгейде стандартталған, ал екінші бөлігі Қазақстанның ерекшелігін есепке алып бейімделген.

Кедейліктің динамикасы

15.3%

2015 жыл

Кедейлік деңгейі

7.4%

2024 жыл

Кедейлік деңгейі

Кедейліктің 2015–2024 жылдар аралығында 15.3%-дан 7.4%-ға дейін қысқаруы — статистикалық бақылаудың мемлекеттік саясатқа тигізетін нақты әсерінің дәлелі. Ұлттық статистикалық платформа арқылы азаматтар, зерттеушілер және саясаткерлер бұл деректерге еркін қолжетімділік алды. Ашық деректер саясаты демократиялық есеп берушіліктің алғышарты болып табылады.

Геоақпараттық технологиялар (ГИС)

Геоақпараттық жүйелер (ГИС) статистикалық деректерге кеңістіктік өлшем қосады. «Деректер картаға түссе, жаңа сұрақтар туындайды» — бұл постулат аймақтық статистиканың тәжірибесімен расталады.

Кеңістіктік талдаудың мүмкіндіктері

ГИС технологиялары аймақтар бойынша өмір сүру сапасын бірнеше өлшемде бір мезгілде бейнелеуге мүмкіндік береді. Мысалы, денсаулық сақтау инфрақұрылымының орналасуы мен халықтың денсаулық көрсеткіштерін қабаттастыру арқылы медициналық қызметке қолжетімділік «ақтаңдақтарын» анықтауға болады. Ауылдық жерлердің тұрмыс деңгейін, инфрақұрылыққа қолжетімділігін және демографиялық өзгерістерін аудан деңгейінде бақылау ел саясатын негіздейді.

14 облыс пен 2 қаланың 1999-2018 жылдардағы динамикасы

Қазақстанның 14 облысы, Алматы және Астана қалаларының 1999–2018 жылдар аралығындағы әлеуметтік-экономикалық динамикасын ГИС арқылы талдау елдің аймақтық теңсіздігін анық суреттеді. Батыс аймақтардың мұнай кірісінен жылдам өсуі, ал кейбір аграрлық аймақтардың баяу дамуы — кеңістіктік деректер сіздің алдыңызда «тірі карта» ретінде ашылады.

Әлеуметтік мәселелерді визуалдау

ГИС деректерін жұртшылыққа ұғынықты картографиялық визуализация арқылы жеткізу — статистикалық мәдениетті дамытудың маңызды элементі. Интерактивті карталар саясаткерлерге, журналистерге және азаматтарға деректерді интуитивті түрде түсінуге мүмкіндік береді. «Бір карта мың сөзден артық» — ГИС тәжірибесі осыны растайды.

Цифрлық трансформация: 2025-2027 жылдар

Қазақстан Статистика агенттігінің 2025–2027 жылдарға арналған стратегиялық жоспары статистикалық экожүйені түбегейлі жаңарту бағдарламасын қамтиды.

85 нақты іс-шара

Статистикалық экожүйені дамытудың 85 нақты іс-шарасы бекітілген. Іс-шаралар үш бағытта топтастырылған: деректер жинаудың цифрлануы, өңдеу процестерін автоматтандыру, және деректерді тарату каналдарын кеңейту. Әрбір іс-шараның орындаушысы, мерзімі және индикаторы белгіленген.

Кабинет респондента

Кәсіпорындар мен ұйымдар статистикалық есептілікті қағаз форматта емес, «Кабинет респондента» порталы арқылы онлайн ұсынады. Бұл тәсіл респонденттер үшін де (уақыт үнемі, қателіктерді автоматты анықтау), агенттік үшін де (деректер сапасы, өңдеу жылдамдығы) тиімдірек. 2027 жылға қарай барлық статистикалық есептілікті қағазсыз форматқа толық көшіру жоспарланған.

ЭЦП және автоматтандырылған дерекқор

Электрондық цифрлық қолтаңба (ЭЦП) деректер берушінің сәйкестігін растайды және жалған есептілікті болдырмайды. Автоматтандырылған деректер базасы әртүрлі көздерден түскен деректерді өзара байланыстырып, жинақтайды. Мемлекеттік деректер базаларының интеграциясы (салық, кеден, зейнетақы) статистиканың жаңа мүмкіндіктерін ашады.

Оқиғалар тобы және шартты ықтималдық

Толық жүйе: теориялық негіз

Толық топ (полная группа событий) — бір мезгілде болатын тәжірибенің барлық мүмкін нәтижелерін қамтитын оқиғалар жиыны. Олардың ықтималдықтарының қосындысы 1-ге тең: $P(A_1) + P(A_2) + \ldots + P(A_n) = 1$. Бұл принцип ықтималдық теориясының «сақталу заңы» іспеттес — ықтималдықтар «жоғалмайды».

Практикалық қолданыс: сауалнаманың жауап нұсқалары толық топ құруы тиіс — яғни барлық мүмкін жауаптарды қамтуы керек. Кез келген нұсқаны «ұмытып кету» деректердің бұрмалануына алып келеді.

Шартты ықтималдық: деректер талдауда

Шартты ықтималдық $P(A|B)$ деректер талдауда өзара байланыстарды ашудың негізгі құралы. B болды деп білгенде A болу ықтималдығы өзгереді — бұл «жаңартылған білім» ұстанымы.

Деректерді талдаудағы мысалдар:

- Жоғары білімі бар адамның жұмыссыз болу ықтималдығы қандай?
- Ауыл тұрғынының қалалық медициналық қызмет пайдалану ықтималдығы?
- Бірінші баласы қыз болған отбасының екінші балаға үміт артуы?

Шартты ықтималдықты байесиандық жаңарту арқылы кезекті деректер түскен сайын болжамды нақтылауға болады — заманауи машиналық оқытудың негізі осында жатыр.

Болашаққа көзқарас: 2027 және одан әрі

Статистика ғылымы мен тәжірибесі жаңа технологиялардың ықпалымен тез трансформацияланып жатыр. Болашақтың статистикасы бүгіннен мүлде өзгеше болады.

Деректер сапасы мен ашықтығы

Мемлекеттік басқарудың тірегі ретіндегі сапалы, ашық деректер маңызы арта береді. Азаматтардың мемлекеттік шешімдерді бақылауы, ғалымдардың деректерге еркін қолжетімділігі, журналистердің деректерге негізделген журналистикасы — барлығы ашық деректер саясатына тәуелді. «Деректер — жаңа мұнай» деген метафора дауларды туғызғанмен, мемлекеттік ашық деректердің бұл аналогиядан бір айырмашылығы бар: ол тауарсоқ тауар, бірін пайдалану екіншісінің пайдалануына кедергі болмайды.

Жасанды интеллект және Big Data

ЖИ алгоритмдері дәстүрлі іріктемені толықтырады немесе алмастырады. Мобильді операторлардың деректері, транзакциялық деректер, әлеуметтік медиа мониторингі — бұлар «үлкен деректер» арқылы дәстүрлі сауалнамалардан жылдамырақ, арзанырақ және жиі ақпарат береді. Алайда жаңа деректер көздерімен бірге жаңа методологиялық мәселелер де туындайды: репрезентативтілік, жеке ақпараттың қорғалуы, алгоритмдік бейімділіктер.

Статистикалық мәдениет

Деректерге негізделген шешім қабылдау дәуірінде статистикалық сауаттылық жалпы мәдениеттің бөлігіне айналуда. Азаматтар диаграммаларды, пайыздарды, қателік шектерін дұрыс оқи алуы тиіс. Мектеп бағдарламасынан университетке дейін, мемлекеттік шенеуніктерден журналистерге дейін — статистикалық мәдениетті жаппай дамыту болашақтың басымдығы.



Іріктеме және статистикалық бағалау: Деректен шешімге дейін

Статистикалық зерттеудің соңғы мақсаты — белгісіздікті азайтып, негізделген шешімге негіз беру

Деректер жинау

Ғылыми негізделген іріктеу және сапалы бақылау

Статистикалық бағалау

Нүктелік және интервалдық бағалау арқылы болжау

Қорытынды жасау

Бас жиынтық туралы сенімді тұжырымдар

Шешім қабылдау

Деректерге негізделген мемлекеттік және жеке шешімдер

Жоспар және бағдарлама: өлшемнің ережесін бекіту

Статистикалық зерттеудің ғылыми негізі — оның жоспарында. Жоспарсыз жүргізілген зерттеу қажетсіз шығын мен сенімсіз нәтижеге алып келеді. Кәсіби статистикада зерттеудің жоспары мен бағдарламасы зерттеуден бұрын жарияланады немесе тіркеледі — бұл ғылыми адалдықтың белгісі.

Мақсатты айқындаудың маңызы

Зерттеу мақсаты нақты, өлшеулі, қолжетімді, өзекті және уақыт шеңберімен шектелген (SMART) болуы тиіс. Жалпы тұжырымдалған мақсат («халықтың жағдайын зерттеу») зерттеушіге бағдар бермейді. Нақты мақсат («Алматы қаласының 2024 жылы 25–40 жас аралығындағы жалдамалы жұмыскерлердің ай сайынғы таза кірісін анықтау») зерттеудің барлық параметрлерін өзінен өзі анықтайды.

Мақсатты айқындау — барлық кейінгі методологиялық шешімдердің негізі. Ол деректер жинаудың тәсілін, іріктеме көлемін, талдау стратегиясын, тіпті нәтижелерді ұсыну форматын да анықтайды.

Өлшем бірліктері мен бақылау бағдарламасы

Өлшем бірліктерін алдын ала бекіту операционалдау деп аталады. Мысалы, «табыс» деген ұғымды анықтамай зерттеу жүргізу мүмкін емес: жалпы кіріс пе, таза кіріс пе, жан басына шаққандағы кіріс пе, барлық кіріс көздері ескеріле ме? Осы сұрақтардың жауабы алдын ала бекітілмесе, зерттеушілер әртүрлі өлшейді, нәтижелер салыстырылмайды.

Бақылау бағдарламасы — жинақталатын деректердің толық тізімі. Бұл тізімде артық та, кем де сұрақ болмауы тиіс. Артық сұрақтар респондент жүктемесін арттырып, жауап сапасын төмендетеді.

Уақыт пен орын: бақылауды дұрыс қою

Статистикалық бақылаудың кеңістік пен уақыт шеңберін дәл анықтау зерттеуге мағыналы контекст береді және нәтижелердің дұрыс интерпретациясын қамтамасыз етеді.



Географиялық шеңберді анықтау

Зерттеудің географиялық шеңбері (жалпы ұлттық, өңірлік, аудандық немесе қалалық) нәтижелердің қолданылу аясын анықтайды. Мысалы, Шымкент қаласында жүргізілген зерттеу нәтижелерін бүкіл Оңтүстік Қазақстанға жалпылауға болмайды — географиялық өкілдік жоқ. Деректердің нақтылығын қамтамасыз ету үшін географиялық шекара нақты анықталуы тиіс: тіркелу мекенжайы бойынша ма, нақты тұрғылықты жер бойынша ма?



Мерзімдік факторлар

Зерттеудің уақыт шеңбері маусымдық ауытқуларды, экономикалық циклдерді және ерекше оқиғаларды есепке алуы тиіс. Жұмыспен қамтуды жазда зерттесе, маусымдық жұмыстар нәтижені жоғарылатып жібереді. Халықтың кірісін мереке маусымында зерттесе, нәтиже жылдық орташадан жоғары болады. Осы себептен зерттеу мерзімін таңдауда жылдың «бейтарап» кезеңдері жиі қолданылады.

Жаппай емес бақылау мағынасы: шағын масштабтағы әлем

Іріктемелі зерттеу тәсілінің мәні — шағын бірлікпен үлкен жиынтық туралы сенімді тұжырым жасай алу. Бұл статистика ғылымының ең зор жетістіктерінің бірі.

Бүкіл популяция туралы қорытынды жасаудың логикасы

Іріктеме арқылы бас жиынтық туралы қорытынды жасаудың математикалық негізі — ықтималдық теориясы мен орталық шекті теорема. Дұрыс таңдалған іріктеме бас жиынтықтың «миниатюрасы» іспеттес: оның негізгі сипаттамалары іріктемеде пропорционал бейнеленеді.

Мысал: Қазақстанның 19 миллион халқы туралы ақпаратты жинаудың орнына дұрыс жоспарланған 5 000 үй шаруашылығын зерттеп, елдің тұтыну үлгілері туралы ±2% дәлдікпен қорытынды жасауға болады. Бұл ресурс үнемінің масштабын елестетіңіз.

Статистикалық бағалаудың экономикалық тиімділігі

Іріктемелі зерттеудің экономикалық тиімділігі бірнеше деңгейде анық байқалады:

- **Тікелей шығын:** Бас жиынтықтың 1%-ін зерттеу жаппай бақылаудың 1% шығынын талап етпейді — тіркеме шығындары (overhead) айтарлықтай аз болады
- **Уақыт:** Іріктемелі зерттеу апталар мен айлар өткізіледі, жаппай бақылау жылдарды алады
- **Сапа:** Аз бірліктерге жоғары сапалы назар аударуға болады
- **Икемділік:** Іріктеме жылдам жүргізіліп, саясаткерлерге дер кезінде ақпарат береді

Осы факторлар жиыны іріктемені заманауи статистиканың негізгі тетігіне айналдырады.

Өкілдік қателер түсінігі: неге қате шығады?

Репрезентативтілік — іріктеменің бас жиынтықтың негізгі сипаттамаларын дұрыс бейнелеу қасиеті. Бұл іріктемелі зерттеудің мақсаты да, оның ең осал нүктесі де.

Репрезентативтілік дегеніміз не?

Іріктеме бас жиынтықты толық сипаттайды дегеніміз: іріктемедегі әйелдер үлесі бас жиынтықтағы үлеске жуық, іріктемедегі жас топтарының бөліну бас жиынтықтың бөлінуіне сәйкес, географиялық өкілдік тепе-тең. Егер іріктеме тек белгілі топты «жинаса», ол бас жиынтықтың дұрыс бейнесін бермейді.

Жүйелі ауытқулар

Жүйелі ауытқу (systematic bias) — іріктеме нәтижесін бас жиынтықтың нақты мәнінен бір бағытқа үнемі ауытқытатын фактор. Бұл ауытқу іріктеме көлемін ұлғайтумен жойылмайды — ол жүйелі, демек мыңдаған бірлікпен зерттесе де сол бағытқа «жалпайып» кетеді. Жүйелі қатені тек методологиялық жолмен — іріктеу дизайнын түзету арқылы жою мүмкін.

Кездейсоқ ауытқулар

Кездейсоқ ауытқу — іріктемеге тән табиғи «тербеліс». Ол бас жиынтықтың нақты мәнінің айналасында хаотикалық түрде таралады. Іріктеме мөлшері артқан сайын кездейсоқ ауытқудың амплитудасы азаяды. Сенімділік интервалы осы кездейсоқ ауытқудың мөлшерін есептейді.

Жақсы зерттеу жүйелі қатені минимумға жеткізіп, кездейсоқ қатені математикалық жолмен бағалайды. Екеуін де есепке алу — кәсіби статистиканың белгісі.

Кездейсоқ қате: өлшеу-тіркеу процесінің тербелісі

Кездейсоқ қате статистикалық зерттеудің ажырамас серігі. Оны толығымен жою мүмкін емес, бірақ оның мөлшерін бақылауға және азайтуға болады — бұл статистикалық бағалаудың негізгі міндеттерінің бірі.

Табиғи ауытқулардың себептері

Кездейсоқ қате бірнеше кезеңде жинақталады:

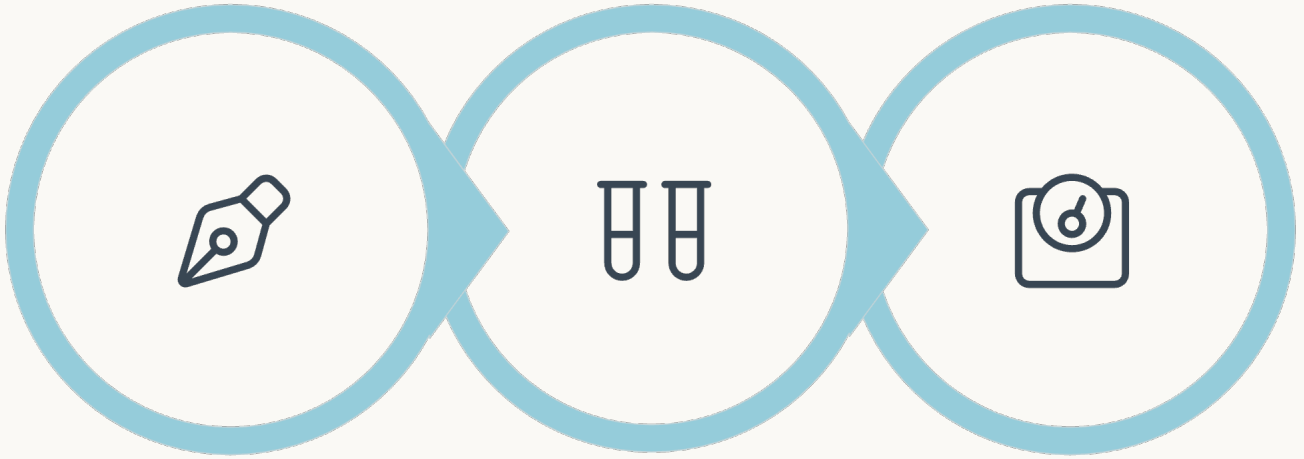
- **Іріктеу кезіндегі:** Кездейсоқ таңдалған бірліктер бас жиынтықты дәл бейнелемеуі мүмкін — бұл «іріктеу вариациясы» деп аталады
- **Өлшеу кезіндегі:** Респондент немесе интервьюер жіберетін кез келген белгісіз қате кездейсоқ болады
- **Тіркеу кезіндегі:** Деректерді нысанға жазу немесе компьютерге енгізудегі типографиялық қателер
- **Өңдеу кезіндегі:** Кодтау мен агрегациялаудағы алгоритмдік ауытқулар

Математикалық азайту әдістері

Статистика кездейсоқ қатені азайтудың бірнеше математикалық тетігін ұсынады:

- **Іріктемені ұлғайту:** n→∞ болғанда іріктеме орташасы бас жиынтық орташасына ұмтылады (үлкен сандар заңы)
- **Стратификация:** Іріктеменің ішкі жүйелілігін арттыру кездейсоқ ауытқуды азайтады
- **Бірнеше рет өлшеу:** Бір нысанды бірнеше рет немесе бірнеше тәсілмен өлшеп, орташа алу
- **Сенімділік интервалы:** Қатенің мөлшерін ашық хабарлап, нәтижеге белгісіздікті қосу

Соңғы тәсіл — белгісіздікті жасырмай, оны ашық хабарлау — ғылыми статистиканың этикалық ұстанымы. «Қате ±3%» деген жазу нәтижені кемітпейді, керісінше оның сенімділігін арттырады.



Дизайнды жақсарту

Іріктемені ұлғайту

Қатені өлшеп хабарлау